UNIVERSITY OF LEEDS

Modelling Pedestrian-Vehicle Interaction for Autonomous Vehicles: From Computational Rationality to Adversarial Testing

Yueyang Wang

University of Leeds

Institute for Transport Studies

Submitted in accordance with the requirements for the degree of
Doctor of Philosophy

January, 2026

Abstract

As autonomous vehicles (AVs) move towards real-world deployment, their ability to interact safely and naturally with human road users has become an important challenge to address. Meeting this challenge requires accurate models of road user behaviour. Existing modelling approaches are typically either mechanistic, which offer interpretability but struggle to generalise across complex environments, or data-driven machine learning (ML) models, which provide strong predictive ability but lack transparency and overlook the mechanisms that generate human behaviour, with performance strongly dependent on data availability and representativeness. This thesis adopts the computational rationality framework, viewing human behaviour as boundedly optimal decision-making under human perceptual and motor constraints, and uses reinforcement learning (RL) to model human-like interactions in road-crossing scenarios.

The first study develops an RL model that explains pedestrian crossing decisions as boundedly optimal behaviour under noisy visual perception. The model captures the observed dependencies of crossing decisions on vehicle time-to-arrival (TTA) and speed, showing that such behaviour can be interpreted as a rational adaptation to perceptual uncertainty.

The second study extends the framework to include motor constraints by integrating a biomechanical representation of walking into the RL problem. This approach enables continuous control of crossing actions and reproduces human-like trade-offs between time pressure and walking effort.

To extend the framework beyond controlled experiments and capture real-world interactions, the third study introduces a multi-agent RL framework in which both pedestrians and vehicles are modelled as adaptive agents interacting under perceptual and motor constraints. The resulting interactions reproduce key behavioural phenomena such as yielding and human-like smoother speed adjustments, closely aligning with patterns observed in a naturalistic dataset.

Finally, an application study evaluates the use of a human-like pedestrian model for AV testing. In closed-loop simulation with a full AV control stack, the interaction outcomes of the rule-based CARLA pedestrian and the behaviourally realistic model are compared. The results show that the human-like model produces more realistic gap acceptance patterns and smoother vehicle

decelerations, and that the generated adversarial scenarios can effectively tune AV braking behaviour to achieve safer and more efficient pedestrian-vehicle interactions.

Together, these studies represent the first application of the computational rationality framework to model and explain pedestrian-vehicle interaction. They show that boundedly optimal decision-making under perceptual and motor constraints can reproduce human interactive behaviour. Moreover, the thesis highlights the effectiveness of human-like behavioural models for AV testing, demonstrating their value in enhancing the realism of simulation-based evaluation and control design.

Intellectual Property Statement

The candidate confirms that the work submitted is his own, except where work which has formed part of jointly authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

Four publications or manuscripts have been produced from research that was undertaken as part of this thesis. Each publication or manuscript is listed below with a full reference and details of its location within this thesis. The work in Chapters 2 to 5 of the thesis has appeared in publications or manuscripts as follows:

The work in Chapter 2 of the thesis has appeared in publication as follows:

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

The candidate developed the main idea for this work, under the guidance of Gustav Markkula, and Aravinda Srinivasan. The empirical data used to validate the assumption and model was open sourced. The candidate performed the modelling work, data analysis, and wrote the manuscript. The results were reviewed by Gustav Markkula, and Aravinda Srinivasan. The manuscript was improved by comments from all the co-authors.

The work in Chapter 3 of the thesis has appeared in publication as follows:

Wang, Y., Srinivasan, A.R., Lee, Y.M., & and Markkula, G. (2025). *Modeling Pedestrian Crossing Behavior: A Reinforcement Learning Approach With Sensory Motor Constraints.* IEEE

Transactions on Intelligent Transportation Systems, vol. 26, no. 10, pp. 16454-16465.

The candidate developed the main idea for this work, under the guidance of Gustav Markkula, and Aravinda Srinivasan. The design of the experiment was led by Yee Mun LEE. The experiment was conducted by the candidate and Hao Qin. The candidate performed the analysis and modelling and wrote the paper. The results were reviewed by Gustav Markkula, and Aravinda Srinivasan. The manuscript was improved by comments from all the co-authors.

The work in Chapter 4 of the thesis has been completed in a manuscript and submitted for possible journal publication, currently under review:

Wang, Y., Dogar, M., & and Markkula, G. (2025) *Realistic pedestrian-driver interaction modelling using multi-agent RL with human perceptual-motor constraints.* arXiv preprint arXiv:2510.27383.

The candidate developed the main idea for this work, under the guidance of Gustav Markkula, and Mehmet Dogar. The empirical data used to validate the assumption and model was provided by Gustav Markkula. The candidate performed the analysis and modelling and wrote the paper. The results were reviewed by Gustav Markkula, and Mehmet Dogar. The manuscript was improved by comments from Gustav Markkula, and Mehmet Dogar.

The work in Chapter 5 of the thesis has been completed in a manuscript intended for submission for possible journal publication:

Wang, Y., Dogar, M., & and Markkula, G. (2025). *Realistic adversarial scenario generation via human-like pedestrian model for autonomous vehicle control parameter optimization.*

The candidate developed the main idea for this work, under the guidance of Gustav Markkula. The pedestrian model in this work were provided by Gustav Markkula. The candidate performed software integration, the modelling work and wrote the manuscript. The results were reviewed by Gustav Markkula, and Mehmet Dogar. The manuscript was improved by comments from Gustav Markkula, and Mehmet Dogar.

Acknowledgements

As this thesis draws to a close, I find that writing these acknowledgements is a task far more challenging than the research itself. Reflecting on the past few years, I am deeply mindful of many people who have guided and supported me along the way. Without their help and encouragement, I would not have found the confidence needed to reach this milestone.

First and foremost, I would like to express my sincere gratitude to my main supervisor, Prof. Gustav Markkula. Our academic connection began during an online lecture in my Master's studies, where I was deeply impressed by your passion for science and dedication to teaching. This experience played an important role in shaping my decision to pursue a PhD, and I feel very fortunate to have later become your doctoral student. Throughout my doctoral studies, you guided me patiently through the complexities of research, always offering support, encouragement, and thoughtful advice. I am also grateful for the many opportunities you provided for academic exchange and collaboration, which greatly broadened my perspective. Your mentorship has had a huge influence on my development as a researcher, and I am sincerely thankful for your guidance throughout this journey.

I would also like to thank my co-supervisors, Prof. Mehmet Dogar and Dr Aravinda Srinivasan, for their valuable insights and constructive feedback. Our discussions have been important in refining this work. In addition, I am grateful to Prof. Antti Oulasvirta and Dr. Jussi P. P. Jokinen. Our collaboration and insightful exchanges formed the foundation of my doctoral research, and I have greatly benefited from your guidance and expertise.

I would like to extend special thanks to Dr. Zihao An and Dr. Kai Tian. From the very beginning of my PhD, you have been important role models to me. Your generosity in sharing your experiences, together with your practical advice and encouragement, made my doctoral journey far more manageable and rewarding.

I am also grateful to my friends at the Institute for Transport Studies (ITS): Songfeng Shen, Jinyi Pan, Siyi Lin, Shuwei Lin, Ruifan Tang, Mengyuan Xu, Victor Cantillo Garcia, Vanessa Ternes, Nadia Voigt da Mata, and Dr. Maximiliano Lizana. Thank you for the companionship, support, and many memorable moments that made my time at ITS both enjoyable and meaningful.

My thanks also go to my friends and colleagues in the Human Factors and Safety Group, including Dr. Chen Peng, Dr. Yue Yang, Dr. Yee Mun Lee, Dr. Ibrahim Ozturk, Dr. Athanasios Tzigieras, Prof. Natasha Merat, Yee Thung Lee, Hao Qin and the entire simulation team. It has been a privilege to work alongside such dedicated and inspiring researchers.

I am deeply grateful to my family for their financial and emotional support, which allowed me to pursue my academic path with confidence and independence.

Finally, I would like to thank my wife, Lanjun Zu, for her constant support, understanding, and patience over the past few years. Having you by my side has meant more to me than words can express. I am also thankful to my cat, Dazhu, and my dog, Xiaohua, whose presence and companionship have brought daily comfort and joy throughout this journey.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

Automated vehicles (AVs) have gained substantial attention from industry, academia, and policymakers for their potential to improve road safety and mobility. By reducing human errors such as distraction and fatigue, AVs promise safer and more efficient transport, particularly benefiting elderly and disabled individuals who face mobility challenges (Van Brummelen et al., 2018). Yet their deployment in real-world traffic remains challenging. A key requirement is reliable operation in mixed traffic, where AVs must anticipate and respond appropriately to the intentions and actions of other road users (Camara et al., 2020). Achieving this depends on successful human-machine interaction, not only between a vehicle and its driver or passengers but also between vehicles and surrounding human drivers, pedestrians, and cyclists (Koopman and Wagner, 2017). Understanding how human road users perceive, decide, and react to one another is therefore fundamental to predicting behaviour and ensuring that AVs can interact safely and smoothly in complex environments (Domeyer et al., 2020).

To develop and validate such interaction capabilities, AVs must be evaluated in controlled yet realistic environments. Direct testing on public roads is costly, time-consuming, and potentially hazardous, especially in rare or safety-critical scenarios. Consequently, simulation-based testing has become a practical and scalable complement to real-world testing (Huang et al., 2016). Compared with real-world testing, simulation enables controlled and repeatable experiments across diverse and potentially dangerous situations (Joisher et al., 2019; Sawitzky et al., 2023). For these reasons, simulation is now recognised as a fundamental tool in the design and assess-

ment of AV control strategies (Wang et al., 2024). For these simulations to support reliable design and safety evaluation, the behaviour of simulated agents must be realistic.

In many simulation environments used for AV evaluation, road user agents remain simplistic. A common choice is scripted agents that follow fixed heuristics for speed, gap acceptance, and yielding; these agents do not sense the vehicle state with uncertainty, adapt their decisions to changing context, or exhibit individual differences, all of which are characteristic of real interactions (Helbing and Molnar, 1995; Chen et al., 2018; Dosovitskiy et al., 2020; Tian et al., 2025). Another common choice is imitation-style agents that learn trajectories from data; while these agents can generate plausible motion in familiar settings, they often fail to adapt effectively to the vehicle in unfamiliar scenarios, leading to accumulating errors over time (Alahi et al., 2016; Codevilla et al., 2019; Lu et al., 2023). Both types can generate interactions that either lack behavioural richness or exhibit unrealistic aggressiveness, thereby distorting how AVs perceive risk and how to react. Ensuring the validity of simulation-based testing requires pedestrian models that capture the richness and realism of human behaviour, allowing AVs to be evaluated under interactions that closely resemble those in the real world.

Together, these challenges illustrate the need for realistic and interpretable models of human behaviour. Such models are not only critical for enabling AVs to understand and predict human actions but also for ensuring that simulation-based testing environments accurately represent real-world interactions.

Among all road users, pedestrians are the most vulnerable, accounting for nearly one quarter of all road fatalities (361/1464) in Great Britain (Department for Transport, 2021). In contrast to drivers, whose behaviour is constrained by vehicle dynamics and traffic regulations, pedestrians have more freedom and fewer physical or regulatory restrictions. This flexibility increases the variability of their decisions and makes behaviour prediction particularly challenging (Jain et al., 2014; Petzoldt, 2014; Mako, 2015; Asaithambi et al., 2016; Tian et al., 2022).

Computational models in the pedestrian behaviour modelling field can be broadly categorised into mechanistic modelling and data-driven machine learning (ML) approaches. Mechanistic models (Fajen and Warren, 2003; Turnwald et al., 2016; Yang et al., 2020; Pekkanen et al., 2022) use psychologically and physically grounded principles, such as social force or perceptual-decision processes, to explain underlying behavioural mechanisms, but are difficult to expand to complex scenarios (Markkula et al., 2023). In contrast, data-driven ML models (Deo and

Trivedi, 2018; Ye et al., 2021; Yang et al., 2025; Divya et al., 2025) focus on trajectory prediction and, when trained on large datasets, can reproduce realistic motion patterns across diverse scenarios. However, their performance strongly depends on data availability, and they often lack interpretability. This thesis aims to bridge the gap between these two paradigms by modelling pedestrian behaviour during interactions with vehicles through computational rationality, an approach that uses reinforcement learning (RL) as a machine learning method to learn boundedly optimal behaviour policies while mechanistically modelling human perceptual and motor constraints. The resulting model aims to generate and explain human-like behaviour, contributing to a deeper understanding of pedestrian-vehicle interaction and advancing the development of behaviourally realistic models of human road users.

The research presented in this thesis consists of two branches, both grounded in the COMMOTIONS modelling framework, which models pedestrian-vehicle interaction through multiple human decision-making mechanisms (Markkula et al., 2023). The first branch focuses on modelling, combining RL with mechanistic principles to model human road user behaviour. The second branch focuses on testing and validation, applying a COMMOTIONS pedestrian model in simulation-based evaluation of AV control. The modelling branch constitutes the main body of the doctoral research, while the testing branch represents a complementary project conducted during the final year.

Following this overview, the remainder of this chapter provides the broader context for the research. Section 1.2 reviews literature relevant to human road user behaviour modelling and AV evaluation. Section 1.3 identifies the key research gaps that motivate this thesis. Section 1.4 then sets out the specific research objectives, and Section 1.5 concludes the chapter by outlining the structure of the thesis and summarising the four studies that form its core.

1.2 Literature review

This literature review introduces research within existing approaches to modelling pedestrian and road user behaviour. It begins with mechanistic modelling approaches that represent human decision-making through explicit behavioural principles, followed by data-driven machine learning methods that prioritise predictive performance. It also reviews RL approaches to road user modelling, as well as cognitively grounded frameworks that explain behaviour through perceptual and motor processes. Simulation-based approaches used in the evaluation of autonomous

vehicles are also discussed.

1.2.1 Mechanistic modelling

Mechanistic models aim to capture pedestrian behaviour through interpretable physical or psychological mechanisms. The following subsections review representative approaches, from simple rule-based models to more complex cognitive models.

Social force models

The social force model represents pedestrian motion as the resultant of physical forces driving goal-oriented movement and repulsion from obstacles and other pedestrians (Helbing and Molnar, 1995). It has proven effective in simulating pedestrian flow and evacuation (Hou et al., 2014; Zhou et al., 2019), and has also been extended to model pedestrian-vehicle interactions such as pedestrian movement at crosswalks (Zeng et al., 2014; Yang et al., 2020). However, the model is fundamentally designed for flow-level simulation, where pedestrians are treated as particles reacting to forces rather than as agents that perceive their environment and make cognitive decisions. Because it lacks an explicit representation of how pedestrians perceive approaching vehicles and evaluate risk, it does not reflect the subtle behavioural patterns observed in microscopic pedestrian-vehicle interactions (Liu et al., 2017; Prédhumeau et al., 2022).

Models emphasising perceptual cues

Perceptual cue-based models assume that road users adjust their movements based on visually available information. Early work reproduced the human route selection pattern in a simple scene by modelling the angular acceleration as a function of goal, obstacle angle and distance according to the pedestrian's visual angle (Fajen and Warren, 2003). Later, Tian et al. (2022) developed a gap acceptance model based on visual looming, which is defined as the expansion rate of the object on the observer's retina (DeLucia, 2015). This study explained the speed-induced unsafe crossing behaviour from the perspective of the human visual perception mechanism, to model gap acceptance frequency as a function of crossing situation. Building on the same visual cues, Tian et al. (2023) extended to more complex continuous traffic flow and a crossing initiation timing model was proposed, which fit experimental data well.

Together, these studies highlight the importance of perceptual information in capturing fundamental aspects of pedestrian decision-making. However, perception-based models do not

capture how motor control and action execution shape crossing behaviour.

Game theoretical models

Game theory provides a framework for analysing interactions between agents whose decisions depend on each other's actions (Von Neumann and Morgenstern, 2007). In road user behaviour, it has been applied to describe how individuals negotiate shared space, such as decisions to yield or proceed (Elvik, 2014). Early models focused on high-level choices like speed or right-of-way, offering insights into strategic reasoning but not continuous trajectories (Hoogendoorn and HL Bovy, 2003; Turnwald et al., 2014; Ma et al., 2017; Johora and Müller, 2018).

Recent work has extended this framework to pedestrian-vehicle interactions, particularly involving AVs. For example, Fox et al. (2018) modelled crossing behaviour as a sequential 'chicken game', showing that pedestrians may exploit the AV's programmed tendency to yield. Camara et al. (2021) and Kalantari et al. (2022) validated these dynamics through virtual reality and naturalistic experiments, demonstrating that strategic reasoning can reproduce key patterns of pedestrian-vehicle interaction.

More recently, Dang et al. (2025) proposed a dynamic game-theoretic model incorporating bounded rationality, relaxing the assumption of perfectly rational agents. While this represents progress towards behavioural realism, several limitations remain. The model assumes perfect sensory information and unconstrained motor execution, and restricts pedestrian behaviour to a discrete set of walking speeds, without continuous control over speed or direction. These simplifications largely stem from the computational challenges of solving game-theoretic formulations in high-dimensional continuous spaces.

Overall, game-theoretic models effectively capture higher-level reasoning and mutual anticipation between agents (Camara et al., 2020), but their simplified perceptual and motor assumptions limit their ability to represent the continuous and uncertain dynamics characteristic of real-world interactions.

Evidence accumulation models

Evidence accumulation models, often formalised as drift-diffusion processes, describe decision-making as the noisy integration of sensory information until a threshold is reached (Ratcliff et al., 2016). Originating in psychology, this framework has been widely used to explain re-

action times and choice variability, and it later inspired applications in interactive road user behaviour (Ratcliff and Strayer, 2014; Markkula et al., 2016; Svärd et al., 2017).

Building on this foundation, Markkula et al. (2018a) applied the evidence accumulation principle to vehicle steering, suggesting that human control consists of discrete 'motor primitives' — short bursts of muscle activation that reflect intermittent adjustments (Giszter, 2015). This idea provided a bridge between perceptual decision-making and motor execution, motivating its extension to pedestrian behaviour. Within this line of work, Giles et al. (2019) demonstrated that pedestrian crossing decisions can be described by a single evidence accumulation process with a dynamic drift rate. Later, Pekkanen et al. (2022) expanded the Variable Drift Diffusion Model (VDDM) by incorporating perceptual inputs such as time-to-arrival (TTA), distance, acceleration and external human-machine interface (eHMI) cues, thereby enhancing its generalisability. The model reproduced qualitative patterns of crossing behaviour but struggled to fit quantitative data across individuals, potentially due to its lack of an explicit representation of variability in perception and motor control.

These developments illustrate how evidence accumulation can unify perception, decision, and action within a single cognitive mechanism. However, most implementations still assume idealised sensory processing and simplified reasoning. To address this gap, Markkula et al. (2023) proposed the *COMMOTIONS* modelling framework, which provides a unified description of how perceptual, cognitive, and motor processes interact during road user decision-making. Noisy visual input is transformed into distance and speed estimates via Bayesian perceptual filtering; these estimates drive affordance-based value computations for both agents; and evidence accumulation determines when an action threshold is reached. By combining these multiple cognitive theories together, the framework reproduces human-like pedestrian-driver interaction patterns.

In COMMOTIONS, a key component is the noisy sensory input. The agent estimates the other road user's distance from a noisy visual angle, which produces distance-dependent uncertainty in position estimation. These noisy observations are then updated through Bayesian perceptual filtering, allowing the agent to form internal beliefs about the other user's speed and future motion. These perceptual and filtering mechanisms are essential for generating human-like behaviour and are also reused in the models developed in this thesis.

Although COMMOTIONS provides a richer cognitive account than earlier mechanistic models,

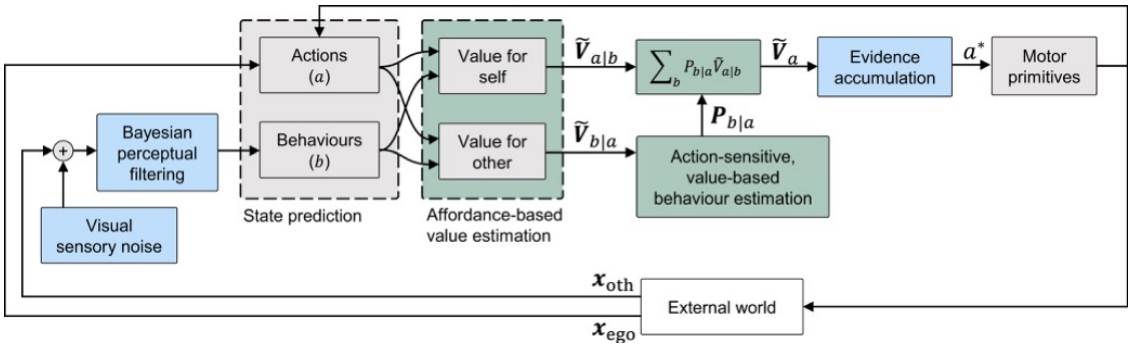

Figure 1.1: Overview of the COMMOTIONS modelling framework (Markkula et al., 2023). Reproduced from Markkula et al. (2023), licensed under CC BY 4.0.

its scope remains constrained. The original implementation focused on a highly simplified interaction scenario involving two agents moving longitudinally along straight, intersecting paths, without lateral movement or broader traffic context. Markkula et al. (2023) themselves note that extending the framework to more complex, realistic environments would be difficult due to the increasing computational and conceptual complexity of integrating multiple cognitive mechanisms.

1.2.2 Data-driven ML models

With increasing computing power and data availability, data-driven models have become a dominant approach to pedestrian behaviour modelling. In this context, data-driven ML models are defined as those that directly learn mappings from input to output using ground-truth data, typically through neural networks trained on observed motion histories, without relying on handcrafted rules or explicit behavioural assumptions. ML-based interaction modelling is commonly formulated as a time-series forecasting task, where the model predicts future trajectories over a fixed horizon based on past motion (Deo and Trivedi, 2018).

Convolutional Neural Networks (CNNs) have been widely used to extract spatial features from scene or map representations for real-time trajectory prediction (Yi et al., 2016; Kumamoto and Yamada, 2017; Doellinger et al., 2018; Hoermann et al., 2018). For instance, Hoermann et al. (2018) combined CNNs with a Bayesian filter to forecast short-horizon trajectories of multiple road users, implicitly capturing interactive patterns. Graph-based extensions such as Graph Convolutional Networks (GCNs) generalise CNNs to relational structures, enabling explicit modelling of inter-agent dependencies (Kipf and Welling, 2016; Ye et al., 2021). Building on this idea, Zhang et al. (2022) proposed a spatial-temporal GCN that predicted crossing intention

from pedestrian skeleton sequences, revealing correlations between head orientation and crossing decisions.

To capture temporal dependencies, Recurrent Neural Networks (RNNs) and their variant networks, Long Short-Term Memory (LSTM), have been widely adopted (Yu et al., 2019). LSTMs have achieved strong performance in pedestrian trajectory forecasting (Alahi et al., 2016; Dai et al., 2019; Rasouli et al., 2019), and extensions such as the Holistic LSTM (Quan et al., 2021) integrate additional memory units for speed, intention, and correlation, achieving state-of-the-art performance from an on-vehicle perspective. More recently, Transformer architectures (Vaswani et al., 2017) have gained attention for their ability to capture long-range dependencies and efficiently process sequential data (Yin et al., 2021; Yuan et al., 2021; Yang et al., 2025; Divya et al., 2025). For example, Yang et al. (2025) introduced a Transformer-based trajectory prediction network that jointly models social interactions and personal intentions through a novel momentum attention mechanism. Their approach achieved state-of-the-art accuracy on metrics such as mean square error (MSE), average displacement error (ADE), and final displacement error (FDE), while also improving computational efficiency through lower parameter counts and faster inference.

Despite their strong predictive accuracy on trajectory-based metrics, several limitations persist. First, data-driven models are often criticised as 'black boxes' due to their lack of interpretability, offering little insight into the mechanisms underlying the behaviours they predict (Madala and Gonzalez, 2023; Roshdi et al., 2024). This opacity raises safety concerns, since unexplained prediction errors can propagate into AV decision-making and undermine trust in safety-critical systems (Doshi-Velez and Kim, 2017; Rudin, 2019). Road user behaviour models can be used in two ways: for real-time prediction of surrounding road users during on-road operation and as behavioural agents in virtual testing simulations. The models developed in this thesis are primarily intended for offline virtual testing rather than real-time prediction.

Second, high accuracy at an aggregate level does not necessarily indicate behaviourally meaningful predictions: models may achieve low trajectory error yet fail to reproduce realistic interaction patterns such as courtesy yielding or collision avoidance (Srinivasan et al., 2023). Third, the absence of theoretical grounding and reliance on large, diverse datasets limit their robustness and generalisability (Xu et al., 2020; Diaz-Ruiz et al., 2022). Collecting sufficient data to represent rare or safety-critical events remains particularly challenging. As a result, while data-driven

ML models have demonstrated high accuracy in modelling and predicting road user behaviour, their lack of interpretability and behavioural realism limits their reliability for prediction and evaluation in safety-critical scenarios.

1.2.3 Reinforcement learning (RL) models

Unlike data-driven ML approaches mentioned in Section 1.2.2 that passively learn mappings from data, RL enables agents to acquire adaptive decision policies through trial-and-error interaction with a dynamic environment in order to maximise cumulative rewards (Kaelbling et al., 1996). Owing to its sequential decision-making nature and ability to capture interactions between agents and their surroundings, RL has been increasingly applied to simulate road user behaviour (Charalambous and Chrysanthou, 2019; Vizzari and Cecconello, 2022). In a typical formulation, the environment is represented as a Markov Decision Process (MDP), defined by a tuple $< S, A, T, R >$, where S and A denote the sets of states and actions, T the transition function, and R the reward function. The agent seeks a policy π that maximises the expected cumulative reward,

$$V^{\pi}(s) = \mathbb{E}_{\pi}\left[\sum_{k=0}^{\infty} \gamma^{k} r_{t+k+1} \,\middle|\, s_t = s\right],$$

where r_{t+k+1} is the reward received $k+1$ steps after time t. The notation $\mathbb{E}_{\pi}[\cdot]$ indicates that the expectation is taken with respect to the action selections prescribed by policy π, and the vertical bar denotes conditioning on the initial state $s_t = s$. The discount factor $\gamma \in [0, 1)$ weights future rewards, allowing the value function $V^{\pi}(s)$ to represent the expected long-term return starting from state s when following policy π.

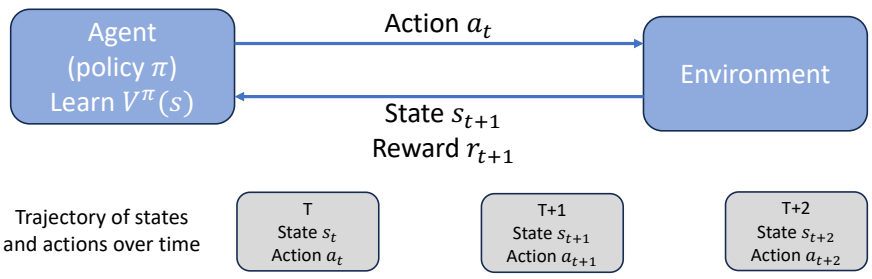

Figure 1.2: The reinforcement learning interaction loop in which the agent selects an action according to its policy, receives a reward from the environment, and transitions to a new state.

At a high level, RL proceeds through an iterative interaction loop in which the agent observes a

state, selects an action according to its policy, receives a reward, and transitions to a new state (Figure 1.2). Through repeated interaction, the agent updates its policy to increase expected long-term reward. A range of algorithmic approaches have been developed to implement this optimisation process. RL algorithms can be broadly grouped into value based, policy gradient, and actor-critic approaches. Chapter 2 uses Deep Q Learning (DQN) (Van Hasselt et al., 2016; Wang et al., 2016), a value-based method that approximates state-action values $Q(s, a)$ and is well suited to modelling discrete pedestrian decisions. Chapter 3 employs Proximal Policy Optimisation (PPO) (Schulman et al., 2017), an on-policy policy-gradient algorithm that improves stability by constraining updates to stay close to the current policy, making it effective for continuous locomotion control. Chapter 4 adopts Soft Actor Critic (SAC) (Haarnoja et al., 2018), an off policy actor-critic algorithm based on a maximum entropy objective that offers stable multi agent learning for coupled pedestrian-vehicle interactions in continuous action spaces.

A major challenge in applying RL to model human behaviour lies in designing reward functions that realistically reflect human preferences. Inverse Reinforcement Learning (IRL) provides a data-driven solution by inferring latent reward functions from observed behaviour (Ng, Russell, et al., 2000). For example, Nasernejad et al. (2021) and Nasernejad et al. (2023) used IRL to model pedestrian-vehicle interactions in near-miss scenarios, recovering reward structures from real-world demonstrations and showing that the derived post-encroachment time (PET), defined as the temporal separation between a pedestrian and a vehicle crossing the same conflict zone, correlated strongly with empirical observations. However, these models are typically formulated as fully observable MDPs, assuming that agents have complete access to environmental information. In reality, human decision-making is subject to perceptual uncertainty and motor limitations, which can be more accurately captured within a Partially Observable Markov Decision Process (POMDP) framework (Kochenderfer, 2015). By neglecting these limitations, MDP-based formulations risk overestimating the precision and rationality of human decisions (Faisal et al., 2008; Kwon et al., 2015; Gorrini et al., 2018).

Another key aspect of natural road user behaviour is its inherently interactive nature: pedestrians and drivers continuously adapt to each other's actions (Domeyer et al., 2022). This has motivated the adoption of Multi-Agent Reinforcement Learning (MARL), in which multiple agents learn concurrently within the same environment. MARL frameworks have been applied

to model cooperative and competitive interactions among vehicles in a variety of driving scenarios (Schmidt et al., 2022; Konstantinidis et al., 2023). These models can reproduce interaction patterns such as gap acceptance, acceleration adjustment, and collision avoidance, reflecting the bidirectional adaptation that is absent in single-agent RL.

However, the majority of MARL work presumes full observability, giving agents access to complete and noise-free state information. A smaller number of studies have begun to relax this assumption by introducing partial observability into MARL-based driving agents. For example, Sackmann (2024) incorporated a restricted field of view, and Cornelisse and Vinitsky (2024) leveraged partial observability to improve agent robustness for AV benchmarking. However, these works did not incorporate human-inspired elements such as distance-dependent visual noise or biomechanically constrained motor control, as included in the models developed in this thesis to better reflect human constraints when interacting with the real world.

In summary, RL provides a flexible framework for modelling adaptive and interactive decision-making. Yet, current implementations often simplify human perception and motor control, highlighting the need for models that integrate bounded information, uncertainty, and physical constraints to achieve greater behavioural realism.

1.2.4 Frameworks for modelling human behaviour

Human behaviour can be studied through a variety of modelling frameworks, many of which have already been discussed in earlier sections. Approaches such as gap acceptance models grounded in choice modelling, evidence accumulation models, supervised learning frameworks, and RL each provide reusable structures for describing specific aspects of road user behaviour. However, these frameworks face limitations when applied to complex, uncertain, and interactive traffic environments. Choice models often rely on simplified decision heuristics, evidence accumulation models typically describe isolated perceptual decisions rather than continuous interaction, supervised learning offers limited interpretability, and RL usually assumes perfect perception and unconstrained motor control.

To examine alternative frameworks that may address some of these limitations, this subsection considers two further perspectives drawn from cognitive science: cognitive architectures and computational rationality. These approaches provide explicit accounts of perceptual, cognitive, and motor processes and therefore offer potential routes toward modelling human behaviour in

more mechanistic and psychologically interpretable ways.

Cognitive architectures

Cognitive architectures aim to model the underlying processes of human cognition within a unified computational framework. Rather than relying solely on statistical associations, they provide mechanistic explanations of perception, memory, reasoning, and action by specifying how these components interact over time. Two of the most influential architectures are ACT-R (Anderson et al., 1997; Anderson, 2009) and Soar (Laird, 2019).

ACT-R (Adaptive Control of Thought—Rational) models human cognition as the interaction between a set of modules responsible for perceptual, declarative, and procedural functions. Behaviour emerges from the competition among production rules operating over memory chunks, where activation strengths determine retrieval likelihood. These mechanisms allow ACT-R to reproduce human patterns in reaction time, learning, and multitasking, and it has been applied to domains such as driving (Salvucci et al., 2001; Salvucci, 2006) and visual attention (Fleetwood and Byrne, 2006).

Soar, in contrast, emphasises problem solving and learning through chunking, where new rules are generated from problem-solving episodes, enabling gradual improvement over time (Laird, 2019). Grounded in the problem space hypothesis, Soar represents goal-oriented behaviour as a search process through states and operators, and automatically creates new procedural knowledge through chunking when problem-solving impasses are successfully resolved. Both frameworks are designed to achieve cognitive plausibility by grounding computation in psychological theory and enabling explicit mapping between model components and cognitive processes.

Despite their explanatory power, cognitive architectures face important limitations when applied to complex and dynamic environments such as traffic interactions. They typically require extensive manual specification of task rules, perceptual inputs, and parameter values, which restricts scalability and generalisation (Brasoveanu and Dotlačil, 2021). Furthermore, although probabilistic components have been introduced, their primarily symbolic representations make it difficult to capture the continuous and uncertain perceptual-motor dynamics that characterise natural human movement (Byrne et al., 2010). For these reasons, cognitive architectures are not well suited as a modelling framework for pedestrian behaviour in autonomous vehicle applications, where agents must adapt flexibly to uncertainty, interact with other road users,

and operate within continuous perceptual and motor spaces.

Computational rationality

These limitations mentioned in the last section have motivated the development of alternative frameworks, such as computational rationality, which seek to retain psychological interpretability while formalising behaviour as optimisation under perceptual, cognitive, and motor constraints.

Computational rationality provides a general framework for modelling behaviour as approximately optimal adaptation under perceptual, cognitive, and motor constraints (Lewis et al., 2014; Oulasvirta et al., 2022). Rather than assuming idealised perception or perfectly rational decision-making, the framework formalises how limitations in sensing, memory, attention, and motor execution shape the space of strategies available to an agent. Behaviour is therefore characterised as the solution to an optimisation problem defined jointly by task goals and human-like constraints, making this approach particularly suitable for dynamic and uncertain settings such as road user interaction. A schematic comparison between a standard RL formulation and a computationally rational agent is shown in Figure 1.3.

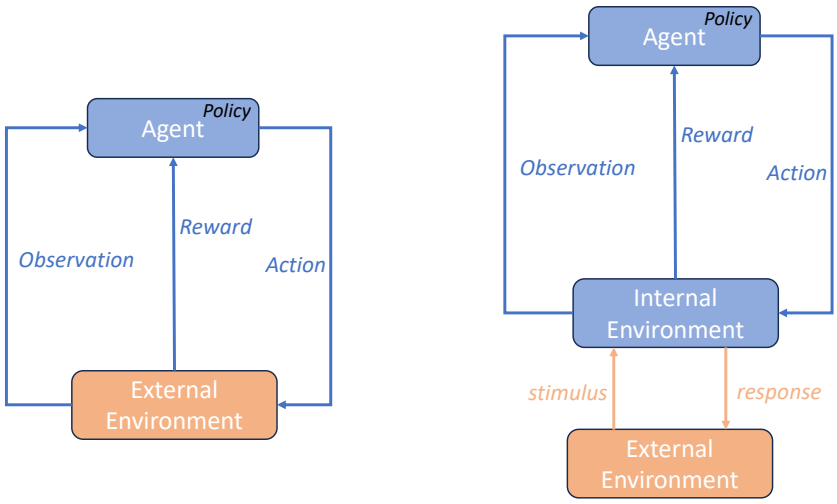

Figure 1.3: Comparison between a standard RL agent (left) and a computationally rational agent (right). In computational rationality, the agent interacts with the external world through an internal environment with perceptual, cognitive, and motor constraints.

In computational rationality models, these human constraints are embedded directly into the environment in which the agent learns. Perceptual limitations are typically represented as partial or noisy observations, such as distance-dependent visual uncertainty, limited visual fields, sampling delays, or imprecise state estimates. Cognitive limitations arise through bounded

memory representations, imperfect belief updating, or attention-switching costs. Motor limitations are captured through execution noise, delayed actuation, bounded acceleration or steering rates, and other physical constraints. Encoding constraints within the environment ensures that learned behaviour reflects realistic human information-processing limits rather than idealised assumptions.

Action policies are not hand crafted; instead they are derived through RL, which progressively discovers strategies that approximate the optimal policy within the specified perceptual and motor constraints. The agent repeatedly observes, acts, and receives rewards under uncertainty, gradually improving its behaviour to maximise long-term utility. Although the policy itself is learned, the parameters that describe perceptual and motor limitations must be specified or fitted to human data, typically by minimising discrepancies between simulated and observed behaviour. In this thesis, we refer to these parameters as non-policy parameters to distinguish them from the trainable parameters of the neural network (weights and biases).

Computational rationality has been widely used in human-computer interaction (HCI), modelling behaviours such as visual search (Tseng and Howes, 2015), typing (Sarcar et al., 2018), and touchscreen interaction (Jokinen et al., 2021). A particularly relevant example for the present thesis is the multitasking driving model by Jokinen et al. (2020), which frames lane keeping and in-vehicle visual search as a joint optimisation problem under uncertainty. The model incorporates visual sampling constraints, noisy motor control, and switching costs, and employs RL to discover strategies that balance road monitoring with secondary-task performance. This work demonstrates how computational rationality can generate human-like adaptive behaviour in traffic settings. Although it represents one of the first applications of the framework to a transportation task, its focus and methodology remain grounded in HCI rather than in modelling road user interaction.

Beyond this early work, there are currently no computational rationality models addressing problems in the transportation field, particularly in road user interaction modelling. In addition, existing applications remain restricted to single-agent settings, whereas real-world road user interaction involves multiple agents whose decisions continuously shape one another's mind. Extending computational rationality to such interactive settings requires a framework capable of modelling boundedly rational agents who learn and adapt in the presence of other adaptive agents. Multi-agent reinforcement learning (MARL) offers a natural foundation for this

extension, yet it has not previously been integrated with computational rationality in the modelling of road user interaction. Addressing this gap forms a central motivation for the approach developed in this thesis.

Despite these advantages, computational rationality should not be viewed as a complete account of pedestrian behaviour. The framework is best suited to behaviours that can be approximated as goal-directed adaptation under constraints, but real pedestrian behaviour may also reflect simple heuristics, habits, or social conventions, which may not always need to be modelled explicitly through a single optimisation-based account. In this thesis, computational rationality is therefore used as a modelling framework for specific pedestrian-vehicle interaction tasks, rather than as a general theory of all pedestrian behaviour.

1.2.5 Virtual testing of AVs

Simulation-based, or virtual, testing has become an indispensable tool in the development and evaluation of AVs. Real-world testing alone is costly, time-consuming, and potentially hazardous, particularly for rare or safety-critical situations. In contrast, simulation allows controlled, repeatable, and scalable evaluation of AV behaviour across diverse environmental and interaction scenarios (Huang et al., 2016; Joisher et al., 2019; Sawitzky et al., 2023). As a result, virtual environments are now widely recognised as a core component of the AV development pipeline (Wang et al., 2024).

The realism and reliability of such evaluations depend critically on how other traffic participants are modelled. Pedestrians pose a particular challenge because of behavioural variability and weaker adherence to traffic rules. Many simulation frameworks still rely on simplified agents based on simplified rule-based models (Papadimitriou et al., 2016; Rashid et al., 2024). Although computationally efficient, these agents, as discussed above, do not represent perceptual uncertainty, adaptive decision making, or inter-individual variability in pedestrian behaviour (Camara et al., 2020). AV evaluations built on such agents risk producing interactions that are oversimplified or implausible, which limits their value for evaluating AV performance in realistic, safety-critical scenarios.

As discussed in previous sections, recent studies have developed human-like behavioural models that incorporate multiple human decision-making mechanisms, providing more realistic representations of how pedestrians perceive vehicles, evaluate risk, and decide when to cross (Markkula

et al., 2018b; Pekkanen et al., 2022; Tian et al., 2025; Wang et al., 2025). However, despite their potential to enhance simulation realism, there is little published work on integrating such models into virtual testing frameworks for AV evaluation. Existing simulation studies typically continue to rely on oversimplified or scripted pedestrian agents (Papadimitriou et al., 2016; Rashid et al., 2024), leaving the benefits of cognitively grounded models largely unexplored in testing contexts.

Parallel to these efforts on realism, another line of research has focused on adversarial scenario generation. This approach deliberately constructs traffic situations that challenge AV decision-making, often by pushing interactions towards unsafe or near-miss outcomes. Such techniques have been developed mainly for robustness testing, aiming to expose AV system failure modes. Common implementations employ RL or black-box optimisation to train pedestrian agents that provoke risky events such as collisions (Song et al., 2023; Hanselmann et al., 2022; Priisalu et al., 2023). For instance, the 'suicidal pedestrian' framework explicitly trains agents to collide with AVs (Yang et al., 2023), while the adversarial jaywalker model uses a multi-state rule-based controller to generate hazardous crossings for robustness testing (Muktadir and Whitehead, 2022). Although effective at identifying vulnerabilities in AV control, these methods prioritise difficulty over behavioural plausibility. The resulting agents often exhibit exaggerated or unrealistic actions that fail to reflect real pedestrian intentions, thereby undermining realism and limiting their applicability to real-world AV testing (Dyro et al., 2024).

Beyond robustness testing, adversarial scenarios also offer opportunities for control optimisation in pedestrian-vehicle interactions. Effective interaction requires balancing safety and efficiency: AVs must avoid collisions while maintaining smooth and timely motion. Optimising AV control against either overly simple or overly aggressive pedestrian agents may yield strategies that fail to generalise to real human interactions, producing behaviours that are either unsafe or excessively conservative (Robert et al., 2021; Crosato et al., 2021). Yet, adversarial scenario generation and human-like behavioural modelling have so far developed as separate research directions with limited integration, which is one of the methodological gaps that this thesis seeks to address.

1.3 Research gaps

- **Gap 1: Lack of an integrated modelling framework that combines mechanistic interpretability with machine learning scalability in modelling pedestrian-vehicle interaction.**

  Existing studies are mainly divided between two paradigms. Mechanistic models, such as the COMMOTIONS framework, explicitly represent human perceptual and cognitive mechanisms and thus provide strong interpretability, but they become highly complex when extended to capture rich interactive behaviours, limiting their scalability to complex real-world scenarios. In contrast, data-driven ML models offer flexibility and strong predictive capability, yet they lack interpretability and rely heavily on large, representative datasets, an issue that becomes critical for safety-related, data-scarce scenarios.

  This trade-off between interpretability and adaptability is not unique to autonomous driving. Similar debates have long existed in the broader field of human-robot interaction (HRI), where researchers face comparable challenges in modelling human behaviour for collaborative, assistive, and interactive robotic systems (Mutlu et al., 2016; Kulkarni et al., 2020; Spitale et al., 2025). Positioning pedestrian-vehicle interaction within this broader HRI context shows the need for integrated frameworks that combine mechanistic grounding with data-driven learning to achieve interpretability, generalisability, and behavioural realism.

- **Gap 2: Limited understanding of pedestrian-vehicle interaction from the perspective of boundedly optimal behaviour under human perceptual-motor constraints.**

  While mechanistic models have begun to incorporate individual mechanisms such as perceptual uncertainty, motor constraints or bounded rationality, few studies have systematically integrated these processes or examined how they jointly shape the human pedestrian behaviour. As a result, the field still lacks an understanding of how human perceptual and motor constraints give rise to the adaptive strategies observed in real pedestrian-vehicle interactions. Recent frameworks such as COMMOTIONS demonstrate the feasibility of combining multiple cognitive mechanisms mathematically to generate and explain pedestrian-vehicle interaction patterns, but their validation has been limited to controlled

experimental settings.

- **Gap 3: Limited use of human-like behavioural models in realistic and adversarial virtual testing.**

  Although virtual testing has become essential for evaluating AV safety, most simulation environments still rely on simplified pedestrian agents based on deterministic or rule-based heuristics. These agents fail to capture the variability and uncertainty observed in human behaviour, leading to unrealistic pedestrian-AV interactions. Meanwhile, adversarial scenario generation methods tend to prioritise difficulty over behavioural plausibility, producing exaggerated pedestrian behaviour. Integrating models developed for behavioural realism could improve both the representativeness of virtual testing and the robustness of AV control optimisation.

1.4 Research objectives

To address the research gaps identified above, this work sets out four main objectives that collectively aim to advance the modelling of pedestrian-vehicle interaction and the application of human-like models in AV evaluation. The specific objectives are as follows:

- **Objective 1:** To identify and quantify key human perceptual and motor constraints that shape pedestrian crossing behaviour. This objective focuses on revealing how perceptual uncertainty, motor effort, and individual variability influence pedestrians' behaviour under both controlled and naturalistic conditions.
- **Objective 2:** To incorporate the identified perceptual and motor constraints into an RL framework and derive boundedly optimal decision policies that reproduce human-like behaviour under uncertainty. This objective formalises pedestrian decision-making as an optimisation problem constrained by perceptual noise and motor limitations, and uses RL to obtain adaptive strategies consistent with observed behavioural patterns.
- **Objective 3:** To extend the developed model from single-agent settings to interactive pedestrian-vehicle scenarios. This objective enables both agents to adapt to one another's actions through multi-agent RL to model the interaction patterns observed in the real-world scenario.

- **Objective 4:** To apply the human-like road user behaviour model to virtual testing of AV. This objective integrates human-like behavioural models into a closed-loop simulation environment to generate behavioural directly informed realistic yet safety-critical scenarios for evaluating and optimising AV control performance.

1.5 Thesis outline

This thesis comprises four main studies that together advance the modelling and evaluation of pedestrian-vehicle interaction. The COMMOTIONS framework serves as an important conceptual point of reference for this thesis, but it is not the modelling framework that is directly developed across the doctoral work. Its main role is to motivate the view that pedestrian behaviour should be understood in relation to perceptual and motor mechanisms. Building on this general perspective, Chapters 2 to 4 develop a separate RL based computational rationality framework for modelling boundedly optimal pedestrian and driver behaviour. Chapter 5 is different in scope, as it applies the pedestrian model derived from the COMMOTIONS framework in AV simulation testing, rather than extending the RL framework introduced in the earlier chapters.

Figure 1.4 presents an overview of the thesis structure, showing how the two research branches, core modelling and applied testing, are organised across the core chapters of the thesis. Chapter 2 introduces the modelling of pedestrian decision-making under visual constraints, based on button-press crossing decisions collected in a controlled VR environment. Chapter 3 extends this framework to model full pedestrian walking trajectories in VR by incorporating both perceptual and motor constraints, moving from discrete decisions to continuous sensory-motor control. Chapter 4 further extends the framework to naturalistic road user interaction by developing a multi-agent RL model in which pedestrians and drivers co-adapt to one another in real-world traffic settings. This progression, from simplified decision tasks, to embodied pedestrian movement, to interactive pedestrian-driver behaviour is illustrated in Figure 1.4. Chapter 5 is a study applying the COMMOTIONS model to AV simulation testing. Finally, Chapter 6 concludes the thesis by summarising the main findings, discussing their implications for human behaviour modelling and AV evaluation, and pointing directions for future research.

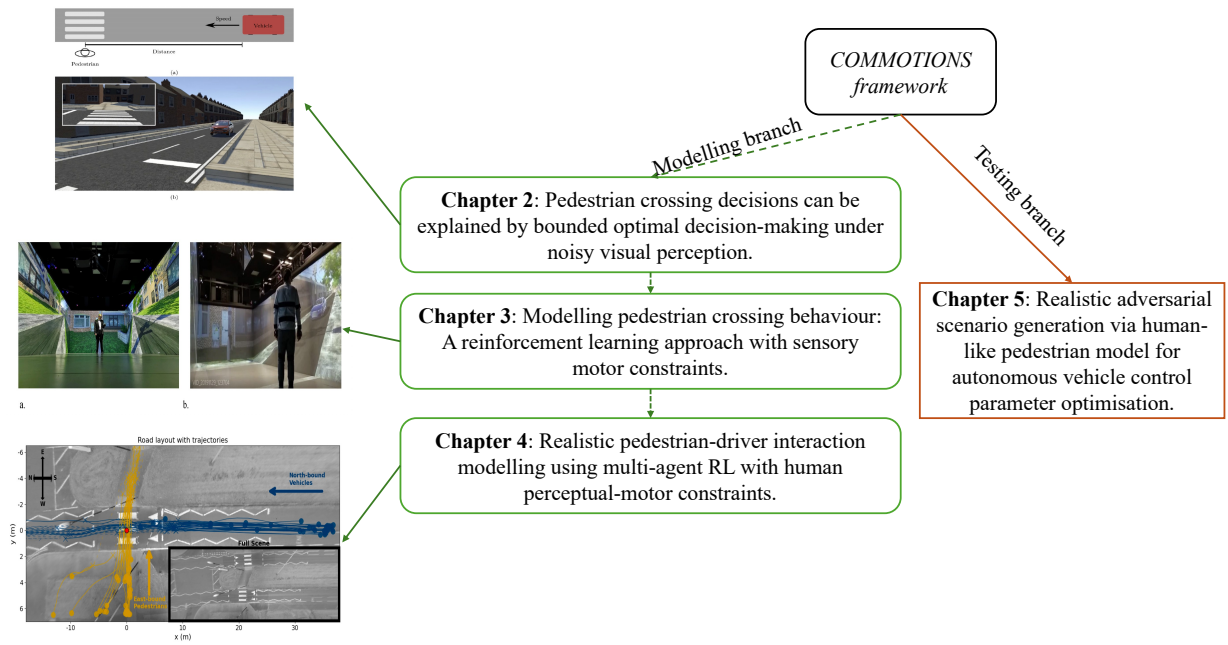

Figure 1.4: Thesis structure and its relation to the COMMOTIONS framework. Chapters 2-4 form the core doctoral branch, progressing from perceptual uncertainty in decision-making to sensory-motor constrained control and interactive multi-agent modelling. These chapters are conceptually informed by key ideas from COMMOTIONS, especially its mechanistic treatment of perceptual uncertainty, but they develop a distinct computational rationality framework based on reinforcement learning. Chapter 5 applies the pedestrian model derived from COMMOTIONS to AV testing and controller optimisation.

Chapter 2, *'Pedestrian crossing decisions can be explained by bounded optimal decision-making under noisy visual perception'* (published in *Transportation Research Part C*), addresses **Gap 1** and **Gap 2**, contributing to **Objectives 1-2**. This chapter develops an RL model that formalises pedestrian crossing as boundedly optimal behaviour under noisy perception. By incorporating the noisy visual information into the RL environment, the model reproduces empirically observed crossing patterns and explains pedestrian decision-making as optimal behaviour under perceptual constraints.

Chapter 3, *'Modelling pedestrian crossing behaviour: A reinforcement learning approach with sensory-motor constraints'* (published in *IEEE Transactions on Intelligent Transportation Systems*), extends this framework from discrete crossing decisions to the full crossing process by incorporating both perceptual and motor constraints, addressing the same research gaps and further extending this thesis' contribution to **Objectives 1-2**. The resulting sensory-motor constrained model reflects how perception, motor effort, and time pressure shape pedestrian crossing behaviour. It produces human-like trajectories and adaptive strategies, thereby estab-

lishing a more comprehensive account of boundedly optimal crossing behaviour.

Chapter 4, *'Realistic pedestrian-driver interaction modelling using multi-agent reinforcement learning with human perceptual-motor constraints'* (under review for journal publication), extends the single-agent framework to a multi-agent setting and contributes to **Objective 3**. In this study, pedestrians and vehicles are both modelled as learning agents that co-adapt to each other's actions under human-like sensory and motor constraints. This work extends computational rationality-based modelling of pedestrian-vehicle interaction beyond controlled single-agent settings to multi-agent scenarios, and reproduces several adaptive behaviours observed in real-world traffic.

Chapter 5, *'Realistic adversarial scenario generation via human-like pedestrian models for autonomous vehicle control optimisation'* (under review for journal publication), addresses **Gap 3** and contributes to **Objective 4**. This study integrates the COMMOTIONS human-like pedestrian model into a full-stack AV simulation environment.

Finally, **Chapter 6** concludes the thesis by synthesising the findings from all studies, discussing theoretical and practical implications for human-like modelling and AV evaluation, and outlining future research directions. Collectively, these chapters form a progression from understanding human perceptual-motor constraints, through modelling boundedly optimal and interactive behaviour, to applying human-like agents for robust and realistic AV testing.

References

Alahi, Alexandre, Goel, Kratarth, Ramanathan, Vignesh, Robicquet, Alexandre, Fei-Fei, Li, and Savarese, Silvio (2016). "Social lstm: Human trajectory prediction in crowded spaces". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 961–971.

Anderson, John R (2009). *How can the human mind occur in the physical universe?* Oxford University Press.

Anderson, John R, Matessa, Michael, and Lebiere, Christian (1997). "ACT-R: A theory of higher level cognition and its relation to visual attention". In: *Human–Computer Interaction* 12.4, pp. 439–462.

Asaithambi, Gowri, Kuttan, Manu O, and Chandra, Sarath (2016). "Pedestrian road crossing behavior under mixed traffic conditions: a comparative study of an intersection before and after implementing control measures". In: *Transportation in developing economies* 2.2, p. 14.

Brasoveanu, Adrian and Dotlačil, Jakub (2021). "Reinforcement Learning for Production-Based Cognitive Models". In: *Topics in Cognitive Science* 13.3, pp. 467–487.

Byrne, Michael D, O'malley, Marcia K, Gallagher, Melissa A, Purkayastha, Sagar N, Howie, Nicole, and Huegel, Joel C (2010). "A preliminary ACT-R model of a continuous motor task". In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Vol. 54. 13. SAGE Publications Sage CA: Los Angeles, CA, pp. 1037–1041.

Camara, Fanta, Bellotto, Nicola, Cosar, Serhan, Weber, Florian, Nathanael, Dimitris, Althoff, Matthias, Wu, Jingyuan, Ruenz, Johannes, Dietrich, André, Markkula, Gustav, et al. (2020). "Pedestrian models for autonomous driving part ii: high-level models of human behavior". In: *IEEE Transactions on Intelligent Transportation Systems* 22.9, pp. 5453–5472.

Camara, Fanta, Dickinson, Patrick, and Fox, Charles (2021). "Evaluating pedestrian interaction preferences with a game theoretic autonomous vehicle in virtual reality". In: *Transportation research part F: traffic psychology and behaviour* 78, pp. 410–423.

Charalambous, Panayiotis and Chrysanthou, Yiorgos (2019). "Why did the human cross the road?" In: *Proceedings of the 12th ACM SIGGRAPH Conference on Motion, Interaction and Games*, pp. 1–2.

Chen, Xu, Treiber, Martin, Kanagaraj, Venkatesan, and Li, Haiying (2018). "Social force models for pedestrian traffic–state of the art". In: *Transport reviews* 38.5, pp. 625–653.

Codevilla, Felipe, Santana, Eder, López, Antonio M, and Gaidon, Adrien (2019). "Exploring the limitations of behavior cloning for autonomous driving". In: *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 9329–9338.

Cornelisse, Daphne and Vinitsky, Eugene (2024). "Human-compatible driving partners through data-regularized self-play reinforcement learning". In: *arXiv preprint arXiv:2403.19648*.

Crosato, Luca, Wei, Chongfeng, Ho, Edmond SL, and Shum, Hubert PH (2021). "Human-centric autonomous driving in an av-pedestrian interactive environment using svo". In: *2021 IEEE 2nd International Conference on Human-Machine Systems (ICHMS)*. IEEE, pp. 1–6.

Dai, Shengzhe, Li, Li, and Li, Zhiheng (2019). "Modeling vehicle interactions via modified LSTM models for trajectory prediction". In: *IEEE Access* 7, pp. 38287–38296.

Dang, Meiting, Zhao, Dezong, Wang, Yafei, and Wei, Chongfeng (2025). "Dynamic Game-Theoretical Decision-Making Framework for Vehicle-Pedestrian Interaction With Human Bounded Rationality". In: *IEEE Transactions on Intelligent Transportation Systems*.

DeLucia, Patricia R. (2015). "Perception of Collision". In: *The Cambridge Handbook of Applied Perception Research*. Ed. by Robert R. Hoffman, Peter A. Hancock, Mark W. Scerbo, Raja Parasuraman, and James L.Editors Szalma. Cambridge Handbooks in Psychology. Cambridge University Press, pp. 568–591. DOI: 10.1017/CBO9780511973017.035.

Deo, Nachiket and Trivedi, Mohan M (2018). "Convolutional social pooling for vehicle trajectory prediction". In: *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pp. 1468–1476.

Department for Transport (2021). *Reported road casualties Great Britain, annual report: 2021*. https://www.gov.uk/government/statistics/reported-road-casualties-great-britain-annual-report-2021. Accessed: 2025-10-22.

Diaz-Ruiz, Carlos A, Xia, Youya, You, Yurong, Nino, Jose, Chen, Junan, Monica, Josephine, Chen, Xiangyu, Luo, Katie, Wang, Yan, Emond, Marc, et al. (2022). "Ithaca365: Dataset and Driving Perception Under Repeated and Challenging Weather Conditions". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 21383–21392.

Divya, J, Divyadharshini, M, and Divyapriya, V (2025). "Gaze-based Pedestrian Safety Prediction using Attention LSTM". In: *2025 3rd International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)*. IEEE, pp. 1838–1845.

Doellinger, Johannes, Spies, Markus, and Burgard, Wolfram (2018). "Predicting occupancy distributions of walking humans with convolutional neural networks". In: *IEEE Robotics and Automation Letters* 3.3, pp. 1522–1528.

Domeyer, Joshua E, Lee, John D, and Toyoda, Heishiro (2020). "Vehicle automation–other road user communication and coordination: Theory and mechanisms". In: *IEEE Access* 8, pp. 19860–19872.

Domeyer, Joshua E, Lee, John D, Toyoda, Heishiro, Mehler, Bruce, and Reimer, Bryan (2022). "Driver-Pedestrian Perceptual Models Demonstrate Coupling: Implications for Vehicle Automation". In: *IEEE Transactions on Human-Machine Systems* 52.4, pp. 557–566.

Doshi-Velez, Finale and Kim, Been (2017). "Towards a rigorous science of interpretable machine learning". In: *arXiv preprint arXiv:1702.08608*.

Dosovitskiy, Alexey, Beyer, Lucas, Kolesnikov, Alexander, Weissenborn, Dirk, Zhai, Xiaohua, Unterthiner, Thomas, Dehghani, Mostafa, Minderer, Matthias, Heigold, Georg, Gelly, Sylvain, et al. (2020). "An image is worth 16x16 words: Transformers for image recognition at scale". In: *arXiv preprint arXiv:2010.11929*.

Dyro, Robert, Foutter, Matthew, Li, Ruolin, Di Lillo, Luigi, Schmerling, Edward, Zhou, Xilin, and Pavone, Marco (2024). "Realistic extreme behavior generation for improved av testing". In: *arXiv preprint arXiv:2409.10669*.

Elvik, Rune (2014). "A review of game-theoretic models of road user behaviour". In: *Accident Analysis & Prevention* 62, pp. 388–396.

Faisal, A Aldo, Selen, Luc PJ, and Wolpert, Daniel M (2008). "Noise in the nervous system". In: *Nature reviews neuroscience* 9.4, pp. 292–303.

Fajen, Brett R and Warren, William H (2003). "Behavioral dynamics of steering, obstable avoidance, and route selection." In: *Journal of Experimental Psychology: Human Perception and Performance* 29.2, p. 343.

Fleetwood, Michael D and Byrne, Michael D (2006). "Modeling the visual search of displays: a revised ACT-R model of icon search based on eye-tracking data". In: *Human-Computer Interaction* 21.2, pp. 153–197.

Fox, Charles, Camara, Fanta, Markkula, Gustav, Romano, Richard, Madigan, Ruth, Merat, Natasha, et al. (2018). "When should the chicken cross the road?: Game theory for autonomous vehicle-human interactions". In.

Giles, Oscar, Markkula, Gustav, Pekkanen, Jami, Yokota, Naoki, Matsunaga, Naoto, Merat, Natasha, and Daimon, Tatsuru (2019). "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, pp. 366–372.

Giszter, Simon F (2015). "Motor primitives—new data and future questions". In: *Current opinion in neurobiology* 33, pp. 156–165.

Gorrini, Andrea, Crociani, Luca, Vizzari, Giuseppe, and Bandini, Stefania (2018). "Observation results on pedestrian-vehicle interactions at non-signalized intersections towards simulation". In: *Transportation research part F: traffic psychology and behaviour* 59, pp. 269–285.

Haarnoja, Tuomas, Zhou, Aurick, Abbeel, Pieter, and Levine, Sergey (2018). "Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor". In: *International conference on machine learning*. Pmlr, pp. 1861–1870.

Hanselmann, Niklas, Renz, Katrin, Chitta, Kashyap, Bhattacharyya, Apratim, and Geiger, Andreas (2022). "King: Generating safety-critical driving scenarios for robust imitation via kinematics gradients". In: *European Conference on Computer Vision*. Springer, pp. 335–352.

Helbing, Dirk and Molnar, Peter (1995). "Social force model for pedestrian dynamics". In: *Physical review E* 51.5, p. 4282.

Hoermann, Stefan, Bach, Martin, and Dietmayer, Klaus (2018). "Dynamic occupancy grid prediction for urban autonomous driving: A deep learning approach with fully automatic labeling". In: *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, pp. 2056–2063.

Hoogendoorn, Serge and HL Bovy, Piet (2003). "Simulation of pedestrian flows by optimal control and differential games". In: *Optimal control applications and methods* 24.3, pp. 153–172.

Hou, Lei, Liu, Jian-Guo, Pan, Xue, and Wang, Bing-Hong (2014). "A social force evacuation model with the leadership effect". In: *Physica A: Statistical Mechanics and its Applications* 400, pp. 93–99.

Huang, WuLing, Wang, Kunfeng, Lv, Yisheng, and Zhu, FengHua (2016). "Autonomous vehicles testing methods review". In: *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, pp. 163–168.

Jain, Akash, Gupta, Ankit, and Rastogi, Rajat (2014). "Pedestrian crossing behaviour analysis at intersections". In: *International Journal for Traffic and Transport Engineering* 4.1, pp. 103–116.

Johora, Fatema T and Müller, Jörg P (2018). "Modeling interactions of multimodal road users in shared spaces". In: *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, pp. 3568–3574.

Joisher, Karan, Khan, Suhaib, and Ranadive, Omkar (2019). "Simulation Environment for Development and Testing of Autonomous Learning Agents". In: *2nd International Conference on Advances in Science & Technology (ICAST)*.

Jokinen, Jussi, Acharya, Aditya, Uzair, Mohammad, Jiang, Xinhui, and Oulasvirta, Antti (2021). "Touchscreen typing as optimal supervisory control". In: *Proceedings of the 2021 CHI conference on human factors in computing systems*, pp. 1–14.

Jokinen, Jussi PP, Wang, Zhenxin, Sarcar, Sayan, Oulasvirta, Antti, and Ren, Xiangshi (2020). "Adaptive feature guidance: Modelling visual search with graphical layouts". In: *International Journal of Human-Computer Studies* 136, p. 102376.

Kaelbling, Leslie Pack, Littman, Michael L, and Moore, Andrew W (1996). "Reinforcement learning: A survey". In: *Journal of artificial intelligence research* 4, pp. 237–285.

Kalantari, Amir Hossein, Markkula, Gustav, Uzondu, Chinebuli, Lyu, Wei, Pedro, Jorge Garcia de, Madigan, Ruth, Lee, Yee Mun, Holmes, Christopher, and Merat, Natasha (2022). *Vehicle-Pedestrian Interactions at Uncontrolled Locations: Leveraging Distributed Simulation to Support Game-Theoretic Modeling*. Tech. rep.

Kipf, Thomas N and Welling, Max (2016). "Semi-supervised classification with graph convolutional networks". In: *arXiv preprint arXiv:1609.02907*.

Kochenderfer, Mykel J (2015). *Decision making under uncertainty: theory and application*. MIT press.

Konstantinidis, Fabian, Sackmann, Moritz, Hofmann, Ulrich, and Stiller, Christoph (2023). "Modeling interaction-aware driving behavior using graph-based representations and multi-agent reinforcement learning". In: *2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, pp. 1643–1650.

Koopman, Philip and Wagner, Michael (2017). "Autonomous vehicle safety: An interdisciplinary challenge". In: *IEEE Intelligent Transportation Systems Magazine* 9.1, pp. 90–96.

Kulkarni, Anagha, Sreedharan, Sarath, Keren, Sarah, Chakraborti, Tathagata, Smith, David E, and Kambhampati, Subbarao (2020). "Designing environments conducive to interpretable robot behavior". In: *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, pp. 10982–10989.

Kumamoto, Koji and Yamada, Keiichi (2017). "CNN-based pedestrian orientation estimation from a single image". In: *2017 4th IAPR Asian conference on pattern recognition (ACPR)*. IEEE, pp. 13–18.

Kwon, Oh-Sang, Tadin, Duje, and Knill, David C (2015). "Unifying account of visual motion and position perception". In: *Proceedings of the National Academy of Sciences* 112.26, pp. 8142–8147.

Laird, John E (2019). *The Soar cognitive architecture*. MIT press.

Lewis, Richard L, Howes, Andrew, and Singh, Satinder (2014). "Computational rationality: Linking mechanism and behavior through bounded utility maximization". In: *Topics in cognitive science* 6.2, pp. 279–311.

Liu, Manxia, Zeng, Weiliang, Chen, Peng, and Wu, Xuyi (2017). "A microscopic simulation model for pedestrian-pedestrian and pedestrian-vehicle interactions at crosswalks". In: *PLoS one* 12.7, e0180992.

Lu, Yiren, Fu, Justin, Tucker, George, Pan, Xinlei, Bronstein, Eli, Roelofs, Rebecca, Sapp, Benjamin, White, Brandyn, Faust, Aleksandra, Whiteson, Shimon, et al. (2023). "Imitation is not enough: Robustifying imitation with reinforcement learning for challenging driving scenarios". In: *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, pp. 7553–7560.

Ma, Wei-Chiu, Huang, De-An, Lee, Namhoon, and Kitani, Kris M (2017). "Forecasting interactive dynamics of pedestrians with fictitious play". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 774–782.

Madala, Kaushik and Gonzalez, Carlos Avalos (2023). "Metrics for machine learning models to facilitate sotif analysis in autonomous vehicles". In: *SAE International Journal of Advances and Current Practices in Mobility* 6.2023-01-0829, pp. 782–790.

Mako, Emese (2015). "Evaluation of human behaviour at pedestrian crossings". In: *2015 6th ieee international conference on cognitive infocommunications (coginfocom)*. IEEE, pp. 443–447.

Markkula, Gustav, Boer, Erwin, Romano, Richard, and Merat, Natasha (2018a). "Sustained sensorimotor control as intermittent decisions about prediction errors: Computational framework and application to ground vehicle steering". In: *Biological cybernetics* 112, pp. 181–207.

Markkula, Gustav, Engström, Johan, Lodin, Johan, Bärgman, Jonas, and Victor, Trent (2016). "A farewell to brake reaction times? Kinematics-dependent brake response in naturalistic rear-end emergencies". In: *Accident Analysis & Prevention* 95, pp. 209–226.

Markkula, Gustav, Lin, Yi-Shin, Srinivasan, Aravinda Ramakrishnan, Billington, Jac, Leonetti, Matteo, Kalantari, Amir Hossein, Yang, Yue, Lee, Yee Mun, Madigan, Ruth, and Merat, Natasha (2023). "Explaining human interactions on the road by large-scale integration of computational psychological theory". In: *PNAS nexus* 2.6, pgad163.

Markkula, Gustav, Romano, Richard, Madigan, Ruth, Fox, Charles W, Giles, Oscar T, and Merat, Natasha (2018b). "Models of human decision-making as tools for estimating and opti-

mizing impacts of vehicle automation". In: *Transportation research record* 2672.37, pp. 153–163.

Muktadir, Golam Md and Whitehead, Jim (2022). "Adversarial jaywalker modeling for simulation-based testing of autonomous vehicle systems". In: *2022 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, pp. 1697–1702.

Mutlu, Bilge, Roy, Nicholas, and Šabanović, Selma (2016). "Cognitive human–robot interaction". In: *Springer handbook of robotics*, pp. 1907–1934.

Nasernejad, Payam, Sayed, Tarek, and Alsaleh, Rushdi (2021). "Modeling pedestrian behavior in pedestrian-vehicle near misses: A continuous Gaussian Process Inverse Reinforcement Learning (GP-IRL) approach". In: *Accident Analysis & Prevention* 161, p. 106355.

– (2023). "Multiagent modeling of pedestrian-vehicle conflicts using Adversarial Inverse Reinforcement Learning". In: *Transportmetrica A: transport science* 19.3, p. 2061081.

Ng, Andrew Y, Russell, Stuart, et al. (2000). "Algorithms for inverse reinforcement learning." In: *Icml*. Vol. 1, p. 2.

Oulasvirta, Antti, Jokinen, Jussi PP, and Howes, Andrew (2022). "Computational rationality as a theory of interaction". In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pp. 1–14.

Papadimitriou, Eleonora, Auberlet, Jean-Michel, Yannis, George, and Lassarre, Sylvain (2016). "Simulation of Pedestrians and Motorised Traffic: existing research and future challenges". In: *Civil and Environmental Engineering: Concepts, Methodologies, Tools, and Applications*. IGI Global Scientific Publishing, pp. 1646–1662.

Pekkanen, Jami, Giles, Oscar Terence, Lee, Yee Mun, Madigan, Ruth, Daimon, Tatsuru, Merat, Natasha, and Markkula, Gustav (2022). "Variable-drift diffusion models of pedestrian road-crossing decisions". In: *Computational Brain & Behavior* 5.1, pp. 60–80.

Petzoldt, Tibor (2014). "On the relationship between pedestrian gap acceptance and time to arrival estimates". In: *Accident Analysis & Prevention* 72, pp. 127–133.

Prédhumeau, Manon, Mancheva, Lyuba, Dugdale, Julie, and Spalanzani, Anne (2022). "Agent-Based Modeling for Predicting Pedestrian Trajectories Around an Autonomous Vehicle". In: *Journal of Artificial Intelligence Research* 73, pp. 1385–1433.

Priisalu, Maria, Paduraru, Ciprian, and Smichisescu, Cristian (2023). "Varied realistic autonomous vehicle collision scenario generation". In: *Scandinavian Conference on Image Analysis*. Springer, pp. 354–372.

Quan, Ruijie, Zhu, Linchao, Wu, Yu, and Yang, Yi (2021). "Holistic LSTM for pedestrian trajectory prediction". In: *IEEE transactions on image processing* 30, pp. 3229–3239.

Rashid, Md Mobasshir, Seyedi, MohammadReza, and Jung, Sungmoon (2024). "Simulation of pedestrian interaction with autonomous vehicles via social force model". In: *Simulation modelling practice and theory* 132, p. 102901.

Rasouli, Amir, Kotseruba, Iuliia, Kunic, Toni, and Tsotsos, John K (2019). "Pie: A large-scale dataset and models for pedestrian intention estimation and trajectory prediction". In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 6262–6271.

Ratcliff, Roger, Smith, Philip L, Brown, Scott D, and McKoon, Gail (2016). "Diffusion decision model: Current issues and history". In: *Trends in cognitive sciences* 20.4, pp. 260–281.

Ratcliff, Roger and Strayer, David (2014). "Modeling simple driving tasks with a one-boundary diffusion model". In: *Psychonomic bulletin & review* 21, pp. 577–589.

Robert, Lionel, Yang, X Jessie, Tilbury, Dawn, et al. (2021). "Automated vehicle behavior design for pedestrian interactions at unsignalized crosswalks". In: *Available at SSRN 3859366*.

Roshdi, Mohamed, Petzold, Julian, Wahby, Mostafa, Ebrahim, Hussein, Berekovic, Mladen, and Hamann, Heiko (2024). "On the Road to Clarity: Exploring Explainable AI for World Models in a Driver Assistance System". In: *2024 IEEE Conference on Artificial Intelligence (CAI)*. IEEE, pp. 1032–1039.

Rudin, Cynthia (2019). "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead". In: *Nature machine intelligence* 1.5, pp. 206–215.

Sackmann, Moritz (2024). *Learning Driver Behavior Models for Predicting Urban Traffic Situations*. Friedrich-Alexander-Universitaet Erlangen-Nuernberg (Germany).

Salvucci, Dario D (2006). "Modeling driver behavior in a cognitive architecture". In: *Human factors* 48.2, pp. 362–380.

Salvucci, Dario D, Boer, Erwin R, and Liu, Andrew (2001). "Toward an integrated model of driver behavior in cognitive architecture". In: *Transportation research record* 1779.1, pp. 9–16.

Sarcar, Sayan, Jokinen, Jussi PP, Oulasvirta, Antti, Wang, Zhenxin, Silpasuwanchai, Chaklam, and Ren, Xiangshi (2018). "Ability-based optimization of touchscreen interactions". In: *IEEE Pervasive Computing* 17.1, pp. 15–26.

Sawitzky, Tamara von, Grauschopf, Thomas, and Riener, Andreas (2023). "A flexible simulation environment for enhanced VRU research". In: *Adjunct Proceedings of the 15th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, pp. 283–286.

Schmidt, Lukas M, Brosig, Johanna, Plinge, Axel, Eskofier, Bjoern M, and Mutschler, Christopher (2022). "An introduction to multi-agent reinforcement learning and review of its application to autonomous mobility". In: *2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, pp. 1342–1349.

Schulman, John, Wolski, Filip, Dhariwal, Prafulla, Radford, Alec, and Klimov, Oleg (2017). "Proximal policy optimization algorithms". In: *arXiv preprint arXiv:1707.06347*.

Song, Qunying, Tan, Kaige, Runeson, Per, and Persson, Stefan (2023). "Critical scenario identification for realistic testing of autonomous driving systems". In: *Software Quality Journal* 31.2, pp. 441–469.

Spitale, Micol, Babu, Srikar, Cakmak, Serhan, Cheong, Jiaee, and Gunes, Hatice (2025). "Exploring Causality for HRI: A Case Study on Robotic Mental Well-being Coaching". In: *arXiv preprint arXiv:2503.11684*.

Srinivasan, Aravinda Ramakrishnan, Lin, Yi-Shin, Antonello, Morris, Knittel, Anthony, Hasan, Mohamed, Hawasly, Majd, Redford, John, Ramamoorthy, Subramanian, Leonetti, Matteo,

Billington, Jac, et al. (2023). "Beyond RMSE: Do machine-learned models of road user interaction produce human-like behavior?" In: *IEEE Transactions on Intelligent Transportation Systems*.

Svärd, Malin, Markkula, Gustav, Engström, Johan, Granum, Fredrik, and Bärgman, Jonas (2017). "A quantitative driver model of pre-crash brake onset and control". In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Vol. 61. 1. SAGE Publications Sage CA: Los Angeles, CA, pp. 339–343.

Tian, Kai, Markkula, Gustav, Wei, Chongfeng, Lee, Yee Mun, Madigan, Ruth, Hirose, Toshiya, Merat, Natasha, and Romano, Richard (2023). "Deconstructing pedestrian crossing decisions in interactions with continuous traffic: An anthropomorphic model". In: *IEEE Transactions on Intelligent Transportation Systems* 25.3, pp. 2466–2478.

Tian, Kai, Markkula, Gustav, Wei, Chongfeng, Lee, Yee Mun, Madigan, Ruth, Merat, Natasha, and Romano, Richard (2022). "Explaining unsafe pedestrian road crossing behaviours using a psychophysics-based gap acceptance model". In: *Safety science* 154, p. 105837.

Tian, Kai, Wei, Chongfeng, Lyu, Wei, Wang, Yueyang, Lee, Yee Mun, Merat, Natasha, Romano, Richard, and Markkula, Gustav (2025). "Interacting With Yielding Vehicles: A Perceptually Plausible Model for Pedestrian Road Crossing Decisions". In: *IEEE Transactions on Intelligent Transportation Systems*.

Tseng, Yuan-Chi and Howes, Andrew (2015). "The adaptation of visual search to utility, ecology and design". In: *International Journal of Human-Computer Studies* 80, pp. 45–55.

Turnwald, Annemarie, Althoff, Daniel, Wollherr, Dirk, and Buss, Martin (2016). "Understanding human avoidance behavior: interaction-aware decision making based on game theory". In: *International journal of social robotics* 8.2, pp. 331–351.

Turnwald, Annemarie, Olszowy, Wiktor, Wollherr, Dirk, and Buss, Martin (2014). "Interactive navigation of humans from a game theoretic perspective". In: *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, pp. 703–708.

Van Brummelen, Jessica, O'brien, Marie, Gruyer, Dominique, and Najjaran, Homayoun (2018). "Autonomous vehicle perception: The technology of today and tomorrow". In: *Transportation research part C: emerging technologies* 89, pp. 384–406.

Van Hasselt, Hado, Guez, Arthur, and Silver, David (2016). "Deep reinforcement learning with double q-learning". In: *Proceedings of the AAAI conference on artificial intelligence*. Vol. 30. 1.

Vaswani, Ashish, Shazeer, Noam, Parmar, Niki, Uszkoreit, Jakob, Jones, Llion, Gomez, Aidan N, Kaiser, Łukasz, and Polosukhin, Illia (2017). "Attention is all you need". In: *Advances in neural information processing systems* 30.

Vizzari, Giuseppe and Cecconello, Thomas (2022). "Pedestrian simulation with reinforcement learning: a curriculum-based approach". In: *Future Internet* 15.1, p. 12.

Von Neumann, John and Morgenstern, Oskar (2007). "Theory of games and economic behavior". In: *Theory of games and economic behavior*. Princeton university press.

Wang, Yueyang, Srinivasan, Aravinda Ramakrishnan, Jokinen, Jussi PP, Oulasvirta, Antti, and Markkula, Gustav (2025). "Pedestrian crossing decisions can be explained by bounded optimal decision-making under noisy visual perception". In: *Transportation Research Part C: Emerging Technologies* 171, p. 104963.

Wang, Ziyu, Ma, Jing, and Lai, Edmund MK (2024). "A Survey of Scenario Generation for Automated Vehicle Testing and Validation". In: *Future Internet* 16.12, p. 480.

Wang, Ziyu, Schaul, Tom, Hessel, Matteo, Hasselt, Hado, Lanctot, Marc, and Freitas, Nando (2016). "Dueling network architectures for deep reinforcement learning". In: *International conference on machine learning*. PMLR, pp. 1995–2003.

Xu, Yiran, Yang, Xiaoyin, Gong, Lihang, Lin, Hsuan-Chu, Wu, Tz-Ying, Li, Yunsheng, and Vasconcelos, Nuno (2020). "Explainable object-induced action decision for autonomous vehicles". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 9523–9532.

Yang, Changzhi, Pan, Huihui, and Wang, Jue (2025). "Interact, Plan, and Go: Transformers With Social Intentions for Trajectory Prediction". In: *IEEE Transactions on Consumer Electronics*.

Yang, Dongfang, Özgüner, Ümit, and Redmill, Keith (2020). "A social force based pedestrian motion model considering multi-pedestrian interaction with a vehicle". In: *ACM Transactions on Spatial Algorithms and Systems (TSAS)* 6.2, pp. 1–27.

Yang, Yuhang, Kujanpää, Kalle, Babadi, I Amin, Pajarinen, Joni, and Ilin, Alexander (2023). "Suicidal pedestrian: Generation of safety-critical scenarios for autonomous vehicles". In: *2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, pp. 1983–1988.

Ye, Luyao, Wang, Zezhong, Chen, Xinhong, Wang, Jianping, Wu, Kui, and Lu, Kejie (2021). "GSAN: Graph self-attention network for learning spatial–temporal interaction representation in autonomous driving". In: *IEEE Internet of Things Journal* 9.12, pp. 9190–9204.

Yi, Shuai, Li, Hongsheng, and Wang, Xiaogang (2016). "Pedestrian behavior understanding and prediction with deep neural networks". In: *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14*. Springer, pp. 263–279.

Yin, Ziyi, Liu, Ruijin, Xiong, Zhiliang, and Yuan, Zejian (2021). "Multimodal Transformer Networks for Pedestrian Trajectory Prediction." In: *IJCAI*, pp. 1259–1265.

Yu, Yong, Si, Xiaosheng, Hu, Changhua, and Zhang, Jianxun (2019). "A review of recurrent neural networks: LSTM cells and network architectures". In: *Neural computation* 31.7, pp. 1235–1270.

Yuan, Ye, Weng, Xinshuo, Ou, Yanglan, and Kitani, Kris (2021). "AgentFormer: Agent-Aware Transformers for Socio-Temporal Multi-Agent Forecasting". In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*.

Zeng, Weiliang, Chen, Peng, Nakamura, Hideki, and Iryo-Asano, Miho (2014). "Application of social force model to pedestrian behavior analysis at signalized crosswalk". In: *Transportation research part C: emerging technologies* 40, pp. 143–159.

Zhang, Xingchen, Angeloudis, Panagiotis, and Demiris, Yiannis (2022). "ST CrossingPose: A Spatial-Temporal Graph Convolutional Network for Skeleton-Based Pedestrian Crossing Intention Prediction". In: *IEEE Transactions on Intelligent Transportation Systems* 23.11, pp. 20773–20782.

Zhou, Jibiao, Guo, Yanyong, Dong, Sheng, Zhang, Minjie, and Mao, Tianqi (2019). "Simulation of pedestrian evacuation route choice using social force model in large-scale public space: Comparison of five evacuation strategies". In: *PLoS one* 14.9, e0221872.

Chapter 2

Pedestrian crossing decisions can be explained by bounded optimal decision-making under noisy visual perception

Abstract

This paper presents a model of pedestrian crossing decisions based on the theory of computational rationality. It is assumed that crossing decisions are boundedly optimal, with bounds on optimality arising from human cognitive constraints. While previous models of pedestrian behaviour have been either 'black-box' machine learning models or mechanistic models with explicit assumptions about cognitive factors, we combine both approaches. Specifically, we mechanistically model noisy human visual perception and model reward considering human constraints in crossing, but we use reinforcement learning to learn boundedly optimal behaviour policy. The model reproduces a larger number of known empirical phenomena than previous models, in particular: (1) the effect of the time to arrival of an approaching vehicle on whether the pedestrian accepts the gap, the effect of the vehicle's speed on both (2) gap acceptance and (3) pedestrian timing of crossing in front of yielding vehicles, and (4) the effect on this crossing timing of the stopping distance of the yielding vehicle. Notably, our findings suggest

that behaviours previously framed as 'biases' in decision-making, such as speed-dependent gap acceptance, might instead be a product of rational adaptation to the constraints of visual perception. Our approach also permits fitting the parameters of cognitive constraints and rewards per individual to better account for individual differences, achieving good quantitative alignment with experimental data. To conclude, by leveraging both RL and mechanistic modelling, our model offers novel insights into pedestrian behaviour and may provide a useful foundation for more accurate and scalable pedestrian models.

2.1 Introduction

2.1.1 Background

Pedestrian crossing behaviour is a key factor in the urban transport system, impacting traffic efficiency and safety (Leu et al., 2012). Developing accurate models for pedestrian crossing decisions is therefore essential for effective urban planning (González-Méndez et al., 2021; Pelorosso, 2020). In particular, with the rise of autonomous vehicles (AVs), understanding pedestrian behaviour is important for the safe and harmonious integration of AVs into the existing transport ecosystem (González-Méndez et al., 2021; Pelorosso, 2020; Duric et al., 2002; Banovic et al., 2019; Al-Shihabi and Mourant, 2001; Li et al., 2016). The unpredictable nature of pedestrian movements, influenced by a complex interplay of cognitive processes and environmental cues, presents a significant challenge in this regard (Camara et al., 2020; Crosato et al., 2023).

In pedestrian behaviour modelling, computational models are predominantly based on either mechanistic modelling or data-driven machine learning (ML) approaches. Mechanistic models, which are grounded in assumptions about underlying cognitive mechanisms, aim to accurately represent the underlying processes causing pedestrian behaviour (Fajen and Warren, 2003; Turnwald et al., 2016; Yang et al., 2020; Pekkanen et al., 2021). However, formulating these mechanisms in a way that generalises across a wide variety of situations and scenarios is highly non-trivial. Therefore, since real-world traffic situations are highly variable and complex, mechanistic models typically suffer from scalability or generalisability issues. On the other hand, data-driven ML models, which predict pedestrian movement by learning from large datasets of pedestrian trajectories, show promise in handling complex real-world situations (Dai et al., 2019; Abughalieh and Alawneh, 2020; Quan et al., 2021; Yin et al., 2021; Yuan et al., 2021; Zhang et al., 2022; Li et al., 2022). However, ML models come with inherent challenges related to interpretability, data dependence, and robustness in diverse conditions (Althoff and Lutz, 2018; Klischat et al., 2020). These issues will be discussed in more detail in Section 2.1.2.

Ideally, one would want a modelling approach combining the strengths of both mechanistic and ML models. One possible candidate for such an approach is available from *computational rationality* (Lewis et al., 2014; Gershman et al., 2015), sometimes also referred to as resource rationality (Lieder and Griffiths, 2020). Both computational rationality and resource rationality are theories about human cognition and behaviour, positing that humans behave *rationally*, i.e.,

optimally with respect to some utility or cost function, but that this optimality is *bounded* by the constraints imposed by the human cognition and body. For example, pedestrians in road traffic have perceptual, cognitive, and motor limits which constrain their behaviour. Rooted in the principle of expected utility maximisation, introduced by Von Neumann and Morgenstern (1947), and expanded by Herbert Simon's concept of bounded rationality (Simon, 1955), computational rationality applies these decision-making principles while considering human constraints. In recent years, the development of models based on these theories has become increasingly attainable due to advances in the field of deep reinforcement learning (RL), enabling the learning of optimal decisions through environmental interactions (Silver et al., 2016; Silver et al., 2017; Sutton and Barto, 2018). Modern deep RL allows us to find boundedly optimal behaviour policy across highly varied scenarios and learn interactive behaviours which-to the extent that the theory of computational rationality is correct-mirror human-like decision processes and behavioural adaptability (Jokinen et al., 2020; Jokinen et al., 2021; Chen et al., 2021a; Chen et al., 2021b). [1] Contrary to traditional mechanistic models, which often struggle with flexibility in adapting to diverse scenarios and integration of various principles, and ML models, which can lack interpretability and adaptability, the computational rationality framework, when combined with RL, has the potential to overcome these constraints (Oulasvirta et al., 2022).

There have been some initial studies on the use of computational rationality in driver behaviour modelling (Jokinen et al., 2021; Jokinen et al., 2022), but so far, none in the context of vulnerable road users. Here, we provide a first demonstration of the value of this modelling framework in the context of pedestrian behaviour. Unlike previous pedestrian behaviour models, which are either mechanistic models, suffering from scalability or generalisability issues due to the complexities in real-world traffic conditions, or data-driven machine learning models, with challenges related to interpretability and data dependence, computational rationality addresses these issues by explaining human behaviour through a causal link between subjective utility, capacities, experience, and observed actions (Howes et al., 2023), optimising utility under defined constraints. The main human constraint we are considering in this work is imperfect human visual perception, characterised by the noise inherent in the visual system's processing of dynamic stimuli. This constraint often leads to mistakes in how people estimate the distance and

[1] Here, policy refers to the agent's decision rule, that is, a mapping from its current state or observation to an action.

speed of oncoming traffic, which we model mechanistically based on work in cognitive neuroscience. Looming aversion, which refers to the instinctive tendency to avoid objects that appear to be rapidly growing in size as a cue for an impending collision (DeLucia, 2008; Tian et al., 2022), is also considered in our model. We test our model on data from a controlled virtual reality experiment on pedestrian road-crossing, and find that our model captures four phenomena observed in the human experiment, which have not been captured simultaneously by one model. Furthermore, our model results also show that speed-dependent gap acceptance–which has previously been described as due to biases in human TTA estimation (Petzoldt, 2014)–is a boundedly optimal behaviour given the particular noise characteristics of the human visual system. Moreover, we show how individual differences in road user behaviour can be efficiently modelled using the computational rationality approach by conditioning the RL on the human constraint parameters. This method, which has not been previously used in road user behaviour modelling, includes parameters which can vary between humans as inputs to the RL, to learn optimal policy across variations in these parameters.

This paper is organised as follows: The rest of this section describes empirical studies of the behavioural phenomena we wish to capture and previous research on pedestrian behaviour modelling. Section 2.2 describes the dataset we used and the modelling approach. The results of our proposed model are presented in Section 2.3, followed by Section 2.4, with a discussion of results and future research plans. Finally, Section 2.5 provides a conclusion.

2.1.2 Related work

Numerous empirical studies have investigated the key factors influencing pedestrian behaviour (Oxley et al., 2005; Jain et al., 2014; Sun et al., 2015; Asaithambi et al., 2016; Gorrini et al., 2018; Tian et al., 2022). We focus on four main empirical phenomena, some but not all of which have been captured by existing models:

(1) TTA-dependent gap acceptance: A key behaviour of interest is *gap acceptance*, where pedestrians decide to cross a road based on the time or spatial distance between them and an approaching vehicle. One crucial factor affecting this decision is the *time to arrival (TTA)*. Researchers have shown that pedestrians are more inclined to accept gaps for crossing when the TTA is higher (Oxley et al., 2005; Lobjois and Cavallo, 2007; Petzoldt, 2014).

(2) Speed-dependent gap acceptance: Additionally, the dynamics of approaching vehicles play

an important role (Schneemann and Gohl, 2016). Pedestrians are more likely to cross in a given TTA when oncoming vehicles travel at higher speeds, i.e., with a larger approach distance (Lobjois and Cavallo, 2007; Tian et al., 2022). This phenomenon has been described as a bias in the human estimation of TTA (Petzoldt, 2014; Sun et al., 2015). Tian et al. (2022) explained this speed-dependent crossing behaviour by visual looming, the perceived growth of an object's size as it approaches (DeLucia, 2008). Specifically, visual looming increases slowly at long distances, indicating that higher-speed vehicles might produce smaller collision threats to pedestrians for a given TTA, thus influencing pedestrians to feel safer when crossing in front of faster vehicles.

(3) Speed-dependent yielding acceptance: In scenarios where vehicles yield, the relationship between speed and pedestrian crossing behaviour tends to reverse. Pedestrians have been found to interpret lower vehicle speeds as an indication of yielding, especially when the vehicle is at a closer distance, thus increasing the probability of crossing (Tian et al., 2023). On the other hand, early crossing decisions (while the vehicle is still some distance away) are similar to the speed-dependent gap acceptance regardless of whether the car is yielding.

(4) Stopping distance-dependent yielding acceptance: Driver use of exaggerated deceleration, or 'short-stopping', serves as a cue for pedestrians to cross (Domeyer et al., 2019). Short-stopping enhances the pedestrian's perception of the driver's intent, increasing their confidence in crossing safely (Domeyer et al., 2019). Risto et al. (2017) argued that the stopping distance of a vehicle correlates with pedestrian willingness to cross in the yielding scenario. Notably, Tian et al. (2023) found that the car's braking behaviour mainly influences late crossing decisions. Specifically, pedestrians tend to cross more readily when a vehicle stops at a greater distance from the crosswalk, seemingly interpreting this as a clear indication of the driver's intent to yield.

In our study, we aim to model pedestrian behaviour and reproduce these four phenomena mentioned above, summarised in Table 2.1. These phenomena provide valuable insights and highlight some of the nuances of human decision-making in traffic interactions that computational models should capture to be useful for practical applications such as simulation environments and AV algorithms. As mentioned, there are many models which describe pedestrian behaviour in terms of hypothesised underlying mechanisms (Fajen and Warren, 2003; Giles et al., 2019; Pekkanen et al., 2021; Tian et al., 2022; Tian et al., 2023). For example, Petzoldt (2014) devel-

oped a logit-based gap acceptance model capturing speed-dependent gap acceptance, described as a bias due to a visual heuristic used by pedestrians. However, this heuristic is truly biased in the sense that it is suboptimal or whether it might be an optimal adaptation to human constraints. Later, Tian et al. (2022) proposed a gap acceptance model utilising the binary choice logit approach. Unlike previous models that relied on traffic gap cues, their model incorporated visual looming cues, resulting in an improved fit to observed data. Another modelling approach is based on the concept of evidence accumulation, describing decision-making as a noisy accumulation of evidence from a stimulus (Ratcliff et al., 2016), which has also been applied to pedestrian crossing behaviour modelling (Pekkanen et al., 2021). However, due to the complexity of this model, the authors fitted it with a single parameterisation across all participants and with limited quantitative goodness of fit to the experimental data. The complexity of this model also makes it difficult to extend it to more sophisticated scenarios. To extend the scope of modelled scenarios, Markkula et al. (2023) integrated a large number of existing psychological and cognitive theories, such as theories of sensory noise, Bayesian perception, evidence accumulation decision-making, and long-term valuation of action affordances. Although this model was capable of reproducing several empirically observed phenomena in human road user interaction, the authors highlighted the limitations of mechanistic modelling and underscored the need for cognitively and behaviourally informed ML. Overall, given the high complexity of human behaviour, relying on a single cognitive, mechanistic model seems insufficient to comprehensively describe pedestrian behaviour. The existing models' success at capturing the targeted phenomena is listed in Table 2.1. However, it is worth noting that the aim of our work is not only to capture these phenomena but also to propose a general framework for modelling pedestrian

Table 2.1: Key phenomena in pedestrian behaviour and corresponding literature

Phenomenon	**Petzoldt (2014) and Tian et al. (2022)**	**Pekkanen et al. (2021)**	**This paper**
(1) TTA-dependent gap acceptance	✓	✓	✓
(2) Speed-dependent gap acceptance	✓	✓	✓
(3) Speed-dependent yielding acceptance			✓
(4) Stopping distance-dependent yielding acceptance		✓	✓

behaviour.

With advances in computing power and machine-learning methods, a growing number of researchers have adopted data-driven ML algorithms for modelling pedestrian behaviour. For instance, many models based on convolutional neural networks have been developed for pedestrian trajectory prediction due to their ability to process spatial inputs, such as images and video frames. (Yi et al., 2016; Abughalieh and Alawneh, 2020; Kumamoto and Yamada, 2017; Doellinger et al., 2018). Moreover, recurrent neural networks have garnered attention given their efficacy in sequence prediction tasks, making them apt for pedestrian trajectory modelling (Alahi et al., 2016; Dai et al., 2019; Quan et al., 2021). Recently, the Transformer model, known for its ability to process sequences in parallel and capture long-range interactions, has also gained popularity in pedestrian behaviour modelling (Lorenzo et al., 2021; Yin et al., 2021; Yuan et al., 2021). While the accuracy of these ML models can be impressive, they share several limitations. First, they often act as 'black boxes', meaning they may have the potential to predict behaviours but fail to reveal the reasons or mechanisms behind those behaviours. This intensive data-driven orientation can be challenging, especially when aiming for interpretability in AV systems, where understanding pedestrian intent is crucial (Srinivasan et al., 2023). Second, the lack of theoretical grounding of the behaviour can sometimes lead to overfitting, where the models might perform well on the training data but fail to generalise well to unseen data, potentially resulting in less reliable predictions in varied real-world scenarios (Xu et al., 2020). Third, their heavy reliance on large training datasets poses another challenge: collecting such extensive data under all possible road conditions is almost impossible, and critical scenarios are missing in most datasets (Diaz-Ruiz et al., 2022).

As mentioned in Section 2.1.1, computational rationality provides a potential middle ground, combining the best of mechanistic and ML models, but has only seen limited use in road user behaviour modelling. Jokinen et al. (2021) proposed a hierarchical RL model of multitasking behaviour in driving and fitted it to observed data (Jokinen et al., 2022). However, their focus was on in-vehicle multitasking rather than pedestrian-vehicle interactions.

2.1.3 Summary of research gaps

Existing models of pedestrian crossing behaviour fail to capture the four mentioned critical empirical phenomena simultaneously, and neglect essential human perceptual and cognitive

constraints like noisy perception and looming aversion. Furthermore, no existing model combines mechanistic and machine learning approaches, or accounts for individual differences in pedestrian behaviour.

Here, we aim to quantitatively model the crossing decisions of human pedestrians when faced with constant-speed and yielding vehicles. There are three research objectives in this study:

(1) To develop a boundedly optimal model, integrating human constraints, with a focus on noisy perception and looming aversion, that is able to capture the four targeted empirical phenomena in human pedestrian crossing behaviours in scenarios where vehicles may or may not yield to the pedestrian.

(2) To fit the boundedly optimal model quantitatively to human crossing decision data on a per-individual level (i.e., with different model parameter values for different individuals).

(3) To assess how the human constraints and preferences integrated into the model influence the crossing decisions.

An early version of this work has been presented as a conference paper (Wang et al., 2023). This conference paper qualitatively explored the effect of only noisy visual perception on crossing decisions and only in scenarios with approaching vehicles at constant speed; in other words, it only partially addressed the first objective above, and not at all the second or third objectives.

2.2 Methods

2.2.1 Dataset

This study utilised a dataset sourced from a previous experiment reported by Giles et al. (2019). A visualisation of the experimental setup is provided in Figure 2.1, which offers a bird's-eye perspective. Twenty participants (age 24-60, average 27.9 years; 11 males and 9 females) took part in the study and were recruited from a university participant pool. We did not record the racial or ethnic identities of our participants in this study. Regarding the age influence, it is important to note that the experimental task only required button press responses, such that participant variation in motor ability was less of a concern than it would have been in a task requiring actual walking. Participants wore an HTC Vive Virtual Reality (VR) headset as part of the setup, immersing them in a virtual crossing task. The rendered VR space included a

straight two-lane road, spanning a width of 5.85 m, and a zebra crossing at the participant's initial location.

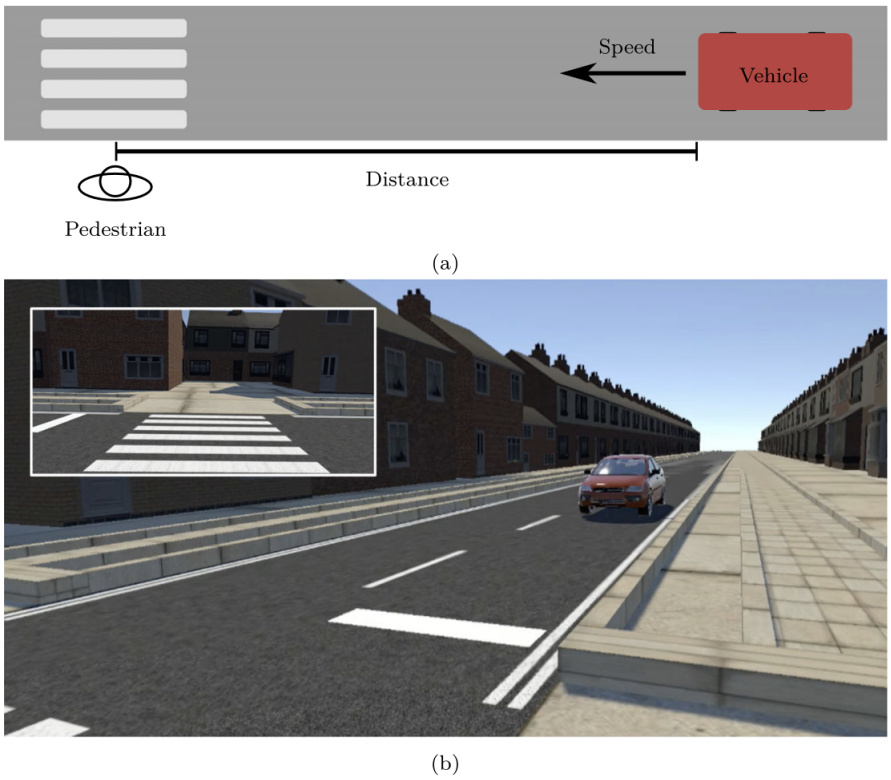

Figure 2.1: (a) Birds-eye view of the experiment. (b) A sample view of the virtual scene, as shown by the head-mounted display, at the beginning of each trial (inset) and as participants turned their heads to look for oncoming traffic. (Source: Pekkanen et al. (2021). This image is available under a Creative Commons Attribution 4.0 International License.)

Table 2.2: Detailed description of vehicle approach scenarios in the experiment. The table lists each scenario type along with key parameters: Initial Vehicle Speed (v_0 in m/s), Initial Distance (d_0 in m) from the pedestrian, Initial TTA (τ_0 in s), and Stopping Distance (d_{stop} in m) for Yielding scenarios. 'N/A' indicates not applicable for constant speed scenarios.

Scenario type	v_0 (m/s)	d_0 (m)	τ_0 (s)	d_{stop} (m)
Constant speed	6.94	15.90	2.29	N/A
	13.89	31.81	2.29	N/A
	6.94	31.81	4.58	N/A
	13.89	63.61	4.58	N/A
	6.94	47.71	6.87	N/A
	13.89	95.42	6.87	N/A
Yielding	6.94	15.90	2.29	4
	13.89	31.81	2.29	4
	13.89	31.81	2.29	8
	6.94	31.81	4.58	4
	13.89	63.61	4.58	4
	13.89	63.61	4.58	8
	6.94	47.71	6.87	4
	13.89	95.42	6.87	4

At the start of each trial, participants were positioned before the zebra crossing, facing straight across it. They had been instructed that they should look to the right for any oncoming traffic when they felt ready to begin the trial. This head movement triggered the start of the trial, with an oncoming vehicle initialised at an initial distance d_0 and speed v_0. The experiment included a mix of scenarios, namely constant-speed and yielding scenarios. The detail of these scenarios is shown in Table 2.2, with the initial TTA $\tau_0 = d_0/v_0$ also listed. In six constant-speed scenarios, the vehicle appeared at a distance d_0 (all distances measured longitudinally along the road from the participant's location to the front of the car) and maintained a constant-speed v_0 while approaching and passing the zebra crossing. In eight yielding scenarios, the vehicle appeared at initial distance and speed v_0 and d_0, and immediately decelerated at a constant rate to stop at a distance d_{stop} from the participant. Participants were instructed to press the HTC Vive's controller button when they felt safe to cross. It should be noted that the vehicle's behaviour was not influenced by the pedestrian's decision. This setup required participants to assess crossing safety independently. Once the button was pressed, the system recorded the *Crossing Initiation Time (CIT)*, defined as the interval between the start of the trial and the moment when the button press occurred. At this point, the position of the participant's point of view in the virtual environment (i.e., the VR 'camera') moved across the zebra crossing at the speed of 1.31 m/s. Each participant faced 6 unique constant-speed trials and 8 unique yielding trials in a randomised order, aggregating to a total of 280 trials that were used for the model's validation.

The low number of repetitions per participant was adopted by Giles et al. (2019) to limit behavioural adaptation effects. The resulting dataset is relatively small in size. However, since all of the main phenomena we target have also been reported in other experiments, the sample size is not a main concern here. We adopted this dataset here because of the button-press paradigm, which provides crossing onset distributions with minimal impact from motor variability, aligning with our present modelling emphasis on perception rather than motor control.

2.2.2 Model variants

In this section, we explain our two main mechanistic assumptions, rooted in cognitive science and neuroscience, and introduce different model variants derived from these assumptions:

- **Assumption of noisy perception:** We assume that the agent's perception of the envi-

ronment is inherently noisy and imperfect. This noisy observation obtained by the agent is according to the principle of the human visual system, i.e., the sensory input received by our human visual system is noisy (Faisal et al., 2008).

- **Assumption of looming aversion:** Our second assumption involves the natural aversion to looming objects (DeLucia, 2008; Tian et al., 2022). Looming aversion refers to the instinctive tendency to avoid objects that appear to be rapidly growing in size, as this is often a cue for an impending collision (DeLucia, 2008; Tian et al., 2022).

These two assumptions guide the development of our model variants. Each variant incorporates the assumptions in different ways to explore how they individually and collectively influence pedestrian behaviour. The following subsection outlines the four main model variants derived from these assumptions, as visualised in Figure 2.2:

(1) **Baseline Model (BM):** The BM serves as our control variant, assuming an ideal observer with neither visual constraints nor looming aversion. This model establishes a baseline against which to measure the impact of our two key assumptions.

(2) **Looming Model (LM):** The LM considers the impact of looming aversion on crossing decisions. It allows us to understand how this aversion, independent of visual constraints, can affect pedestrian behaviour.

(3) **Visual constraints Model (VM):** Conversely, the VM examines the role of visual constraints alone. It helps us investigate how the noisy sensory information influences pedestrian crossing.

(4) **Visual constraints and Looming Model (VLM):** The VLM combines both visual constraints and looming aversion to provide insights into their joint effect.

Each combination of model assumptions mentioned above defines an RL problem. In the next section, we describe this RL problem in detail for the different model variants.

2.2.3 Reinforcement learning problem

Our model based on computational rationality used RL to derive the near-optimal behaviour under constraints. In RL, the agent chooses the action by following a policy π, which yields a probability $\pi(s, a) = p(a \mid s)$ of taking a particular action from the given state. The optimal

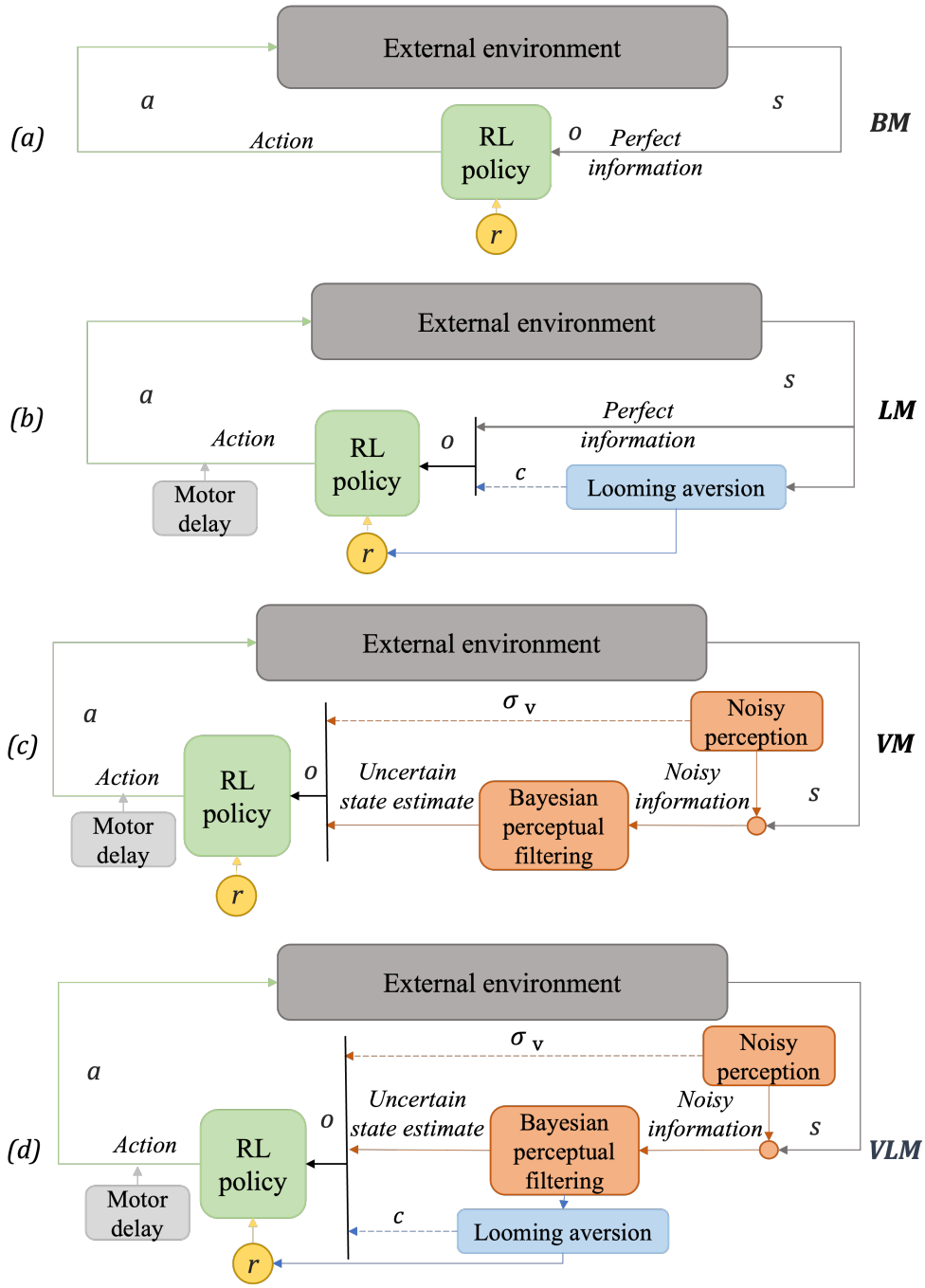

Figure 2.2: Comparison of models. (a) BM: Baseline Model. (b) LM: Looming aversion only. (c) VM: Visual constraints only. (d) VLM: Visual constraints and looming aversion. σ_{v} and c represent the sensory noise and looming aversion weight respectively. s, o, a, and r represent the state, observation, action and reward respectively.

policy π^*, maximises the value function:

$$V^*(s) = \max_a \left[R(s,a) + \gamma \sum_{s' \in S} T(s,a,s') V^*(s') \right] \tag{2.1}$$

where γ is the discount factor, a value between 0 and 1, which balances the importance of

immediate rewards versus future ones. Here, $V^*(s)$ represents the optimal value function at state s, and $V^*(s')$ represents the optimal value at state s'.

In our study, we assumed that humans make decisions under uncertainties from the noisy perception system, such that the overall task can be understood as a Partially Observable Markov Decision Process (POMDP). In a POMDP, the agent does not have direct access to the actual state S; instead, 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.

State space S At each time step t, the environment is in a state $s_t \in S$. This is the positions and velocities of the pedestrian (the RL agent) and the vehicle. All model variants in our study share the same state space, encapsulating these five critical variables: the pedestrian's position $x_{\mathrm{p}}, y_{\mathrm{p}}$, the vehicle's position $x_{\mathrm{veh}}, y_{\mathrm{veh}}$, the vehicle's velocity v, and time step t. The simulation state update time step is set to 0.1 s, a duration suitable for our dataset and model.

Action space A At every time step t, the agent executes an action a_t from the set A. For the purposes of this study, and consistent with the experimental button press, the agent's choices are binary: either to 'Go' or to 'Not Go'. Upon selecting 'Go', the agent proceeds straight at the speed of 1.31 m/s, as in the experiment. Recognising that human reaction times vary, we introduced a motor delay to simulate this aspect. This delay, implemented after the 'Go' decision, was sampled from a Gaussian distribution with a mean of 0.6 s and a standard deviation of 0.2 s. After the decision was made, the simulation progressed until the agent either crossed the road safely or a collision happened.

Transition T The transition function defines how the current state s_t changes to the next state s_{t+1} based on the action a_t. In our model, when the 'Not Go' action is selected, the vehicle's movement follows kinematic equations with the given speed, and the agent's position remains unchanged. Conversely, when the 'Go' action is chosen, whether the collision happens is calculated, and the corresponding reward is given to the agent. Then, the simulation finishes.

Reward R In the experiment where the dataset was collected, the participant's task was to cross the road as soon as they felt safe to do so, either before or after the car had passed

them (Giles et al., 2019). Therefore, we assume that the participants in this experiment (presumably similar to pedestrians in real traffic) wanted to cross safely but with minimal loss of time. We designed the reward structure accordingly: The agent will be given a reward of 20 when crossing the road without collision and a reward of -20 if a collision happens. These reward parameters were shared across model variants. After some initial manual testing, they were fixed to yield reasonable crossing behaviour of the BM agent but before structured fitting of any other model parameters to the human data. A time penalty, a negative reward of $0.01 \times t$, will also be given to the agent when the episode terminates.[2]

Regarding the visual looming, this phenomenon is represented mathematically as inverse τ—the ratio of a vehicle's optical expansion rate to its size on the observer's retina, which serves as an estimate of the inverse TTA (DeLucia, 2015; Markkula et al., 2016). We incorporated this concept into the reward function of the *LM* and *VLM* models, as shown in panels (b) and (d) of Figure 2.2, to account for the looming aversion in pedestrian decision-making.

The reward function r is defined as follows:

$$r = \begin{cases} \max(-20, \min(+20, 20 - 0.01 \cdot t - c \cdot \frac{1}{\hat{\tau}})), & \text{if arrival} \\ -20, & \text{if collision} \end{cases} \tag{2.2}$$

where c is the weight of the looming aversion. For model variants with non-noisy perception, $\tau = \frac{x_{\text{veh}} - x_{\text{p}}}{v}$, and for those with noisy perception, $\hat{\tau} = \frac{\hat{x}_{\text{veh}} - x_{\text{p}}}{\hat{v}}$, where $\hat{x}_{\text{veh}}$ and $\hat{v}$ are noisy estimates of vehicle position and speed (see further below). The reward r is bounded within the range $[-20, +20]$ to prevent extreme values from influencing the model training. As the focus of this study is on the potential effect of noisy perception and the looming aversion to the crossing decision, we have kept the reward function simple; future work can further refine it to better capture human preferences.

Observation space O The agent receives observation $o_t \in O$ at each time step. We tested different formulations of the observation space: In *BM*—our simplest, baseline model as shown in panel (a) of Figure 2.2—the exact position and velocity of both the vehicle and the pedestrian are presented as inputs to the agent. Conversely, in the models with visual constraints

[2]Further information not present in the published article: the baseline reward values were later examined through a reward sensitivity analysis across all four model variants. This analysis, reported in Appendix A.1, showed that the baseline setting of $R = 20$ and $tp = 0.01$ lies within a stable and behaviourally reasonable region, and also clarified how changes in these values affect model behaviour.

assumption, *VM* and *VLM*, illustrated in panels (c) and (d) of Figure 2.2, the position and velocity details provided to the agent are subject to noise, emulating the inherent uncertainty and imperfection in human visual perception, with an angular noise introduced at the level of the human retina (Kwon et al., 2015). We assumed that the agent observes the position of the other agent along its line of travel by observing the angle below the horizon of the other agent (Ooi et al., 2001; Markkula et al., 2023), with a constant Gaussian angular noise of standard deviation, σ_{v}, which could vary between pedestrians. In practice, this means that the pedestrian observes the vehicle's distance along the road with a distance-dependent noise of standard deviation (Markkula et al., 2023):

$$\sigma_x(t) = |d_{\mathrm{l}}(t)| \left(1 - \frac{h}{d(t) \cdot \tan(\arctan \frac{h}{d(t)} + \sigma_{\mathrm{v}})}\right), \tag{2.3}$$

where $d_{\mathrm{l}}(t)$ is the longitudinal distance between the vehicle and the crossing point, $d(t)$ is the distance between the agent and the approaching vehicle, and h is the eye height over the ground of the ego agent, which is set to 1.6 m for all pedestrian agents for simplicity.

Additionally, there is evidence that the human perception system interprets its noisy input in a Bayes-optimal manner, and Bayesian methods have been successful in modelling perception and sensorimotor control (Kwon et al., 2015; Knill and Pouget, 2004; Stocker and Simoncelli, 2006). Therefore, we used a Kalman filter as a model of the human visual perception to perceive the environment (Kwon et al., 2015; Markkula et al., 2023). In the initialisation of the Kalman filter, we used distinct prior distributions for position and velocity. Each distribution is centred on the actual vehicle state, with its standard deviation set equal to that of the initial set of values. At each time step, the Kalman filter receives noisy positional data regarding the other agent (vehicle). The filter then produces estimates of the vehicle's position and velocity, along with their respective uncertainties. These filtered estimates represent the agent's belief state, reflecting the Bayes-optimal inference of the vehicle's state based on the noisy observations. These belief estimates are then fed into the RL policy as inputs, together with the precise self-position and velocity of the agent itself and the visual noise parameter σ_{v}.

Furthermore, the *LM* and *VLM* models incorporate the concept of visual looming. This looming aversion, modelled through inverse τ as shown in Equation 2.2, affects the agent's perception of potential collision risks. We gave the weight for the looming aversion, c, as input to the RL policy.

Table 2.3: Observation space variables for each model variant. '✓' indicates the variable is directly observed. Here, $x_{\mathrm{p}}, y_{\mathrm{p}}$ denote the pedestrian's position, $x_{\mathrm{veh}}, y_{\mathrm{veh}}$ the vehicle's position, v the vehicle's velocity, $\hat{x}_{\mathrm{veh}}, \hat{v}$ the Kalman filter's estimates of the vehicle's position and velocity, P_{p} and P_{v} are the variances in the Kalman filter's estimate of position and speed of the approaching vehicle, σ_{v} represents the standard deviation of sensory noise, c the looming aversion weight, and t the time step.

Model	x_{p}	y_{p}	x_{veh}	y_{veh}	v	Additional Variables
BM	✓	✓	✓	✓	✓	t
LM	✓	✓	✓	✓	✓	c, t
VM	✓	✓	Est.($\hat{x}_{\mathrm{veh}}$)	Est.($\hat{y}_{\mathrm{veh}}$)	Est.($\hat{v}$)	$P_{\mathrm{p}}, P_{\mathrm{v}}, \sigma_{\mathrm{v}}, t$
VLM	✓	✓	Est.($\hat{x}_{\mathrm{veh}}$)	Est.($\hat{y}_{\mathrm{veh}}$)	Est.($\hat{v}$)	$P_{\mathrm{p}}, P_{\mathrm{v}}, \sigma_{\mathrm{v}}, c, t$

To distinguish σ_{v} and c from the parameters of the policy neural network (connection weights and biases), we will refer to these two parameters as non-policy parameters. By feeding the non-policy parameters as inputs to the RL policy, we are not implying that the human agent 'observes' its own parameter values. Instead, we are just conditioning the RL policy on these non-policy parameters, as a more convenient alternative to learning entirely different policies for different parameter combinations.

2.2.4 Reinforcement learning algorithm

Deep Q-Networks (DQNs) optimise the Q-function to estimate the expected utility of actions in given states, balancing immediate and future rewards. The Q-function is defined as:

$$Q(s, a) = r + \gamma \max_{a'} Q(s', a'), \tag{2.4}$$

where s represents the current state, a the action taken in state s, and r the immediate reward received after taking action a in state s. The DQN algorithm is suitable for problems with a continuous state space and a discrete action space, such as in our case. We adopted Dueling DQN, which enhances standard DQN by separating the action selection from evaluation with its unique structure of a value function $V(s)$ and an action advantage function $A(s, a)$. This allows for more precise evaluations in scenarios where actions have similar outcomes, improving decision-making accuracy by focusing on both state value and action advantages (Van Hasselt et al., 2016; Wang et al., 2016).

We represented the $Q(s, a)$ values using a fully connected feedforward neural network with two hidden layers of 512 and 256 nodes. It has been observed that larger networks tend to yield more stable policies, and policies are more robust to noise (Xie et al., 2019). The learning rate

and discount factor were set to 0.0001 and 0.99, respectively. To encourage exploration, we implemented an ϵ - greedy algorithm: at each time step t, the system either randomly selects an action with probability ϵ or chooses the action with the highest Q value with probability $1 - \epsilon$ (Wunder et al., 2010). We initialised ϵ at 1 and decreased it by 5^{-5} in each learning step. The minimum value of ϵ was set to 0.001. We trained model M, VM and LM over 25,000 episodes for convergence. Convergence here means that rewards stabilise and no longer increase over time. The choice of this criterion is rooted in the understanding that the model has likely learned an optimal or near-optimal policy once the rewards stop improving significantly. For the model VLM, we extended the training to 45,000 episodes due to its added complexities from both visual constraints and looming aversion assumptions.

To avoid the agent learning a simplistic strategy of always crossing immediately, we also included scenarios with an initial TTA of 1 s, in which safe crossing was only feasible after the vehicle's arrival.

2.2.5 Fitting of non-policy parameters

Our most complex model variant, VLM, has two free model parameters, σ_{v} and c. Testing different values for these non-policy parameters is essential because they alter the environment in the RL problem, thus influencing the agent's crossing decision.

Learning a separate RL policy for each possible combination of non-policy parameter values is computationally expensive (Howes et al., 2023; Li et al., 2023). Recent studies have adopted a more efficient approach within the computational rationality framework to address this challenge. This method involves conditioning the RL policy on non-policy parameters by integrating them as additional inputs to the model (Keurulainen et al., 2023). In this case, we provided σ_{v} and c as additional inputs to the RL policy during learning, as illustrated by the dashed line in Figure 2.2 (d). Utilising this approach, the model can learn a boundedly optimal policy for arbitrary values of the non-policy parameters.

We do not know the correct values for the noise magnitude parameter, σ_{v}, and the weight for the looming aversion, c, during the training phase. Moreover, these values might differ for individual participants in the experiment. Therefore, during RL training, for each episode we sampled values for these parameters from a range of σ_{v} values (from 0 to 1 in increments of 0.1) and c values (from 0 to 100 in increments of 10). Then, after RL training was complete,

we tested this learned RL policy across the entire 10×10 grid of non-policy parameters to make predictions about model crossing decisions in all fourteen experiment scenarios for each parameter value combination.

Building on this exhaustive parameter search, we determined the σ_v and c that fitted the experimental data best for each participant by likelihood maximisation. We estimated the probability density function of CIT predicted for each model by kernel density estimation separately for each of the fourteen scenarios (Table 2.2). This allowed us to calculate the model likelihood of each σ_v and c for each participant as the product of multiplying the model-predicted probability density at the participant's observed button press times.

2.3 Results

2.3.1 Empirical results

Panel (a) of Figure 2.3 shows that the pedestrian gap acceptance rate increases with the initial TTA ($p < 0.05$ and $p < 0.05$ for 6.9 m/s and 13.9 m/s respectively; Chi-squared test), which aligns with previous research (Oxley et al., 2005; Lobjois and Cavallo, 2007; Petzoldt, 2014). From Figure 2.3, we can also see a speed-dependent gap acceptance rate, suggesting that pedestrians were more likely to accept the gap as the vehicle's speed increased, given the same initial TTA, again aligning with findings from previous studies (Lobjois and Cavallo, 2007; Tian et al., 2022). However, while this speed effect was particularly evident at initial TTA of 4.6 s and 6.9 s, the associated p-values (0.06 and 0.08) are not statistically significant ($p<0.05$), possibly due to our limited sample size.

As shown in panel (b) of Figure 2.3, this speed-dependency is also visible in the CIT metric (Tian et al., 2022), which in our case denotes the time from the start of each trial to the pedestrian's crossing decision. This phenomenon was most evident in the initial TTA of 4.6 s ($p < 0.05$; Wilcoxon Signed-Rank test).

As for the yielding scenarios, where the car stopped to let the participant cross, it can be seen in Figure 2.3 (c) that the car speed's impact on CIT was most pronounced under a higher initial TTA condition of 6.9 s ($p < 0.05$; Wilcoxon Signed-Rank test). Specifically, the third graph in panel (c) of Figure 2.3 shows the orange curve ascending more rapidly than the blue curve, which means that more pedestrians cross earlier when the vehicle's initial speed is higher at

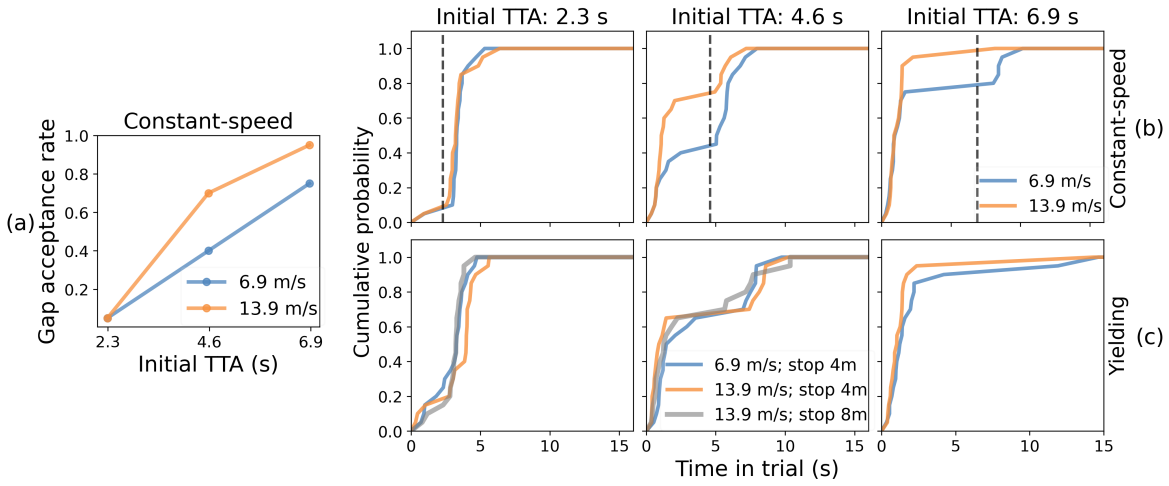

Figure 2.3: Empirical results. (a) Gap acceptance rate in constant-speed scenarios. (b) Cumulative probability of Crossing Initiation Time (CIT) in constant-speed scenarios; black dashed vertical lines indicate the times vehicles passed pedestrians. The X-axis means the time elapsed from the beginning of each trial. (c) Cumulative probability of CIT in yielding scenarios. Note: In yielding scenarios, there are only two conditions in the initial TTA of 6.9 s as shown in Table 2.2, so there is no grey line in the third panel of yielding scenarios.

higher TTA conditions. This speed-dependent trend is the same as in constant-speed scenarios. Conversely, an inverse pattern emerged for a lower initial TTA of 2.3 s. More pedestrians chose to cross earlier when the approaching vehicle's initial speed was low during yielding ($p < 0.05$; Wilcoxon Signed-Rank test), indicating a speed-dependent yielding acceptance. This speed-dependent yielding behaviour is illustrated by the blue curve's quicker increase compared to the orange curve in the first graph of the panel (c) in Figure 2.3. This finding is in line with the work by Ackermann et al. (2019) and Tian et al. (2023), suggesting that pedestrians tend to interpret low speed in itself as indicative of vehicle yielding. Interestingly, a combination of these two phenomena is observed at the initial TTA of 4.6 s, with a speed-dependent gap acceptance in the early phase of the scenario and a speed-dependent yielding acceptance behaviour emerging in the late phase of the trial. Therefore, across the entire scenario, the mean CIT is not significantly affected by car speed ($p = 0.21$; Wilcoxon Signed-Rank test).

Finally, stopping distance-dependent yield acceptance can also be seen in Figure 2.3. More pedestrians tended to cross early when the vehicle decelerated at a higher rate, as shown by the orange line lagging behind the grey line in the yield acceptance phase of the scenario (from about 2 s in the TTA 2.3 s scenario and about 5 s in the TTA 4.6 s scenario), indicating that greater vehicle deceleration encourages early crossing (Ackermann et al., 2019). This effect of stopping distance on CIT was statistically significant at the initial TTA of 2.3 s ($p < 0.05$;

Wilcoxon Signed-Rank test).

In sum, the empirical observations show indications of the four main phenomena we are targeting, even if some are present as trends rather than as statistically significant effects in this relatively small dataset.

2.3.2 Model results

After observing these key behaviours in empirical settings, we now investigate whether our model reproduces these phenomena.

All model variants converged; a plot of the reward curves can be found in A.2.

Evaluation of model outcomes against observed phenomena

We evaluated our models based on their ability to replicate our four targeted key pedestrian behaviours identified in empirical studies.

First, we analysed the results of the baseline model (BM). The results of BM for the constant-speed scenarios are shown in the upper-left panel of Figure 2.4 and the second row in Figure 2.5. As shown in Figure 2.4, the BM agent decided not to cross at the initial TTA of 1 s but consistently accepted all safe crossing opportunities available during initial TTA conditions 2.3 s, 4.6 s and 6.9 s, where it crossed at the very start of the trial (shown in Figure 2.5 and Figure 2.6). This behaviour can be attributed to the agent's perfect information about the vehicle's state, permitting it to achieve optimal behaviour for an agent whose goal is to cross the road safely in the least possible time. The BM model variant does not capture any of the four targeted phenomena. Below, we will consider these phenomena one at a time.

TTA-dependent gap acceptance The VM, LM, and VLM models all qualitatively captured the TTA-dependent gap acceptance phenomenon, showing increasing gap acceptance rates with rising initial TTA conditions. This pattern is particularly evident in the LM model, as seen in the upper-right panel of Figure 2.4. The VLM model, integrating both visual constraints and looming aversion (shown in the bottom-right panel of Figure 2.4), most accurately reflected the human data, demonstrating the most human-like sensitivity to initial TTA conditions.

Speed-dependent gap acceptance While including just looming aversion (model LM) was enough to capture the effect of TTA on gap acceptance, this model did not show any effect of

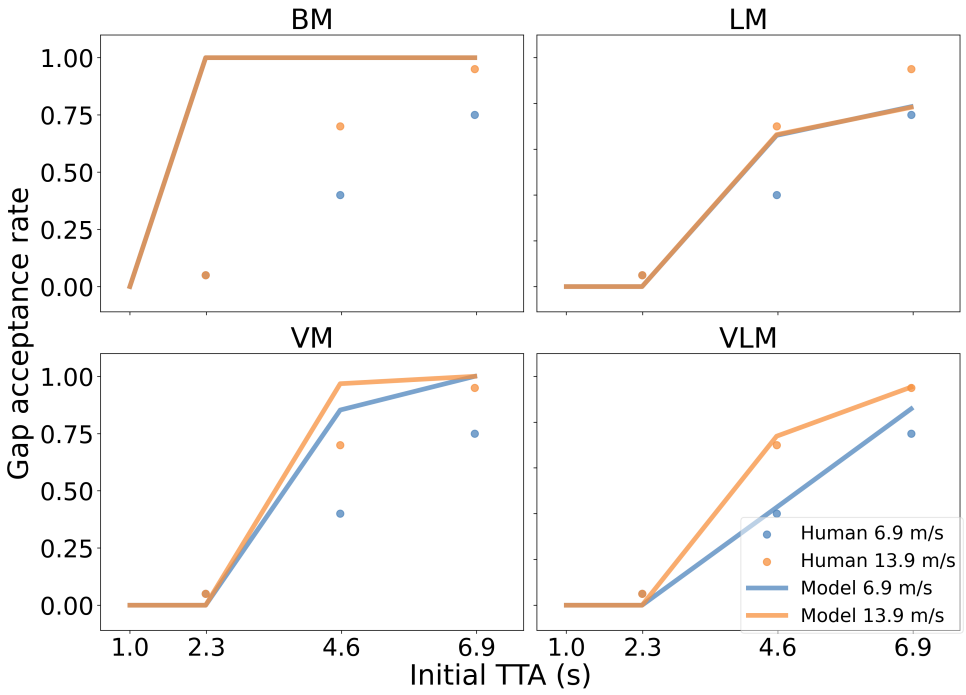

Figure 2.4: Gap acceptance rate by human participants and different models. BM: Baseline model. LM: Model with the looming aversion assumption. VM: Model with the visual constraints assumption. VLM: Model with both visual constraints and looming aversion assumptions. Note: the lines of the model LM and BM overlap each other.

Table 2.4: Quantitative assessment of model performance

Model type	**Log Lik.**	**Params.**	**AIC**	**MAD(s)**
BM	−1289	0	2578	2.90
LM	−594	20	1227	0.94
VM	−588	20	1215	1.30
VLM	−533	40	1146	0.34

Note: Log Lik. refers to Log Likelihood, Params. refers to Free Parameters, and MAD refers to the mean absolute deviation across all scenarios.

vehicle speed on gap acceptance. This can be seen in Figure 2.4, where the LM lines for two speed conditions overlap entirely. On the other hand, the inclusion of visual constraints (model VM) caused the model to exhibit speed-dependent gap acceptance behaviour by separating the gap acceptance curves at the initial TTA of 4.6 s. The VLM model again exhibited the best similarity with human data, showing a clear speed dependency in the gap acceptance rate.

Speed-dependent yielding acceptance As shown in the bottom row in Figure 2.6, at the initial TTA condition of 2.3 s, speed-dependent yielding acceptance emerged in the results of VLM, similar to the human data, visible as a quicker increase of the blue curve than the orange curve. In other words, the agent was more likely to cross the road earlier when the initial speed

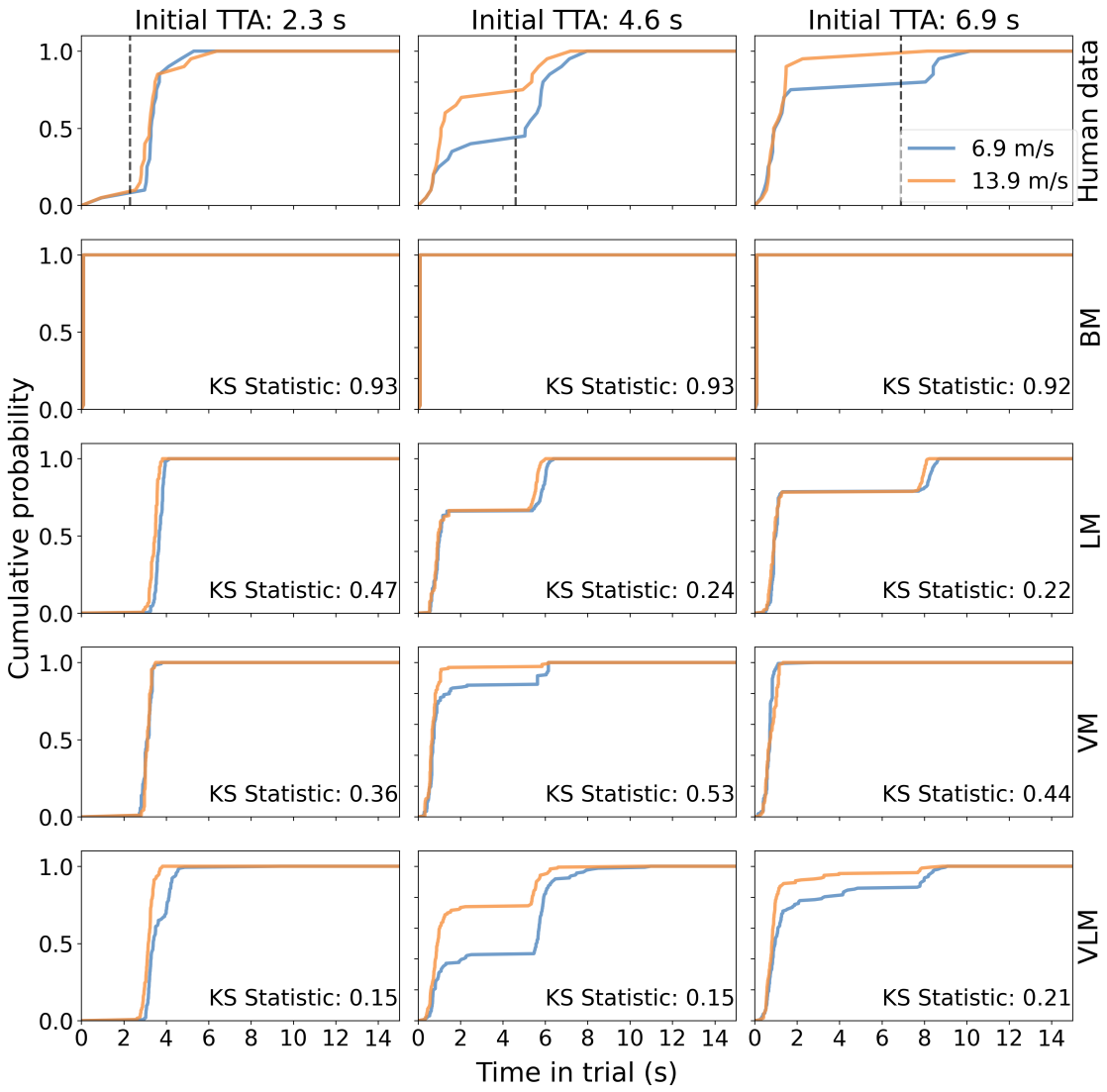

Figure 2.5: Cumulative probability for Crossing Initiation Time (CIT) in constant-speed scenarios. Black dashed vertical lines in the human data indicate the times vehicles passed pedestrians. Vehicles passed pedestrians at the same time in the trial for each initial TTA condition. The X-axis represents the time elapsed from the beginning of each trial. The KS statistic quantifies the maximum divergence between the cumulative distribution functions of human data and model results, indicating their distributional similarity. Model abbreviations are as defined in Figure 2.4.

of the approaching vehicle was lower. Furthermore, and again in line with the human data, at the initial TTA condition of 6.9 s, the initial speed of the vehicle has the opposite effect on CIT (speed-dependent gap acceptance), and in the 4.6 s initial TTA condition both of the two phenomena are visible. These speed effects can, to some extent, be observed in the results of LM and VM at the initial TTA of 2.3 s as well, but not for the 4.6 s and 6.9 s initial TTA conditions.

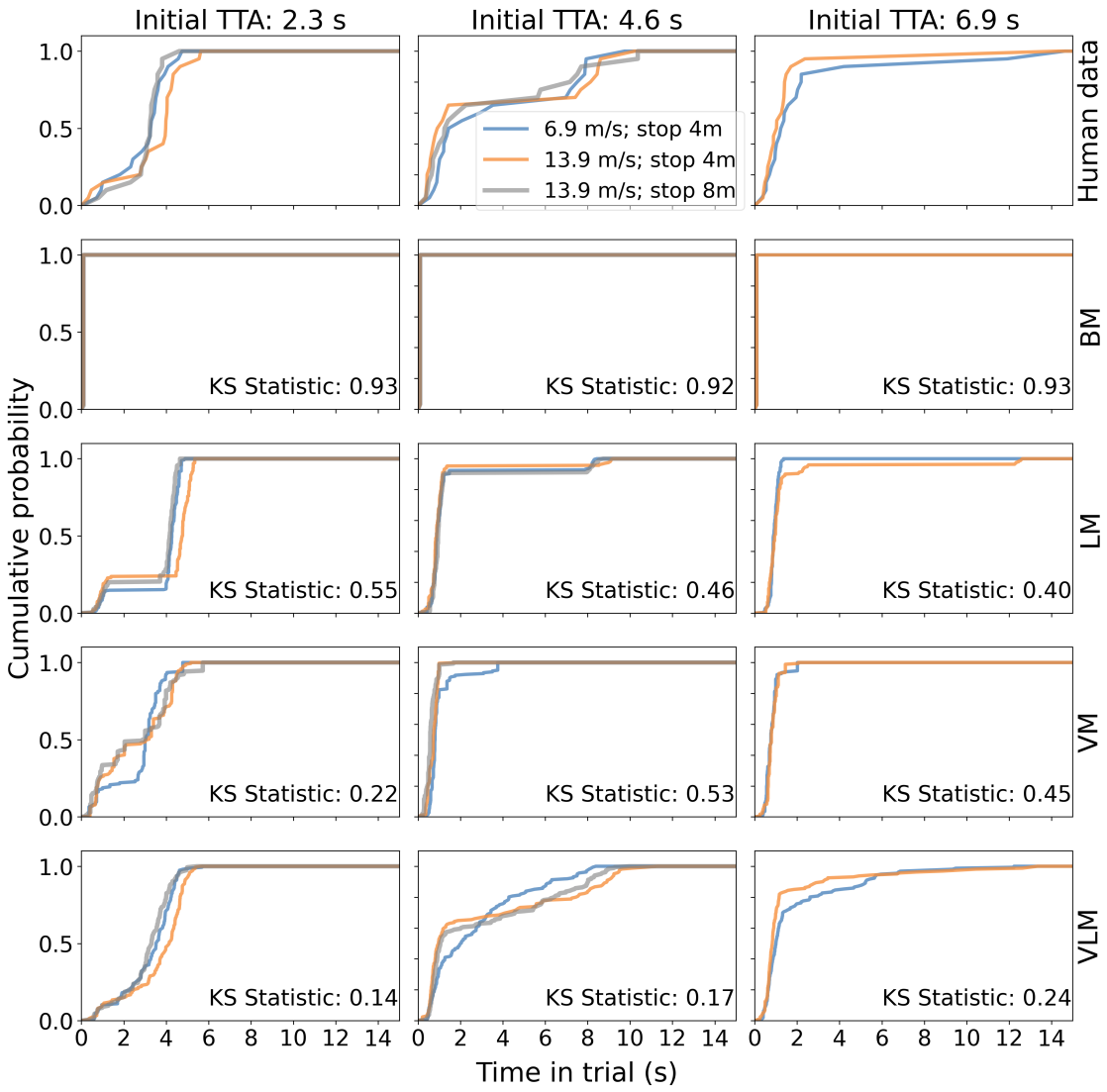

Figure 2.6: Cumulative probability for CIT in yielding scenarios. See Figure 2.5 for further details. .

Stopping distance-dependent yielding acceptance The VLM agent replicated this phenomenon both at the initial TTA condition of 2.3 s and 4.6 s, i.e., the agent had a greater tendency to cross earlier when the vehicle stopped at a greater distance, visible as the orange line lagging the grey line in the CIT plots. Notably, at the initial TTA of 4.6 s, this tendency appeared only in the late crossing decisions (from about 5 s). This is probably because the deceleration information has a stronger effect on the crossing decision when the vehicle is not far from the pedestrian, in line with (Tian et al., 2023). The LM model also captured this pattern, especially at the initial TTA condition of 2.3 s. However, this pattern did not appear in the results of the VM model.

Quantitative Assessment of Model Performance

We assessed model performance using log-likelihood, Akaike Information Criterion (AIC), and Mean Absolute Deviation (MAD). Higher log-likelihood and lower AIC values indicate better model fit and simplicity, respectively. At the same time, a smaller MAD suggests a closer

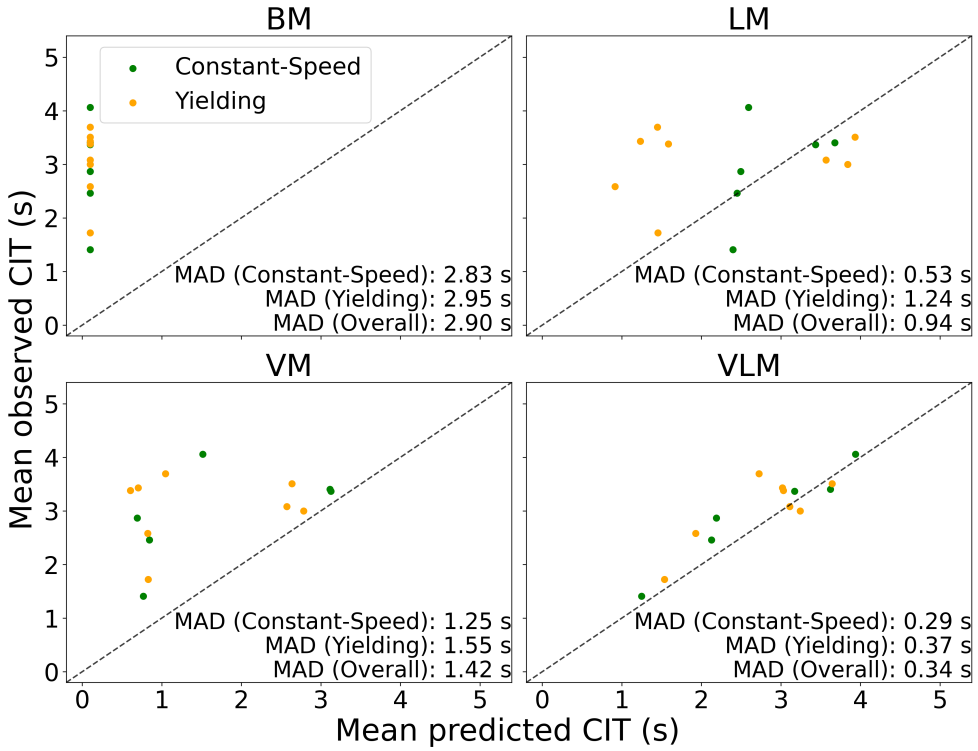

Figure 2.7: Predicted vs observed mean Crossing Initiation Time (CIT) across the different scenarios. This comparison illustrates the model's performance in estimating CIT in relation to actual measurements, with the dotted line representing an ideal prediction where observed and predicted values match perfectly.

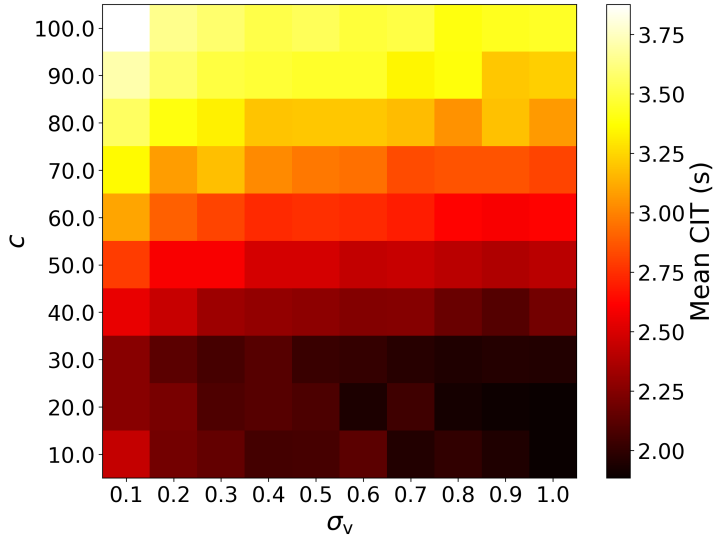

Figure 2.8: Heatmap of mean CIT for different values of the two non-policy parameters. Different colours represent the average CIT values across all different scenarios.

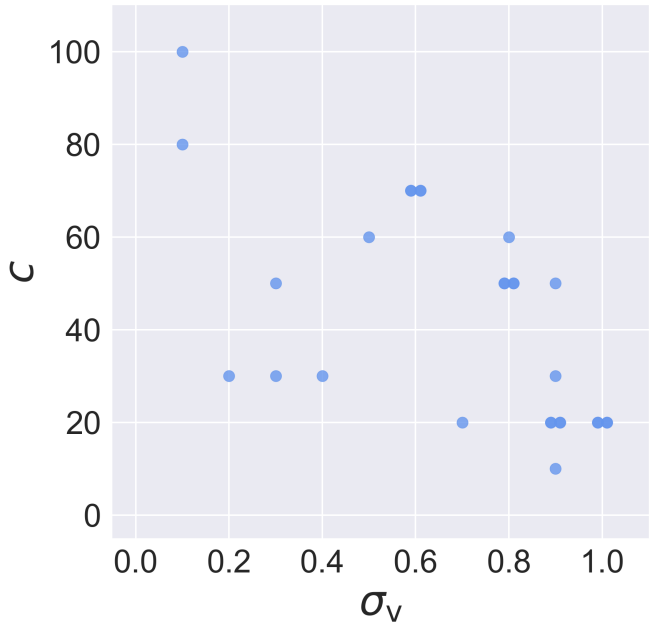

Figure 2.9: Assessment of Fitted Parameters for the model with visual constraints and looming aversion assumptions (VLM). Points represent non-policy parameter combinations, with overlapping points displayed as slightly horizontally offset to indicate multiple occurrences.

match between predicted and observed crossing times. As shown in Table 2.4, the VLM model demonstrated superior performance across all metrics: it achieved the highest log-likelihood (-533), the lowest AIC (1146), and the smallest MAD (0.34 s), indicating it most accurately represents real-world pedestrian behaviour.

Finally, we explored the influence of free parameters, σ_v and the looming aversion weight c, on the outcomes of our most effective model, the VLM. As illustrated in Figure 2.8, we computed the mean CIT of VLM across all test scenarios for every combination of parameter values. There was a noticeable increase in CIT with a larger looming aversion weight c. We observed a reverse trend with regard to the visual noise parameter, σ_v. An increment in σ_v generally led to a decrease in CIT.

Figure 2.9 presents a scatter plot that visualises two selected non-policy parameter combinations for the VLM model, with some data points slightly offset horizontally to denote multiple occurrences. The best-fit parameters lie inside the range of the search parameters, indicating that the selected parameter range we used was sufficiently broad.

2.3.3 Additional model tests

In addition to our main analyses, we conducted tests on four additional alternative models with different assumptions or different modelling approaches, but none of these performed better than the VLM model. Detailed results and figures from the tests of these alternative model variants can be found in A.3.

2.4 Discussion

2.4.1 Main findings

In this study, we developed a pedestrian crossing decision model. This model is based on the computational rationality framework and uses reinforcement learning (RL) to derive boundedly optimal behaviour policies under assumptions about human constraints and preferences. We see four main findings:

First, the VLM model variant, which incorporates both visual constraints and looming aversion assumptions, captured all four of our targeted behavioural phenomena. To our knowledge, no previous models have captured all four of these phenomena. In addition, these phenomena have not previously been addressed using a bounded optimality modelling framework.

Second, by testing different model variants, we demonstrated how specific human constraints affect pedestrian behaviour. In particular, our investigation of the model with visual constraints revealed that speed-dependent gap acceptance, previously interpreted as a 'bias' in decision-making, may instead reflect a rational adaptation to visual perception constraints.

Third, our research underscores the necessity of including both visual constraints and looming aversion assumptions. The model variants with only one of these assumptions, VM and LM, each only reproduced some of the phenomena. This finding highlights the complexity of pedestrian decision-making processes. Previous research, such as the study by Tian et al. (2022), has typically focused on single aspects of human constraints—like looming aversion—whereas our results suggest that both visual constraints and looming aversion play important roles, and both need to be integrated into the same model to explain the observed human behaviour.

Fourth, another important finding is the effectiveness of our parameter conditioning method. Applying this method to the VLM model allowed us to fit the model to the empirical data and

demonstrate good quantitative fits with just two free model parameters per participant.

In the following subsections, we will discuss the impact of our mechanistic assumptions about human constraints and preferences on crossing decisions and explore our study's broader implications and limitations.

2.4.2 Impact of mechanistic assumptions on crossing decisions

In this subsection, we will first discuss how assumptions about human constraints and preference affect the generation of the four targeted phenomena:

TTA-dependent gap acceptance While both the model with the looming aversion assumption (LM) and the model with the visual constraints assumption (VM) demonstrate increased gap acceptance with rising initial TTA, this effect is more evident in the LM model compared to the VM model. This finding suggests that looming aversion plays an essential role in assessing safe crossing opportunities. In the LM model, a greater TTA reduces $\frac{1}{\tau}$, resulting in a lower looming aversion reward for crossing at higher TTA conditions. This phenomenon aligns with the model's inherent logic, where a higher TTA is associated with a lower perceived threat, encouraging the agent to cross.

Regarding VM, the qualitative pattern of TTA-dependent gap acceptance for this model is caused by the perceptual noise. Given that the agent cannot observe the vehicle's state accurately, there is a higher risk of collision during the crossing, especially in the lower TTA conditions. To avoid the penalty from the collision, the agent is more conservative and less likely to cross at the low initial TTA condition (2.3 s). However, at higher initial TTAs, the impact of the perceptual noise is less detrimental to making safe crossing decisions, leading to higher gap acceptance rates.

Speed-dependent gap acceptance Our findings indicate that models incorporating the visual constraints assumption can generate speed-dependent gap acceptance behaviour. Understanding why speed-dependent gap acceptance arises from the visual constraints is not entirely trivial. Our best interpretation is that it is related to the more dispersed distribution of estimated TTA at lower vehicle speeds for a given time gap, as shown in Figure A.13. Therefore, from a reward maximisation perspective, the agent makes more cautious decisions when interacting with a lower-speed vehicle.

This finding is interesting in relation to past explanations of speed-dependent gap acceptance as being caused by biases in human TTA estimation, due to humans making inappropriate use of distance information (Petzoldt, 2014) or visual looming information (Tian et al., 2022) when judging crossing safety. Our findings instead suggest that speed-dependent gap acceptance is an optimal behaviour, but that this is an optimality which is bounded by the particular noise characteristics of the human visual system. Differently put, given the nature of human visual perception, speed-dependent gap acceptance is an entirely rational behaviour. It is, of course, still possible that human pedestrians make use of visual looming information when making these decisions, as suggested by Tian et al. (2022), but if so, this seems to be part of a boundedly optimal strategy rather than a biased (i.e., suboptimal) heuristic.

Speed-dependent yielding acceptance Our results show that both the model with the noisy perception assumption and the model with the looming aversion assumption can independently generate the speed-dependent yielding acceptance behaviour. This can be understood as follows: Under the noisy perception assumption, perceptual noise decreases as the distance between the vehicle and the agent reduces. When a vehicle approaches at a slower speed, for a given TTA, it is closer to the pedestrian, resulting in less perceptual noise. Consequently, the agent can more accurately estimate the vehicle's position, leading to an earlier decision to cross. For the model with looming aversion, the slower vehicle speed near the agent reduces the $\frac{1}{\tau}$ value. This lower value increases the reward for crossing, thus encouraging the agent to cross earlier during the vehicle's yielding.

Stopping distance-dependent yielding acceptance This behaviour is captured effectively by the looming aversion assumption (but not by the noisy perception assumption). This can be understood by considering that when a vehicle decelerates from the same initial distance and speed, stopping at a greater deceleration, further away from the pedestrian, results in a smaller $\frac{1}{\tau}$, thereby reducing the penalty to the agent for crossing in front of the vehicle. Therefore, the agent's tendency to cross increases when the vehicle stops further away.

Exploration of non-policy parameters As illustrated in Figure 2.8, there is a noticeable increase in average *Crossing Initiation Time (CIT)* with a larger looming aversion weight c. This is reasonable since a higher value of c amplifies the perceived threat, thus leading the agent to make more cautious crossing decisions.

Conversely, higher noise levels σ_v in the visual system are associated with slight reductions in mean CIT. This outcome is somewhat counterintuitive, especially when contrasted with the smaller CIT exhibited by BM, which operates without visual noise. This disparity suggests a complex interaction between the looming aversion and visual noise parameters in VLM's decision-making process. A possible explanation could be that lower levels of visual noise allow the agent to predict future looming more accurately and, thus, more strategically delay crossing the road to reduce the looming aversion penalty during a yielding vehicle's approach. Conversely, when the level of visual noise is high, the agent's risk assessment of potential collisions becomes less reliable, which may lead to a tendency to cross the road more quickly to avoid the collision and time penalties despite the penalty for looming aversion, and thus maintain a lower overall penalty from the perspective of reward maximisation.

2.4.3 Implications and future work

In contrast to previous studies in modelling work (Pekkanen et al., 2021; Tian et al., 2022), we do not specify in our model how the agent should make decisions. Instead, our model allows the agent to learn boundedly optimal decision-making, given its environment and constraints. By continually encountering varied scenarios in the dynamic environment, the agent learns its policy through an iterative process of trial and error.

This dynamic learning approach offers several advantages. For one, it imparts a level of adaptability to our model, enabling it to be attuned to different possible real-world conditions. Furthermore, given the great success of modern deep RL in learning behaviour policy for highly challenging tasks, the approach we have taken here should in principle be extensible to modelling road user behaviour and interaction also in more general, high-complexity traffic scenarios. This is an interesting direction for future work, which will require both the adoption of more advanced RL methods than we have used here, and importantly also further research to establish more complete mechanistic models of the involved human perceptual, cognitive, and motor constraints.

While our learning-based approach models the tendency of humans to naturally adapt and refine their judgments and actions based on their experiences, it is crucial to note that the RL algorithm by which our model learns boundedly optimal policy is not intended to mimic the learning process in humans. Our model converges towards a policy that could be similar to

boundedly optimal human behaviour, but the path it takes to reach this convergence is almost certainly not human-like. The value of our RL approach lies in its ability to explore complex scenarios and generate behaviour that aligns with the principles of computational rationality, thereby providing a partly mechanistic and more explainable alternative to purely data-driven ML models.

Furthermore, our findings contribute to improving the predictive accuracy of pedestrian crossing models. We reveal that two parameters – the looming aversion weight and the noise of agents' visual systems – have an important influence on the modelled CIT. In this regard, using real-world data from a large representative population to ascertain these parameters and subsequently incorporating them into RL processes may be important steps towards enhanced model accuracy.

The method of training the RL model conditioned on non-policy parameters is established in computational rationality within the human-computer interaction domain (Keurulainen et al., 2023). Here, we demonstrated that we could successfully apply this method to pedestrian modelling. This parameter-conditioning is particularly beneficial if the parameter space is large, so an interesting future direction will be to investigate pedestrian models conditioned also on a wider range of non-policy parameters. Again, applying this model to real-world pedestrian data presents a promising research direction. For example, we could infer non-policy parameters, such as risk perception or walking speed, which are unique to different individuals, which could allow more precise prediction of pedestrian behaviours in various traffic scenarios.

As mentioned in Section 2.1, while the sample size of only 20 participants might seem small, it is important to note that our main findings concern the model's ability to replicate the observed qualitative phenomena, and that these have been observed in many other datasets, including naturalistic datasets (Oxley et al., 2005; Lobjois and Cavallo, 2007; Petzoldt, 2014; Tian et al., 2022; Ackermann et al., 2019; Tian et al., 2023). This alignment across different studies supports the validity of our findings despite the limited sample. Furthermore, the main direct use of the dataset here is to demonstrate that we can do parameter inference with our approach, including fitting to individual participants. It is important to note that we are not suggesting that the specific parameter values we have obtained here will necessarily generalise, e.g., to larger naturalistic samples. Instead, we reproduce qualitative behaviour phenomena that have also been observed in naturalistic data, and we show that our model can be fitted quantitatively

to data. In future work, we plan to extend our model to address naturalistic datasets where more diverse scenarios are available. This can not only improve the applicability of our model to real-world autonomous vehicle (AV) applications but also increase the generalisability of the observed behaviour.

In this study, we used VR to simulate real-world conditions. VR offers a controlled, safe, and replicable environment that is helpful for studying pedestrian behaviour. Kalantari et al. (2023) have shown that pedestrian crossing behaviours observed in VR settings are comparable to those in real-world environments, lending credence to the use of VR as a reliable method for behavioural studies. We believe this is an important step toward applying this modelling method to more complex real-world scenarios.

The scenario involving one pedestrian crossing at an unsignalised road was specifically chosen to isolate and model pedestrian reactions to one oncoming vehicle without the influence of traffic control devices such as traffic lights. This specific focus was selected to simplify the crossing environment, ensuring that the observed behaviours were predominantly influenced by the vehicle's approach to provide a clear view of the factors influencing pedestrian behaviour under these conditions. In future work, we could include more complex scenarios, such as roads with traffic lights or interactions between pedestrians and multiple vehicles, to enhance the practical applicability of our research.

In addition, the scenarios presented here focused primarily on constant-speed and deceleration conditions. Although these two scenarios are commonly seen in pedestrian-vehicle interaction, other scenarios may also be interesting; for example, drivers may accelerate when refusing to give pedestrians priority (Markkula et al., 2023). To include more diverse scenarios in future studies will be helpful for a more comprehensive pedestrian behaviour model.

In the field of AV simulation, the potential applications of our model are promising. For instance, it could improve the realism and effectiveness of AV testing simulations (Markkula and Dogar, 2022). Furthermore, by incorporating a detailed understanding of pedestrian behaviour, real-time AV algorithms can be fine-tuned to anticipate and respond to pedestrian actions more accurately. For example, by combining our model's predictions with real-time data from vehicle sensors, AV systems could achieve a higher level of situational awareness, enabling them to make safer decisions in complex urban environments (Camara et al., 2020). It should be noted that it utilises only pedestrian and vehicle trajectory information—specifically, their positions

and velocities. This type of data is readily captured by the sensor suites in modern vehicles for real-time sensing. Similar uses of these models in, for example, infrastructure design can also benefit pedestrian traffic safety more broadly, also in conventional human-driven traffic (Kraidi and Evdorides, 2020; Zhu et al., 2022).

Regarding the real-world applicability of our model in autonomous vehicle contexts, it is important to note that it utilises only pedestrian and vehicle trajectory information—specifically, their positions and velocities. This type of data is readily captured by the sensor suites in modern vehicles for real-time sensing.

There is clear and ample scope for further improvements to our model. For example, there is more variability in the human crossing initiation time than in the model. This could be due to both between- and within-individual variability of decision-influencing factors. In addition, our reward function is simple, and so far, we have fine-tuned only parts of it (the looming aversion). This was sufficient for our purposes, to show qualitative patterns of human road-crossing, and contrary to our expectations, it even allowed us to achieve relatively good quantitative fits, but there is clear room for further work. For example, following a similar approach to what we did here with σ_{v}, we could also tune, for example, the time-loss penalty in the reward function. This adjustment would allow the model to more accurately reflect how different pedestrians value the trade-off between time and safety.

Another aspect to consider for future improvement of our model is its current focus exclusively on the pedestrian's reaction to vehicles, without accounting for the driver's reaction to the pedestrian as discussed in Section 2.2. Real-world scenarios often involve a dynamic interplay between both parties, influencing each other's decisions. To address this gap, future studies could employ multiagent RL to model the interactions between pedestrians and drivers more realistically (Hu, Wellman, et al., 1998).

Finally, it is interesting to note that we face the challenge of fully understanding the complexity of our model's behaviour. This issue arises both in relation to the speed-dependent gap acceptance (we have shown that this behaviour is boundedly optimal, but we have not been able to fully explain why) and the effect of visual noise levels on CITs (zero noise causes rapid decisions but low levels of noise cause slower decisions than high levels of noise). While our computational rationality model is more interpretable than fully data-driven ML models, it only partially eliminates the challenge of explicating the 'why' behind its decision-making processes. On a

high level, the behaviour of our model, governed by the principle of computational rationality and informed by RL, can be explained quite simply as boundedly optimal, but formulating a more detailed explanation of *why* a certain behaviour is boundedly optimal may not always be straightforward.

2.5 Conclusions

In conclusion, this study contributes to modelling human pedestrian crossing decisions, utilising a computational rationality framework, and using RL to identify boundedly optimal behaviour policy. Under our chosen human-like constraints, we demonstrate that our model can emulate human-like behaviour in road-crossing scenarios with both non-yielding and yielding vehicles. Interestingly, our findings suggest that the previously reported speed-dependent gap acceptance behaviour in pedestrian decision-making can be understood as a rational adaptation to visual perception constraints. Our best model not only reproduces this behaviour but also qualitatively and quantitatively captures three other key phenomena: TTA-dependent gap acceptance, speed-dependent yielding acceptance, and stopping distance-dependent yielding acceptance. In addition, we identified two parameters—the weight of looming aversion and the noise in agents' visual systems—which can provide a better understanding of human decision-making in real-world pedestrian crossing scenarios. While we recognise the need to investigate and develop our model further, the insight from this study provides a promising foundation for improved pedestrian models. Not least, the inherent adaptability of RL suggests the potential for future extensions to more complex traffic scenarios, a challenge that traditional mechanistic models may find difficult to address. The potential applications of our model are notable. In the field of AVs, our results can contribute to more realistic simulated road users in virtual testing environments and guide vehicular decision-making algorithms to better anticipate human behaviour, ensuring safer human-AV co-existence.

References

Abughalieh, Karam M and Alawneh, Shadi G (2020). "Predicting pedestrian intention to cross the road". In: *IEEE Access* 8, pp. 72558–72569.

Ackermann, Claudia, Beggiato, Matthias, Bluhm, Luka-Franziska, Löw, Alexandra, and Krems, Josef F (2019). "Deceleration parameters and their applicability as informal communication signal between pedestrians and automated vehicles". In: *Transportation research part F: traffic psychology and behaviour* 62, pp. 757–768.

Alahi, Alexandre, Goel, Kratarth, Ramanathan, Vignesh, Robicquet, Alexandre, Fei-Fei, Li, and Savarese, Silvio (2016). "Social lstm: Human trajectory prediction in crowded spaces". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 961–971.

Althoff, Matthias and Lutz, Sebastian (2018). "Automatic generation of safety-critical test scenarios for collision avoidance of road vehicles". In: *2018 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, pp. 1326–1333.

Asaithambi, Gowri, Kuttan, Manu O, and Chandra, Sarath (2016). "Pedestrian road crossing behavior under mixed traffic conditions: a comparative study of an intersection before and after implementing control measures". In: *Transportation in developing economies* 2.2, p. 14.

Banovic, Nikola, Oulasvirta, Antti, and Kristensson, Per Ola (2019). "Computational modeling in human-computer interaction". In: *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1–7.

Camara, Fanta, Bellotto, Nicola, Cosar, Serhan, Weber, Florian, Nathanael, Dimitris, Althoff, Matthias, Wu, Jingyuan, Ruenz, Johannes, Dietrich, André, Markkula, Gustav, et al. (2020). "Pedestrian models for autonomous driving part ii: high-level models of human behavior". In: *IEEE Transactions on Intelligent Transportation Systems* 22.9, pp. 5453–5472.

Chen, Haiyang, Chang, Hyung Jin, and Howes, Andrew (2021a). "Apparently irrational choice as optimal sequential decision making". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 35. 1, pp. 792–800.

Chen, Xiuli, Acharya, Aditya, Oulasvirta, Antti, and Howes, Andrew (2021b). "An adaptive model of gaze-based selection". In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–11.

Crosato, Luca, Tian, Kai, Shum, Hubert PH, Ho, Edmond SL, Wang, Yafei, and Wei, Chongfeng (2023). "Social Interaction-Aware Dynamical Models and Decision-Making for Autonomous Vehicles". In: *Advanced Intelligent Systems*, p. 2300575.

Dai, Shengzhe, Li, Li, and Li, Zhiheng (2019). "Modeling vehicle interactions via modified LSTM models for trajectory prediction". In: *IEEE Access* 7, pp. 38287–38296.

DeLucia, Patricia R (2008). "Critical roles for distance, task, and motion in space perception: Initial conceptual framework and practical implications". In: *Human Factors* 50.5, pp. 811–820.

– (2015). "Perception of Collision". In: *The Cambridge Handbook of Applied Perception Research*. Ed. by Robert R. Hoffman, Peter A. Hancock, Mark W. Scerbo, Raja Parasuraman, and James L.Editors Szalma. Cambridge Handbooks in Psychology. Cambridge University Press, pp. 568–591. DOI: 10.1017/CBO9780511973017.035.

Diaz-Ruiz, Carlos A, Xia, Youya, You, Yurong, Nino, Jose, Chen, Junan, Monica, Josephine, Chen, Xiangyu, Luo, Katie, Wang, Yan, Emond, Marc, et al. (2022). "Ithaca365: Dataset and Driving Perception Under Repeated and Challenging Weather Conditions". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 21383–21392.

Doellinger, Johannes, Spies, Markus, and Burgard, Wolfram (2018). "Predicting occupancy distributions of walking humans with convolutional neural networks". In: *IEEE Robotics and Automation Letters* 3.3, pp. 1522–1528.

Domeyer, Joshua, Dinparastdjadid, Azadeh, Lee, John D, Douglas, Grace, Alsaid, Areen, and Price, Morgan (2019). "Proxemics and kinesics in automated vehicle–pedestrian communication: Representing ethnographic observations". In: *Transportation research record* 2673.10, pp. 70–81.

Duric, Zoran, Gray, Wayne D, Heishman, Ric, Li, Fayin, Rosenfeld, Azriel, Schoelles, Michael J, Schunn, Christian, and Wechsler, Harry (2002). "Integrating perceptual and cognitive modeling for adaptive and intelligent human-computer interaction". In: *Proceedings of the IEEE* 90.7, pp. 1272–1289.

Faisal, A Aldo, Selen, Luc PJ, and Wolpert, Daniel M (2008). "Noise in the nervous system". In: *Nature reviews neuroscience* 9.4, pp. 292–303.

Fajen, Brett R and Warren, William H (2003). "Behavioral dynamics of steering, obstable avoidance, and route selection." In: *Journal of Experimental Psychology: Human Perception and Performance* 29.2, p. 343.

Gershman, Samuel J, Horvitz, Eric J, and Tenenbaum, Joshua B (2015). "Computational rationality: A converging paradigm for intelligence in brains, minds, and machines". In: *Science* 349.6245, pp. 273–278.

Giles, Oscar, Markkula, Gustav, Pekkanen, Jami, Yokota, Naoki, Matsunaga, Naoto, Merat, Natasha, and Daimon, Tatsuru (2019). "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, pp. 366–372.

González-Méndez, Mauricio, Olaya, Camilo, Fasolino, Isidoro, Grimaldi, Michele, and Obregón, Nelson (2021). "Agent-based modeling for urban development planning based on human needs. Conceptual basis and model formulation". In: *Land Use Policy* 101, p. 105110.

Gorrini, Andrea, Crociani, Luca, Vizzari, Giuseppe, and Bandini, Stefania (2018). "Observation results on pedestrian-vehicle interactions at non-signalized intersections towards simulation". In: *Transportation research part F: traffic psychology and behaviour* 59, pp. 269–285.

Howes, Andrew, Jokinen, Jussi PP, and Oulasvirta, Antti (2023). "Towards machines that understand people". In: *AI Magazine* 44.3, pp. 312–327.

Hu, Junling, Wellman, Michael P, et al. (1998). "Multiagent reinforcement learning: theoretical framework and an algorithm." In: *ICML*. Vol. 98, pp. 242–250.

Jain, Akash, Gupta, Ankit, and Rastogi, Rajat (2014). "Pedestrian crossing behaviour analysis at intersections". In: *International Journal for Traffic and Transport Engineering* 4.1, pp. 103–116.

Jokinen, Jussi, Remes, Ulpu, Kujala, Tuomo, and Corander, Jukka (2022). "Bayesian Parameter Inference for Cognitive Simulators". In: *Bayesian Methods for Interaction and Design*. Ed. by

John H. Williamson, Antti Oulasvirta, Per Ola Kristensson, and Nikola Banovic. Cambridge, United Kingdom: Cambridge University Press, pp. 308–334. ISBN: 978-1-108-79270-7. DOI: 10.1017/9781108874830.016.

Jokinen, Jussi PP, Kujala, Tuomo, and Oulasvirta, Antti (2021). "Multitasking in driving as optimal adaptation under uncertainty". In: *Human factors* 63.8, pp. 1324–1341.

Jokinen, Jussi PP, Wang, Zhenxin, Sarcar, Sayan, Oulasvirta, Antti, and Ren, Xiangshi (2020). "Adaptive feature guidance: Modelling visual search with graphical layouts". In: *International Journal of Human-Computer Studies* 136, p. 102376.

Kalantari, Amir Hossein, Lin Yi-Shin, yslin, Mohammadi, Ali, Merat, Natasha, and Markkula, Gustav (Sept. 2023). *Testing the validity of multi participant distributed simulation for understanding and modeling road user interaction*. DOI: 10.31234/osf.io/gk9af. URL: osf.io/preprints/psyarxiv/gk9af.

Keurulainen, Antti, Westerlund, Isak Rafael, Keurulainen, Oskar, and Howes, Andrew (2023). "Amortised Experimental Design and Parameter Estimation for User Models of Pointing". In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pp. 1–17.

Klischat, Moritz, Liu, Edmond Irani, Holtke, Fabian, and Althoff, Matthias (2020). "Scenario factory: Creating safety-critical traffic scenarios for automated vehicles". In: *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, pp. 1–7.

Knill, David C and Pouget, Alexandre (2004). "The Bayesian brain: the role of uncertainty in neural coding and computation". In: *TRENDS in Neurosciences* 27.12, pp. 712–719.

Kraidi, R and Evdorides, H (2020). "Pedestrian safety models for urban environments with high roadside activities". In: *Safety Science* 130, p. 104847.

Kumamoto, Koji and Yamada, Keiichi (2017). "CNN-based pedestrian orientation estimation from a single image". In: *2017 4th IAPR Asian conference on pattern recognition (ACPR)*. IEEE, pp. 13–18.

Kwon, Oh-Sang, Tadin, Duje, and Knill, David C (2015). "Unifying account of visual motion and position perception". In: *Proceedings of the National Academy of Sciences* 112.26, pp. 8142–8147.

Leu, George, Curtis, Neville J, and Abbass, Hussein (2012). "Modeling and evolving human behaviors and emotions in road traffic networks". In: *Procedia-Social and Behavioral Sciences* 54, pp. 999–1009.

Lewis, Richard L, Howes, Andrew, and Singh, Satinder (2014). "Computational rationality: Linking mechanism and behavior through bounded utility maximization". In: *Topics in cognitive science* 6.2, pp. 279–311.

Li, Lihuan, Pagnucco, Maurice, and Song, Yang (2022). "Graph-based spatial transformer with memory replay for multi-future pedestrian trajectory prediction". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 2231–2241.

Li, Lin, Liu, Yanheng, Wang, Jian, Deng, Weiwen, and Oh, Heekuck (2016). "Human dynamics based driver model for autonomous car". In: *IET Intelligent Transport Systems* 10.8, pp. 545–554.

Li, Zhi, Ko, Yu-Jung, Putkonen, Aini, Feiz, Shirin, Ashok, Vikas, Ramakrishnan, Iv, Oulasvirta, Antti, and Bi, Xiaojun (2023). "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*. CHI '23. Hamburg, Germany: Association for Computing Machinery, p. 18. ISBN: 9781450394215. DOI: 10.1145/3544548.3580640. URL: https://doi.org/10.1145/3544548.3580640.

Lieder, Falk and Griffiths, Thomas L (2020). "Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources". In: *Behavioral and brain sciences* 43, e1.

Lobjois, Régis and Cavallo, Viola (2007). "Age-related differences in street-crossing decisions: The effects of vehicle speed and time constraints on gap selection in an estimation task". In: *Accident analysis & prevention* 39.5, pp. 934–943.

Lorenzo, Javier, Alonso, Ignacio Parra, Izquierdo, Rubén, Ballardini, Augusto Luis, Saz, Álvaro Hernández, Llorca, David Fernández, and Sotelo, Miguel Ángel (2021). "Capformer: Pedestrian crossing action prediction using transformer". In: *Sensors* 21.17, p. 5694.

Markkula, Gustav and Dogar, Mehmet (2022). "Models of human behavior for human–robot interaction and automated driving: How accurate do the models of human behavior need to be?" In: *IEEE Robotics & Automation Magazine* 31.3, pp. 115–120.

Markkula, Gustav, Engström, Johan, Lodin, Johan, Bärgman, Jonas, and Victor, Trent (2016). "A farewell to brake reaction times? Kinematics-dependent brake response in naturalistic rear-end emergencies". In: *Accident Analysis & Prevention* 95, pp. 209–226.

Markkula, Gustav, Lin, Yi-Shin, Srinivasan, Aravinda Ramakrishnan, Billington, Jac, Leonetti, Matteo, Kalantari, Amir Hossein, Yang, Yue, Lee, Yee Mun, Madigan, Ruth, and Merat, Natasha (2023). "Explaining human interactions on the road by large-scale integration of computational psychological theory". In: *PNAS nexus* 2.6, pgad163.

Ooi, Teng Leng, Wu, Bing, and He, Zijiang J (2001). "Distance determined by the angular declination below the horizon". In: *Nature* 414.6860, pp. 197–200.

Oulasvirta, Antti, Jokinen, Jussi PP, and Howes, Andrew (2022). "Computational rationality as a theory of interaction". In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pp. 1–14.

Oxley, Jennifer A, Ihsen, Elfriede, Fildes, Brian N, Charlton, Judith L, and Day, Ross H (2005). "Crossing roads safely: an experimental study of age differences in gap selection by pedestrians". In: *Accident Analysis & Prevention* 37.5, pp. 962–971.

Pekkanen, Jami, Giles, Oscar Terence, Lee, Yee Mun, Madigan, Ruth, Daimon, Tatsuru, Merat, Natasha, and Markkula, Gustav (2021). "Variable-drift diffusion models of pedestrian road-crossing decisions". In: *Computational Brain & Behavior*, pp. 1–21.

Pelorosso, Raffaele (2020). "Modeling and urban planning: A systematic review of performance-based approaches". In: *Sustainable cities and society* 52, p. 101867.

Petzoldt, Tibor (2014). "On the relationship between pedestrian gap acceptance and time to arrival estimates". In: *Accident Analysis & Prevention* 72, pp. 127–133.

Quan, Ruijie, Zhu, Linchao, Wu, Yu, and Yang, Yi (2021). "Holistic LSTM for pedestrian trajectory prediction". In: *IEEE transactions on image processing* 30, pp. 3229–3239.

Ratcliff, Roger, Smith, Philip L, Brown, Scott D, and McKoon, Gail (2016). "Diffusion decision model: Current issues and history". In: *Trends in cognitive sciences* 20.4, pp. 260–281.

Risto, Malte, Emmenegger, Colleen, Vinkhuyzen, Erik, Cefkin, Melissa, and Hollan, Jim (2017). "Human-vehicle interfaces: The power of vehicle movement gestures in human road user coordination". In: *Driving assessment conference*. Vol. 9. 2017. University of Iowa.

Schneemann, Friederike and Gohl, Irene (2016). "Analyzing driver-pedestrian interaction at crosswalks: A contribution to autonomous driving in urban environments". In: *2016 IEEE intelligent vehicles symposium (IV)*. IEEE, pp. 38–43.

Al-Shihabi, Talal and Mourant, Ronald R (2001). "A framework for modeling human-like driving behaviors for autonomous vehicles in driving simulators". In: *Proceedings of the fifth international conference on Autonomous agents*, pp. 286–291.

Silver, David, Huang, Aja, Maddison, Chris J, Guez, Arthur, Sifre, Laurent, Van Den Driessche, George, Schrittwieser, Julian, Antonoglou, Ioannis, Panneershelvam, Veda, Lanctot, Marc, et al. (2016). "Mastering the game of Go with deep neural networks and tree search". In: *nature* 529.7587, pp. 484–489.

Silver, David, Schrittwieser, Julian, Simonyan, Karen, Antonoglou, Ioannis, Huang, Aja, Guez, Arthur, Hubert, Thomas, Baker, Lucas, Lai, Matthew, Bolton, Adrian, et al. (2017). "Mastering the game of go without human knowledge". In: *nature* 550.7676, pp. 354–359.

Simon, Herbert A (1955). "A behavioral model of rational choice". In: *The quarterly journal of economics*, pp. 99–118.

Srinivasan, Aravinda Ramakrishnan, Lin, Yi-Shin, Antonello, Morris, Knittel, Anthony, Hasan, Mohamed, Hawasly, Majd, Redford, John, Ramamoorthy, Subramanian, Leonetti, Matteo, Billington, Jac, et al. (2023). "Beyond RMSE: Do machine-learned models of road user inter-

action produce human-like behavior?" In: *IEEE Transactions on Intelligent Transportation Systems*.

Stocker, Alan A and Simoncelli, Eero P (2006). "Noise characteristics and prior expectations in human visual speed perception". In: *Nature neuroscience* 9.4, pp. 578–585.

Sun, Rouxian, Zhuang, Xiangling, Wu, Changxu, Zhao, Guozhen, and Zhang, Kan (2015). "The estimation of vehicle speed and stopping distance by pedestrians crossing streets in a naturalistic traffic environment". In: *Transportation research part F: traffic psychology and behaviour* 30, pp. 97–106.

Sutton, Richard S and Barto, Andrew G (2018). *Reinforcement learning: An introduction*. MIT press.

Tian, Kai, Markkula, Gustav, Wei, Chongfeng, Lee, Yee Mun, Madigan, Ruth, Merat, Natasha, and Romano, Richard (2022). "Explaining unsafe pedestrian road crossing behaviours using a psychophysics-based gap acceptance model". In: *Safety science* 154, p. 105837.

Tian, Kai, Tzigieras, Athanasios, Wei, Chongfeng, Lee, Yee Mun, Holmes, Christopher, Leonetti, Matteo, Merat, Natasha, Romano, Richard, and Markkula, Gustav (2023). "Deceleration parameters as implicit communication signals for pedestrians' crossing decisions and estimations of automated vehicle behaviour". In: *Accident Analysis & Prevention* 190, p. 107173.

Turnwald, Annemarie, Althoff, Daniel, Wollherr, Dirk, and Buss, Martin (2016). "Understanding human avoidance behavior: interaction-aware decision making based on game theory". In: *International journal of social robotics* 8.2, pp. 331–351.

Van Hasselt, Hado, Guez, Arthur, and Silver, David (2016). "Deep reinforcement learning with double q-learning". In: *Proceedings of the AAAI conference on artificial intelligence*. Vol. 30. 1.

Von Neumann, John and Morgenstern, Oskar (1947). "Theory of games and economic behavior, 2nd rev". In.

Wang, Yueyang, Srinivasan, Aravinda Ramakrishnan, Jokinen, Jussi PP, Oulasvirta, Antti, and Markkula, Gustav (2023). "Modeling human road crossing decisions as reward maxi-

mization with visual perception limitations". In: *2023 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, pp. 1–6.

Wang, Ziyu, Schaul, Tom, Hessel, Matteo, Hasselt, Hado, Lanctot, Marc, and Freitas, Nando (2016). "Dueling network architectures for deep reinforcement learning". In: *International conference on machine learning*. PMLR, pp. 1995–2003.

Wunder, Michael, Littman, Michael L, and Babes, Monica (2010). "Classes of multiagent q-learning dynamics with epsilon-greedy exploration". In: *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, pp. 1167–1174.

Xie, Zhaoming, Clary, Patrick, Dao, Jeremy, Morais, Pedro, Hurst, Jonathan, and Panne, Michiel van de (2019). "Iterative reinforcement learning based design of dynamic locomotion skills for cassie". In: *arXiv preprint arXiv:1903.09537*.

Xu, Yiran, Yang, Xiaoyin, Gong, Lihang, Lin, Hsuan-Chu, Wu, Tz-Ying, Li, Yunsheng, and Vasconcelos, Nuno (2020). "Explainable object-induced action decision for autonomous vehicles". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 9523–9532.

Yang, Dongfang, Özgüner, Ümit, and Redmill, Keith (2020). "A social force based pedestrian motion model considering multi-pedestrian interaction with a vehicle". In: *ACM Transactions on Spatial Algorithms and Systems (TSAS)* 6.2, pp. 1–27.

Yi, Shuai, Li, Hongsheng, and Wang, Xiaogang (2016). "Pedestrian behavior understanding and prediction with deep neural networks". In: *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14*. Springer, pp. 263–279.

Yin, Ziyi, Liu, Ruijin, Xiong, Zhiliang, and Yuan, Zejian (2021). "Multimodal Transformer Networks for Pedestrian Trajectory Prediction." In: *IJCAI*, pp. 1259–1265.

Yuan, Ye, Weng, Xinshuo, Ou, Yanglan, and Kitani, Kris M (2021). "Agentformer: Agent-aware transformers for socio-temporal multi-agent forecasting". In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 9813–9823.

Zhang, Xingchen, Angeloudis, Panagiotis, and Demiris, Yiannis (2022). "ST CrossingPose: A Spatial-Temporal Graph Convolutional Network for Skeleton-Based Pedestrian Crossing Intention Prediction". In: *IEEE Transactions on Intelligent Transportation Systems* 23.11, pp. 20773–20782.

Zhu, Manman, Sze, NN, and Newnam, Sharon (2022). "Effect of urban street trees on pedestrian safety: A micro-level pedestrian casualty model using multivariate Bayesian spatial approach". In: *Accident Analysis & Prevention* 176, p. 106818.

Chapter 3

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

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.

3.1 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 (Rasouli and Tsotsos, 2019). 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 (Xu et al., 2024; Man et al., 2025), and their effectiveness has been proven (Izquierdo et al., 2024; Alhawiti et al., 2024). Other researchers have tried to develop models that can help us understand and predict pedestrian behavior (Camara et al., 2020). 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 (Hollmann, 2015; Markkula et al., 2023). 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 (Papathanasopoulou et al., 2022; Quan et al., 2021). 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 (Benrachou et al., 2022).

Our previous work has focused on addressing these challenges by developing a pedestrian model that integrates the strengths of both cognitive and ML approaches (Wang et al., 2025). 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 experiment on pedestrian road-crossing.

3.2 Background

3.2.1 Pedestrian crossing behaviors

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 (Wang et al., 2025), 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 (Tian et al., 2022). Another important behavioral patterns regarding the pedestrian crossing speed, was not captured by our previous crossing decision model (Wang et al., 2025). This metric tends to decrease as the time gap increases, suggesting a compensatory behavior where pedestrians walk faster when accepting shorter gaps (Kalantarov et al., 2018).

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.

Noisy perception Human perception of the world is inherently noisy and imperfect (Faisal et al., 2008). This noisiness can be due to several factors, such as individual differences in sensory acuity (Manning et al., 2022) and environmental conditions (e.g., poor lighting, weather). It has been argued that this visual limitation affected the pedestrian crossing decision (Kotseruba

and Rasouli, 2023). 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.

Looming aversion Visual looming describes the perceived growth of an object's size as it approaches (DeLucia, 2008). The aversion to looming refers to the tendency of individuals to react more strongly to objects that appear to be rapidly approaching (Mulier et al., 2024). This phenomenon is rooted in the perceptual system's sensitivity to motion cues that signal potential threats, and influences pedestrian crossing decision. For example, Tian et al. (2022) explained speed-dependent gap acceptance behavior by using looming aversion.

Time pressure Time pressure has a significant effect on pedestrian crossing behavior (Tian et al., 2022). For example, when pedestrians are under time constraints, either due to the rapid approach of a vehicle or a short signal phase (Kalantarov et al., 2018), they tend to initiate crossing faster and adopt higher walking speeds.

Walking effort Individuals choose their walking speed accounting for not only the time but also the energy spent on walking (Carlisle and Kuo, 2023). Researchers have considered the walking effort in gait modeling works (Choi, 1997; Faraji et al., 2018). 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 (Oxley et al., 2004).

Ballistic speed control Early researchers noted that humans tend to adjust their sensorimotor movements through intermittent, ballistic control, rather than continuous adjustments (Tustin, 1947; Craik, 1948). 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 (Barton et al., 2019), but in the present model we neglect this behavior.

3.2.2 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.

Logistic regression models

Logistic regression models, including more recent cognitive cue-based models like the one proposed in Tian et al. (2022), aim to predict whether a pedestrian will cross the street under certain conditions. These models typically consider factors such as vehicle speed, and distance (Lobjois and Cavallo, 2007). 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.

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 (Giles et al., 2019; Pekkanen et al., 2022). 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 (Markkula et al., 2023). 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 (Markkula et al., 2023).

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 (Quan et al., 2021). 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 (Madala and Gonzalez, 2023). In addition, the

'black box' nature of ML models in autonomous driving poses risks, as their lack of transparency can lead to unexplainable mistakes (Roshdi et al., 2024). Second, high-level statistical accuracy does not guarantee that the models capture those behaviors that matter to humans (Srinivasan et al., 2023). 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 (Xu et al., 2020b). 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 (Diaz-Ruiz et al., 2022).

Models based on (bounded) optimality

Another approach to modeling human behavior is *bounded optimality* (Wang et al., 2025; Hoogendoorn and HL Bovy, 2003; Wang et al., 2023), 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 (Gershman et al., 2015; Oulasvirta et al., 2022). 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 (Kaelbling et al., 1996). 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 (Wang et al., 2025). 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

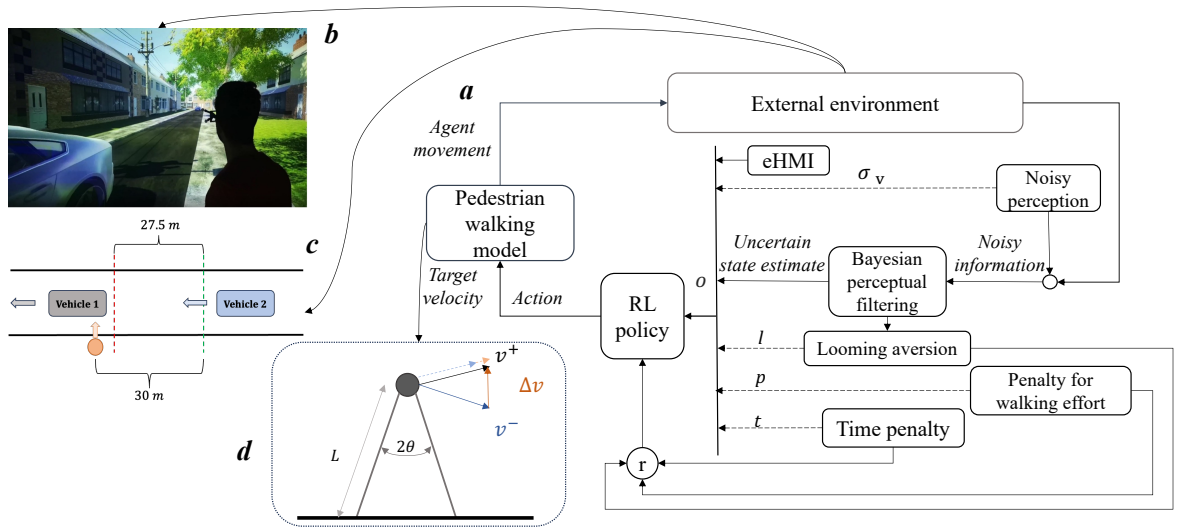

Figure 3.1: Panel (a) is the Model framework. Panel (b) is the virtual environment. Panel (c) is a schematic of the deceleration procedure used in this study. Panel (d) is the walking model, adopted from (Carlisle and Kuo, 2023).

allow us to more accurately represent the physical execution of walking actions during road crossing.

3.3 Methods

3.3.1 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 Lee et al. (2024), below we provide a summary for completeness[1]. In this study, the task of the participant was to cross the road between two approaching vehicles safely – as shown in panel (b) of Figure 3.1. 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) (Lee et al., 2024). 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

[1]More detailed description of the experiment is reported in Appendix B.1.

Table 3.1: 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

agreeing to take part in the study. Ethical approval was obtained from the University of Leeds Research Committee [2].

The recorded data used from the experiment include participants' crossing decisions, CIT, and walking speeds; these quantities were 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 behaviours in yielding scenarios as either early or late. This classification relies on the concept of bimodal crossing, where pedestrian responses can typically be categorised into two distinct modes based on their timing relative to vehicle behaviour, with pedestrians choosing to cross either well before a vehicle slows down or waiting until it is clearly safe to do so. In our study, early and late crossings were defined based on CIT, using as threshold the moment at which the vehicle speed had dropped to two-thirds of its initial speed, which approximately corresponded to the minimum between the two peaks in the bimodal CIT distribution. It should be noted that the distinction between early and late crossings is applicable only in yielding scenarios.

3.3.2 Model

Sensory-motor mechanisms

In this paper, we explore the influence of four sensory-motor mechanisms on the crossing behavior as detailed in Section 3.2.1, noisy perception, visual looming, time pressure, walking effort and ballistic speed control; as illustrated in Figure 3.1) (a) [3].

[2]Ethical approval was obtained from the University of Leeds Research Committee (Ref: 0536).

[3]In the published version, 'time pressure' was mistakenly included in this list. The model described in this chapter in fact included four mechanisms: noisy perception, visual looming, walking effort, and ballistic speed control.

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 to approximate the state of the environment refer to the previous paper for mathematical details (Wang et al., 2025).

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 (Markkula et al., 2016; DeLucia, 2015). 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.

Walking effort To account for walking effort, we adopt the biomechanical model of walking from Carlisle and Kuo (2023). The equation for the new velocity v_i^+ is:

$$v_i^+ = v_i^- \cos 2\theta + \sqrt{2u_i} \sin 2\theta \tag{3.1}$$

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 3.1 starts with the assumption that the pedestrian's leg motion can be described as a pendulum (see panel (d) of Figure 3.1), and that the speed change is caused by the effort u_i. This effort required for this change is related to the kinetic energy:

$$u_i = \frac{1}{2} m (\Delta v)^2 \tag{3.2}$$

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:

$$u_i = \frac{\left(v^- \cos(2\theta) - v^+\right)^2}{2 \cdot \sin^2(2\theta)} \tag{3.3}$$

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.

Ballistic speed control As mentioned in Section 3.2.1, 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 $s = {v_{\text{ped}}}^{0.42}$ (Grieve, 1968). Consequently, the duration of a walking step is given b $T_{\text{step}} = {v_{\text{ped}}}^{-0.58}$.

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.

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.

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.

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.

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.

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 (Howes et al., 2023; Li et al., 2023), which we refer to as non-policy parameters to distinguish them from the parameters of the policy neural network (connection

weights and biases).

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 (Vizzari and Cecconello, 2022; Sun et al., 2019; Xu et al., 2020a) was used.

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

Reinforcement learning

Proximal Policy Optimization (PPO) is a policy gradient RL algorithm that improves training stability and reliability by using a clipped objective function (Schulman et al., 2017). 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 [4].

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 (Raffin et al., 2021). The model was trained for 3 million environment time steps using the PPO algorithm. The parameters for training are shown in Table 3.2. The reward and total loss during training are shown in Fig. 3.2. Due to differences in POMDP structure between models, their final reward

[4]Further information not present in the published article: the choice of PPO in this chapter reflected the relatively constrained optimisation problem, in which only the pedestrian policy was learned and the action space consisted of discrete walking-speed choices. A more detailed discussion of RL algorithm choice is provided in Section 6.2.2.

Table 3.2: 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.

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.

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

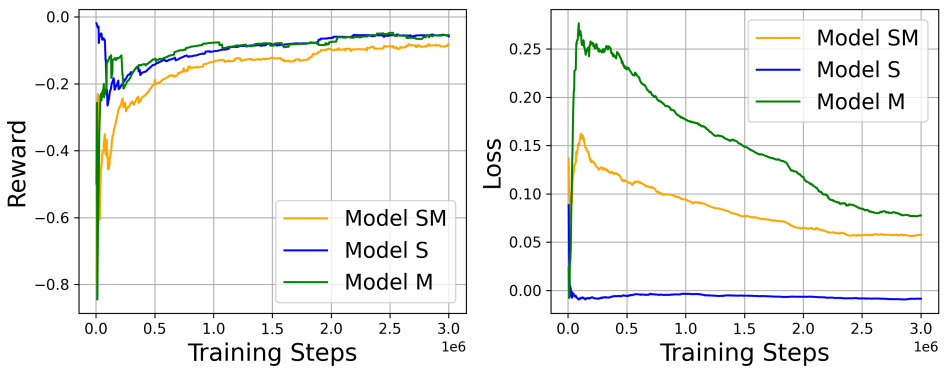

Figure 3.2: Reward and training loss of different model variants.

Figure 3.3: 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 (Gutmann, Cor, et al., 2016). This inference method has been previously applied to RL-based simulation models of human behavior (Kangasrääsiö et al., 2017; Moon et al., 2022). 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:

$$\begin{aligned}\text{Discrepancy} = \ \log\Bigg(&\left(\sum_{i=1}^{n_{\text{ny}}} \frac{|\text{g}_i - \hat{\text{g}}_i|}{\text{g}_{\max}}\right) + \left(\sum_{i=1}^{n_y} \frac{|\text{e}_i - \hat{\text{e}}_i|}{\text{e}_{\max}}\right) \\ &+ \left(\sum_{i=1}^{n} \frac{\left|\text{CIT}_i - \hat{\text{CIT}}_i\right|}{\text{CIT}_{\max}}\right) + \left(\sum_{i=1}^{n} \frac{|\text{v}_i - \hat{\text{v}}_i|}{\text{v}_{\max}}\right)\Bigg)\end{aligned} \tag{3.4}$$

where the ˆ 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 3.4 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 3.3.2, and ran 80 optimization iterations, which was sufficient to achieve convergence for all model variants.

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.

Table 3.3: 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	

3.4 Results

All the behavioral patterns we focus on in this study are summarized in Figure 3.3, with the gray cells indicating the statistically significant effects of the independent variable on pedestrian

Table 3.4: 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	$< .001$	***

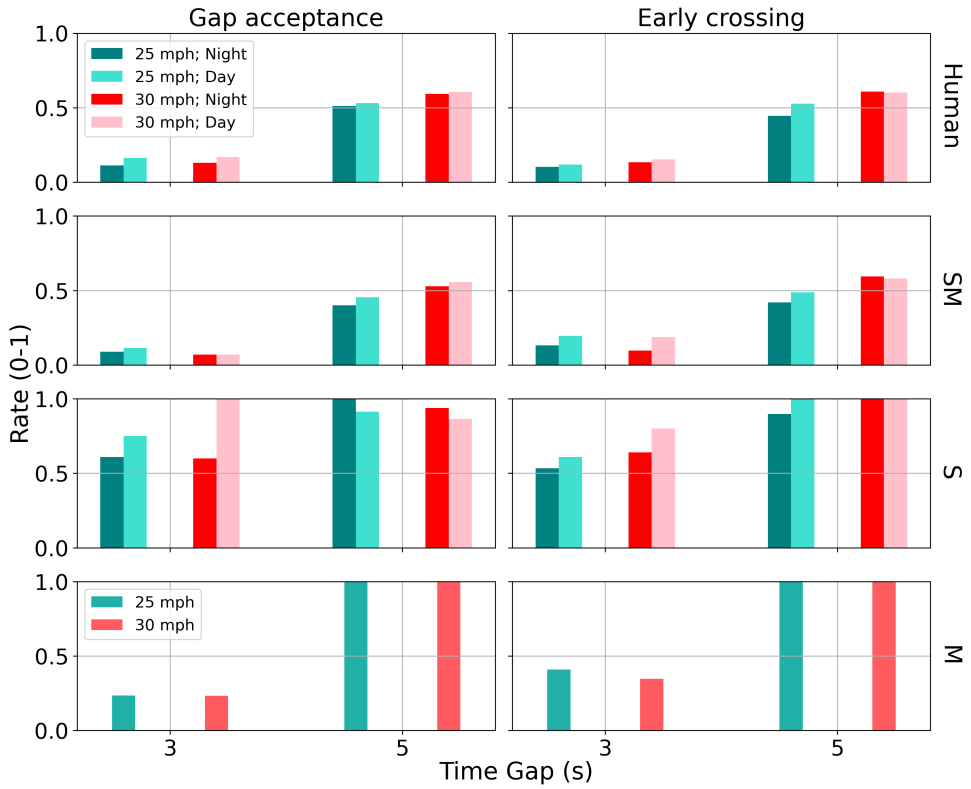

Figure 3.4: Gap acceptance rate in non-yielding scenarios and early crossing rate in yielding scenarios as observed in human behavior and predicted by the models.

behaviors observed in the experiment. The cells with symbols mean that the corresponding model variant captures that behavioral pattern.

3.4.1 Experimental results

In this section, we present the experimental findings for the four behavioral metrics introduced in Section 3.3.2: 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.

Gap acceptance and early crossing rate

As can be seen in the first row of Figure 3.4, 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 3.2. Additionally, pedestrians were more inclined to cross during the daytime compared to nighttime ($p < 0.05$).

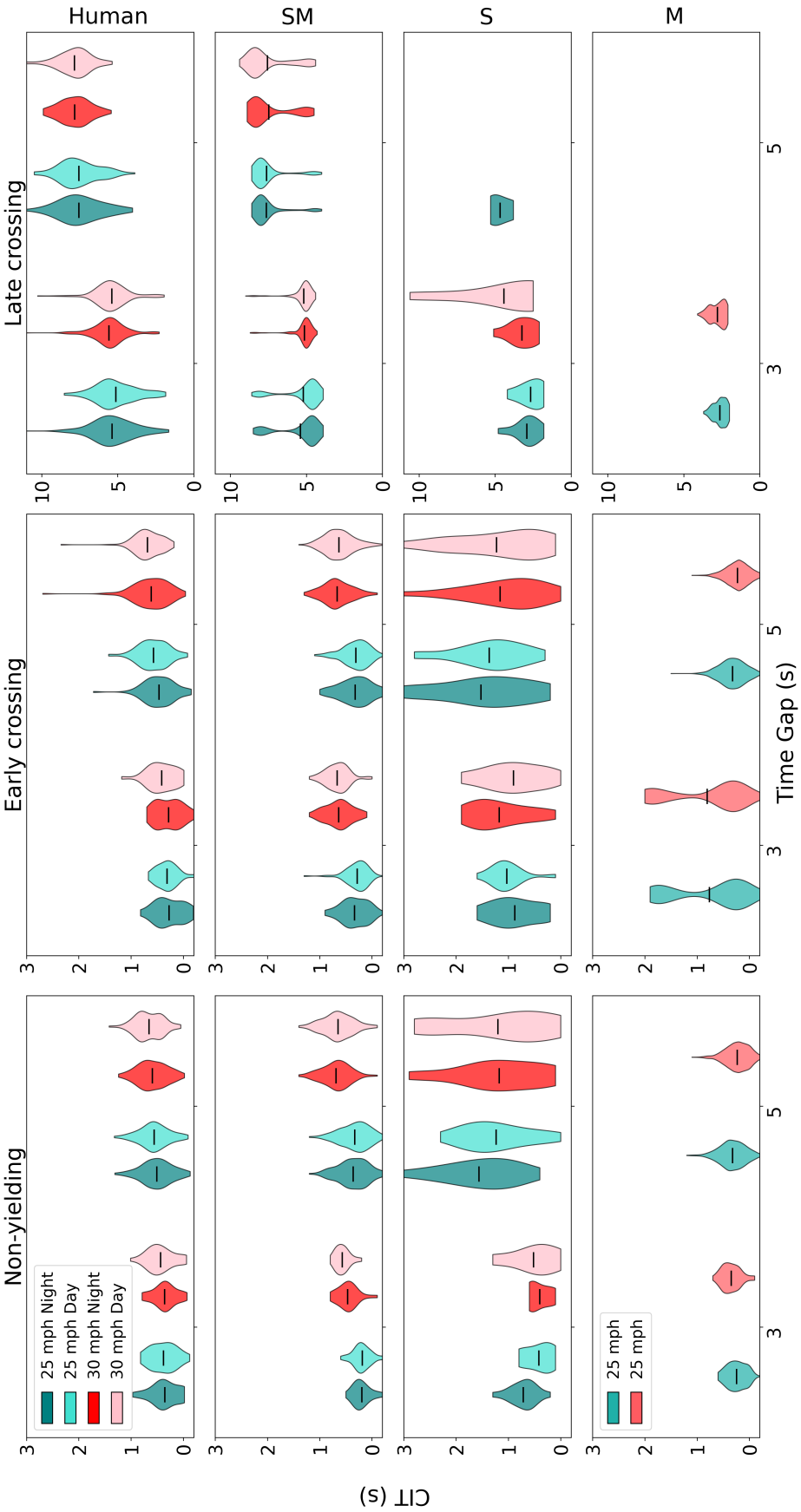

Figure 3.5: Crossing Initiation Time (CIT) under different conditions. This figure shows the distribution of CIT under different conditions. The first row shows human data, the second row presents the VLW model result (accounting for walking effort), and the third row shows the VL model result (without 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 3.5: 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	*

Crossing initiation time (CIT)

As can be seen in the first row of Figure 3.5, 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 3.4, eHMI had an effect on CIT in late crossing scenarios ($p < .001$). Specifically, when eHMI was on, pedestrians started to cross earlier.

Average walking speed

From the first row of Figure 3.6, 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 Figure 3.3.

3.4.2 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 Figure 3.3.

The overall collision rate in the experiment was 0.002, indicating that collisions were rare [5]. The collision rates for the SM, S, and M models were somewhat higher but still small, at 0.027,

[5]Here 'rare' should be understood only in relation to this specific experimental setting. A broader discussion is provided in Section 6.3.

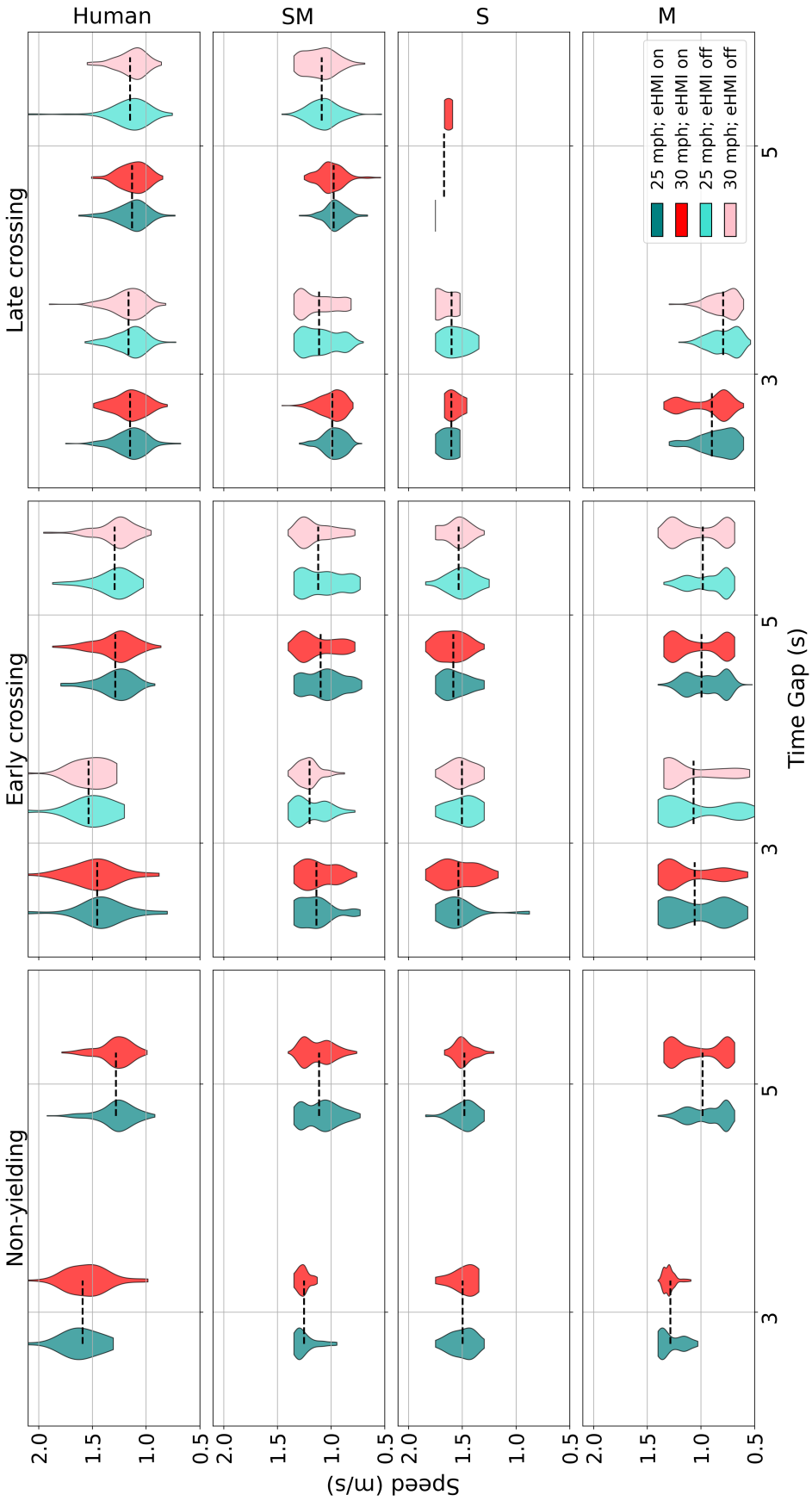

Figure 3.6: Average walking speed. This figure illustrates the average walking speeds for different crossing decisions—non-yielding, early crossing, and late crossing—under different vehicle speeds (25 mph and 30 mph) and eHMI conditions (on and off). Each pair of violins represents different vehicle 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.

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.

Gap acceptance and early crossing rate

Rows 2-4 of Figure 3.4 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 Figure 3.4, 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.

CIT

Rows 2-4 of Figure 3.5 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.

Average walking speed

Rows 2-4 of Figure 3.6 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.

The pedestrian speed profile

Figure 3.7 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 Figure 3.7, using direct speed control (Vizzari and Cecconello, 2022; Sun et al., 2019; Xu et al., 2020a) 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.

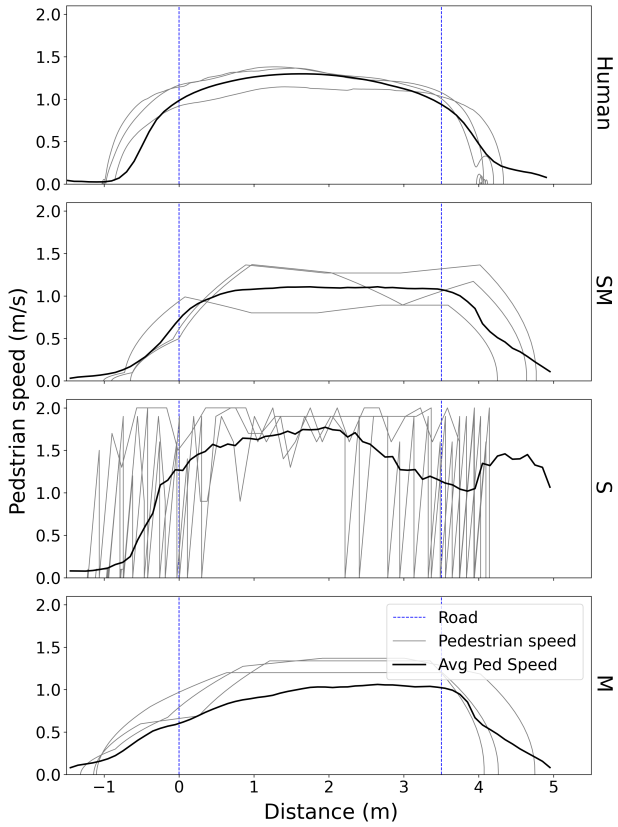

Figure 3.7: Average walking speed profile of the experiment, and different model variants. Gray lines are the randomly selected speed curves and black lines are the averaged speed curves of all conditions. The blue dashed lines denote the road curbs, and the walking direction is from left to right.

3.5 Discussion

3.5.1 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 Figure 3.3, 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.

3.5.2 Impact of sensory-motor constraints on the pedestrian crossing locomotion

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 (Song et al., 2021).

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.

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 Figure 3.3, 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 (Wang et al., 2025). 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.

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.

3.5.3 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 (Camara et al., 2020). Furthermore, the model can be integrated into virtual testing environments, providing a valuable tool for evaluating AV systems under various pedestrian crossing scenarios (Rasouli and Tsotsos, 2019). 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 (Rasouli and Tsotsos, 2019; Camara et al., 2020).

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 (Lee et al., 2026). 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 (Kalantari et al., 2023), 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.

3.6 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

Alhawiti, Ali, Kwigizile, Valerian, Oh, Jun-Seok, Asher, Zachary D., Hakimi, Obaidullah, Aljohani, Saad, and Ayantayo, Sherif (2024). "The Effectiveness of eHMI Displays on Pedestrian–Autonomous Vehicle Interaction in Mixed-Traffic Environments". In: *Sensors* 24.15, pp. 5018–5018. DOI: 10.3390/s24155018.

Barton, Sean L, Matthis, Jonathan S, and Fajen, Brett R (2019). "Control strategies for rapid, visually guided adjustments of the foot during continuous walking". In: *Experimental brain research* 237, pp. 1673–1690.

Benrachou, Djamel Eddine, Glaser, Sebastien, Elhenawy, Mohammed, and Rakotonirainy, Andry (2022). "Use of social interaction and intention to improve motion prediction within automated vehicle framework: A review". In: *IEEE Transactions on Intelligent Transportation Systems* 23.12, pp. 22807–22837.

Camara, Fanta, Bellotto, Nicola, Cosar, Serhan, Weber, Florian, Nathanael, Dimitris, Althoff, Matthias, Wu, Jingyuan, Ruenz, Johannes, Dietrich, André, Markkula, Gustav, et al. (2020). "Pedestrian models for autonomous driving part ii: high-level models of human behavior". In: *IEEE Transactions on Intelligent Transportation Systems* 22.9, pp. 5453–5472.

Carlisle, Rebecca Elizabeth and Kuo, Arthur D (2023). "Optimization of energy and time predicts dynamic speeds for human walking". In: *Elife* 12, e81939.

Choi, Tae Ho (1997). *Development of a mathematical model of gait dynamics*.

Craik, Kenneth JW (1948). "Theory of the human operator in control systems. II. Man as an element in a control system". In: *British journal of psychology* 38.3, p. 142.

DeLucia, Patricia R (2008). "Critical roles for distance, task, and motion in space perception: Initial conceptual framework and practical implications". In: *Human Factors* 50.5, pp. 811–820.

DeLucia, Patricia R. (2015). "Perception of Collision". In: *The Cambridge Handbook of Applied Perception Research*. Ed. by Robert R. Hoffman, Peter A. Hancock, Mark W. Scerbo, Raja Parasuraman, and James L.Editors Szalma. Cambridge Handbooks in Psychology. Cambridge University Press, pp. 568–591. DOI: 10.1017/CBO9780511973017.035.

Diaz-Ruiz, Carlos A, Xia, Youya, You, Yurong, Nino, Jose, Chen, Junan, Monica, Josephine, Chen, Xiangyu, Luo, Katie, Wang, Yan, Emond, Marc, et al. (2022). "Ithaca365: Dataset and Driving Perception Under Repeated and Challenging Weather Conditions". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 21383–21392.

Faisal, A Aldo, Selen, Luc PJ, and Wolpert, Daniel M (2008). "Noise in the nervous system". In: *Nature reviews neuroscience* 9.4, pp. 292–303.

Faraji, Salman, Wu, Amy R., and Ijspeert, Auke Jan (2018). "A simple model of mechanical effects to estimate metabolic cost of human walking". In: *Scientific Reports* 8.1, pp. 10998–10998. DOI: 10.1038/S41598-018-29429-Z.

Gershman, Samuel J, Horvitz, Eric J, and Tenenbaum, Joshua B (2015). "Computational rationality: A converging paradigm for intelligence in brains, minds, and machines". In: *Science* 349.6245, pp. 273–278.

Giles, Oscar, Markkula, Gustav, Pekkanen, Jami, Yokota, Naoki, Matsunaga, Naoto, Merat, Natasha, and Daimon, Tatsuru (2019). "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, pp. 366–372.

Grieve, DW (1968). "Gait patterns and the speed of walking." In: *Bio-Med. Eng.* 3, pp. 119–112.

Gutmann, Michael U, Cor, Jukka, et al. (2016). "Bayesian optimization for likelihood-free inference of simulator-based statistical models". In: *Journal of Machine Learning Research* 17.125, pp. 1–47.

Hollmann, Claudia (2015). "A cognitive human behaviour model for pedestrian behaviour simulation". Doctoral dissertation. University of Greenwich.

Hoogendoorn, Serge and HL Bovy, Piet (2003). "Simulation of pedestrian flows by optimal control and differential games". In: *Optimal control applications and methods* 24.3, pp. 153–172.

Howes, Andrew, Jokinen, Jussi PP, and Oulasvirta, Antti (2023). "Towards machines that understand people". In: *AI Magazine* 44.3, pp. 312–327.

Izquierdo, Rubén, Alonso, Javier, Benderius, Ola, Sotelo, Miguel Ángel, and Llorca, David Fernández (2024). "Pedestrian and Passenger Interaction with Autonomous Vehicles: Field Study in a Crosswalk Scenario". In: *International Journal of Human-computer Interaction*, pp. 1–19. DOI: 10.1080/10447318.2024.2426856.

Kaelbling, Leslie Pack, Littman, Michael L, and Moore, Andrew W (1996). "Reinforcement learning: A survey". In: *Journal of artificial intelligence research* 4, pp. 237–285.

Kalantari, Amir Hossein, Lin Yi-Shin, yslin, Mohammadi, Ali, Merat, Natasha, and Markkula, Gustav (Sept. 2023). *Testing the validity of multi participant distributed simulation for understanding and modeling road user interaction*. DOI: 10.31234/osf.io/gk9af. URL: osf.io/preprints/psyarxiv/gk9af.

Kalantarov, Semyon, Riemer, Raziel, and Oron-Gilad, Tal (2018). "Pedestrians' road crossing decisions and body parts' movements". In: *Transportation research part F: traffic psychology and behaviour* 53, pp. 155–171.

Kangasrääsiö, Antti, Athukorala, Kumaripaba, Howes, Andrew, Corander, Jukka, Kaski, Samuel, and Oulasvirta, Antti (2017). "Inferring cognitive models from data using approximate Bayesian computation". In: *Proceedings of the 2017 CHI conference on human factors in computing systems*, pp. 1295–1306.

Kotseruba, Iuliia and Rasouli, Amir (2023). "Intend-Wait-Perceive-Cross: Exploring the Effects of Perceptual Limitations on Pedestrian Decision-Making". In: *2023 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, pp. 1–6.

Lee, Yee Mun, Madigan, Ruth, Wang, Yueyang, Garcia, Jorge, Qin, Hao, Srinivasan, Aravinda, Markkula, Gustav, and Merat, Natasha (2026). "Crossing in the dark: Investigating the effect

of vehicle kinematics and eHMI on older pedestrians' crossing behavior in a virtual reality experiment". In: *Journal of Safety Research* 96, pp. 315–328.

Lee, Yee Mun, Madigan, Ruth, Wang, Yueyang, Qin, Hao, Srinivasan, Aravinda, Markkula, Gustav, and Merat, Natasha (2024). "Under the Moonlight: The Effect of Vehicle Kinematics and eHMI on Older Pedestrians' Crossing Behaviour". In: *PsyArXiv preprint osf.io/preprints/psyarxiv/a2m9w*.

Li, Zhi, Ko, Yu-Jung, Putkonen, Aini, Feiz, Shirin, Ashok, Vikas, Ramakrishnan, Iv, Oulasvirta, Antti, and Bi, Xiaojun (2023). "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*. CHI '23. Hamburg, Germany: Association for Computing Machinery, p. 18. ISBN: 9781450394215. DOI: 10.1145/3544548.3580640. URL: https://doi.org/10.1145/3544548.3580640.

Lobjois, Régis and Cavallo, Viola (2007). "Age-related differences in street-crossing decisions: The effects of vehicle speed and time constraints on gap selection in an estimation task". In: *Accident analysis & prevention* 39.5, pp. 934–943.

Madala, Kaushik and Gonzalez, Carlos Avalos (2023). "Metrics for machine learning models to facilitate SOTIF analysis in autonomous vehicles". In: *SAE International Journal of Advances and Current Practices in Mobility* 6.2023-01-0829, pp. 782–790.

Man, Siu Shing, Huang, Chuyu, Ye, Qing, Chang, Fangrong, and Chan, Alan Hoi Shou (2025). "Pedestrians' Interaction with eHMI-equipped Autonomous Vehicles: A Bibliometric Analysis and Systematic Review". In: *Accident Analysis & Prevention* 209, p. 107826.

Manning, Tyler S, Pillow, Jonathan W, Rokers, Bas, and Cooper, Emily A (2022). "Humans make non-ideal inferences about world motion". In: *Journal of Vision* 22.14, pp. 4054–4054.

Markkula, Gustav, Engström, Johan, Lodin, Johan, Bärgman, Jonas, and Victor, Trent (2016). "A farewell to brake reaction times? Kinematics-dependent brake response in naturalistic rear-end emergencies". In: *Accident Analysis & Prevention* 95, pp. 209–226.

Markkula, Gustav, Lin, Yi-Shin, Srinivasan, Aravinda Ramakrishnan, Billington, Jac, Leonetti, Matteo, Kalantari, Amir Hossein, Yang, Yue, Lee, Yee Mun, Madigan, Ruth, and Merat,

Natasha (2023). "Explaining human interactions on the road by large-scale integration of computational psychological theory". In: *PNAS nexus* 2.6, pgad163.

Moon, Hee-Seung, Do, Seungwon, Kim, Wonjae, Seo, Jiwon, Chang, Minsuk, and Lee, Byungjoo (2022). "Speeding up inference with user simulators through policy modulation". In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pp. 1–21.

Mulier, Lana, Slabbinck, Hendrik, and Vermeir, Iris (2024). "Face your fears: direct and indirect measurement of responses to looming threats". In: *Cognition and Emotion* 38.1, pp. 187–197.

Oulasvirta, Antti, Jokinen, Jussi PP, and Howes, Andrew (2022). "Computational rationality as a theory of interaction". In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pp. 1–14.

Oxley, Jennifer, Corben, Bruce, Fildes, Brian, and Charlton, Judith (2004). "Older pedestrians: Meeting their safety and mobility needs". In: *Proceedings of the Road Safety Research, Policing and Education Conference, Perth, Australia*, pp. 14–16.

Papathanasopoulou, Vasileia, Spyropoulou, Ioanna, Perakis, Harris, Gikas, Vassilis, and Andrikopoulou, Eleni (2022). "A data-driven model for pedestrian behavior classification and trajectory prediction". In: *IEEE Open Journal of Intelligent Transportation Systems* 3, pp. 328–339.

Pekkanen, Jami, Giles, Oscar Terence, Lee, Yee Mun, Madigan, Ruth, Daimon, Tatsuru, Merat, Natasha, and Markkula, Gustav (2022). "Variable-drift diffusion models of pedestrian road-crossing decisions". In: *Computational Brain & Behavior* 5.1, pp. 60–80.

Quan, Ruijie, Zhu, Linchao, Wu, Yu, and Yang, Yi (2021). "Holistic LSTM for pedestrian trajectory prediction". In: *IEEE transactions on image processing* 30, pp. 3229–3239.

Raffin, Antonin, Hill, Ashley, Gleave, Adam, Kanervisto, Anssi, Ernestus, Maximilian, and Dormann, Noah (2021). "Stable-Baselines3: Reliable Reinforcement Learning Implementations". In: *Journal of Machine Learning Research* 22.268, pp. 1–8. URL: `http://jmlr.org/papers/v22/20-1364.html`.

Rasouli, Amir and Tsotsos, John K (2019). "Autonomous vehicles that interact with pedestrians: A survey of theory and practice". In: *IEEE transactions on intelligent transportation systems* 21.3, pp. 900–918.

Roshdi, Mohamed, Petzold, Julian, Wahby, Mostafa, Ebrahim, Hussein, Berekovic, Mladen, and Hamann, Heiko (2024). "On the Road to Clarity: Exploring Explainable AI for World Models in a Driver Assistance System". In: *2024 IEEE Conference on Artificial Intelligence (CAI)*, pp. 1032–1039. DOI: 10.1109/CAI59869.2024.00187.

Schulman, John, Wolski, Filip, Dhariwal, Prafulla, Radford, Alec, and Klimov, Oleg (2017). "Proximal policy optimization algorithms". In: *arXiv preprint arXiv:1707.06347*.

Song, Seungmoon, Kidziński, Łukasz, Peng, Xue Bin, Ong, Carmichael, Hicks, Jennifer, Levine, Sergey, Atkeson, Christopher G, and Delp, Scott L (2021). "Deep reinforcement learning for modeling human locomotion control in neuromechanical simulation". In: *Journal of neuroengineering and rehabilitation* 18, pp. 1–17.

Srinivasan, Aravinda Ramakrishnan, Lin, Yi-Shin, Antonello, Morris, Knittel, Anthony, Hasan, Mohamed, Hawasly, Majd, Redford, John, Ramamoorthy, Subramanian, Leonetti, Matteo, Billington, Jac, et al. (2023). "Beyond RMSE: Do machine-learned models of road user interaction produce human-like behavior?" In: *IEEE Transactions on Intelligent Transportation Systems*.

Sun, Libo, Zhai, Jinfeng, and Qin, Wenhu (2019). "Crowd Navigation in an Unknown and Dynamic Environment Based on Deep Reinforcement Learning". In: *IEEE Access* 7, pp. 109544–109554. DOI: 10.1109/ACCESS.2019.2933492.

Tian, Kai, Markkula, Gustav, Wei, Chongfeng, Lee, Yee Mun, Madigan, Ruth, Merat, Natasha, and Romano, Richard (2022). "Explaining unsafe pedestrian road crossing behaviours using a psychophysics-based gap acceptance model". In: *Safety science* 154, p. 105837.

Tustin, Arnold (1947). "The nature of the operator's response in manual control, and its implications for controller design". In: *Journal of the Institution of Electrical Engineers-Part IIA: Automatic Regulators and Servo Mechanisms* 94.2, pp. 190–206.

Vizzari, Giuseppe and Cecconello, Thomas (2022). "Pedestrian simulation with reinforcement learning: a curriculum-based approach". In: *Future Internet* 15.1, p. 12.

Wang, Yueyang, Srinivasan, Aravinda Ramakrishnan, Jokinen, Jussi PP, Oulasvirta, Antti, and Markkula, Gustav (2023). "Modeling human road crossing decisions as reward maximization with visual perception limitations". In: *2023 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, pp. 1–6.

– (2025). "Pedestrian crossing decisions can be explained by bounded optimal decision-making under noisy visual perception". In: *Transportation Research Part C: Emerging Technologies* 171, p. 104963.

Xu, Chen, Dong, Changyin, Yang, Wenzhang, Hou, Yuxuan, and Wang, Hao (2024). "Platoon control and external human–machine interfaces: innovations in pedestrian–autonomous vehicle interactions". In: *Transportmetrica*, pp. 1–32. DOI: 10.1080/23249935.2024.2398594.

Xu, Dong, Huang, Xiao, Li, Zhenlong, and Li, Xiang (2020a). "Local motion simulation using deep reinforcement learning". In: *Transactions in GIS* 24.3, pp. 756–779.

Xu, Yiran, Yang, Xiaoyin, Gong, Lihang, Lin, Hsuan-Chu, Wu, Tz-Ying, Li, Yunsheng, and Vasconcelos, Nuno (2020b). "Explainable object-induced action decision for autonomous vehicles". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 9523–9532.

Chapter 4

Realistic pedestrian-driver interaction modelling using multi-agent RL with human perceptual-motor constraints

Abstract

Modelling pedestrian-driver interactions is critical for understanding human road user behaviour and developing safe autonomous vehicle systems. Existing approaches often rely on rule-based logic, game-theoretic models, or 'black-box' machine learning methods. However, these models typically lack flexibility or overlook the underlying mechanisms, such as sensory and motor constraints, which shape how pedestrians and drivers perceive and act in interactive scenarios. In this study, we propose a multi-agent reinforcement learning (RL) framework that integrates both visual and motor constraints of pedestrian and driver agents. Using a real-world dataset from an unsignalised pedestrian crossing, we evaluate four model variants—one without constraints, two with either motor or visual constraints, and one with both—across behavioural metrics of interaction realism. Results show that the combined model with both visual and motor constraints performs best. Motor constraints lead to smoother movements that resemble human speed adjustments during crossing interactions. The addition of visual constraints introduces

perceptual uncertainty and field-of-view limitations, leading the agents to exhibit more cautious and variable behaviour, such as less abrupt deceleration. In this data-limited setting, our model outperforms a supervised behavioural cloning model, demonstrating that our approach can be effective without large training datasets. Finally, our framework accounts for individual differences by modelling parameters controlling the human constraints as population-level distributions, a perspective that has not been explored in previous work on pedestrian-vehicle interaction modelling. Overall, our work demonstrates that multi-agent RL with human constraints is a promising modelling approach for simulating realistic road user interactions.

4.1 Introduction

4.1.1 Background

The development of autonomous vehicles (AVs) has gained significant momentum in recent years, positioning them as a key component of future mobility systems. One major challenge in deploying AVs is the lack of realistic models that can capture the subtle and dynamic interactions between human road users in mixed traffic (Rasouli and Tsotsos, 2019). Pedestrians deserve particular attention because they are vulnerable road users and often behave less predictably than drivers, who are constrained by vehicle dynamics and traffic regulations (González-Méndez et al., 2021). To develop realistic pedestrian models, it is therefore crucial to account for the mutual adaptations between pedestrians and drivers, which fundamentally shape crossing decisions and interaction patterns in real-world traffic.

Current approaches to modelling road user interaction can be broadly categorised into two types: mechanistic models and data-driven machine learning (ML) models. Mechanistic models include cognitive models, physics-based models, and game-theoretic models. These models rely on strong behavioural assumptions. For example, cognitive models are based on hypotheses about underlying mental processes and aim to represent how human behaviour is generated (Pekkanen et al., 2022; Markkula et al., 2023). Such models offer high interpretability and can provide deeper insights into the mechanisms behind observed behaviours. However, the high complexity of these models limits their scalability and generalisability.

In contrast, data-driven models, such as LSTMs and Transformers, learn behavioural patterns directly from data (Yang et al., 2025; Divya et al., 2025). Given sufficient data and computational resources, they can achieve high predictive accuracy on trajectory-based metrics such as root mean square error (RMSE). Nevertheless, they often lack interpretability, which is a key requirement in safety-critical applications like autonomous driving (Aradi, 2020; Itkina and Kochenderfer, 2023). In addition, their performance tends to degrade in data-limited or rare situations, which often correspond to safety-critical events where accurate behavioural prediction is important.

Another class of algorithms with the potential to bridge the aforementioned gap is reinforcement learning (RL). Modelling human behaviour using RL builds on the theoretical assumption that humans act to maximise expected rewards, while deep RL provides a powerful computational

approach to approximate such reward-maximising behaviour policies (Sutton and Barto, 2018; Howes et al., 2023). This framework offers the flexibility to generalise across a wider range of contexts and scenarios (Chen et al., 2015; Chen et al., 2021a; Jokinen et al., 2021).

In our previous work, we developed the pedestrian crossing model using RL that incorporated human perceptual and motor constraints, resulting in human-like crossing timing and speed (Wang et al., 2025a; Wang et al., 2025b). However, these studies were limited in two important ways. First, only the pedestrian was modelled with a single-agent policy, while the vehicle followed a preprogrammed trajectory, preventing realistic mutual adaptation between the two agents. Second, they were evaluated using data from controlled experiments under simplified laboratory conditions, which limited the diversity of interaction contexts. Third, although between-individual variability was considered by fitting fixed non-policy parameters for each participant across repeated trials, this approach relied on multiple observations per individual and does not generalise well to real-world situations.

In this work, we address these limitations by modelling both pedestrian and vehicle behaviours within a multi-agent RL framework, where both agents adapt to each other's actions using a real-world dataset of unsignalised crossing scenarios. Building on prior efforts to incorporate human perceptual and motor constraints, this study is the first to embed such constraints within a fully interactive, two-agent setting grounded in real-world data. Furthermore, we introduce a novel population-level parameter fitting procedure that captures between-individual variability by modelling human constraint-related parameters as distributions across the population, from which each individual's parameters are drawn. This approach makes it possible to model individual differences in road user behaviour even when only a small amount of data is available.

4.1.2 Related work

Given the importance of pedestrian-vehicle interaction in the safe deployment of AVs, various modelling approaches have been developed to capture and predict such behaviour. In this section, we review relevant studies, organised by methodological category.

Mechanistic models

Mechanistic models build on physics, psychology, or cognitive science, and explain road user behaviour through explicit mathematical formulations. These approaches offer interpretability

and explanation of the behaviour but often struggle to capture the flexibility of human behaviour in dynamic traffic environments.

Early models often used logistic regression to predict binary crossing decisions, typically based on variables such as vehicle speed and distance (Lobjois and Cavallo, 2007; Tian et al., 2022b). Extensions of these models incorporated perceptual cues to explain why certain decisions are made (Tian et al., 2022a). However, these models are inherently limited in scope, as their binary outputs and reliance on simple decision thresholds make them unsuitable for simulating the continuous and adaptive nature of real-world road behaviour.

Game-theoretic models, which simulate strategic reasoning between interacting agents, have also been applied. For example, Fox et al. (2018) modelled pedestrian-AV interactions to support safe vehicle planning. More recently, Dang et al. (2025) proposed a dynamic game-theoretic framework incorporating bounded rationality, improving upon classical models by relaxing the assumption of fully rational agents. However, several practical limitations persist. For instance, the model assumed that agents have perfect sensory information and can execute actions without motor constraints. Moreover, pedestrian actions were restricted to a low-dimensional, discrete set of walking speeds, without the possibility to continuously adjust speed or direction, primarily due to the computational difficulty in solving game-theoretic formulations in high-dimensional continuous spaces.

Evidence accumulation models, grounded in psychology and cognitive science, have also been applied to road user behaviour modelling (Giles et al., 2019; Pekkanen et al., 2022). In this approach, decision-making is modelled as a noisy integration of sensory and contextual information over time, terminating in an action once a decision threshold is reached. These models aim to capture the gradual nature of human decision processes and offer insight into the underlying cognitive mechanisms. However, modelling complex road-user interactions using evidence accumulation alone remains challenging. To address this, Markkula et al. (2023) proposed an integrated model that combines noisy perception, evidence accumulation, and game-theoretic reasoning, enabling the reproduction of several empirically observed behaviours. The authors showed that to simulate even highly simplified pedestrian-vehicle interactions (e.g., requiring both agents to travel along straight perpendicular paths) requires the integration of multiple theoretical frameworks. Due to their structural complexity and reliance on strong assumptions, such mechanistic models are difficult to generalise and scale to more varied and complex

real-world traffic environments.

Data-driven models

In this paper, we define pure data-driven models as those that directly learn the mapping from input to output using ground-truth data, typically through neural networks trained on observed motion histories, without relying on handcrafted rules or explicit behavioural assumptions.

A common strategy for modelling road user interaction is to formulate trajectory prediction as a time-series forecasting problem, in which the model predicts future trajectories over a fixed horizon based on preceding motion. For example, Deo and Trivedi (2018) proposed an LSTM-based architecture that captures spatial relationships between vehicles to forecast human-driven vehicle trajectories. Other architectures have also been explored for interaction modelling. Graph-based neural networks represent the relational structure among agents in a scene (Yi et al., 2016; Ye et al., 2021), while more recently, Transformer-based models have gained attention for pedestrian behaviour modelling due to their ability to capture long-range dependencies and efficiently process sequential data (Yang et al., 2025; Divya et al., 2025).

Although these models achieve strong performance on trajectory-based accuracy metrics, several limitations remain. First, data-driven models are often criticised as 'black boxes': they lack interpretability and provide little insight into the underlying mechanisms of the behaviours they predict (Madala and Gonzalez, 2023). This lack of transparency is concerning in the context of autonomous driving, since behaviour models are often used within AV decision-making pipelines, and unexplained prediction errors may propagate into AV planning and pose safety risks (Roshdi et al., 2024). Second, high predictive accuracy on an aggregate level does not necessarily imply that the model captures behaviourally meaningful features of an interaction (Srinivasan et al., 2023). For example, as shown by Srinivasan et al. (2023), machine-learned trajectory predictors with low RMSE can still fail to reproduce human-like interaction patterns such as courtesy lane changes or collision aversion. Third, the absence of theoretical grounding can reduce the robustness of such models. Although they may perform well on training data, their generalisability to rare or unseen scenarios is often limited (Xu et al., 2020), and they depend heavily on large, diverse training datasets. In practice, collecting sufficient data to cover the full range of road user behaviours is challenging, as many safety-critical or rare events are underrepresented, or entirely missing, from existing datasets (Diaz-Ruiz et al., 2022).

Reinforcement learning models

Different from data-driven models, RL agents learn decision policies through trial-and-error interactions with the environment in order to maximise cumulative reward. Owing to its sequential decision-making nature and its ability to capture interactions between agents and their environment, RL has been applied to simulate road user interactions (Charalambous and Chrysanthou, 2019; Vizzari and Cecconello, 2022).

One major challenge in behaviour modelling with RL lies in designing reward functions that realistically reflect human intentions. Inverse reinforcement learning (IRL) addresses this by inferring latent reward functions from demonstrated behaviour (Ng, Russell, et al., 2000). For instance, Nasernejad et al. (2021) and Nasernejad et al. (2023) used IRL to model pedestrian-vehicle interactions in near-miss scenarios, recovering reward structures from real-world demonstrations. These models are typically formalised as Markov Decision Processes (MDPs), which assume full observability of the environment. However, human decision-making often occurs under uncertainty, making Partially Observable MDPs (POMDPs) a more appropriate framework. Consequently, by assuming full access to environmental information, MDP-based RL models overlook key human limitations such as perceptual uncertainty, leading to reward functions that may misrepresent the decision-making processes of real road users.

One line of work which has adopted the POMDP perspective and used RL to develop human models is computational rationality, which assumes that humans behave (near-) optimally given internal utility functions and cognitive or physical constraints (Gershman et al., 2015; Oulasvirta et al., 2022). This approach was initially applied in human-computer interaction (Chen et al., 2015; Tseng and Howes, 2015; Jokinen et al., 2020). More recently, we have applied it to pedestrian crossing decisions, incorporating cognitive mechanisms such as perceptual noise (Wang et al., 2023; Wang et al., 2025a). We subsequently extended this approach to simulate the full pedestrian crossing process, including responses to lighting conditions and external human-machine interfaces (Wang et al., 2025b). However, these models were developed and validated in controlled experimental settings: Participants always started from standstill at the kerb, omitting the pedestrian approach phase; participants experienced multiple similar interactions during the experiment, which may have caused some learning effects. Moreover, the car was preprogrammed and did not respond to the pedestrian's behaviour, whereas two-sided interaction is known to be important for real-world driver-pedestrian interaction (Schneemann

and Gohl, 2016).

Two-sided, or joint, behaviour is a key characteristic of natural traffic interactions, as pedestrians and drivers continuously adapt to each other's actions (Domeyer et al., 2022). This has motivated the use of multi-agent reinforcement learning (MARL) to model interactive road user behaviour. Recent MARL studies have explored cooperative and competitive interactions among vehicles or between pedestrians and vehicles (Cornelisse and Vinitsky, 2024; Sackmann, 2024), showing that multi-agent frameworks can capture the bidirectional adaptation that is absent in single-agent formulations. However, these approaches typically focus on optimising interaction outcomes or reproducing empirical patterns, without accounting for human perceptual or motor constraints that shape how such adaptations emerge. In contrast, past computational rationality models have focused on human perceptual and motor limitations but have rarely extended to fully interactive, two-agent settings. The present study bridges this gap by combining computational rationality with a MARL framework, enabling both pedestrian and driver agents to adapt to each other under human-like sensory and motor constraints.

4.1.3 Contribution of this work

This study addresses key limitations in existing models of pedestrian-vehicle interaction by formulating the problem as a multi-agent RL task, in which both the pedestrian and vehicle are modelled as agents subject to visual and motor constraints. We focus specifically on one-to-one interactions between a single pedestrian and a single vehicle. In addition, we introduce gaze-dependent acuity, moving beyond the simplifying assumption in our earlier controlled-experiment modelling that pedestrians continuously fixate on the vehicle. Unlike such prior assumptions, our model allows pedestrians to control their own gaze direction, making perceptual uncertainty dependent on both distance and gaze allocation. This represents a novel contribution that is essential for modelling realistic road user interaction, where pedestrians approach from different positions and are not always looking directly at the vehicle

To enable this modelling approach, we carefully selected and preprocessed a real-world trajectory dataset that enables the extraction of such one-to-one interaction events while minimising the influence of other road users and road layout. This allows us to construct a focused learning environment that aligns closely with the modelling assumptions and objectives.

Importantly, to the best of our knowledge, this is the first study to bring a bounded optimality

perspective into a multi-agent RL framework for modelling human-human and human-machine interaction. By integrating cognitive constraints, real-world data, and multi-agent interaction, this work represents a meaningful step towards making the framework applicable to AV development, including both testing and optimisation.

4.2 Methods

This section presents our modelling framework for pedestrian-vehicle interaction. Section 4.2.1 provides a high-level overview of the framework. Section 4.2.2 outlines the cognitive and motor constraints that shape the agents' behaviour, and Section 4.2.3 introduces the parameters specifying these human constraints. Section 4.2.4 defines the model variants used for evaluation. Section 4.2.5 then describes the real-world dataset and the extraction of interaction trajectories. The reinforcement learning problem and algorithm are detailed in Sections 4.2.6 and 4.2.7, followed by parameter fitting (Section 4.2.8) and the behavioural cloning model (Section 4.2.9).

4.2.1 Overview of the method

We formulate pedestrian-vehicle interaction as a multi-agent reinforcement learning (RL) problem. The framework includes two policies: one for the pedestrian, π_{ped}, and one for the vehicle, π_{veh}. These policies are trained through trial-and-error interaction in a simulated environment that replicates the geometry of the real-world location from which we collected our dataset (as shown in Fig. 4.1), and the initial conditions in that dataset.

These take their simplest form in the No-Constraint (NC) variant of our model, where the pedestrian and the vehicle are represented as agents with ideal, non-constrained perception and motor execution. In this baseline model variant, the pedestrian observes its own position and speed together with the vehicle's position and speed, and chooses continuous values of walking speed and direction. The vehicle observes its own position and speed and the pedestrian's position and speed, and selects a continuous acceleration value. This NC formulation is straightforward but unrealistic, as it ignores many of the perceptual and motor constraints that shape real human behaviour.

To make the RL agent more human-like, we extended the NC framework with additional assumptions, leading to three constrained model variants: the Visual-Constraint (VC) model, the Motor-Constraint (MC) model, and the full Visual-and-Motor-Constraint (VMC) model, as

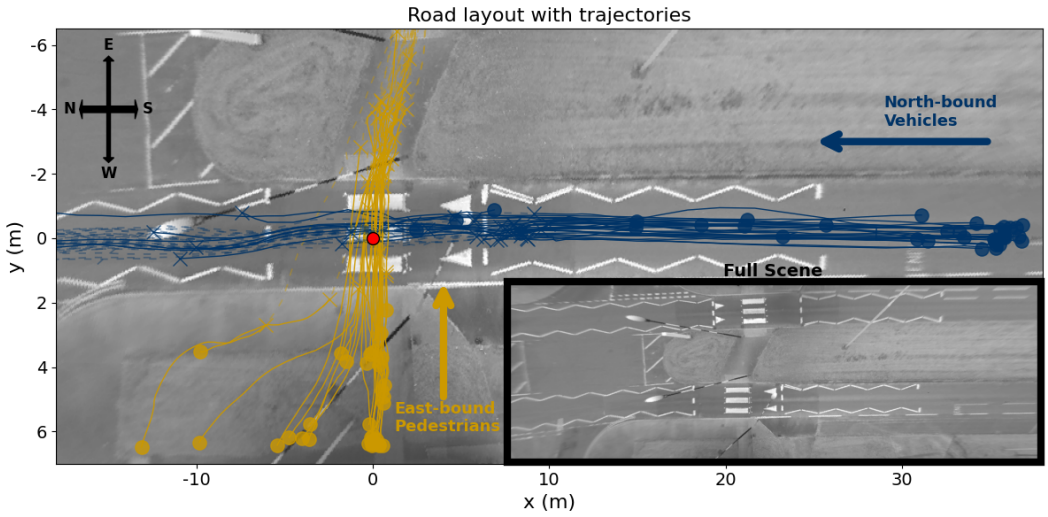

Figure 4.1: Real-world pedestrian (yellow) and vehicle (blue) trajectories overlaid on an image of the road layout at the study site. Solid lines represent trajectory segments within the first 6 seconds of the interaction, while dashed lines indicate segments beyond this window. The 'o' markers denote the starting points and the '×' markers denote the end points of the 6-second segments. The red dot marks the centre of the zebra crossing, referred to in the text as the crossing point. A zoomed-out view of the full scene is shown in the bottom right for context. The compass rose at the top left indicates the cardinal directions.

further described below.

4.2.2 Mechanistic assumptions of human constraints

This section outlines the key mechanistic assumptions in our model of pedestrian-vehicle interactions. These assumptions reflect perceptual, motor, and cognitive limitations derived from empirical and theoretical research, which shape how agents perceive the environment and select actions. Some of these mechanisms, such as noisy perception, ballistic speed control, and walking effort, were introduced in our prior work (Wang et al., 2025a; Wang et al., 2025b), while others, including gaze-dependent acuity and constraints on driver acceleration control, are new additions aimed at supporting more realistic two-agent interactions. The remainder of this section describes each of these mechanisms in detail.

Noisy visual input

Human perception is inherently uncertain and subject to visual noise (Faisal et al., 2008). Prior studies have suggested that these perceptual limitations influence pedestrian crossing decisions (Kotseruba and Rasouli, 2023). To capture this effect, we introduced angular noise at the level of the human retina (Kwon et al., 2015), which was also implemented in our previous pedestrian

behaviour models (Wang et al., 2025a; Wang et al., 2025b). Specifically, the perceived position of the other agent is derived from the visual angle below the horizon (Ooi et al., 2001; Markkula et al., 2023), with a constant Gaussian angular noise of standard deviation, ν, which could vary between agents. In practice, this means that the ego agent perceives the other agent's position along its line of approach with a noise of standard deviation (Markkula et al., 2023):

$$\sigma_x(t) = |d_\mathrm{l}(t)| \left(1 - \frac{h}{d(t) \cdot \tan(\arctan \frac{h}{d(t)} + \nu)}\right), \tag{4.1}$$

where $d_\mathrm{l}(t)$ is the longitudinal distance from the other agent to the crossing point (i.e., along the y-axis for the pedestrian and the x-axis for the vehicle in Fig. 4.1), $d(t)$ is the distance between two agents, and h is the eye height over the ground of the ego agent, set to 1.6 m for the pedestrian agent and 1.2 m for the vehicle agent, assuming homogeneous heights for simplicity.

While the formulation above captures perception uncertainty, our earlier modelling based on controlled experimental data assumed that pedestrians continuously fixated on the approaching vehicle, which was reasonable since participants interacted with a single car and needed to monitor it to ensure safe crossing. In contrast, the present work focuses on real-world data, where pedestrians approach the crossing from different positions, sometimes such that the vehicle is initially even behind the pedestrian, and are therefore not necessarily fixating on the vehicle at all times. To account for this, we modelled the decline in visual acuity as a function of gaze eccentricity, making visual uncertainty dependent not only on distance but also on whether the vehicle falls within or outside the centre of the pedestrian's gaze. Human perception is most accurate at the fovea and becomes less precise with increasing distance from the centre of gaze. We adopted the midget retinal ganglion cell (RGC) density model proposed by Watson (2014), which relates receptive field density $d_{gf}(\epsilon)$ to the eccentricity angle ϵ in degrees.

Following Watson (2014), we treated visual acuity as proportional to the square root of RGC density. To normalise acuity to 1 at the fovea, we define: $\alpha(\epsilon) = \frac{\sqrt{d_{gf}(\epsilon)}}{\sqrt{d_{gf}(0)}}$, where $d_{gf,0}$ is the peak density at zero eccentricity. The RGC density function is defined as:

$$d_{gf}(\epsilon) = d_{gf,0} \cdot \left[a_k \cdot \left(1 + \frac{|\epsilon|}{r_{2,k}}\right)^{-2} + (1 - a_k) \cdot \exp\left(-\frac{|\epsilon|}{r_{e_k}}\right)\right]$$

using parameters from Watson (2014): $a_k = 0.9729$, $r_{2,k} = 1.084$, $r_{e_k} = 7.633$, and $d_{gf,0} = 33163.2\,\mathrm{deg}^{-2}$. For modelling purposes, we approximate the perceptual noise level at a given eccentricity as the inverse of acuity: $\mathrm{noise}(\epsilon) = \frac{1}{\alpha(\epsilon)}$. This effect was incorporated by scaling the visual noise from Equation 4.1 by an acuity-based factor:

$$\sigma_x^{\mathrm{mod}}(t, \epsilon) = \sigma_x(t) \cdot \left(\frac{1}{\alpha(\epsilon)} + \delta \right) \tag{4.2}$$

where ϵ is the angular deviation between the pedestrian's gaze direction and the vehicle, $\alpha(\epsilon)$ is the relative visual acuity at angle ϵ as defined above, and δ is a very small constant (10^{-5}) added as a minor implementation safeguard to ensure numerical robustness during simulation; it is not theoretically required and does not affect the results. Note that for simplicity, we only used Equation 4.2 for the pedestrian agent; for the driver agent we used Equation 4.1.

Bayesian visual perception

Beyond the introduction of noise in the visual input, it is also important to consider how humans interpret such uncertain information. There is evidence that the human perception system interprets its noisy input in a Bayes-optimal manner, and Bayesian inference has been widely used to model perception and sensorimotor decision-making (Kwon et al., 2015; Knill and Pouget, 2004; Stocker and Simoncelli, 2006). In line with our previous work on modelling pedestrian perception under uncertainty (Wang et al., 2025b; Wang et al., 2025a), we employed a Kalman filter to model visual perception (Kwon et al., 2015; Markkula et al., 2023). At initialisation, perceived position and speed are drawn from Gaussian distributions centred at the true values, with standard deviations determined by the visual noise and by the variability of speed observed in the data, respectively. The Kalman filter then receives a new noisy position observation at each simulation time step, with noise as described above, and produces estimates of the other agent's position and velocity, along with their respective uncertainties. These filtered estimates represent the agent's belief state, serving as a Bayes-optimal inference about the environment under uncertainty.

Ballistic walking speed control

Human motor control is often characterised by intermittent, ballistic adjustments rather than continuous fine-tuning (Tustin, 1947; Craik, 1948). In line with this view, we assumed that each

walking step is ballistic, such that once initiated, the pedestrian maintains a fixed acceleration throughout the step. While some studies suggest humans may adjust trajectories mid-step (Barton et al., 2019), this behaviour is omitted here for simplicity. In our model, after selecting an action, the pedestrian applies a fixed acceleration to reach the desired speed, calculated as: $a_{\text{ped},t^*} = \frac{v_{\text{ped}}(t^*) - v_{\text{ped}}(t^*-1)}{T_{\text{step}}}$, where $v_{\text{ped}}(t^*)$ and $v_{\text{ped}}(t^*-1)$ represent the velocities of the agent at two consecutive decision points. Here, t^* indexes decision points to distinguish from the simulation time step t in the RL environment. T_{step} is the duration of a walking step, which is derived from empirical relations between step length and speed, with step length s estimated by Grieve (1968): $s = {v_{\text{ped}}}^{0.42}$. This yields $T_{\text{step}} = v_{\text{ped}}^{-0.58}$.

Walking effort

Human walking behaviour reflects a trade-off between travel time and energetic cost (Carlisle and Kuo, 2023). Prior research has incorporated walking effort into biomechanical models of gait (Choi, 1997; Faraji et al., 2018). In our model, we adopted the biomechanical framework from Carlisle and Kuo (2023) to represent walking effort as a cost associated with speed change during a step. The equation for the new velocity v_i^+ is: $v_i^+ = v_i^- \cos 2\theta + \sqrt{2u_i} \sin 2\theta$, 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 speed change due to exerted effort. This formulation stems from modelling the leg as a pendulum, where effort u_i corresponds to the kinetic energy required to alter walking speed: $u_i = \frac{1}{2}m(\Delta v)^2$ and solving for Δv: $\Delta v = \sqrt{2u_i/m}$. If we instead express effort per unit mass, we define: $U_i = u_i/m$, which simplifies to: $\Delta v = \sqrt{2U_i}$. Therefore, the effort U_i required for this change is given by:

$$U_i = \frac{(v^- \cos(2\theta) - v^+)^2}{2 \cdot \sin^2(2\theta)} \tag{4.3}$$

The total walking effort is defined as: $E_w = w_{\text{ped}} \cdot U_i$, where w_{ped} is a parameter that scales the walking effort for different individuals. A higher value of w_{ped} reflects a stronger preference for conserving energy during crossing.

Acceleration control of the driver

Human driving behaviour is subject to perceptual, cognitive, and motor delays, meaning that, in contrast with most RL models of driving (Sackmann, 2024; Cornelisse and Vinitsky, 2024),

control actions such as steering or acceleration are not applied instantaneously. In the broader driver modelling literature, these delays are often modelled using a low-pass filter to reflect gradual motor execution (Sharp, 2005). Following this approach, we modelled vehicle acceleration as a smoothed response toward a target acceleration value. At each time step, the driver selects a target acceleration a_{target} through its policy, and the actual acceleration $a_{\text{veh},t}$ evolves incrementally as:

$$a_{\text{veh},t} = a_{\text{veh},t-1} + \frac{a_{\text{target}} - a_{\text{veh},t-1}}{w_{\text{veh}}}, \tag{4.4}$$

where w_{veh} is the motor non-policy parameter representing the vehicle's responsiveness. This formulation assumes that the driver cannot instantaneously achieve their intended acceleration, but must gradually adjust due to both human neuromuscular limitations and vehicle dynamics. The resulting $a_{\text{veh},t}$ is used to update the vehicle's speed and position at each time step.

4.2.3 Non-policy parameters of human constraints

We refer to the parameters that specify the human constraints as non-policy parameters, to distinguish them from the trainable weights and biases of the policy network. We considered four such parameters: ν_{ped} and ν_{veh}, which are *visual non-policy parameters* related to noisy visual input, and w_{ped} and w_{veh}, which are *motor non-policy parameters* that control motor execution effects, i.e., walking effort scaling for the pedestrian and acceleration smoothing for the vehicle.

In our previous work, we conditioned each agent's RL policy on non-policy parameters that captured its perceptual and motor constraints (Wang et al., 2025a; Wang et al., 2025b). These parameters were provided as additional inputs to the policy network during training, allowing the learned policy to adapt its behaviour according to different combinations of perceptual and motor limitations. After training, we optimised the non-policy parameter values for each participant in the dataset, by minimising the difference between simulated and observed trajectories, thereby obtaining per-individual estimates of perceptual and motor constraints.

In this work, our dataset includes only a single observed interaction for each human agent, which does not permit us to fit per-individual parameter values. Instead, we directly model the population variability of the four non-policy parameters, as four separate normal distributions, and estimate the mean and standard deviation of these distributions from our data. Therefore, during RL policy training, at the start of each simulated interaction episode, we

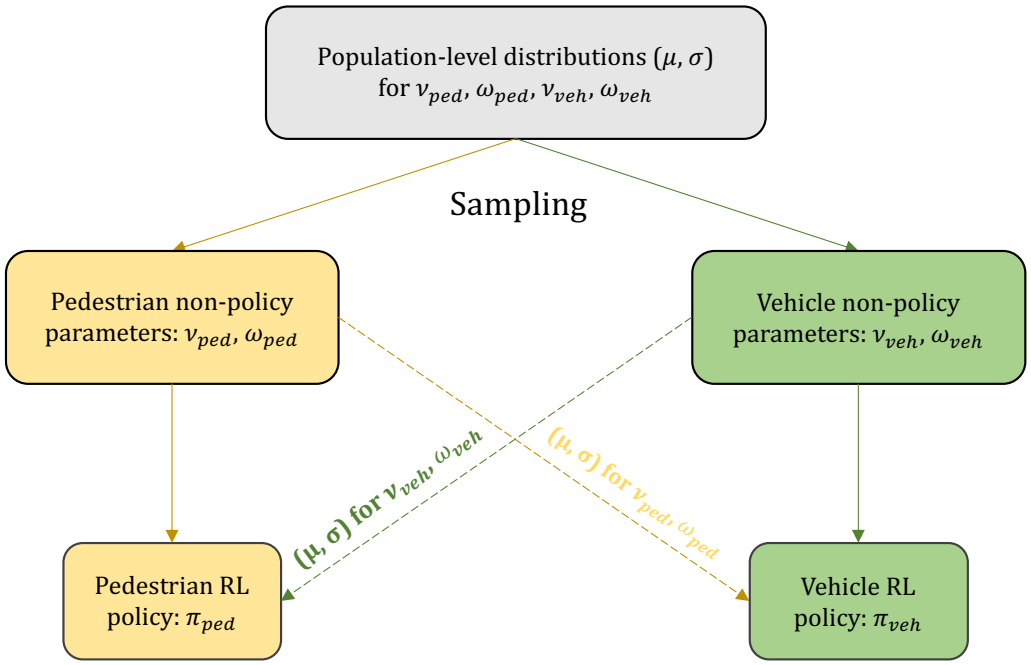

Figure 4.2: Overview of the two-stage sampling and conditioning scheme for non-policy parameters. Population-level means and standard deviations (μ, σ) are defined for $(\nu_{\text{ped}}, w_{\text{ped}}, \nu_{\text{veh}}, w_{\text{veh}})$. At the start of each episode, agent-specific values are sampled from these distributions. Each RL policy (π_{ped}, π_{veh}) is conditioned on its own parameters and on the population-level (μ, σ) of the other agent. Each observation o also includes this information for conditioning.

first sampled random values for these means and standard deviations. Specifically, the *visual* non-policy parameters (ν_{ped}, ν_{veh}) use means sampled uniformly from [0.01, 0.1] and standard deviations from [0.001, 0.01] (in radians). The *motor* non-policy parameters are dimensionless: for the vehicle (w_{veh}), means are sampled from [1, 10] and standard deviations from [0.1, 1]; for the pedestrian (w_{ped}), means from [0.05, 0.5] and standard deviations from [0.005, 0.05]. Thereafter, individual parameter values were drawn from the resulting normal distributions and assigned to the agents in that episode (see Figure 4.2 for an overview of this sampling and conditioning process). As in our previous work, each agent's policy is conditioned on its own non-policy parameter values, but here we also condition the policy on the population mean and standard deviation of the other agent's non-policy parameters. This reflects our assumption that each road user behaves in a way that is boundedly optimal given its own constraints, and given the distribution of constraints in the population of other road users.

4.2.4 Model variants

To investigate how our different model assumptions influence the interaction between the pedestrian and vehicle, we categorised the assumptions into two types: **visual constraints** and **motor constraints**. The visual constraints include *noisy visual input*, *Bayesian visual perception*, while the motor constraints include *walking effort*, *pedestrian ballistic speed control*, and *driver acceleration control*. Comparison of different model variants is shown in Figure 4.3.

Based on the inclusion or exclusion of these constraints, we defined the following model variants:

- **No-Constraint (NC) model**: A baseline model without visual or motor constraints.
- **Motor-Constraint (MC) model**: A model incorporating motor constraints only.
- **Visual-Constraint (VC) model**: A model incorporating visual constraints only.
- **Visual and Motor-Constraint (VMC) model**: A full model incorporating both visual and motor constraints.

4.2.5 Dataset

This study focuses on modelling the interaction between one pedestrian and one vehicle. We used the dataset reported by Kalantari et al. (2025), which fits our aim of extracting one-one pedestrian-vehicle interaction with minimal effects from other road users and road layout. To obtain clean and representative pedestrian-vehicle interactions from real-world data, we employed a multi-stage trajectory extraction and filtering pipeline.

Figure 4.1 shows a top-down view of the study site in Leeds, UK, from which our data were collected. As shown in the bottom-right inset of Figure 4.1, the site features a staggered pedestrian crossing with two single-lane zebra crossings. In this work, we only considered interactions at the *Western* crossing, between North-bound vehicles (note that the UK has left-hand traffic) and East-bound pedestrians. The reasons for only including these interactions were the South-bound vehicles were also influenced by turning traffic from the side street, which introduces variability in vehicle speed when approaching the zebra crossing.

The dataset includes pedestrians crossing in both directions. Since pedestrians with different walking directions correspond to different origins and destinations, modelling both would require training two separate pedestrian policies. We excluded West-bound (from up to bottom in

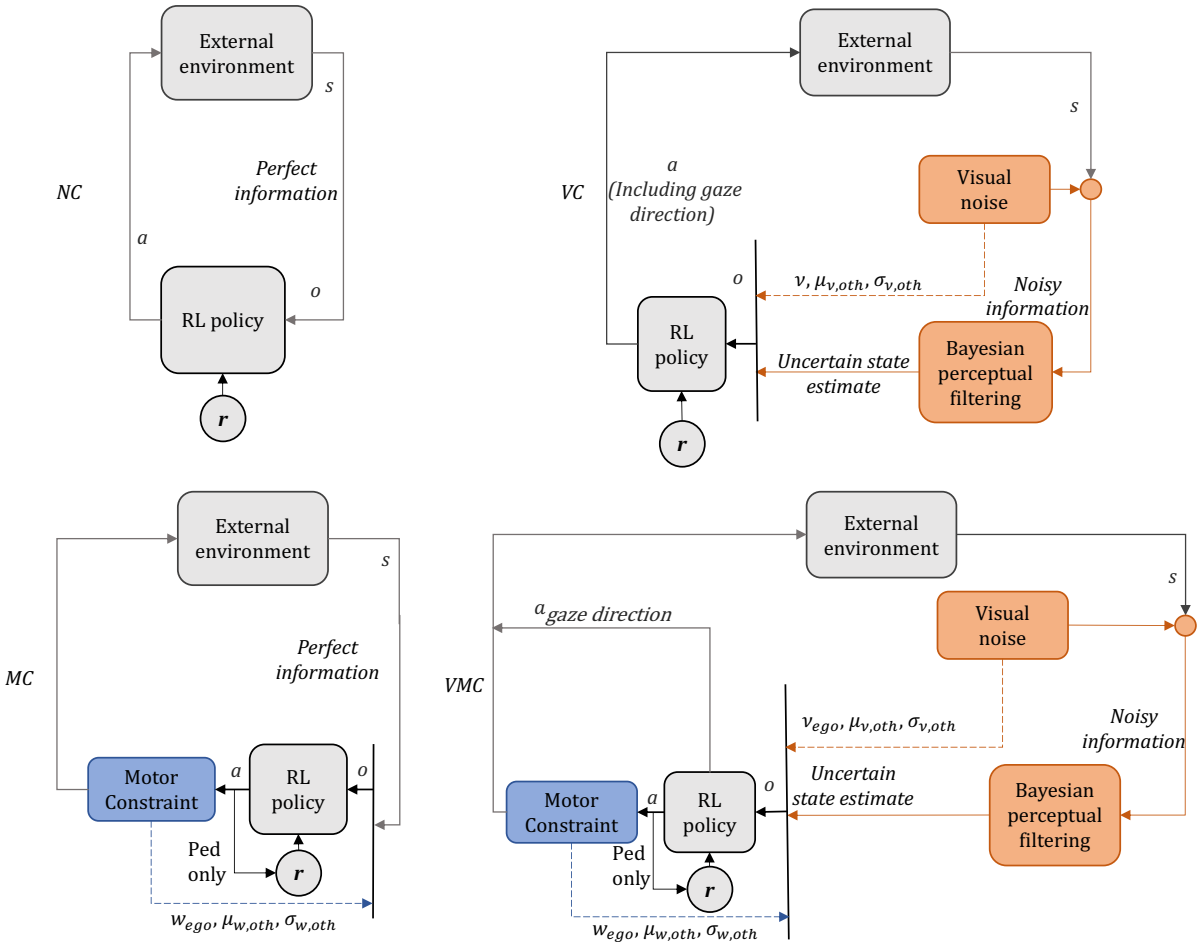

Figure 4.3: Comparison of model variants. Arrows follow the reinforcement learning loop: s is the true environment state, a is the agent action (including gaze direction in the VC and VMC models), r is the received reward, and o is the observation fed to the RL policy. In the MC and VMC models, the action a contributes to the reward only for the pedestrian agent (walking effort penalty), as indicated by the 'Ped only' label.

Figure 4.1) pedestrians because they have already crossed another road and passed through the central refuge island, which exposes them to additional influences from South-bound vehicles. To maintain a clear focus on one-to-one interactions, we retained only interactions where the pedestrian walks from bottom to top in Figure 4.1 (labelled as East-bound in the dataset). In the original data, 311 East-bound pedestrian trajectories and 49,167 North-bound vehicles remained. To ensure that the retained pedestrian-vehicle pairs were not influenced by pedestrians walking in the opposite direction, we removed any North-bound vehicle approaching from the right (within 25 metres before the zebra crossing) if it temporally overlapped with at least one West-bound pedestrian in the central refuge island.

For each pedestrian, we identified potential interacting vehicles based on their spatial proximity and temporal overlap. Specifically, a vehicle was considered part of the interaction if it was

located within 25 m upstream of the zebra crossing, and a pedestrian was considered if they were within 20 m of the crossing along the x-axis and within 5 m of the kerb along the y-axis. Both agents also had to be present in the scene during an overlapping time interval. An interaction pair was retained only if exactly one vehicle satisfied these conditions for a given pedestrian. A second filtering step was then applied to ensure that each vehicle was matched with at most one pedestrian. Interactions involving multiple potential matches were excluded to maintain strict one-to-one interaction. After this step, 77 interaction pairs were left (each interaction pair consists of one East-bound pedestrian and one North-bound vehicle).

Following the above filtering steps, we further truncated each pedestrian trajectory to include only the portion where the pedestrian was within 10 metres of the crossing point, measured along the x-axis in Figure 4.1. This ensured that the retained segments contained only the parts of the trajectories relevant to the crossing interaction. Each pedestrian-vehicle pair was then temporally aligned by retaining only the overlapping portion of their trajectories within the interaction zone. For each agent, the entry and exit times were defined as the moments of entering and leaving that zone, respectively, and the overlap was defined from the later entry to the earlier exit. From this synchronised window, we extracted the first 6 seconds, measured from the synchronised start time, in order to capture the early phase of interaction when both agents were simultaneously present. Interaction pairs with a total overlap shorter than 6 seconds were discarded to ensure sufficient behavioural content for modelling. This procedure left 34 interaction pairs, which are visualised in Figure 4.1.

4.2.6 Reinforcement learning problem

Reinforcement learning (RL) provides a computational framework in which an agent interacts with an environment by selecting actions to maximise cumulative rewards. In cases such as ours, where the agent does not have full information about its environment, the problem is typically formulated as a partially observable Markov decision process (POMDP), defined by a tuple $(\mathcal{S}, \mathcal{A}, \mathcal{T}, \mathcal{R}, \mathcal{O})$, where:

State space ($\mathcal{S}$):

At each time step t, the environment state s_t includes the true positions and velocities of both the pedestrian and the vehicle, as well as the simulation time. Specifically, the state comprises the pedestrian's position $(x_{\text{ped}}, y_{\text{ped}})$ and walking speed v_{ped}, the vehicle's position $(x_{\text{veh}}, y_{\text{veh}})$

and speed v_{veh}, the yaw angle of the pedestrian's walking direction θ_{ped}, and the vehicle's acceleration a_{veh}. The walking angle θ_{ped} is measured relative to the upward y-axis of the map (see Figure 4.1), with 0° indicating motion along the positive y-direction.

During training, initial conditions were sampled from a joint distribution fitted to the initial states of the pedestrians and vehicles at the start of each interaction segment in the real-world dataset. This distribution was fitted using kernel density estimation (KDE), covering start positions and speeds for both agents. This approach allowed the model to train on a diverse set of realistic interaction scenarios. During model testing, each simulation was initialised directly from the observed start state of the corresponding real-world interaction.

Action space ($\mathcal{A}$):

The pedestrian's action a_{ped} consists of forward walking speed v and a change in walking angle $\Delta\theta$, and optionally a gaze angle relative to the body angle ϕ (included in visual-constrained models, such as VC and VMC). Gaze orientation was not associated with any biomechanical effort or explicit cost, meaning the agent could freely direct its gaze. The driver's action $a_{\text{veh},target}$ is the target acceleration. It should be noted that in models with motor constraints (MC and VMC), the pedestrian speed action is not applied instantaneously. Instead, the selected speed only takes effect at the onset of a new walking step and remains constant during the execution of that step, consistent with the ballistic control assumption described in Section 4.2.2. This reflects the view that pedestrians cannot adjust their velocity continuously within a step but only between successive steps.

Transition function ($\mathcal{T}$):

The transition function defines how the current state s_t evolves into the next state s_{t+1} given the agents' actions a_t.

In the NC and VC models, actions affect agent speeds and positions directly. In the MC and VMC models, pedestrian transitions are governed by ballistic execution: the walking speed is updated using a fixed acceleration over the step duration as explained above, and the walking angle is updated according to the selected change in walking angle. For the vehicle, the speed is updated using a smoothed acceleration $a_{\text{veh},t}$, as described in Section 4.2.2.

The simulation terminates when both agents reach their respective goal regions or a collision

occurs. The pedestrian's goal region is defined as having crossed the driving lane, while the vehicle's goal region is defined as having passed beyond the zebra crossing.

Reward function ($\mathcal{R}$):

The reward function is designed to remain as simple as possible while encouraging successful and safe crossing and discouraging undesirable behaviours such as collisions, and rule violations. Most components build upon those used in our previous study (Wang et al., 2025b), which produced realistic and interpretable crossing behaviour. The present formulation retains these components while introducing additional terms to accommodate new action definitions, such as directional walking choices. To reflect the influence of time pressure on road-user decisions (Tian et al., 2022a; Kong et al., 2023), a time cost is included in the arrival rewards for both agents, making delayed arrivals less rewarding.

Pedestrian agent The pedestrian receives a positive reward upon successfully reaching the sidewalk on the opposite side: $r^{\text{ped}}_{\text{arrive}} = 40 - 0.5 \cdot t$, where t denotes the simulation time. The constant value of 40 was chosen based on initial tests with the NC model, which showed that this setting led to reasonable crossing behaviour. A penalty of $r_{\text{collision}} = -40$ is applied in the event of a collision with the vehicle. In addition, the pedestrian is rewarded for making forward progress before reaching the sidewalk on the opposite side: $r_{\text{move}} = \Delta y$, where Δy is the change in longitudinal (up y-direction in Figure 4.1) position over the previous time step. Unlike in the previous models (Wang et al., 2025b), this term is included here because the pedestrian can choose its walking direction. The forward progress reward therefore guides the pedestrian agent to cross the road in the desired direction rather than wandering laterally.

In models with motor constraints (MC and VMC), as previously described, a *walking effort penalty* is introduced to represent the physical cost of accelerating from one walking speed to another (Carlisle and Kuo, 2023; Wang et al., 2025b). At each action update, the ballistic model computes the required speed change, and the energy cost is penalised as: $r_{\text{walk}} = -E_w = -w_{\text{ped}} \cdot U_i$, where w_{ped} is the motor effort parameter, and U_i is the walking effort work calculated as described in Section 4.2.2. This penalty is only applied at the start of each step. An additional penalty is applied for stepping onto the road outside the crosswalk at each time step: $r_{\text{off}} = -0.2$.

Summing over the terms, the total pedestrian reward is given by:

$$r_{\text{ped}} = r^{\text{ped}}_{\text{arrive}} + r_{\text{move}} + r_{\text{walk}} + r_{\text{off}} + r_{\text{collision}} \tag{4.5}$$

Vehicle agent Like the pedestrian, the vehicle receives a positive reward upon successfully crossing the designated zone: $r^{\text{veh}}_{\text{arrive}} = 40 - 0.5 \cdot t$. A collision incurs a penalty of -40. To represent the traffic rule requiring drivers to yield to pedestrians at zebra crossings, a penalty of $r_{\text{nonyield}} = -30$ is imposed when the vehicle enters the crossing area without yielding to a nearby pedestrian. Specifically, this penalty is triggered once at the time step when the vehicle first enters a zone extending 5 m from the zebra crossing, provided that the pedestrian is located within 2 m in x direction of the zebra crossing and within 3 m of the curb in the y direction.

The total vehicle reward is:

$$r_{\text{veh}} = r^{\text{veh}}_{\text{arrive}} + r_{\text{nonyield}} + r_{\text{collision}} \tag{4.6}$$

Observation space ($\mathcal{O}$):

At each time step, both agents receive a normalised observation vector comprising (i) their own state, (ii) an estimate of the other agent's state, and (iii) relevant non-policy parameters. Each observation also includes the current time step t, allowing the agent to account for the passage of time given that the reward function incorporates a time cost. All features are normalised to the range $[0, 1]$ by linearly scaling each variable between its approximate minimum and maximum values, ensuring that negative values are also mapped within this range.

The **pedestrian** observes its own x and y coordinates, forward walking speed, walking direction, gaze yaw angle, and the vehicle's y coordinate and speed. In models without visual constraints, the vehicle state is observed without noise, whereas in models with visual constraints they are estimated from noisy input as described in Section 4.2.2. In the latter case, the observation additionally includes uncertainty measures from the Kalman filter, specifically the estimated variances of the vehicle's position and speed. Execution-related features capture the pedestrian's step cycle: in models with motor constraints, the observation provides the remaining execution time of the current walking step, indicating the time remaining until a new speed command can take effect. The observation also includes the pedestrian's own non-policy parameters, as

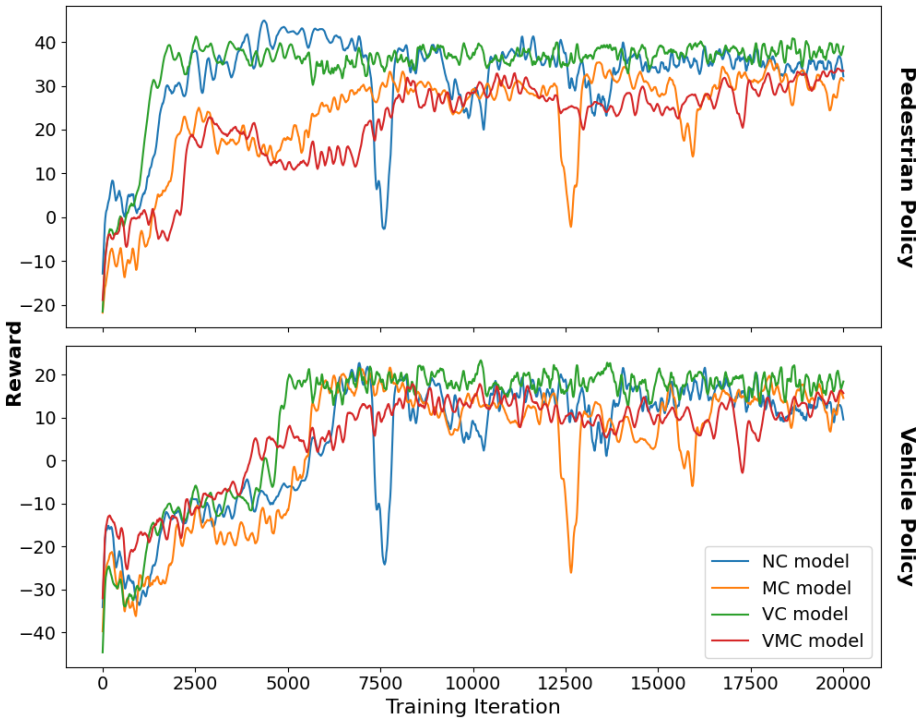

Figure 4.4: Reward plot.

well as the mean and standard deviation of the vehicle's non-policy parameter distributions, as explained in Section 4.2.3.

The **vehicle** observes its own x coordinate, speed, acceleration, and, in models with motor constraints, also the target acceleration. In models with visual constraints, the vehicle does not observe the pedestrian state directly but instead receives Kalman-filtered estimates of the pedestrian's x and y coordinates and speeds. Along with these estimates, the observation also includes the corresponding uncertainty measures from the Kalman filter, namely the estimated variances of the pedestrian's position and speed. The pedestrian's walking direction is also provided. The observation further includes the vehicle's non-policy parameters, along with the mean and standard deviation of the pedestrian's non-policy parameter distribution.

Table 4.1: Hyperparameters for the RL algorithm.

Parameter	**Value**
Algorithm	Soft Actor-Critic (SAC)
Training iterations	20,000
Hidden layers	[256, 128, 64]
Activation	ReLU
Learning rate	0.001
Discount factor (γ)	0.995
Batch size	8192

4.2.7 Reinforcement learning algorithm

We used the Soft Actor-Critic (SAC) algorithm as implemented in Ray RLlib for all RL experiments (Liang et al., 2018). SAC optimises a stochastic policy in continuous action spaces and supports efficient, stable learning through off-policy updates (Haarnoja et al., 2018). The reuse of past experiences may offer some benefits in our multi-agent setting, where interactions are complex (Christianos et al., 2020). In our case, two policies were trained simultaneously, one for the pedestrian agent and one for the vehicle agent.

Compared with PPO, SAC offered two practical advantages in this context. First, its off-policy formulation allows past experience to be reused through a replay buffer, which can improve sample efficiency in multi-agent interaction settings where informative events may be relatively sparse (Christianos et al., 2020). Second, SAC learns a stochastic policy while explicitly encouraging exploration through entropy regularisation. This can reduce the tendency to converge to overly rigid behavioural strategies. These properties made SAC more suitable for learning flexible interactive behaviours in this chapter. This choice also had implications for the learned behaviour. In preliminary experiments, PPO often led the agents to select actions near the boundaries of the action space, suggesting poorer optimisation behaviour in this more complex multi-agent setting. Such optimisation artefacts are not only a technical concern, but can also reduce behavioural realism by producing overly extreme or mechanically implausible actions. SAC was therefore adopted as a more suitable algorithm for capturing nuanced interaction patterns such as yielding, gradual speed adjustment, and behavioural variability across repeated interactions.

We trained the SAC model for 20,000 iterations. The reward curves of different model variants during training is shown in Figure 4.4, illustrating the convergence trends of the different model variants. The hyperparameters used in our implementation are summarised in Table 4.1.

4.2.8 Fitting of non-policy parameters

The values for the non-policy parameters that best describe the population are not learned through the RL training but are instead fitted by comparing model trajectories of the trained RL policy with real-world trajectory data.

To quantify how well the model behaviour matches real-world data, we adopted a composite Negative Log-Likelihood (NLL) metric inspired by recent work on evaluation of the performance

of simulation-based agents (Montali et al., 2023). In this approach, comparison between model and real-world trajectories is conducted over a diverse set of behavioural metrics encompassing both motion dynamics and interaction outcomes, such as linear speed, acceleration, angular velocity, angular acceleration, distance to nearest object, time-to-collision (TTC), distance to road edge, offroad indication, and collisions. In Montali et al. (2023), each of these metrics is first converted into a likelihood score by comparing simulated rollouts with logged data using kernel density estimates. NLLs are then averaged across time and agents, and finally aggregated into a single composite score via a weighted average. Safety-critical metrics such as collisions and offroad departures are assigned higher weights in their formulation.

Building on this framework, we used the same likelihood-based approach and adapted the set of metrics to our specific pedestrian-vehicle interaction scenario. For simplicity, we did not introduce different weights for different metrics, effectively setting all weights to one. Following Montali et al. (2023), we included metrics describing motion dynamics and interactions, namely the linear speed and linear acceleration of both agents, the pedestrian's angular velocity and angular acceleration, and the inter-agent distance. In addition, we added a few metrics to better capture behaviours relevant to our task, including the x positions of both agents, the pedestrian's y position and heading angle, and the projected Post-Encroachment Time (PET) (Lin et al., 2024). The projected PET is defined as the predicted time difference, based on current trajectories, between the pedestrian passing the y coordinate of the crossing point and the vehicle passing its x coordinate (the red dot in Figure 4.1). A positive PET indicates that the pedestrian crosses first, while a negative PET indicates that the vehicle crosses first.

Given the scoring method described above, model fitting amounts to running simulations for different values of a non-policy parameter vector $\boldsymbol{\phi}$ and finding the one that minimises the composite NLL. The vector

$$\boldsymbol{\phi} = [\mu_{\nu,\text{ped}}, \sigma_{\nu,\text{ped}}, \mu_{\nu,\text{veh}}, \sigma_{\nu,\text{veh}}, \mu_{w,\text{ped}}, \sigma_{w,\text{ped}}, \mu_{w,\text{veh}}, \sigma_{w,\text{veh}}],$$

defines the means and standard deviations of the population distributions for the four non-policy parameters $(\nu_{\text{ped}}, \nu_{\text{veh}}, w_{\text{ped}}, w_{\text{veh}})$, giving eight free parameters in total.

Since the model contains stochasticity arising from both noisy visual input and the trained SAC policy, likelihood estimates based on a single rollout would be overly sensitive to inciden-

tal variation. We therefore ran five rollouts for each parameter vector. In practice, examination of the resulting behavioural metric distributions suggested that five repetitions were sufficient to capture the main spread of the simulated behaviour and to provide a reasonably stable basis for the composite likelihood estimates. This should be understood as a pragmatic sufficiency judgement for the present fitting procedure, rather than a claim that five rollouts are universally optimal. To mitigate the compounding divergence between model and real-world trajectories, model rollouts during testing were limited to 2 s in duration. Each 6 s real-world trajectory was therefore divided into three 2 s segments, with initial conditions taken from the real-world data and shared non-policy parameter values used across all three segments.

For each behavioural metric, a KDE was fitted to the real-world data using Scott's rule for bandwidth selection (Scott, 2015). The model distributions were then evaluated under the fitted KDEs to obtain per-feature likelihoods. The overall composite NLL was computed by averaging the NLL across all behavioural features and trajectory segments.

To optimise the non-policy parameter vector $\boldsymbol{\phi}$, we used Bayesian optimisation with Gaussian processes via `gp_minimize` from `scikit-optimize`. During optimisation, each element of $\boldsymbol{\phi}$ was constrained within the ranges specified in Section 4.2.3. The optimisation was run for 500 iterations, with 100 initial random samples, using Expected Improvement as the acquisition function. The optimisation returns point estimates of these eight values, which define the population-level distributions from which individual agents sample their non-policy parameters during simulation. In addition to the composite NLL used as the optimisation objective, we also reported per-metric NLLs and the Kolmogorov-Smirnov (KS) statistic as complementary distributional measures, since KS is a widely used method for comparing empirical distributions and may be more familiar to some readers.

4.2.9 Behavioural cloning model

Behavioural cloning (BC) (Bain and Sammut, 1995) was implemented as a supervised learning baseline for comparison with our RL framework. Its purpose was to evaluate to what extent a non-interactive imitation learning model, trained purely on the real-world demonstrations in our small dataset, could replicate realistic pedestrian and vehicle behaviours. Given the data limited setting, we trained and evaluated the BC baseline on the entire dataset to keep the training and evaluation conditions consistent with the RL models. This choice maximised access

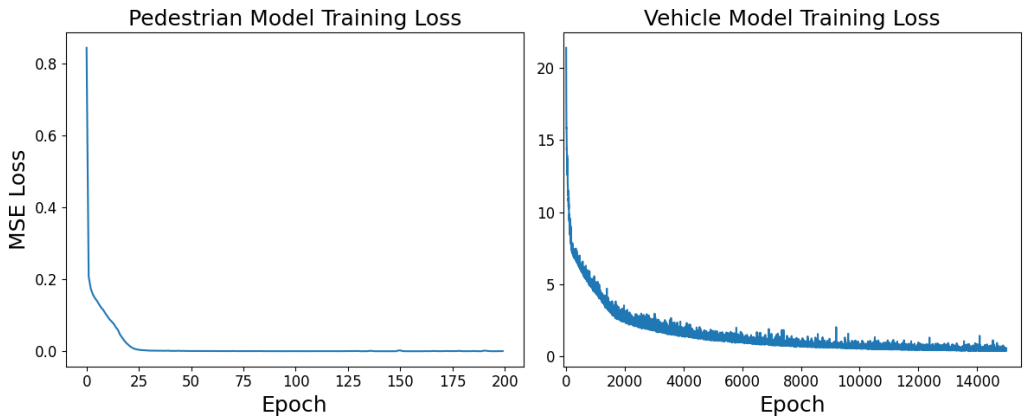

Figure 4.5: Training loss curves for the BC model.

to available data and was expected to yield strong one-step (open-loop) predictions, whereas our main comparisons focus on closed-loop rollout performance on the same interaction set.

The BC models for both pedestrian and vehicle agents were trained separately, using the same data as for our RL model, but instead using supervised learning to predict the next agent action given the current state. The input features included both self and other-agent states: x and y position, speed, heading angle for the pedestrian; and x position and speed for the vehicle, as well as time. These features correspond to the observation space of the NC variant of our RL model. The BC models produced the same action outputs as the NC model. Specifically, the pedestrian model predicted walking speed and heading angle, while the vehicle model predicted linear acceleration.

The pedestrian and vehicle BC models shared the same network architecture: two hidden layers with 64 units each and ReLU activation. They were trained using mean squared error (MSE) loss and the Adam optimiser with a learning rate of 0.001, until subjectively judged convergence, reached after 200 epochs and 15,000 epochs for the pedestrian and vehicle models, respectively. Training losses are shown in Figure 4.5.

4.3 Results

The results presented in this section are based on simulations using the trained RL model, with the fitted non-policy parameter distributions. For each real-world interaction, we initialised the simulation with the corresponding human starting conditions and ran the model five times, the same as in the fitting procedure, to account for variability arising from the visual noise

and from the RL policy. The evaluation was performed on full interaction trajectories, without segmentation.

The same procedure was applied to the BC model, which is deterministic in its action predictions. In this case, the five runs per interaction yielded identical trajectories, but were repeated for consistency with the evaluation protocol used for the RL model.

Table 4.2: Quantitative comparison of model performance across three behavioural metrics: negative log-likelihood (NLL), average displacement error (ADE), and final displacement error (FDE).

Metric	BC	NC	VC	MC	VMC
NLL	3.92	3.57	3.28	3.19	**2.42**
ADE (m)	5.92	5.17	4.60	5.61	**2.87**
FDE (m)	11.54	11.79	11.36	9.41	**5.41**

4.3.1 Comparing model and real-world behaviour metric distributions

To assess distributional similarity between model behaviours and real-world data, we report per-metric NLLs (Montali et al., 2023), together with KS statistics, as introduced in Section 4.2.8.

As shown in Table 4.2, the VMC model achieved the lowest overall NLL (2.42), indicating

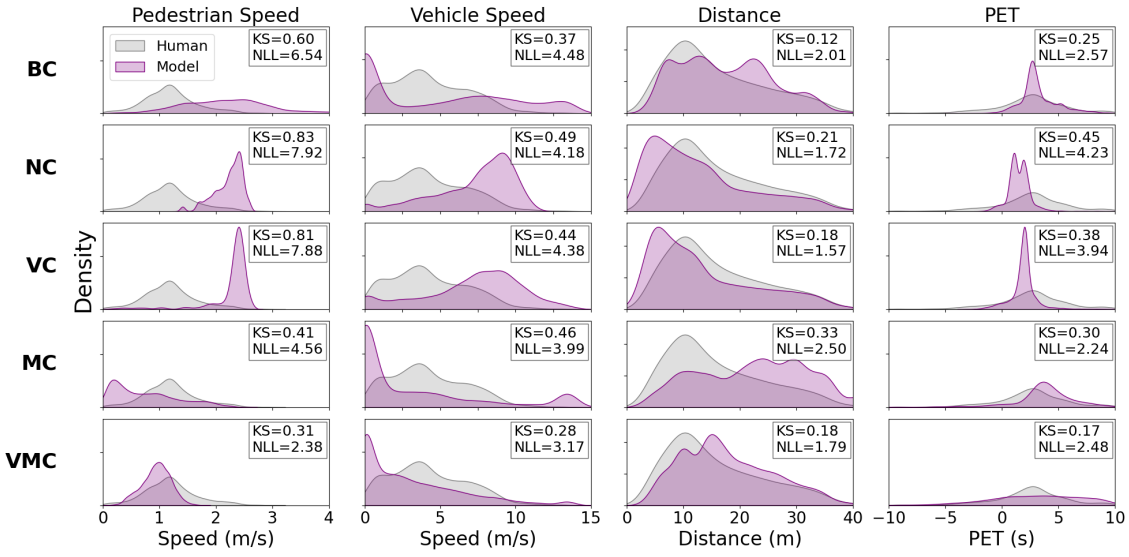

Figure 4.6: Comparison of model and real-world behavioural distributions across all time steps and trajectories for four key interaction metrics: pedestrian speed, vehicle speed, inter-agent distance, and projected post-encroachment time (PET). Each row corresponds to a different model variant. The Kolmogorov-Smirnov (KS) statistic shown in each subplot quantifies the similarity between model and real-world distributions, with lower values indicating closer alignment.

the best overall alignment with real-world trajectories. This was followed by the MC and VC models, while the NC and BC models showed relatively higher NLL scores, reflecting poorer fit to the real-world data.

To provide more detail, Figure 4.6 shows per-feature comparisons for four representative metrics: pedestrian speed, vehicle speed, inter-agent distance, and projected PET. For each feature, we report the NLL together with the KS statistic as complementary measures. Positional features such as x and y coordinates are not shown in Figure 4.6; however see the trajectory-level comparisons in the next section Additional behavioural metrics used in fitting, are presented in Appendix B (FigureC.1).

Across the four plotted metrics, the VMC model provided the closest match to real-world data, showing the lowest NLL and KS values across all metrics, as illustrated in Figure 4.6. Although the MC model attained a slightly lower NLL for PET, (2.24 vs. 2.48), the VMC still achieved the lowest KS, indicating the best overall distributional similarity.

In contrast, the NC and VC models yielded much poorer fits for pedestrian speed, i.e., high per-feature NLLs (7.92 and 7.88) and large KS values (0.83 and 0.81), likely because they lack a walking-effort cost, as shown in the second and third rows of the first column in Figure 4.6. The BC model aligned relatively well in distance (KS = 0.12, NLL = 2.01) and PET (KS = 0.25, NLL = 2.57), but failed to reproduce the real-world pedestrian-speed distribution.

4.3.2 Comparing model and real-world trajectories

We next evaluated the spatial and temporal similarity between model and real-world trajectories using average displacement error (ADE) and final displacement error (FDE). ADE measures the average point-wise distance between model and real-world trajectories over time, while FDE quantifies the distance at the final timestep. Lower values indicate better alignment with real-world human behaviour. Summary results are shown in Table 4.2, with full trajectory visualisations provided in Appendix A.

Figure 4.7 presents three representative interactions: a vehicle-first case (interaction 18), a pedestrian-first case with vehicle yielding (interaction 33), and a pedestrian-first case without vehicle yielding (interaction 67). In this figure, the BC model is deterministic, so the model trajectories are identical and appear as a single purple line. In the NC model, there is no visual noise and the only stochasticity comes from the policy network, which is small, so the

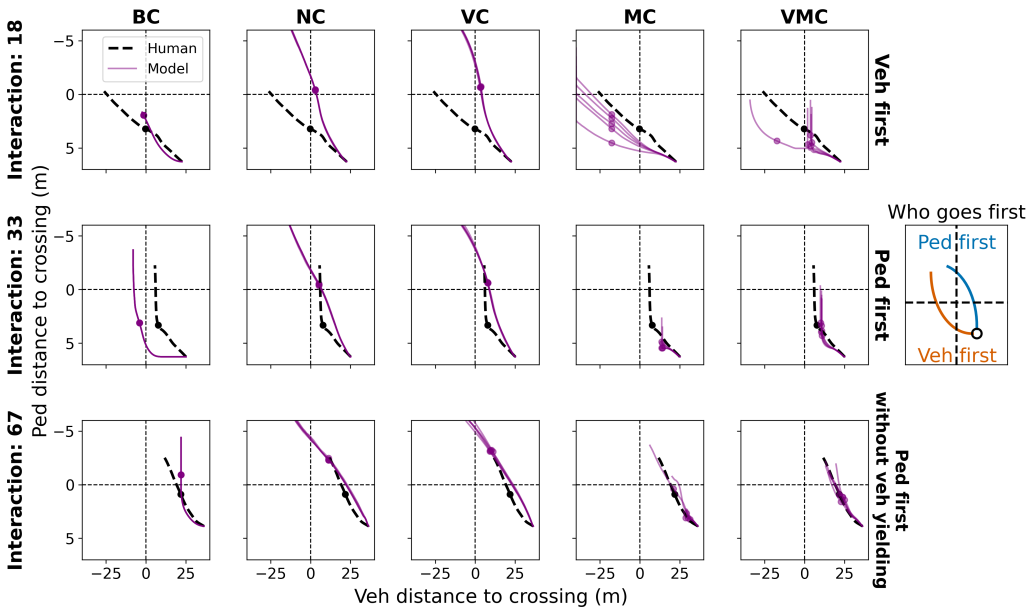

Figure 4.7: Comparison between model and real-world trajectories for three representative pedestrian-vehicle interactions (IDs 18, 33, and 67). Model trajectories are shown in purple. The x-axis represents the vehicle's longitudinal distance to the crossing point; the y-axis shows the pedestrian's lateral distance. Purple and black dots mark the 3-second positions of the model and real-world trajectories, respectively. Each row corresponds to a different interaction type: vehicle-first (ID 18), pedestrian-first with vehicle yielding (ID 33), and pedestrian-first without vehicle yielding (ID 67). The rightmost panel provides a schematic comparison between pedestrian-first (blue) and vehicle-first (orange) scenarios, with both trajectories originating from the same starting point (black circle).

trajectories look as if they have merged into one. As schematically shown in the rightmost panel, in this type of plot, if the trajectory passes above the origin, this indicates that the pedestrian crossed before the vehicle, and if the trajectory passes below the origin, the vehicle went first. In Figure 4.7, model performance can be assessed by how close the model trajectories are to the real-world trajectory, with closer alignment indicating higher fidelity. Good agreement can also be reflected by the model position at 3 s (purple dot) coinciding with the human position at 3 s (black dot).

In Table 4.2, across both ADE and FDE, the VMC model achieved the lowest errors, demonstrating its ability to reproduce human-like trajectories throughout the interaction. In contrast, the BC model and RL variants lacking motor constraints (NC and VC) showed larger errors, particularly in FDE, indicating cumulative divergence from real-world trajectories.

The BC model generally performed well in vehicle-first scenarios, as shown in interaction 18. In contrast, it failed to generate human-like pedestrian-first behaviours. In interaction 33 (second row of Figure 4.7), the vehicle decelerated to a stop after passing the crossing, an unrealistic action not observed in real-world data, likely due to the BC model's lack of contextual un-

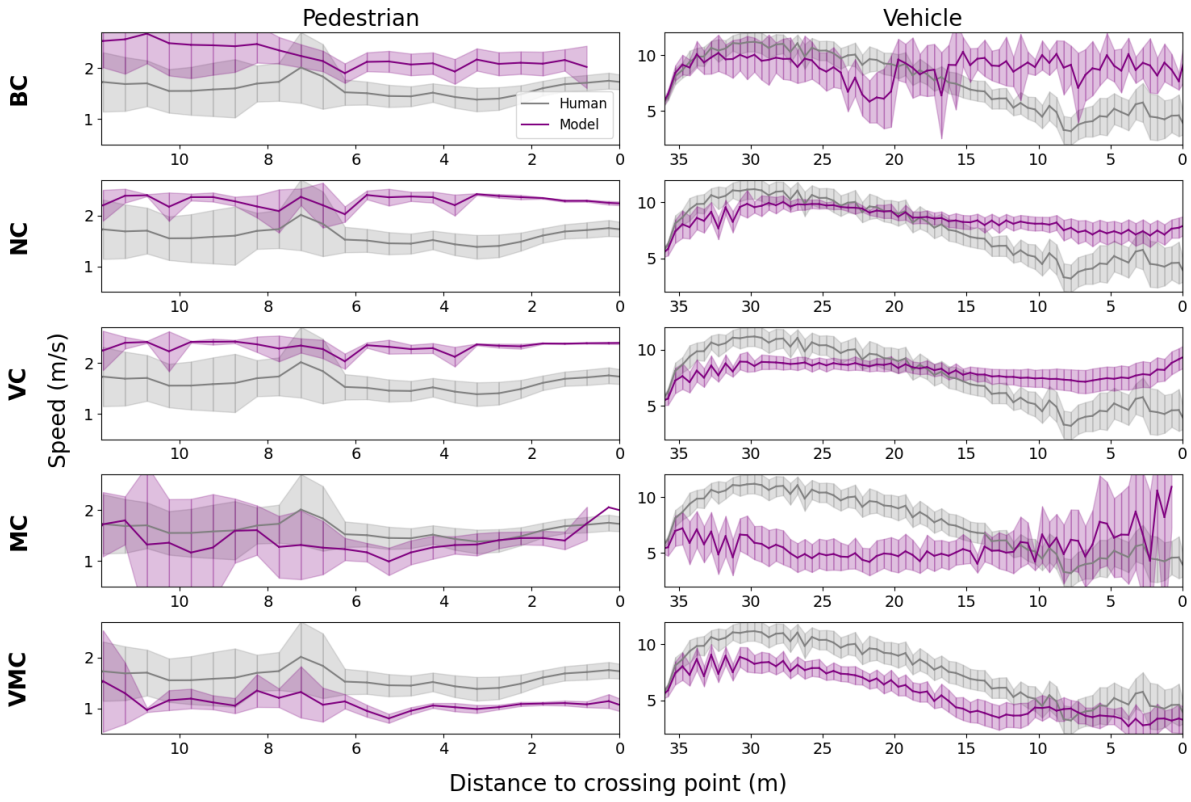

Figure 4.8: Average speed profiles of pedestrians (left) and vehicles (right). Real-world data are shown in black, and simulation results in purple. For each model, the solid line represents the mean speed, and the shaded area indicates the 95% confidence interval. The x-axis represents the agent's distance to the crossing point: larger values indicate positions farther away from the crossing, while smaller values (towards the right) represent positions closer to or at the crossing. Model trajectories were first averaged across repetitions sharing the same non-policy parameters before computing group-level statistics.

derstanding of interaction. Similarly, in interaction 67, the vehicle stopped even though the pedestrian had already crossed and posed no threat, again highlighting the BC model's limited understanding of the interaction dynamics.

Models without motor constraints (NC and VC) allowed pedestrians to cross quickly without cost, leading to the absence of vehicle-first events. As shown in the second and third columns of Figure 4.7, with these models, pedestrians consistently crossed first, and vehicle-yielding behaviours were not observed.

The MC model performed well in vehicle-first cases, but driver behaviour was often more assertive than in the real-world data. In interaction 18 in Figure 4.7, vehicles crossed with earlier and faster motion. While the MC model also captured pedestrian-first cases (rows 2 and 3), vehicle responses in yielding scenarios appeared more abrupt and occurred further away from the pedestrian compared to real-world trajectories (shown in row 2).

The VMC model reproduced all three interaction types observed in the real-world data, as shown in the fifth column of Figure 4.7. This model also captured variability within the vehicle-first category, with some vehicles yielding while others crossed first, visible in the first row of the fifth column. A strength of the VMC model is that the simulated positions at 3 s (purple dots) lie close to the real-world human positions (black dots), indicating good spatial and temporal alignment with real-world trajectories. However, a limitation is that in vehicle-first scenarios the model typically produced pedestrian-first outcomes, and when it did generate vehicle-first behaviour the vehicle showed the same aggressive driving behaviour as the MC model. It can be noted that all RL-based models were able to reproduce pedestrian-first scenarios without vehicle yielding (interaction 67), suggesting that such non-interactive patterns were relatively easier to learn regardless of constraint assumptions. However, only the VMC model showed close alignment with the real-world trajectory, as shown by the purple and black dots being close at 3 s.

4.3.3 Speed of the pedestrian and driver agent

We further examined the average speed profiles of both the pedestrian and the vehicle as a function of their distance to the crossing point, as shown in Figure 4.8. These profiles are informative for assessing how agents adjust their motion around the crossing, for example slowing down before entry, and yielding. In Figure 4.8, both agents approached the crossing from left to right along the x-axis: larger values indicate positions farther from the crossing, while smaller values represent proximity to the crossing point. In the real-world data, vehicle speed gradually decreased when approaching the crossing, while pedestrians typically slowed down slightly before entering.

The BC model failed to reproduce the deceleration behaviour for vehicles. Pedestrian speed also deviated from the empirical pattern, showing only a weak reduction before crossing.

In the NC and VC models, vehicle agents exhibited minimal deceleration, while pedestrian agents maintained consistently higher speeds than observed in real-world data. This behaviour stems from the absence of motor constraints, which allowed agents to accelerate or maintain high speeds without penalty. As a result, both models failed to replicate the typical slowing-down pattern seen in real-world interactions.

In contrast, the MC and VMC models more accurately captured the deceleration patterns of

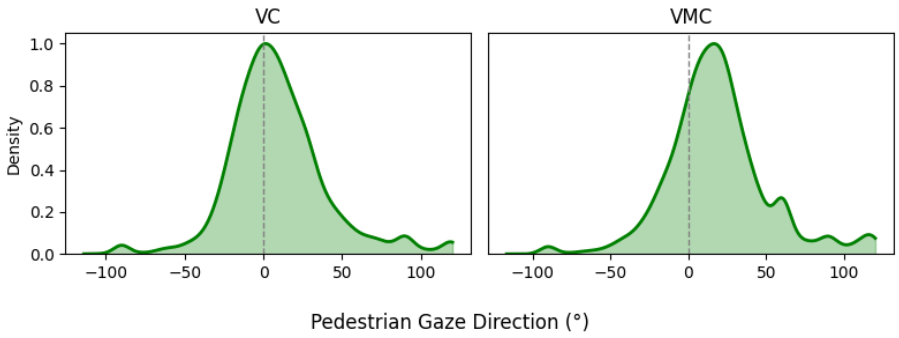

Figure 4.9: Distribution of pedestrian gaze orientation in the VC and VMC models, measured relative to the yy-axis (i.e., the upward direction on the map in Fig. 4.1). A gaze direction of 0° indicates that the pedestrian is looking straight along the yy-axis (perpendicular to the road), while positive values indicate gaze directed toward the approaching vehicle.

both agents. Pedestrian agents in these models slowed down before crossing (around $x = 6$ m) and accelerated afterwards, aligning more closely with real-world human behaviour. Vehicle agents in the MC model applied stronger and earlier braking than observed in the real-world data, followed by a more abrupt acceleration, as shown in the second column of the fourth row of Figure 4.8. The VMC model better reproduced the observed trends, although it still exhibited slightly earlier braking than human drivers (around $x = 15$ m) and both agents maintained overall lower speeds compared to the real-world data, as illustrated in the second column of the fifth row.

4.3.4 Gaze Direction under Visual Constraints

From Figure 4.9, and 4.10, it can be observed that in the VMC model, pedestrians exhibited a clear tendency to direct their gaze toward the oncoming vehicle to their right. In contrast, in the VC model the gaze direction was more centrally distributed.

A closer examination of Figure 4.10 reveals more nuanced, situation-adaptive gaze patterns. For instance, in interaction 67, the pedestrian mainly looked forward, whereas in interactions 18 and 33, the gaze shifted rightward early in the interaction when the vehicle approached, and some leftward looking late in interaction 18. These patterns suggest that the VMC pedestrian learned to monitor the vehicle when it was relevant to crossing safety, leading to a situation-adaptive gaze distribution.

This behavioural difference likely arises from the reward structure of the model. The VMC model includes motor constraints, which means the pedestrian cannot cross the road fast without

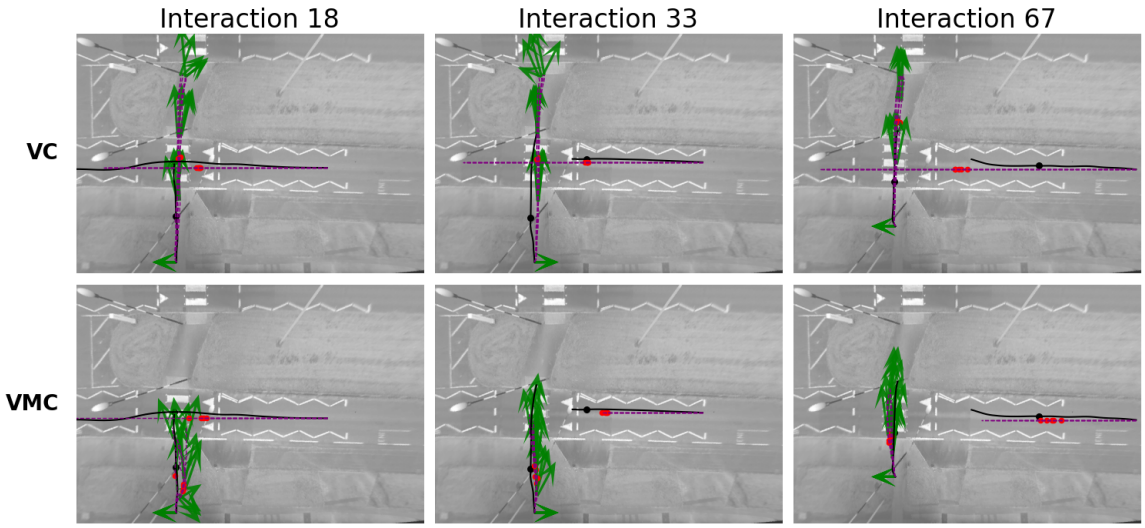

Figure 4.10: Comparison between model and real-world trajectories for the VC and VMC models across three representative interactions. Real-world trajectories are shown in black, and model trajectories in purple. Green arrows indicate the pedestrian's gaze orientation at 2 s intervals. Red and black dots denote the position of the pedestrian after 3 seconds in the model and real-world data, respectively. The background image corresponds to the real-world scene of the crossing. Note that the VMC model shows less variability than the VC model, resulting in overlapping gaze orientation arrows across repeated simulations.

penalty. As a result, the pedestrian must carefully monitor the vehicle to cross safely and avoid the collision penalty. In contrast, the VC model lacks such constraints: pedestrians can cross quickly and are rarely affected by the vehicle. As discussed in Section 4.3.2, the VC model failed to reproduce vehicle-first or vehicle-yielding scenarios, reducing the need for vehicle monitoring. Since gaze direction only matters when avoiding collisions, the learned policy may default to centre-looking or random gaze movements when vehicle monitoring is unnecessary.

4.4 Discussion

4.4.1 Main findings

In this work, we present a novel multi-agent RL framework that explicitly incorporates human-like perceptual and motor constraints. Some prior studies have explored partial observability in RL-based road user agents. For example, the multi-agent RL framework by Vinitsky et al. (2022) included a restricted field of view, and Cornelisse et al. (2025) leveraged partial observability to improve agent robustness for AV benchmarking. However, these works did not incorporate human-inspired elements such as distance-dependent visual noise or biomechanically constrained motor control, as included in our model to better reflect human constraints when

interacting with the real world.

To make the most effective use of the limited real-world data, all models were trained and evaluated on the full dataset. Given that the RL models contain relatively few free parameters (non-policy parameters in this study), the risk of overfitting was considered low. Nevertheless, two potential concerns may arise from this setup. First, one might question whether the superior performance of the VMC model compared to the other RL-based models could be due to its greater flexibility arising from a larger number of non-policy parameters. Second, the training setup for the BC baseline does not follow the conventional practice in supervised learning, of separating training and validation data. To address these issues, we conducted an additional three-fold training/validation procedure across all 2 s trajectory segments. In each fold, all models were trained (supervised learning of the BC model; Bayesian fitting of non-policy parameter for the RL model variants) on two-thirds of the data and evaluated on the remaining one-third, and the models were ranked based on their average performance across these three folds. The relative ranking of model performance using this fitting and evaluation method was exactly the same as the results in Table 4.2, confirming that the improved performance of the VMC model was not simply a consequence of additional flexibility or the atypical training setup for the BC baseline.

A key result is that the VMC model, which includes both visual and motor constraints, achieved the highest similarity to real-world data across behavioural, kinematic, and trajectory-based metrics, outperforming variants omitting one or both. Importantly, the VMC model also outperformed the behavioural cloning (BC) model, which was trained directly on real-world human demonstrations and contained a larger number of trainable parameters. This highlights the advantage of reward-driven learning under structured human-like constraints, particularly in data-limited settings.

Another methodological innovation of our work lies in modelling pedestrian-vehicle interactions at the population level through non-policy parameter distributions. Prior work in computational rationality modelling of human behaviour has often focused on individual-level modelling, where each agent is assigned its own parameter set based on observed behaviour (Chen et al., 2021b; Jokinen et al., 2021; Wang et al., 2025a; Wang et al., 2025b). While such approaches can capture inter-individual variability in controlled experiments, where each participant is observed across multiple trials, they are less suited to real-world road user modelling, where individuals are

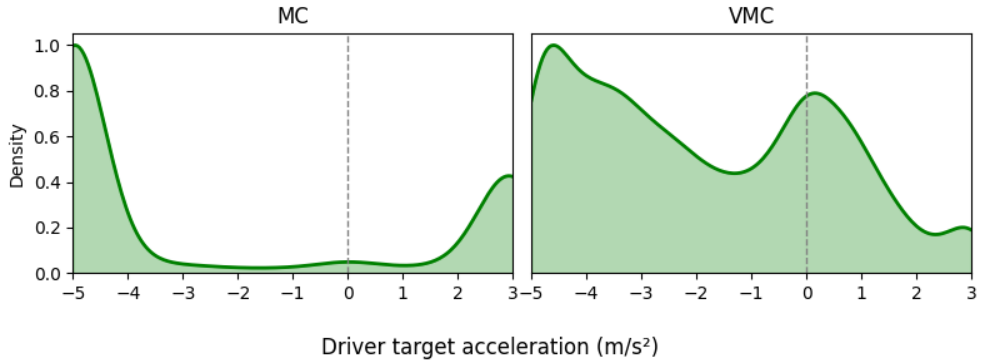

Figure 4.11: Distribution of driver target acceleration values for MC and VMC models.

typically observed only once for a short period of time, e.g., as in our case during a single interaction event. Therefore, we instead estimated population-level distributions for motor and perceptual non-policy parameters. Our simulations are then run by drawing individual agent parameters from these fitted distributions, enabling the model to capture population-level variability rather than optimising a single parameter set for each agent.

4.4.2 Impact of human constraint on the pedestrian-vehicle interaction

Here, we analyse how each type of human-inspired constraint, visual and motor, contributes to the simulated interactions. Comparing the MC and VC models indicates that introducing motor constraints (NC → MC) produces a larger improvement in interaction realism than introducing visual constraints alone (NC → VC). Although both models achieve similar NLL scores (Table 4.2), only the MC model reproduces key behavioural patterns such as vehicle-first and vehicle-yielding scenarios. This reflects the role of motor effort in regulating movement timing: by penalising abrupt acceleration or fast crossings, motor constraints lead to smoother and more human-like behaviours. In contrast, the VC model allows the pedestrian to cross rapidly without cost, which reduces the need for the adaptation to the vehicle behaviour. As a result, interaction patterns observed in real-world data are not reproduced.

Adding visual constraints on top of motor ones (i.e., MC → VMC) further improves realism, particularly in driver responses. As shown in Figure 4.11, the MC vehicle agent often brakes early and sharply, potentially due to overconfidence in its predictions about pedestrian behaviour, to avoid the not yielding penalty. In the VMC model, perceptual uncertainty introduced by visual noise encourages more cautious deceleration, consistent with the human driver deceleration profiles shown in Figure 4.8.

Together, these findings suggest that both visual and motor constraints are necessary for realistic pedestrian-vehicle interaction modelling, but in different ways. Motor constraints primarily shape the timing and smoothness of actions, while visual constraints enhance behavioural realism by inducing more hesitant responses and better capturing the variability observed in human decision-making under uncertainty. However, it should be noted that our ablations are relatively coarse: the MC variant bundles pedestrian ballistic step execution and effort cost with driver acceleration filtering, and the VC variant bundles distance-dependent retinal noise, Bayesian estimation, and gaze-dependent acuity. Below, we discuss further opportunities to build on our work.

4.4.3 Implications and future work

The current model extends our earlier framework to a multi-agent setting using real-world pedestrian-vehicle interaction data. This allows both agents to adaptively respond to each other in more realistic scenarios, bringing the framework closer to practical AV applications.

For example, with regard to AV testing, realistic human behaviour models can support the development of socially compatible AV planners, by providing training environments where AV policies learn to interact with pedestrians in ways that align with human expectations. In our case, the proposed framework enhances realism in two ways. First, its multi-agent reinforcement learning formulation enables agents to adaptively co-evolve during training, rather than only reproducing pre-specified trajectories. Second, the VMC variant not only reproduces realistic motion patterns but also simulates gaze orientation as a proxy for head movement behaviour, thereby enabling analysis of how AV systems might interpret pedestrian intent from observable cues such as head pose and facing direction. This aspect connects directly to research on pedestrian intent recognition, where algorithms commonly use head orientation to infer crossing intention or attention state (Rasouli et al., 2017; Perdana et al., 2021). While several empirical studies have examined pedestrians' head and gaze orientation prior to crossing and their relation to crossing decisions (Zhao et al., 2023; Theisen et al., 2025), few modelling approaches have explicitly coupled gaze direction with the full crossing process. The present framework therefore contributes toward bridging this gap by endogenously generating gaze behaviour as part of the decision-making policy. Nonetheless, the current account of visual attention remains simplified. Future work could incorporate mechanisms such as noisy perception of both the pedestrian's goal and surrounding agents, to capture attention sharing between task goals and other road

users, as well as effort-related costs associated with head or eye movements (Koevoet et al., 2025).

Methodologically, our ablation analysis should be interpreted at the mechanism-family level rather than as single-component attributions, since each variant combines multiple perceptual or motor elements. Future work could include finer ablations, for example separating pedestrian-only and driver-only motor constraints and disentangling visual subcomponents such as noise-only, Kalman-only, and gaze-acuity-only.

In addition to such ablation-focused refinements, several aspects of the environment design could also be improved. The current driver model lacks explicit penalties for abrupt acceleration or deceleration, which in our simulations sometimes led to unrealistic velocity fluctuations, particularly visible at the beginning of the vehicle speed profiles (left side of the right-hand panels in Figure 4.8). Future work could incorporate explicit smoothness-related rewards, such as penalties on acceleration or jerk, to better reflect the biomechanical and comfort constraints observed in human driving behaviour (Todorov and Jordan, 2002). Moreover, our simulated environment did not include a representation of the fact that the real-world zebra crossing was raised, functioning as a 'speed bump' for the driver. We experimented with adding a speed bump as a contextual factor intended to elicit braking, but our model drivers in the simulation either ignored it or stopped before it, and the modification did not provide further improvements in model human-likeness.

A further limitation is that we manually fixed several non-policy parameters, such as the time cost coefficient, rather than estimating them from data. We experimented with including these as additional free parameters in the modelling pipeline, but this substantially increased training time and reduced convergence stability. These results suggest that the specific RL approach we used here may struggle to handle more complex RL environments, such as those including contextual features like speed bumps or variable time costs, without additional architectural or algorithmic advances. Future work could address this by exploring more advanced policy representations, for example recurrent or attention-based neural networks that can better encode history and context, or by applying more explicit forms of inverse reinforcement learning (IRL), going beyond the parameter fitting used here to directly infer reward functions from human demonstrations. In addition, in this study, we used the SAC algorithm rather than the Proximal Policy Optimisation (PPO) algorithm (Schulman et al., 2017). Initial experiments using PPO

revealed that the agent frequently selected boundary values in the action space. While the primary focus of this work was on RL environment design and human-like constraints, future research should also consider the role of RL algorithm choice and optimisation in shaping policy behaviour. The selection of RL algorithms may also affect the stability, and realism of learned behaviours.

Another aspect relating to evaluation robustness concerns the BC baseline. As noted above, we trained and evaluated BC on the entire dataset to align conditions with the RL setup in a data-limited setting. This supports high one-step, open-loop accuracy on the same data, but closed-loop rollouts can accumulate errors once the policy drifts away from the demonstrated trajectory. Future work could address this limitation by training the BC model on a larger and more diverse dataset to better examine its behaviour across varied pedestrian-vehicle interactions.

We argue that the superior performance of the VMC model, relative to the other model variants, primarily reflects structural changes introduced by the perceptual and motor constraints rather than the larger number of non-policy parameters. This interpretation is supported by the training/validation split analysis we mentioned above, where the relative performance ranking remained consistent across folds, indicating that the improved performance of the VMC model is unlikely to result from overfitting due to additional parameters. However, we did not test this using metrics such as the Akaike Information Criterion (AIC), which balances model fit and complexity by penalising the number of free parameters (Akaike, 2003). Our present evaluation followed Montali et al. (2023) in relying on a composite NLL, whereas AIC requires likelihoods computed at the level of individual observations. Future work could therefore either develop a way to compute AIC from our setting, for example by deriving standardised individual-level likelihoods, or explore alternative criteria that explicitly balance goodness of fit against model complexity. Such approaches would help distinguish improvements due to human constraints from those arising from additional flexibility.

Finally, this study focused on simplified one-to-one interactions between a single pedestrian and vehicle. While this setup supports simpler training and interpretable analysis, extending the framework to model groups of pedestrians, and dense traffic environments remains an important next step for real-world scalability and generalisation.

4.5 Conclusions

This study developed a multi-agent reinforcement learning framework that integrates human-like sensory and motor constraints to model pedestrian-vehicle interactions at unsignalised crossings. We evaluated four model variants—No Constraint (NC), Motor Constraint (MC), Visual Constraint (VC), and Visual and Motor Constraint (VMC)—on a real-world dataset using a comprehensive set of behavioural, kinematic, and trajectory-based metrics. The VMC model achieved the highest similarity to human data across all metrics, outperforming other RL variants and a behavioural cloning (BC) model, demonstrating the value of RL-based human behaviour modelling in data-limited conditions. In addition, the VMC model provided a simple representation of gaze orientation, offering a first step toward capturing how pedestrians use gaze in interactive decision-making. By comparing models with different human constraints, we found that motor constraints result in smoother movements that reflect human-like speed adjustments during crossing interactions. The inclusion of visual constraints introduces perceptual uncertainty and field-of-view limitations, leading the agents to exhibit more cautious and variable behaviour, such as less abrupt deceleration. These findings highlight the importance of including both motor and visual constraints in interactive behaviour modelling. Another contribution of our approach lies in its use of population-level model fitting of RL-based human behaviour models. Instead of optimising a fixed parameter set for each simulated agent, we infer population-level distributions over sensory and motor constraints. This allows the model to reflect group-level behavioural tendencies while preserving individual variability. Such an approach is particularly valuable in real-world road user modelling, where repeated measurements per individual are rare, and stands in contrast to prior work primarily conducted in controlled experimental settings. The proposed framework provides a promising direction for both behavioural modelling and the development of socially acceptable AVs. Future work could extend the approach to more complex scenarios, such as group pedestrian dynamics and multi-vehicle interactions, and refine the reward design to better capture social and contextual norms in road user behaviour.

References

Akaike, Hirotugu (2003). “A new look at the statistical model identification”. In: *IEEE Transactions on Automatic Control* 19.6, pp. 716–723.

Aradi, Szilárd (2020). "Survey of deep reinforcement learning for motion planning of autonomous vehicles". In: *IEEE Transactions on Intelligent Transportation Systems* 23.2, pp. 740–759.

Bain, Michael and Sammut, Claude (1995). "A Framework for Behavioural Cloning." In: *Machine intelligence 15*, pp. 103–129.

Barton, Sean L, Matthis, Jonathan S, and Fajen, Brett R (2019). "Control strategies for rapid, visually guided adjustments of the foot during continuous walking". In: *Experimental brain research* 237, pp. 1673–1690.

Carlisle, Rebecca Elizabeth and Kuo, Arthur D (2023). "Optimization of energy and time predicts dynamic speeds for human walking". In: *Elife* 12, e81939.

Charalambous, Panayiotis and Chrysanthou, Yiorgos (2019). "Why did the human cross the road?" In: *Proceedings of the 12th ACM SIGGRAPH Conference on Motion, Interaction and Games*, pp. 1–2.

Chen, Haiyang, Chang, Hyung Jin, and Howes, Andrew (2021a). "Apparently irrational choice as optimal sequential decision making". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 35. 1, pp. 792–800.

Chen, Xiuli, Acharya, Aditya, Oulasvirta, Antti, and Howes, Andrew (2021b). "An adaptive model of gaze-based selection". In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–11.

Chen, Xiuli, Bailly, Gilles, Brumby, Duncan P, Oulasvirta, Antti, and Howes, Andrew (2015). "The emergence of interactive behavior: A model of rational menu search". In: *Proceedings of the 33rd annual ACM conference on human factors in computing systems*, pp. 4217–4226.

Choi, Tae Ho (1997). *Development of a mathematical model of gait dynamics*.

Christianos, Filippos, Schäfer, Lukas, and Albrecht, Stefano (2020). "Shared experience actor-critic for multi-agent reinforcement learning". In: *Advances in neural information processing systems* 33, pp. 10707–10717.

Cornelisse, Daphne, Pandya, Aarav, Joseph, Kevin, Suárez, Joseph, and Vinitsky, Eugene (2025). "Building reliable sim driving agents by scaling self-play". In: *arXiv preprint arXiv:2502.14706*.

Cornelisse, Daphne and Vinitsky, Eugene (2024). "Human-compatible driving partners through data-regularized self-play reinforcement learning". In: *arXiv preprint arXiv:2403.19648*.

Craik, Kenneth JW (1948). "Theory of the human operator in control systems. II. Man as an element in a control system". In: *British journal of psychology* 38.3, p. 142.

Dang, Meiting, Zhao, Dezong, Wang, Yafei, and Wei, Chongfeng (2025). "Dynamic Game-Theoretical Decision-Making Framework for Vehicle-Pedestrian Interaction With Human Bounded Rationality". In: *IEEE Transactions on Intelligent Transportation Systems*.

Deo, Nachiket and Trivedi, Mohan M (2018). "Convolutional social pooling for vehicle trajectory prediction". In: *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pp. 1468–1476.

Diaz-Ruiz, Carlos A, Xia, Youya, You, Yurong, Nino, Jose, Chen, Junan, Monica, Josephine, Chen, Xiangyu, Luo, Katie, Wang, Yan, Emond, Marc, et al. (2022). "Ithaca365: Dataset and Driving Perception Under Repeated and Challenging Weather Conditions". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 21383–21392.

Divya, J, Divyadharshini, M, and Divyapriya, V (2025). "Gaze-based Pedestrian Safety Prediction using Attention LSTM". In: *2025 3rd International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)*. IEEE, pp. 1838–1845.

Domeyer, Joshua E, Lee, John D, Toyoda, Heishiro, Mehler, Bruce, and Reimer, Bryan (2022). "Driver-Pedestrian Perceptual Models Demonstrate Coupling: Implications for Vehicle Automation". In: *IEEE Transactions on Human-Machine Systems* 52.4, pp. 557–566.

Faisal, A Aldo, Selen, Luc PJ, and Wolpert, Daniel M (2008). "Noise in the nervous system". In: *Nature reviews neuroscience* 9.4, pp. 292–303.

Faraji, Salman, Wu, Amy R., and Ijspeert, Auke Jan (2018). "A simple model of mechanical effects to estimate metabolic cost of human walking". In: *Scientific Reports* 8.1, pp. 10998–10998. DOI: 10.1038/S41598-018-29429-Z.

Fox, Charles, Camara, Fanta, Markkula, Gustav, Romano, Richard, Madigan, Ruth, Merat, Natasha, et al. (2018). "When should the chicken cross the road?: Game theory for autonomous vehicle-human interactions". In.

Gershman, Samuel J, Horvitz, Eric J, and Tenenbaum, Joshua B (2015). "Computational rationality: A converging paradigm for intelligence in brains, minds, and machines". In: *Science* 349.6245, pp. 273–278.

Giles, Oscar, Markkula, Gustav, Pekkanen, Jami, Yokota, Naoki, Matsunaga, Naoto, Merat, Natasha, and Daimon, Tatsuru (2019). "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, pp. 366–372.

González-Méndez, Mauricio, Olaya, Camilo, Fasolino, Isidoro, Grimaldi, Michele, and Obregón, Nelson (2021). "Agent-based modeling for urban development planning based on human needs. Conceptual basis and model formulation". In: *Land Use Policy* 101, p. 105110.

Grieve, DW (1968). "Gait patterns and the speed of walking." In: *Bio-Med. Eng.* 3, pp. 119–112.

Haarnoja, Tuomas, Zhou, Aurick, Abbeel, Pieter, and Levine, Sergey (2018). "Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor". In: *International conference on machine learning*. Pmlr, pp. 1861–1870.

Howes, Andrew, Jokinen, Jussi PP, and Oulasvirta, Antti (2023). "Towards machines that understand people". In: *AI Magazine* 44.3, pp. 312–327.

Itkina, Masha and Kochenderfer, Mykel (2023). "Interpretable self-aware neural networks for robust trajectory prediction". In: *Conference on Robot Learning*. PMLR, pp. 606–617.

Jokinen, Jussi PP, Kujala, Tuomo, and Oulasvirta, Antti (2021). "Multitasking in driving as optimal adaptation under uncertainty". In: *Human factors* 63.8, pp. 1324–1341.

Jokinen, Jussi PP, Wang, Zhenxin, Sarcar, Sayan, Oulasvirta, Antti, and Ren, Xiangshi (2020). "Adaptive feature guidance: Modelling visual search with graphical layouts". In: *International Journal of Human-Computer Studies* 136, p. 102376.

Kalantari, Amir Hossein, Lin, Yi-Shin, Mohammadi, Ali, Merat, Natasha, and Markkula, Gustav (2025). "Testing the Validity of Multiparticipant Distributed Simulation for Understanding and Modeling Road User Interaction". In: *IEEE Transactions on Human-Machine Systems*.

Knill, David C and Pouget, Alexandre (2004). "The Bayesian brain: the role of uncertainty in neural coding and computation". In: *TRENDS in Neurosciences* 27.12, pp. 712–719.

Koevoet, Damian, Van Zantwijk, Laura, Naber, Marnix, Mathôt, Sebastiaan, Van der Stigchel, Stefan, and Strauch, Christoph (2025). "Effort drives saccade selection". In: *ELife* 13, RP97760.

Kong, Xiaoqiang, Das, Subasish, Zhang, Yunlong, Wei, Zihang, and Yuan, Chi-hao (2023). "In-depth understanding of pedestrian–vehicle near-crash events at signalized intersections: An interpretable machine learning approach". In: *Transportation research record* 2677.5, pp. 747–759.

Kotseruba, Iuliia and Rasouli, Amir (2023). "Intend-Wait-Perceive-Cross: Exploring the Effects of Perceptual Limitations on Pedestrian Decision-Making". In: *2023 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, pp. 1–6.

Kwon, Oh-Sang, Tadin, Duje, and Knill, David C (2015). "Unifying account of visual motion and position perception". In: *Proceedings of the National Academy of Sciences* 112.26, pp. 8142–8147.

Liang, Eric, Liaw, Richard, Nishihara, Robert, Moritz, Philipp, Fox, Roy, Goldberg, Ken, Gonzalez, Joseph, Jordan, Michael, and Stoica, Ion (2018). "RLlib: Abstractions for distributed reinforcement learning". In: *International conference on machine learning*. PMLR, pp. 3053–3062.

Lin, Tengfeng, Jin, Zhixiong, Choi, Seongjin, and Yeo, Hwasoo (2024). "A Real-time Evaluation Framework for Pedestrian's Potential Risk at Non-Signalized Intersections Based on Predicted Post-Encroachment Time". In: *arXiv preprint arXiv:2404.15635*.

Lobjois, Régis and Cavallo, Viola (2007). "Age-related differences in street-crossing decisions: The effects of vehicle speed and time constraints on gap selection in an estimation task". In: *Accident analysis & prevention* 39.5, pp. 934–943.

Madala, Kaushik and Gonzalez, Carlos Avalos (2023). "Metrics for machine learning models to facilitate sotif analysis in autonomous vehicles". In: *SAE International Journal of Advances and Current Practices in Mobility* 6.2023-01-0829, pp. 782–790.

Markkula, Gustav, Lin, Yi-Shin, Srinivasan, Aravinda Ramakrishnan, Billington, Jac, Leonetti, Matteo, Kalantari, Amir Hossein, Yang, Yue, Lee, Yee Mun, Madigan, Ruth, and Merat, Natasha (2023). "Explaining human interactions on the road by large-scale integration of computational psychological theory". In: *PNAS nexus* 2.6, pgad163.

Montali, Nico, Lambert, John, Mougin, Paul, Kuefler, Alex, Rhinehart, Nicholas, Li, Michelle, Gulino, Cole, Emrich, Tristan, Yang, Zoey, Whiteson, Shimon, et al. (2023). "The waymo open sim agents challenge". In: *Advances in Neural Information Processing Systems* 36, pp. 59151–59171.

Nasernejad, Payam, Sayed, Tarek, and Alsaleh, Rushdi (2021). "Modeling pedestrian behavior in pedestrian-vehicle near misses: A continuous Gaussian Process Inverse Reinforcement Learning (GP-IRL) approach". In: *Accident Analysis & Prevention* 161, p. 106355.

– (2023). "Multiagent modeling of pedestrian-vehicle conflicts using Adversarial Inverse Reinforcement Learning". In: *Transportmetrica A: transport science* 19.3, p. 2061081.

Ng, Andrew Y, Russell, Stuart, et al. (2000). "Algorithms for inverse reinforcement learning." In: *Icml*. Vol. 1, p. 2.

Ooi, Teng Leng, Wu, Bing, and He, Zijiang J (2001). "Distance determined by the angular declination below the horizon". In: *Nature* 414.6860, pp. 197–200.

Oulasvirta, Antti, Jokinen, Jussi PP, and Howes, Andrew (2022). "Computational rationality as a theory of interaction". In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pp. 1–14.

Pekkanen, Jami, Giles, Oscar Terence, Lee, Yee Mun, Madigan, Ruth, Daimon, Tatsuru, Merat, Natasha, and Markkula, Gustav (2022). "Variable-drift diffusion models of pedestrian road-crossing decisions". In: *Computational Brain & Behavior* 5.1, pp. 60–80.

Perdana, Muhammad Ilham, Anggraeni, Wiwik, Sidharta, Hanugra Aulia, Yuniarno, Eko Mulyanto, and Purnomo, Mauridhi Hery (2021). "Early warning pedestrian crossing intention from its head gesture using head pose estimation". In: *2021 International seminar on intelligent technology and its applications (ISITIA)*. IEEE, pp. 402–407.

Rasouli, Amir, Kotseruba, Iuliia, and Tsotsos, John K (2017). "Understanding pedestrian behavior in complex traffic scenes". In: *IEEE Transactions on Intelligent Vehicles* 3.1, pp. 61–70.

Rasouli, Amir and Tsotsos, John K (2019). "Autonomous vehicles that interact with pedestrians: A survey of theory and practice". In: *IEEE transactions on intelligent transportation systems* 21.3, pp. 900–918.

Roshdi, Mohamed, Petzold, Julian, Wahby, Mostafa, Ebrahim, Hussein, Berekovic, Mladen, and Hamann, Heiko (2024). "On the Road to Clarity: Exploring Explainable AI for World Models in a Driver Assistance System". In: *2024 IEEE Conference on Artificial Intelligence (CAI)*. IEEE, pp. 1032–1039.

Sackmann, Moritz (2024). *Learning Driver Behavior Models for Predicting Urban Traffic Situations*. Friedrich-Alexander-Universitaet Erlangen-Nuernberg (Germany).

Schneemann, Friederike and Gohl, Irene (2016). "Analyzing driver-pedestrian interaction at crosswalks: A contribution to autonomous driving in urban environments". In: *2016 IEEE intelligent vehicles symposium (IV)*. IEEE, pp. 38–43.

Schulman, John, Wolski, Filip, Dhariwal, Prafulla, Radford, Alec, and Klimov, Oleg (2017). "Proximal policy optimization algorithms". In: *arXiv preprint arXiv:1707.06347*.

Scott, David W (2015). *Multivariate density estimation: theory, practice, and visualization*. John Wiley & Sons.

Sharp, Robin S (2005). "Driver steering control and a new perspective on car handling qualities". In: *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science* 219.10, pp. 1041–1051.

Srinivasan, Aravinda Ramakrishnan, Lin, Yi-Shin, Antonello, Morris, Knittel, Anthony, Hasan, Mohamed, Hawasly, Majd, Redford, John, Ramamoorthy, Subramanian, Leonetti, Matteo, Billington, Jac, et al. (2023). "Beyond RMSE: Do machine-learned models of road user interaction produce human-like behavior?" In: *IEEE Transactions on Intelligent Transportation Systems*.

Stocker, Alan A and Simoncelli, Eero P (2006). "Noise characteristics and prior expectations in human visual speed perception". In: *Nature neuroscience* 9.4, pp. 578–585.

Sutton, Richard S and Barto, Andrew G (2018). *Reinforcement learning: An introduction*. MIT press.

Theisen, Max, Bergen, Melina, Einhäuser, Wolfgang, and Schießl, Caroline (2025). "Head movements predict pedestrian crossing decisions earlier than walking speed". In: *Transportation Research Part F: Traffic Psychology and Behaviour* 113, pp. 586–608.

Tian, Kai, Markkula, Gustav, Wei, Chongfeng, Lee, Yee Mun, Madigan, Ruth, Merat, Natasha, and Romano, Richard (2022a). "Explaining unsafe pedestrian road crossing behaviours using a psychophysics-based gap acceptance model". In: *Safety science* 154, p. 105837.

Tian, Kai, Markkula, Gustav, Wei, Chongfeng, Sadraei, Ehsan, Hirose, Toshiya, Merat, Natasha, and Romano, Richard (2022b). "Impacts of visual and cognitive distractions and time pressure on pedestrian crossing behaviour: A simulator study". In: *Accident Analysis & Prevention* 174, p. 106770.

Todorov, Emanuel and Jordan, Michael I (2002). "Optimal feedback control as a theory of motor coordination". In: *Nature neuroscience* 5.11, pp. 1226–1235.

Tseng, Yuan-Chi and Howes, Andrew (2015). "The adaptation of visual search to utility, ecology and design". In: *International Journal of Human-Computer Studies* 80, pp. 45–55.

Tustin, Arnold (1947). "The nature of the operator's response in manual control, and its implications for controller design". In: *Journal of the Institution of Electrical Engineers-Part IIA: Automatic Regulators and Servo Mechanisms* 94.2, pp. 190–206.

Vinitsky, Eugene, Lichtlé, Nathan, Yang, Xiaomeng, Amos, Brandon, and Foerster, Jakob (2022). "Nocturne: a scalable driving benchmark for bringing multi-agent learning one step closer to the real world". In: *Advances in Neural Information Processing Systems* 35, pp. 3962–3974.

Vizzari, Giuseppe and Cecconello, Thomas (2022). "Pedestrian simulation with reinforcement learning: a curriculum-based approach". In: *Future Internet* 15.1, p. 12.

Wang, Yueyang, Srinivasan, Aravinda Ramakrishnan, Jokinen, Jussi PP, Oulasvirta, Antti, and Markkula, Gustav (2023). "Modeling human road crossing decisions as reward maximization with visual perception limitations". In: *2023 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, pp. 1–6.

– (2025a). "Pedestrian crossing decisions can be explained by bounded optimal decision-making under noisy visual perception". In: *Transportation Research Part C: Emerging Technologies* 171, p. 104963.

Wang, Yueyang, Srinivasan, Aravinda Ramakrishnan, Lee, Yee Mun, and Markkula, Gustav (2025b). "Modeling Pedestrian Crossing Behavior: A Reinforcement Learning Approach With Sensory Motor Constraints". In: *IEEE Transactions on Intelligent Transportation Systems*.

Watson, Andrew B (2014). "A formula for human retinal ganglion cell receptive field density as a function of visual field location". In: *Journal of vision* 14.7, pp. 15–15.

Xu, Yiran, Yang, Xiaoyin, Gong, Lihang, Lin, Hsuan-Chu, Wu, Tz-Ying, Li, Yunsheng, and Vasconcelos, Nuno (2020). "Explainable object-induced action decision for autonomous vehicles". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 9523–9532.

Yang, Changzhi, Pan, Huihui, and Wang, Jue (2025). "Interact, Plan, and Go: Transformers With Social Intentions for Trajectory Prediction". In: *IEEE Transactions on Consumer Electronics*.

Ye, Luyao, Wang, Zezhong, Chen, Xinhong, Wang, Jianping, Wu, Kui, and Lu, Kejie (2021). "GSAN: Graph self-attention network for learning spatial–temporal interaction representation in autonomous driving". In: *IEEE Internet of Things Journal* 9.12, pp. 9190–9204.

Yi, Shuai, Li, Hongsheng, and Wang, Xiaogang (2016). "Pedestrian behavior understanding and prediction with deep neural networks". In: *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14*. Springer, pp. 263–279.

Zhao, Qingxian, Zhuang, Xiangling, Zhang, Tong, He, Yunqi, and Ma, Guojie (2023). "Pedestrian gaze pattern before crossing road in a naturalistic traffic setting". In: *European Transport Research Review* 15.1, p. 31.

Chapter 5

Realistic adversarial scenario generation via human-like pedestrian model for autonomous vehicle control parameter optimization

Abstract

Autonomous vehicles (AVs) are rapidly advancing and are expected to play a central role in future mobility. Ensuring their safe deployment requires reliable interaction with other road users, not least pedestrians. Direct testing on public roads is costly and unsafe for rare but critical interactions, making simulation a practical alternative. Within simulation-based testing, adversarial scenarios are widely used to probe safety limits, but many prioritise difficulty over realism, producing exaggerated behaviours which may result in AV controllers that are overly conservative. We propose an alternative method, instead using a cognitively inspired pedestrian model featuring both inter-individual and intra-individual variability to generate behaviourally plausible adversarial scenarios. We provide a proof of concept demonstration of this method's potential for AV control optimisation, in closed-loop testing and tuning of an AV controller. Our results show that replacing the rule-based CARLA pedestrian with the human-like model yields more realistic gap acceptance patterns and smoother vehicle decelerations. Unsafe interactions

occur only for certain pedestrian individuals and conditions, underscoring the importance of human variability in AV testing. Adversarial scenarios generated by this model can be used to optimise AV control towards safer and more efficient behaviour. Overall, this work illustrates how incorporating human-like road user models into simulation-based adversarial testing can enhance the credibility of AV evaluation and provide a practical basis to behaviourally informed controller optimisation.

5.1 Introduction

Autonomous vehicles (AVs) have received significant attention from both academia and industry due to their potential to shape the next generation of mobility. They promise benefits such as reduced labour costs and improved safety by mitigating human error, which accounts for the majority of traffic accidents (Muralidhar et al., 2023). Major industrial actors such as Waymo and Baidu Apollo have already deployed AV pilots in urban environments, highlighting both the maturity of this technology and the urgent need for its safe evaluation (WebProNews, 2025; Reuters, 2025).

However, several challenges remain before AVs can be safely deployed at scale in diverse real-world environments. A central requirement is reliable operation in mixed-traffic conditions, where AVs must anticipate and respond to the actions of other road users (Camara et al., 2020). Evaluating AV behaviour in such interactions is therefore essential prior to large-scale deployment.

Testing these behaviours directly on public roads is costly, time consuming, and potentially hazardous, particularly for rare or safety-critical scenarios. It is therefore impractical to rely solely on real-world testing for validation. Simulation-based testing has consequently emerged as a practical and scalable alternative (Huang et al., 2016). Compared with road testing, virtual environments allow controlled and repeatable experiments across a wide range of conditions, including dangerous or rare interactions (Joisher et al., 2019; Sawitzky et al., 2023). For these reasons, simulation is now recognised as a fundamental tool in the design and assessment of AV controller (Wang et al., 2024).

Within simulation environments, the fidelity of the models that represent other road users is particularly critical, since their behaviour directly shapes the realism and rigour of evaluation. Vulnerable road users, and pedestrians in particular, require special attention as they account for approximately 22% of global road traffic fatalities (World Health Organization, 2013). However, modelling interactions with pedestrians is particularly challenging due to their high variability and inherent uncertainty. Unlike interactions between vehicles, which often follow traffic regulations and exhibit relatively constrained dynamics, pedestrian behaviour is shaped by a wider set of contextual, psychological, environmental, and motor factors (Rasouli and Tsotsos, 2019). For instance, differences in motor abilities across individuals influence pedestrians' walking speed

and movement initiation during road crossing (Wang et al., 2025b). Accurately modelling these subtleties in simulation is therefore essential for avoiding overly simplistic assumptions that limit the real-world applicability of AV control algorithms (Mirzabagheri et al., 2025).

Despite this need, most existing simulation frameworks rely on relatively simple pedestrian models. These are often rule-based or deterministic, relying on fixed patterns or basic heuristics such as social-force or cellular automata formulations (Helbing and Molnar, 1995; Okazaki, 1979; Blue and Adler, 2001). While computationally efficient, such approaches oversimplify pedestrian behaviour, failing to capture the natural variability and adaptability that characterise real human decision-making (Camara et al., 2020). As a result, the realism of pedestrian-vehicle interactions in these simulations remains limited, constraining the validity of the evaluation outcomes.

To improve the realism of pedestrian behaviour representation, more advanced pedestrian models have recently been developed. These include cognitively inspired models that capture perceptual uncertainty and decision-making processes (Markkula et al., 2018; Pekkanen et al., 2022; Tian et al., 2025), game-theoretic models that represent interactive intent inference (Dang et al., 2025), and reinforcement learning (RL) models that learn bounded optimal behaviour under human-like constraints (Wang et al., 2025b; Wang et al., 2025a). The COMMOTIONS framework, in particular, integrates sensory, motor, and cognitive mechanisms to reproduce more naturalistic pedestrian behaviour (Markkula et al., 2023). However, despite these advances, such human-like agents have not yet been systematically integrated into simulation-based AV testing. It therefore remains untested whether applying these human-like pedestrian models within AV simulation testing improves behavioural realism, or how their use might influence the safety-critical behaviours of AVs.

Parallel to these efforts on realism, another line of research has investigated adversarial scenario generation. This refers to the deliberate construction of traffic situations designed to challenge AV decision-making, often by pushing interactions towards unsafe or near-miss outcomes. Such approaches have been investigated primarily as a means to expose failure modes in AV control systems. Typical methods employ RL or black-box optimisation to train pedestrian agents that provoke unsafe interactions, such as collisions with vehicles (Song et al., 2023; Hanselmann et al., 2022; Priisalu et al., 2023). For example, the 'suicidal pedestrian' framework explicitly trains agents to collide with AVs (Yang et al., 2023), while the adversarial jaywalker model uses a

multi-state rule-based controller to generate hazardous crossings, and has been proposed for AV robustness testing (Muktadir and Whitehead, 2022). Although effective for robustness testing of the AV control algorithm, these methods typically prioritise scenario difficulty over behavioural plausibility. The resulting agents often exhibit exaggerated or unrealistic behaviours that fail to reflect real pedestrian intentions, thereby undermining realism and limiting their value for informing real-world deployment (Dyro et al., 2024).

Beyond their role in robustness testing, adversarial scenarios also have the potential to inform the optimisation of AV control in pedestrian interaction scenarios. Effective interaction requires balancing safety and efficiency: AVs must avoid collisions and maintain sufficient safety margins, while also preventing unnecessary braking and excessive delays. Optimising AV control against overly simple or overly aggressive pedestrian models may lead to control algorithms that fail to generalise to real human interactions, resulting in behaviours that are either unsafe or overly defensive. In practice, however, adversarial scenario generation has so far been employed mainly as a robustness testing tool, relying on non-human-like pedestrian models that produce exaggerated or unrealistic behaviours. This disconnect reveals a methodological gap: human-like pedestrian models and adversarial testing have largely developed in parallel, and their integration for improving both the realism and practical value of AV evaluation remains limited.

Rather than attempting to replicate the full complexity of real urban environments, the present study focuses on illustrating a conceptual point: that integrating human-like pedestrian models can yield behaviourally plausible adversarial interactions more suitable for AV optimisation than existing rule-based or overly aggressive adversarial agents. As a first step in this line of research, we focus on a simplified one-to-one pedestrian-vehicle interaction at an unsignalised zebra crossing. Building on this foundation, we employ the COMMOTIONS pedestrian model, which integrates human-like sensory, motor, and decision-making mechanisms to reproduce more naturalistic crossing behaviours (Markkula et al., 2023). Using this model, we first test whether incorporating a human-like pedestrian agent improves the behavioural realism of AV simulation testing by benchmarking it against the rule-based CARLA pedestrian. We then employ the model to generate adversarial yet behaviourally realistic scenarios, creating challenging but plausible interactions that go beyond prior approaches. Finally, we use these realistic adversarial scenarios to optimise AV control, targeting a trade-off between safety and efficiency, and we

benchmark the optimisation outcomes against the adversarial jaywalker baseline (Muktadir and Whitehead, 2022). It should be emphasised that the purpose of the simulation method in this chapter is not to generate crashes or near crashes as an end in itself, nor to approximate real-world crash frequencies. Instead, the aim is to expose the AV to relatively safety-critical but non-collision interactions. The method should therefore be understood as a way of creating challenging but behaviourally plausible test scenarios, rather than as a direct model of crash occurrence.

The novelty of this work lies in unifying human-like behavioural realism with adversarial scenario generation and demonstrating its practical value for AV control. Unlike previous studies that either introduced cognitively inspired pedestrian models without embedding them in AV systems, or used adversarial approaches focused solely on robustness testing, the present study offers a proof of concept demonstration of how closed-loop AV testing and optimisation with human-like pedestrian models can be achieved. Such end-to-end evaluations, where pedestrian models and AV controllers are jointly simulated, remain limited in existing AV research. Through this integration, we analyse how two different AV controllers interact with diverse pedestrian individuals, revealing that safety-critical outcomes emerge only under specific pedestrian instances and conditions. This finding highlights the importance of accounting for both inter- and intra-individual variability in pedestrian behaviour when evaluating AV safety and performance, and demonstrates how behaviourally realistic adversarial scenarios can guide the optimisation of AV control parameters.

5.2 Methods

At the core of our framework is the COMMOTIONS pedestrian model (Markkula et al., 2023), which integrates sensory, motor, and cognitive mechanisms to capture human-like road-crossing behaviour. It should be noted that the COMMOTIONS model is not the modelling framework developed across the main doctoral work in this thesis, but is instead used here as an existing mechanistic benchmark for comparison; as explained in Section 1.5. The model was validated by demonstrating that it qualitatively reproduced key interaction phenomena and quantitatively matched pedestrian-driver interaction outcomes observed in a distributed simulator study. It captures inter-individual variability through distinct parameter sets for different individuals, and its perception and decision-making mechanisms are stochastic, i.e., the model also exhibits

Figure 5.1: Illustration of the zebra crossing interaction scenario used to evaluate the impact of different pedestrian models on AV behaviour. The pedestrian begins at a standstill, positioned 4 m from the centreline of the vehicle lane and 2 m from the kerb, while the AV approaches at the constant speed under different Time-to-Arrival (TTA) settings. The pedestrian began at a standstill, positioned 4 m from the centreline of the vehicle lane and 2 m from the kerb, and waited at the kerb before initiating a crossing

intra-individual variability. These properties make it suitable both for generating adversarial scenarios that emerge from plausible human variability and for evaluating how AV controllers respond to more realistic pedestrian behaviour. It should also be noted that the COMMOTIONS model is not designed to generate crashes. Rather, it is a human behaviour model developed to reproduce human-like pedestrian decision making and interaction dynamics, and is therefore expected to successfully avoid collision in most cases. In the present chapter, its role is to generate relatively safety-critical but mainly non-collision interactions that can be used to evaluate AV behaviour, rather than deliberately crash scenarios. For more detailed information about the model, readers are referred to Markkula et al. (2023). Building on this foundation, our method comprises three main components, which will be described below: (i) evaluation of pedestrian model impact on AV interaction, (ii) generation of adversarial scenarios, and (iii) optimisation of AV control using diverse scenario sets.

5.2.1 Evaluation of pedestrian model impact on AV interaction

The first research question addressed in this study is whether a human-like pedestrian model changes AV-pedestrian interaction outcomes and improves the ecological validity of simulation testing. To support realistic testing of autonomous vehicle behaviour, we integrated the open-source Autoware autonomous driving stack into the CARLA simulation environment via the Autoware-CARLA bridge. This setup allowed the AV to perceive and respond to pedestrians using its full planning and control pipeline, providing a stronger basis for evaluating interactions under different behavioural assumptions.

Using this simulation setup, we compared two pedestrian models and two AV control architectures. The baseline CARLA pedestrian was controlled using CARLA's built-in walker controller (CARLA Team, 2026). After being assigned a destination, the CARLA pedestrian moves automatically towards the target under CARLA's internal navigation system, with an optional speed setting. The default CARLA pedestrian is therefore governed by simulator navigation rather than by an explicit behavioural model of road crossing.

More specifically, walking is not initiated through a decision about whether the approaching vehicle leaves a safe or acceptable gap. Instead, once the controller is started, the pedestrian moves towards the assigned destination whenever that destination requires crossing the road. If CARLA's built-in walker logic detects a nearby vehicle along the pedestrian's path, the pedestrian stops temporarily; when this condition no longer holds, movement towards the destination resumes. The stop-go behaviour is therefore governed by destination following plus a proximity-based vehicle check, rather than by an explicit behavioural model in which the pedestrian evaluates vehicle speed, time gap, or gap acceptability before deciding whether to cross. These aspects of the default CARLA pedestrian clearly distinguishes it from the COMMOTIONS model.

Similarly, two AV controllers were considered. The CARLA AV was controlled using the built-in Carla `BehaviorAgent`. This agent follows a planned route, tracks waypoints, and applies low-level vehicle control while reacting to traffic lights and road obstacles through built-in rule-based logic. Its interactions with pedestrians are governed by pre-set behavioural parameters, such as braking distance and related safety thresholds, rather than by an explicit model of pedestrian-vehicle interaction.

The Autoware AV, by contrast, uses a full open-source autonomous driving stack integrated with CARLA via the Autoware-CARLA bridge. Its behaviour is generated through a layered architecture for perception, planning, and control, designed to produce safe and rule-compliant vehicle motion. In pedestrian interaction contexts such as crosswalks, the system explicitly detects relevant road users and adjusts speed or stopping behaviour accordingly. The comparison therefore contrasts a built-in rule-based CARLA driving agent with a full-stack AV architecture in which pedestrian interaction is handled through explicit perception and staged planning.

We then assessed the influence of pedestrian behavioural realism and AV control architecture using a 2×2 design: pedestrian model (CARLA vs. COMMOTIONS) crossed with AV controller (CARLA planner vs. Autoware).

All experiments were conducted in the CARLA Town 1 map, using a controlled zebra crossing scenario. The pedestrian began at a standstill, positioned 4 m from the centreline of the vehicle lane and 2 m from the kerb, and waited at the kerb before initiating a crossing (Figure 5.1). The autonomous vehicle approached along a straight road at a constant speed of 30 km/h. Vehicles were spawned at a predefined Time-to-Arrival (TTA), defined as the time remaining until the vehicle reached the crossing point. We tested TTAs of 6, 10, 14, and 18 s to systematically vary the level of temporal pressure. By fixing the road layout, vehicle speed, and pedestrian trigger point, the design controlled the temporal and spatial structure of each interaction, ensuring that safety-related outcomes such as collision occurrence and abrupt deceleration events could be evaluated consistently. No other road users were present in the environment, allowing us to focus solely on pedestrian-vehicle interaction.

For the COMMOTIONS pedestrian model, which captures inter-individual variability through distinct parameter sets representing different perceptual/ cognitive/ motor characteristics, five such parameter sets were randomly sampled for each TTA condition. Each individual-TTA combination was then simulated four times to account for intra-individual variability, resulting in 20 simulations per TTA condition. In contrast, the CARLA pedestrian model follows a deterministic behaviour, providing a baseline comparison without inter-individual variation. For comparability, since there is some variability in the AV controller behaviour, the number of runs was nonetheless matched across models: 20 simulations were performed for each TTA condition with both COMMOTIONS and CARLA pedestrians. This ensured that differences in the results reflected behavioural model differences rather than unequal sampling effort.

To evaluate each interaction, we used four behavioural metrics. *Collision rate* captured the proportion of episodes ending in a collision. *Gap acceptance rate* measured how often the pedestrian crossed the road before the car. A crossing was counted as 'accepted' if the pedestrian crossed the road before the vehicle reached the conflict point; if the vehicle passed first and the pedestrian crossed afterwards, the trial was counted as 'not accepted'. *Post-encroachment time* (PET), calculated as the time difference between the two agents passing the conflict point (Peesapati et al., 2018). In this study, the conflict point is defined as the geometric intersection of the vehicle's path along the road and the pedestrian's crossing path. *Sudden speed change rate* measured how often the vehicle exhibited at least one abrupt deceleration exceeding 2.5 m/s^2 during a trial, a threshold commonly used in studies of automated driving comfort and safety (Carlowitz et al., 2024). These four metrics were chosen to capture complementary aspects of interaction outcomes: safety (collisions, PET), efficiency (gap acceptance), and comfort (abrupt braking).

5.2.2 Generation of adversarial scenarios via parameter-TTA search

We used the COMMOTIONS pedestrian model together with a parameter-TTA search method to identify individual behavioural profiles and timings that produced critical interactions with low PET. For this experiment, the pedestrian model was paired with the CARLA AV controller rather than the Autoware stack. The CARLA controller can execute large batches of simulations automatically, which is necessary for Bayesian optimisation across hundreds of runs.

For each pedestrian parameter set p, the adversarial search problem was formulated as:

$$\min_{\mathrm{TTA}} \quad \mathrm{PET}(p, \mathrm{TTA}) \quad \text{s.t.} \quad \mathrm{TTA} \in [6, 18] \text{ s}, \tag{5.1}$$

where $\mathrm{PET}(p, \mathrm{TTA})$ denotes the PET from a simulation with pedestrian parameters p and trigger time TTA, with lower values indicating more critical interactions.

In practice, we randomly sampled 100 pedestrian parameter sets from COMMOTIONS, each representing a distinct individual with cognitive, motor, and perceptual characteristics. For each individual p, Bayesian optimisation was performed independently to identify the TTA value within $[6, 18]$ s that minimised PET. Although in this one-dimensional setting a simple grid search over TTA values would also be feasible, we adopted Bayesian optimisation to use an optimisation procedure that remains effective beyond this specific setup. In higher-dimensional

Algorithm 1: Per-individual Bayesian Optimisation of TTA

Input: Set of $N = 100$ pedestrian parameter sets $\mathcal{P} = \{p_1, \ldots, p_N\}$; TTA range $[6, 18]$ s; iteration budget $T = 60$ (with $k = 15$ random initial samples)
Output: Optimal TTA values $\{\text{TTA}_i^*\}_{i=1}^N$
foreach $p_i \in \mathcal{P}$ **do**
  Randomly sample k initial TTA values from $[6, 18]$;
  for $t = 1$ **to** T **do**
    Select next candidate $\text{TTA}_{i,t}$ using Bayesian optimisation;
    Run simulation with pedestrian parameters p_i and $\text{TTA}_{i,t}$;
    Compute PET and update the optimisation model;
  $\text{TTA}_i^* = \arg\min_{\text{TTA}} \text{PET}$ observed for p_i;
return $\{\text{TTA}_i^*\}_{i=1}^N$

scenario spaces, uniform grid search becomes computationally expensive, whereas Bayesian optimisation is well suited to exploring such spaces efficiently. In addition, PET in our simulations is stochastic due to the variability in the COMMOTIONS model, so it is advantageous to use an optimisation method that can handle noisy evaluations.

During each optimisation run, a Gaussian-process surrogate model was initialised and updated iteratively with observed PET outcomes. At each iteration, the acquisition function proposed a candidate TTA, the corresponding simulation was executed using the fixed CARLA AV controller and the given pedestrian parameters, and the resulting PET was recorded. The process was repeated until the iteration budget was exhausted, yielding a worst case TTA^* for that individual. The full procedure is summarised in Algorithm 1. Each optimisation was performed using the `gp_minimize` function from `scikit-optimize`, with 60 iterations and 15 random initial samples.

The resulting set of TTA^* values—one per individual—defined a set of per-individual minima. Cases with PET < 1.5 s were treated as adversarial test sets derived from plausible human behaviour. These adversarial cases formed the basis for the scenario sets used in the subsequent optimisation experiments.

5.2.3 Optimisation of AV control using diverse scenario sets

Building on the setup used for adversarial scenario generation, and using the same CARLA AV controller for consistency, we conducted a second experiment to investigate whether risky but realistic scenarios can inform AV control design by optimising a single AV control parameter: the braking distance. This parameter defines the threshold distance at which the controller ini-

tiates an emergency stop in response to a pedestrian. Smaller values delay braking and increase risk, whereas larger values trigger earlier braking, improving safety but reducing efficiency. The aim was not to fine tune a complete AV stack, but to demonstrate a general optimisation framework in which behaviourally realistic scenarios can be used to identify a braking distance that balances safety and efficiency within a simplified controller setting.

Scenario sets and rationale We optimised the braking distance separately on three complementary scenario sets:

1. **COMMOTIONS low-PET:** The *low-PET subset* from Sec. 5.2.2, i.e., the set of individual-TTA pairs that achieved PET < 1.5 s at the per-individual optimum TTA*. Let its size be N_H. These cases correspond to safety-critical interactions arising from human-like pedestrian behaviour.

2. **COMMOTIONS random:** We drew the same number, N_H, of COMMOTIONS individuals and assigned each a TTA sampled independently from a uniform distribution over $[6, 18]$ s. This controls for behavioural realism *without* adversarial selection, isolating the added value of targeting low-PET situations.

3. **Jaywalker model:** Using the adversarial jaywalker model of Muktadir and Whitehead (2022), we generated kinematically feasible but not necessarily behaviourally plausible behaviours as a rule-based adversarial baseline. The model cycles through six states (*initialising*, *waiting*, *crossing*, *frozen*, *survival*, *finished*) with social-force dynamics within states, enabling tactics such as sudden dash, mid-road "freeze", and "rewind" retreat. The jaywalker model was run using the default *risk* mode, which determines its behavioural style in crossing interactions. To make comparisons fair, we matched the number of scenarios to the High-risk set and sampled the jaywalker TTA uniformly from $[6, 18]$ s. For further implementation details, readers are referred to Muktadir and Whitehead (2022).

Together, these sets allow us to compare optimisation driven by (1) high-risk but human-like interactions, (2) realistic yet untargeted behaviour, and (3) the rule-based adversarial jaywalker model.

Algorithm 2: Bayesian Optimisation of AV Braking Distance

Input: Scenario set $\mathcal{S}$; PET threshold $\tau_{\text{PET}} = 1.5$ s; deceleration limit $a_{\text{th}} = 2.5\,\text{m/s}^2$; search range $d_{\text{brake}} \in [4.0, 25.0]$ m; iteration budget $T = 60$ with $k = 15$ random initial samples

Output: Optimal braking distance d^*_{brake}

Sample k initial d_{brake} values from $[4.0, 25.0]$;

for $t = 1$ **to** T **do**

- Select next candidate d_t using Bayesian optimisation;
- feasible $\leftarrow$ **true**;
- Initialise empty list L;
- **foreach** $(id, \text{TTA}) \in \mathcal{S}$ **do**
  - Run free-flow baseline (no pedestrian) with d_t and measure T^{free};
  - Run interaction with d_t and pedestrian spawned at TTA; measure T^{int}, compute PET and $a^{\max}$;
  - **if** PET $< \tau_{\text{PET}}$ ***or*** $a^{\max} > a_{\text{th}}$ **then**
    - feasible $\leftarrow$ **false**; **break**;
  - $TL \leftarrow \max(0,\, T^{\text{int}} - T^{\text{free}})$; append TL to list L;
- **if** *feasible* **then**
  - objective $\leftarrow$ average(L)
- **else**
  - objective $\leftarrow$ large_penalty;
- Update optimisation model with $(d_t, \text{objective})$;

return $d^*_{\text{brake}} = \arg\min_d$ objective;

AV controller optimisation The optimisation problem was designed to balance safety and efficiency. PET was not maximised directly, because very high PET values do not necessarily indicate safer interactions: they may instead reflect that the AV braked excessively early, reducing efficiency or avoiding interaction altogether. Instead, PET was enforced as a constraint with a threshold of 1.5 s, below which a scenario was deemed unsafe. Likewise, excessive decelerations ($a^{\max} > 2.5$ m/s^2) were treated as violations of passenger comfort. Within these constraints, the objective was to minimise vehicle time lost, defined as the additional travel time in an interaction run compared to a corresponding free-flow run without the pedestrian.

The optimisation problem can be expressed as:

$$\begin{aligned} \min_{d_{\text{brake}}} \quad & \frac{1}{N_{\text{feasible}}} \sum_{i=1}^{N_{\text{feasible}}} T_i \\ \text{s.t.} \quad & \text{PET}_i \geq \tau_{\text{PET}}, \quad \tau_{\text{PET}} = 1.5 \text{ s}, \\ & a_i^{\max} \leq a_{\text{th}}, \quad a_{\text{th}} = 2.5 \text{ m/s}^2, \\ & d_{\text{brake}} \in [4.0, 25.0] \text{ m}, \end{aligned} \tag{5.2}$$

where T_i is the vehicle time lost for scenario i, PET_i is the PET from the interaction run, and $a_i^{\max}$ is the maximum deceleration observed in that run.

The optimisation was run separately for each of the three scenario sets defined above. For every candidate braking distance d_{brake} proposed by the Bayesian optimiser, the AV was evaluated across all scenarios in the set. Each evaluation involved paired simulations: a free-flow baseline to measure T^{free}, and an interaction run with the pedestrian spawned at the given TTA to measure T^{int}, PET, and $a^{\max}$. A candidate was feasible only if all scenarios met the PET and acceleration thresholds; otherwise, it was penalised with a large objective value. This procedure is summarised in Algorithm 2.

The optimisation employed a Bayesian approach using the Gaussian-process-based `gp_minimize` routine from `scikit-optimize`, with a search space for d_{brake} defined as $[4.0, 25.0]$ m. Each optimisation ran for 60 iterations with 15 random initial samples to balance exploration and exploitation.

5.3 Results

5.3.1 Impact of pedestrian models on AV behaviour

From the left panel of Figure 5.2, we can observe that the COMMOTIONS pedestrian model produced no collisions for either the Autoware AV or the CARLA AV across all tested TTA conditions. By contrast, the CARLA pedestrian resulted in multiple collisions, most frequently at intermediate TTAs (10-14 s).

Gap acceptance also differed between the models. For COMMOTIONS, acceptance increased with TTA, starting from about 10% at 6 s and reaching close to 100% at 18 s, a progression consistent with findings from earlier studies of pedestrian crossing behaviour (Oxley et al., 2005; Lobjois and Cavallo, 2007). It is worth noting that the TTA values in the present setup are not directly comparable to those reported in the studies where pedestrians make crossing decisions from the kerb, as pedestrians in our simulations start approximately 2 m from the kerb and therefore require additional time before entering the road. This design choice reflects the fact that real-world interactions often begin before pedestrians reach the kerb (Bandini et al., 2017). In contrast, the CARLA pedestrian showed no clear trend. When paired with the CARLA AV in particular, the CARLA pedestrian's gap acceptance remained high even at short TTAs such

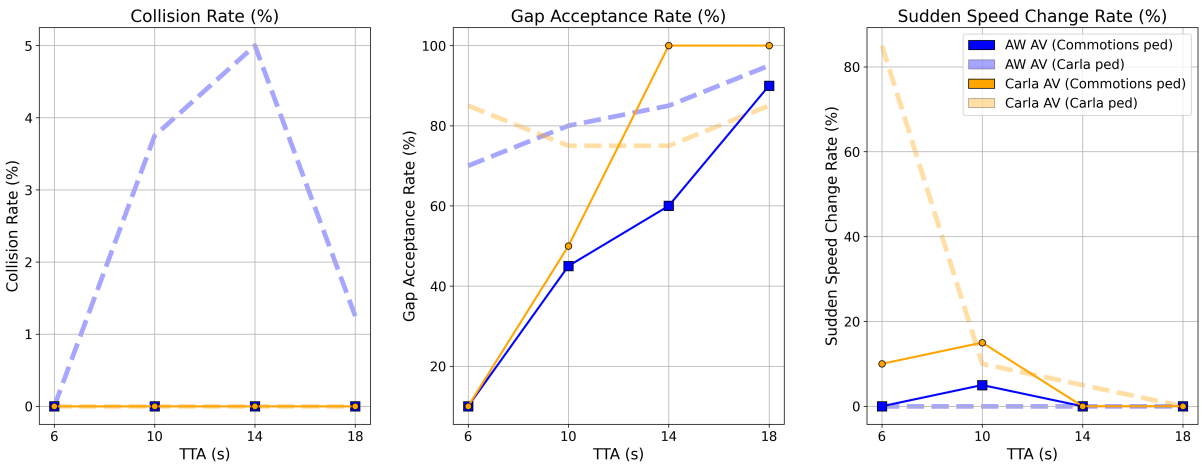

Figure 5.2: Comparison of AV-pedestrian interaction metrics across different pedestrian and AV model combinations at varying Time-to-Arrival (TTA) values. The three panels from left to right show: collision rate, gap acceptance rate, and sudden speed change rate.

as 6 s, indicating behaviour that does not reflect realistic human sensitivity to the time gap (the middle panel of Figure 5.2).

Sudden speed changes, defined as peak decelerations above 2.5 m/s^2, occurred more often with the CARLA pedestrian testing against Carla AV. At TTA = 6 s, these events exceeded 80% of trials with the CARLA AV, while with the COMMOTIONS pedestrian they stayed below 20% across all TTAs and were almost absent with the Autoware AV (the right panel of Figure 5.2).

Figure 5.3 shows the PET across pedestrian-AV combinations and TTA values. For the COMMOTIONS pedestrian, PET values increased with TTA and remained relatively narrow, with almost no samples below the 1.5 s threshold. In contrast, the CARLA pedestrian produced wider variability, most notably at TTA = 14 s, where several PETs dropped below 1.5 s. The mean PET trend lines further illustrate this difference: The COMMOTIONS pedestrian showed a relatively stable increase across TTAs, whereas the CARLA pedestrian exhibited a rise, with much lower mean PET at short TTAs and much higher mean PET at long TTAs.

Figure 5.4 compares the trajectory patterns of pedestrians and vehicles across different combinations of AV controllers and pedestrian models. Each line represents one interaction, with the vehicle's distance to the crossing point on the x-axis and the pedestrian's distance on the y-axis. In this plot, each interaction starts from the upper-right corner. The offset in the starting positions along the y-axis for the CARLA pedestrian model is due to implementation constraints: although the pedestrian was intended to be spawned at the same roadside coordinates as in our setup, the CARLA pedestrian can only be placed on the nearest predefined navigation

waypoint, which does not always coincide exactly with the specified location. In contrast, the COMMOTIONS pedestrian can be spawned at any specified location.

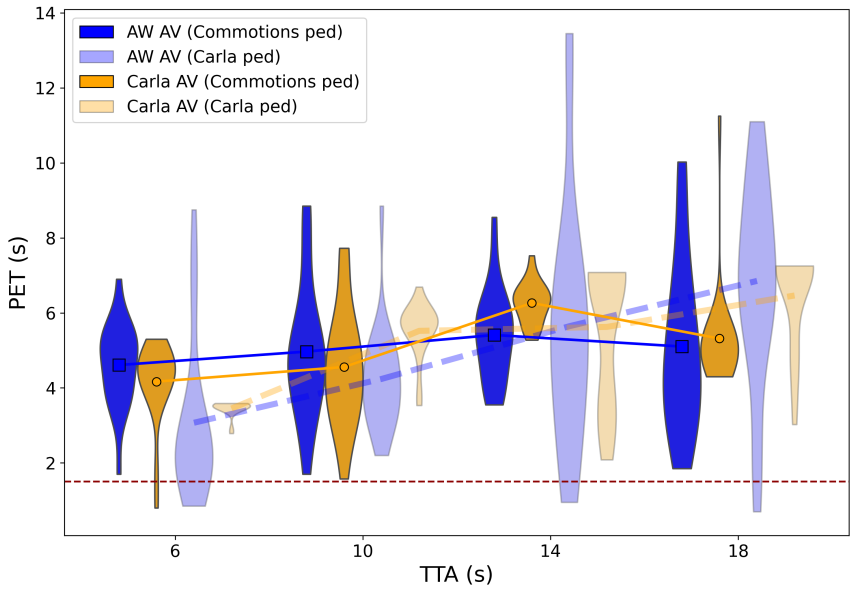

Figure 5.3: Distribution of post-encroachment times (PET) across pedestrian-AV combinations and TTA values. The red dashed line marks the 1.5 s safety-critical threshold.

For both AV controllers, interactions involving COMMOTIONS pedestrians (top panels) show more curved and diverse trajectories, reflecting inter- and intra-individual variability as well as adaptive behaviour and mutual adjustment between the agents. By comparison, trajectories involving the CARLA pedestrian (bottom panels) are mostly straight, indicating limited responsiveness to the approaching vehicle. In addition, the CARLA pedestrian shows more trajectories clustered near the (0, 0) region, suggesting that encounters occurred with shorter temporal and spatial margins. Several horizontal traces around $y = 0$ correspond to the interaction where the CARLA pedestrian became stationary after a collision or a very near miss.

Overall, the above results show that the COMMOTIONS pedestrian generates fewer collisions, smoother AV decelerations, and more consistent PET compared to the CARLA pedestrian, which yields fluctuating gap acceptance and more unsafe outcomes. Together, these results confirm that the COMMOTIONS pedestrian induces richer and more human-like interaction dynamics than the default CARLA pedestrian.

5.3.2 Identification of safety-critical events

As shown in Figure 5.3, although the COMMOTIONS pedestrian generally produced safe interactions, an important property of the model is that it doesn't always: a few interactions

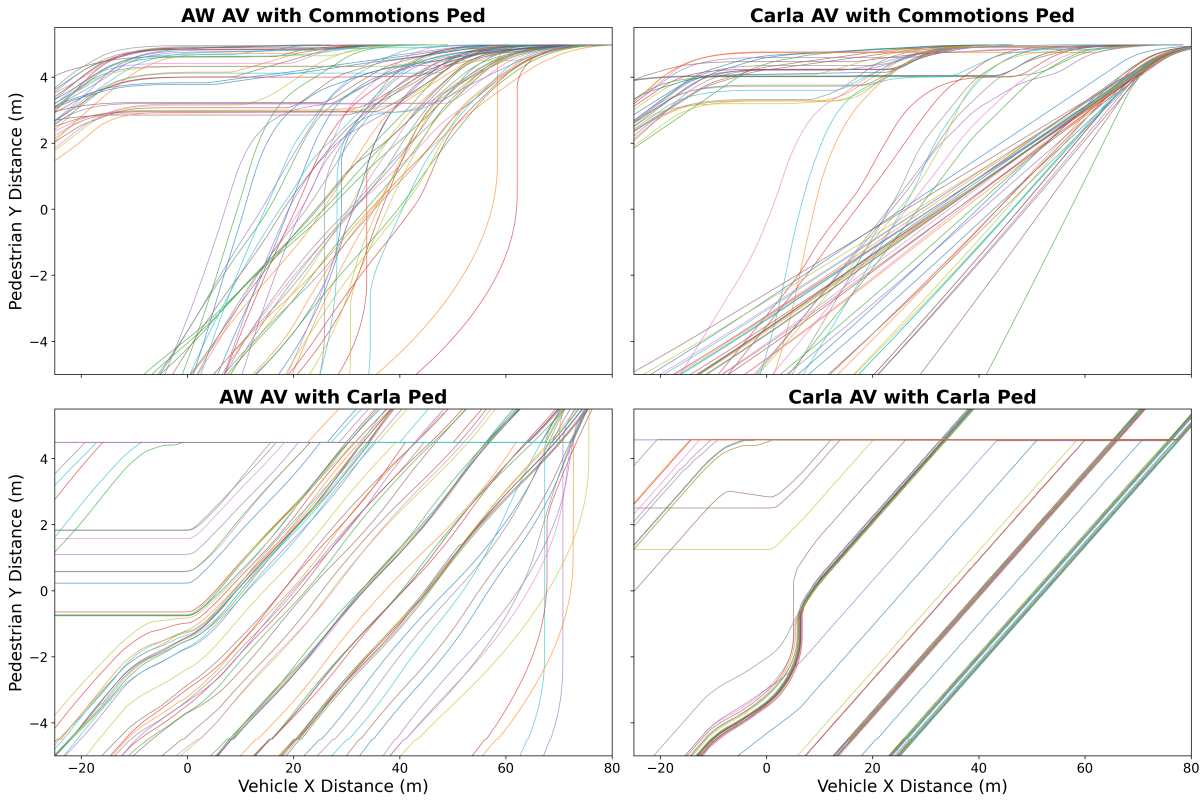

Figure 5.4: Comparison of pedestrian-vehicle trajectory patterns across different combinations of AV controllers and pedestrian models. Each line represents one interaction, plotted by the vehicle's and pedestrian's distances to the crossing point.

exhibited PET below the 1.5 s safety threshold. As described in Sec. 5.2.2, we used the COMMOTIONS pedestrian model in combination with the CARLA AV controller in a parameter-TTA search procedure to generate low-PET scenarios. The resulting outcomes are shown in Figure 5.5.

Figure 5.5 plots PET as a function of TTA for the individuals whose minimum PET fell below 1.5 s. Each line represents one pedestrian individual, with shading indicating the variability across repeated runs arising from variability in the COMMOTIONS model. The plot shows that low PET values typically occurred within the short-TTA range (6-10 s), and it also shows that these low PET values only occurred sometimes for these individuals at these TTAs. In other words, unsafe interactions happen more often for some individuals (inter-individual variability), but not all of the time for any individuals (intra-individual variability).

In Figure 5.5, a clear gap is visible in the 11–15 s TTA range, where individual minima of PET were observed. For TTAs greater than 15 s, some individuals exhibited their minimal PET values at around 4 s. These individuals are characterised by conservative crossing behaviour

and slower walking speeds. Their PET minima occur at large initial TTA values because at these safer TTAs the pedestrians initiate the crossing later and walk more slowly, causing a local minimum in PET.

5.3.3 Improvement in the interaction safety and efficiency through AV Parameter optimisation

Table 5.1 summarises the results across four conditions: before optimisation, and three post-optimisation approaches based on COMMOTION low-PET, COMMOTION random, and jaywalker model. The values in the table show the results of the different AV controller variants interacting with the same 100 individual-TTA combinations of the COMMOTIONS pedestrian as identified in Sec. 5.3.2, ensuring comparability across AV controllers. The rows differ only in the braking distance parameter setting: the default CARLA AV setting for the before-optimisation row, and the optimised values obtained from the three approaches.

When optimisation was guided by low-PET individuals, performance improved across all metrics compared with the before-optimisation baseline. Mean PET almost doubled (4.96 s vs 2.48 s),

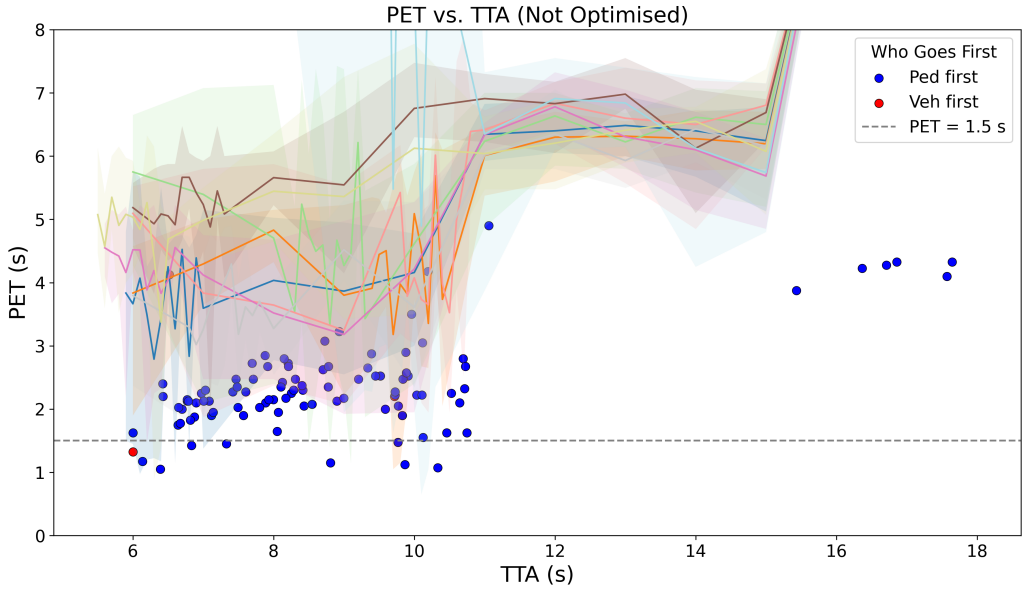

Figure 5.5: Post-encroachment time (PET) versus Time-to-Arrival (TTA) for the 100 individual pedestrians from the COMMOTIONS model under their respective low-PET individual-TTA combinations prior to optimisation. Markers show the minimum PETs for each individual, and marker colour indicates the order of crossing. Each coloured line shows the mean PET across TTAs for one of the eight pedestrians with PET below the safety threshold of 1.5 s (dashed line), and the shaded envelopes denote the range across repeated runs for these pedestrians, to illustrate stochastic intra-individual variability.

the frequency of sudden speed changes fell from 83% to 17%, and average vehicle time lost was reduced from 0.12 s to 0.05 s. The minimum PET also increased from 1.05 s to 1.80 s, i.e., safety-critical interactions no longer occurred. These results show that tuning on human-like safety-critical interactions enabled the AV to maintain larger temporal margins while braking more smoothly and efficiently.

Random-case optimisation also reduced time lost to a similar level (0.04 s) and achieved the lowest sudden speed change rate at 11%. It further raised mean PET to 5.32 s, higher than the low-PET approach, but its minimum PET (1.23 s) still dropped below the safety threshold, indicating that this optimisation approach did not succeed in eliminating safety-critical interactions. Beyond the summary statistics reported in Table 5.1, Figure 5.6 reveals clear differences in the distributional patterns across conditions. In particular, the random approach produces a denser cluster of PET values close to the 1.5 s threshold, whereas the low-PET approach shifts the entire distribution upward, reducing the near-critical interactions. All three optimisation approaches result in noticeably higher PET values than the before-optimisation baseline.

In contrast to the COMMOTIONS-based optimisation results, optimisation against the jaywalker model raised mean PET to 5.52 s but at a clear cost. Sudden speed changes remained

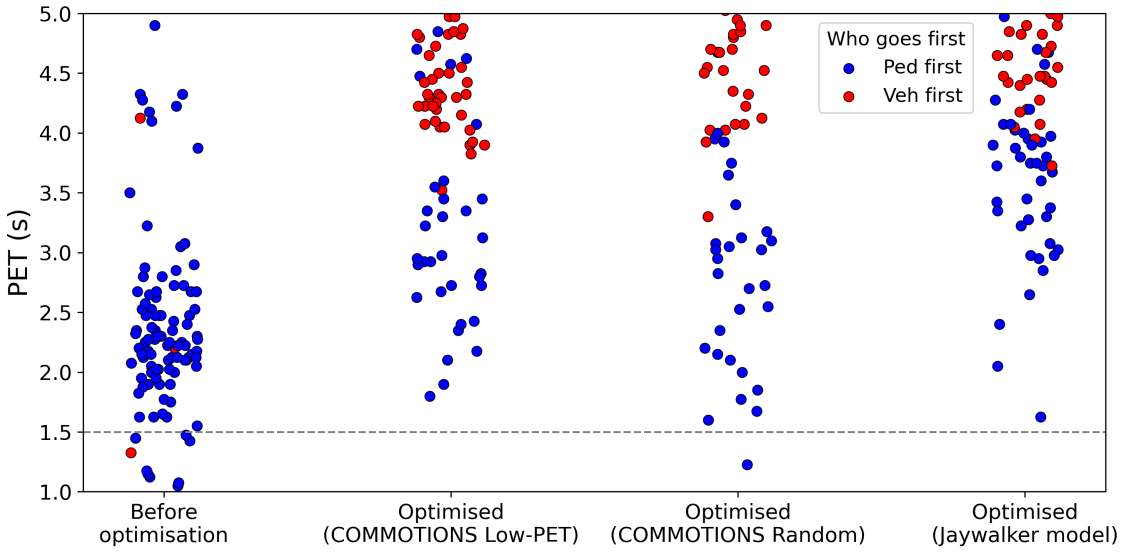

Figure 5.6: Distribution of PET values across four scenario conditions: Before optimisation, Optimised (Low-PET), Optimised (Random), and Optimised (Jaywalker model). Each point represents an individual scenario, with colours indicating whether the pedestrian or the vehicle passed first. The horizontal dashed line marks the PET safety threshold of 1.5 s. The x axis represents different categories only, and the horizontal spread within each category has no quantitative meaning but is used to improve visual separation of points.

Table 5.1: Comparison of AV-pedestrian interaction metrics before and after control parameter optimisation. Optimisation was performed using either low-PET individuals, randomly sampled individuals, or the jaywalker pedestrian model.

Condition	**Mean PET (s)**	**Min PET (s)**	**Sudden speed change rate (%)**	**Average vehicle time lost (s)**
Before optimisation	2.48	1.05	83	0.12
Optimised (COMMOTIONS Low-PET)	4.96	1.80	17	0.05
Optimised (COMMOTIONS Random)	5.32	1.23	11	0.04
Optimised (Jaywalker model)	5.52	1.62	38	0.25

high at 38%, and average vehicle time lost increased to 0.25 s, more than twice the baseline and far above the results of the COMMOTIONS-based optimisation approaches. Although its minimum PET (1.62 s) indicated that all interactions stayed above the 1.5 s safety threshold, these apparent safety gains arose because the jaywalker model's exaggerated, highly adversarial behaviour forced the optimiser to select a much larger braking distance. This resulted in an over-cautious AV that braked excessively early, producing artificially high PET values at the expense of comfort and efficiency.

5.4 Discussion

5.4.1 Influence of human-like pedestrian model on AV testing

The results show that using a human-like pedestrian model in AV simulation has a clear impact on the interaction outcomes. Compared with the rule-based CARLA pedestrian, the COMMOTIONS model exhibited gap acceptance behaviour that increased with TTA in line with observed human crossing patterns. We also found that this was accompanied by more human-like vehicle responses: smoother decelerations, fewer abrupt braking events, and PET values that were more consistent across conditions. These findings demonstrate that replacing simplistic scripted agents with human-like agents produces interactions that better reflect real-world dynamics.

This finding has direct implications for simulation-based AV evaluation. The goal of virtual testing is to expose AVs to situations that resemble those encountered in the real world; using unrealistic pedestrian models risks creating interactions that either exaggerate dangers or overlook real risks, leading to results that do not transfer well to deployment.

5.4.2 Implications for AV testing pipelines

Another contribution of this work is the demonstration that human-like pedestrian models can be used not only to reproduce realistic interactions, but also to generate challenging scenarios that remain behaviourally plausible. The adversarial interactions emerging from the COMMOTIONS model did not require exaggerated or 'suicidal' behaviours; rather, they arose naturally from inter- and intra-individual variability in the decision-making process, as well as from the initial spatial–temporal relationship between the pedestrian and the vehicle. This finding suggests that unsafe interactions are not universal but occur with higher probability for certain pedestrian individuals and under certain timing conditions, highlighting the importance of capturing such variability in human behaviour in AV testing. Yet, these naturally occurring interactions were sufficient to expose limitations in AV control logic, highlighting the potential of such models to support stress testing. Importantly, the role of these scenarios is not to maximise collision occurrence as an end in itself. Rather, they are intended to expose the AV to relatively safety-critical but mainly non-collision interactions, so as to assess how pedestrian behavioural realism affects AV performance and evaluation outcomes, and to support controller tuning under realistic conditions.

The optimisation experiments show that adversarial interactions can also be used to improve AV control performance. Relative to the pre-optimisation controller, braking distance tuning on COMMOTIONS scenarios increased PET, reduced sudden braking, and lowered time lost, demonstrating gains in both safety and efficiency. Compared with optimisation using COMMOTIONS random approach, optimisation using the low-PET approach further raised the minimum PET and reduced the number of near safety-critical interactions near the 1.5 s PET threshold, showing that focusing optimisation on safety-critical situations helps the AV avoid rare but high-risk interactions.

A comparison with the jaywalker model further underscores this point. Jaywalker-based optimisation also increased PET but at the expense of comfort and efficiency, as the AV adopted overly defensive action triggered by unrealistic behaviours. In contrast, optimisation using COMMOTIONS adversarial scenarios achieved safety gains without sacrificing comfort or efficiency, yielding smoother decelerations and lower time lost. This demonstrates that human-like adversarial interactions can support controller tuning that balances safety, comfort, and operational efficiency, an outcome directly relevant to AV development pipelines.

By shifting adversarial testing from pure robustness checks towards optimisation, this approach highlights the potential of human-like adversarial scenarios for simulation-based evaluation and controller tuning.

5.4.3 Limitations and future work

Although the present study provides a controlled and systematic evaluation of pedestrian-AV interactions, several limitations remain that point to avenues for future research. First, the experimental setup was restricted to a simplified zebra crossing with a single pedestrian and a single vehicle. While this scenario provides analytical clarity, it does not capture the complexity of urban traffic environments where multiple agents, occlusions, and contextual cues such as road geometry or traffic signals shape decision-making. Extending the framework to multi-pedestrian, multi-vehicle, or shared-space settings would allow a richer assessment of AV performance under realistic conditions.

Second, the AV side of the interaction was limited to two controllers: the CARLA AV controller and the Autoware autonomous driving stack. Although these provide useful baselines, they do not represent the diversity of state-of-the-art AV architectures, such as deep RL-based controllers (Lopez Pulgarin et al., 2018; Kurzer et al., 2021; Zhan et al., 2023), or commercial proprietary stacks such as including Tesla Autopilot, Mercedes-Benz Drive Pilot, and the Waymo's Driver (Shadab Siddiqui et al., 2025). Future work would therefore test whether the benefits of behaviourally realistic adversarial scenarios generalise across different control algorithms.

Third, the optimisation targeted only a single parameter—braking distance. Real-world AV safety and comfort depend on a broader set of parameters, including acceleration and jerk limits, gap acceptance thresholds, and the interaction between longitudinal and lateral control. Exploring multi-parameter or adaptive optimisation approaches could yield richer insights into the trade-offs between safety, comfort, and efficiency. Moreover, incorporating more advanced optimisation methods could better support multi-parameter optimisation while reducing computational cost.

A further point concerns the scope of the present application study. The aim here is not to reproduce real-world crash rates directly. The pedestrian models used in this chapter were fitted to non-collision interactions, rather than to crash and near-crash cases. Approximating real-world

crash probabilities more closely would likely require larger and more representative naturalistic datasets, including crash and near-crash events, together with calibration procedures designed specifically for these rare events. The same general optimisation framework could also be extended to more complex situations, such as multiple pedestrians or pedestrian groups. However, this would require the virtual road-user model to be adapted to the target context, for example by incorporating mechanisms for social coordination, occlusion, and group-level interaction dynamics, rather than simply transferring the present one-to-one interaction model unchanged.

Finally, although the COMMOTIONS model captures perceptual, motor, and cognitive variability, the work in Chapters 2, 3, and 4 highlights a fundamental limitation of this mechanistic modelling approach. Because behaviour is defined through hand-crafted structure, the model is difficult to extend to higher-dimensional observation spaces, richer action spaces, and more complex interactive settings. It also requires considerable computational resources for parameter sampling and Bayesian optimisation, which may limit scalability to very large test suites. Future work could therefore focus on developing more human-like behavioural models, leveraging parallel simulation infrastructure, and validating against larger naturalistic datasets, in order to strengthen both the robustness and the generalisability of the proposed approach.

5.5 Conclusions

In this study, we investigated how behaviourally realistic pedestrian models can enhance autonomous vehicle (AV) testing and optimisation. We developed a framework that integrates the COMMOTIONS pedestrian model with adversarial scenario generation and AV controller optimisation in the CARLA simulation environment.

Replacing the rule-based CARLA pedestrian with the COMMOTIONS model yielded more naturalistic interactions, including smoother decelerations, fewer abrupt braking events, and gap acceptance patterns consistent with human data. The adversarial scenario search further identified safety-critical yet plausible interactions arising from inter- and intra-individual variability in pedestrian decision processes and timing, without relying on exaggerated or unrealistic behaviours.

Optimising AV control on these human-like adversarial scenarios improved both safety and efficiency, while avoiding the excessive caution induced by artificially adversarial agents. These

results demonstrate that human-like pedestrian models can provide realistic challenges and informative feedback for controller tuning. The proposed framework offers a foundation for scalable and human-centred simulation pipelines, supporting realistic virtual evaluation of AV behaviour across diverse scenarios and control architectures.

References

Bandini, Stefania, Crociani, Luca, Feliciani, Claudio, Gorrini, Andrea, and Vizzari, Giuseppe (2017). "Collision avoidance dynamics among heterogeneous agents: the case of pedestrian/vehicle interactions". In: *Conference of the Italian Association for Artificial Intelligence*. Springer, pp. 44–57.

Blue, Victor J and Adler, Jeffrey L (2001). "Cellular automata microsimulation for modeling bi-directional pedestrian walkways". In: *Transportation Research Part B: Methodological* 35.3, pp. 293–312.

Camara, Fanta, Bellotto, Nicola, Cosar, Serhan, Weber, Florian, Nathanael, Dimitris, Althoff, Matthias, Wu, Jingyuan, Ruenz, Johannes, Dietrich, André, Markkula, Gustav, et al. (2020). "Pedestrian models for autonomous driving part ii: high-level models of human behavior". In: *IEEE Transactions on Intelligent Transportation Systems* 22.9, pp. 5453–5472.

CARLA Team (2026). *Python API reference*. URL: `https://carla.readthedocs.io/en/latest/python_api/` (visited on 03/25/2026).

Carlowitz, Stefanie, Madigan, Ruth, Goodridge, Courtney M, Hilz, Jana, Marberger, Claus, Alt, Philipp, Schulz, Michael, Osswalt, Sabine, Engeln, Arnd, and Merat, Natasha (2024). "Balancing Comfort Deceleration Rate and Environmental Impact at Crosswalks and Intersections in Automated Driving". In: *Available at SSRN 5063808*.

Dang, Meiting, Zhao, Dezong, Wang, Yafei, and Wei, Chongfeng (2025). "Dynamic Game-Theoretical Decision-Making Framework for Vehicle-Pedestrian Interaction With Human Bounded Rationality". In: *IEEE Transactions on Intelligent Transportation Systems*.

Dyro, Robert, Foutter, Matthew, Li, Ruolin, Di Lillo, Luigi, Schmerling, Edward, Zhou, Xilin, and Pavone, Marco (2024). "Realistic extreme behavior generation for improved av testing". In: *arXiv preprint arXiv:2409.10669*.

Hanselmann, Niklas, Renz, Katrin, Chitta, Kashyap, Bhattacharyya, Apratim, and Geiger, Andreas (2022). "King: Generating safety-critical driving scenarios for robust imitation via kinematics gradients". In: *European Conference on Computer Vision*. Springer, pp. 335–352.

Helbing, Dirk and Molnar, Peter (1995). "Social force model for pedestrian dynamics". In: *Physical review E* 51.5, p. 4282.

Huang, WuLing, Wang, Kunfeng, Lv, Yisheng, and Zhu, FengHua (2016). "Autonomous vehicles testing methods review". In: *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, pp. 163–168.

Joisher, Karan, Khan, Suhaib, and Ranadive, Omkar (2019). "Simulation Environment for Development and Testing of Autonomous Learning Agents". In: *2nd International Conference on Advances in Science & Technology (ICAST)*.

Kurzer, Karl, Schörner, Philip, Albers, Alexander, Thomsen, Hauke, Daaboul, Karam, and Zöllner, J Marius (2021). "Generalizing decision making for automated driving with an invariant environment representation using deep reinforcement learning". In: *2021 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, pp. 994–1000.

Lobjois, Régis and Cavallo, Viola (2007). "Age-related differences in street-crossing decisions: The effects of vehicle speed and time constraints on gap selection in an estimation task". In: *Accident analysis & prevention* 39.5, pp. 934–943.

Lopez Pulgarin, Erwin Jose, Irmak, Tugrul, Paul, Joel Variath, Meekul, Arisara, Herrmann, Guido, and Leonards, Ute (2018). "Comparing model-based and data-driven controllers for an autonomous vehicle task". In: *Annual Conference Towards Autonomous Robotic Systems*. Springer, pp. 170–182.

Markkula, Gustav, Lin, Yi-Shin, Srinivasan, Aravinda Ramakrishnan, Billington, Jac, Leonetti, Matteo, Kalantari, Amir Hossein, Yang, Yue, Lee, Yee Mun, Madigan, Ruth, and Merat,

Natasha (2023). “Explaining human interactions on the road by large-scale integration of computational psychological theory”. In: *PNAS nexus* 2.6, pgad163.

Markkula, Gustav, Romano, Richard, Madigan, Ruth, Fox, Charles W, Giles, Oscar T, and Merat, Natasha (2018). “Models of human decision-making as tools for estimating and optimizing impacts of vehicle automation”. In: *Transportation research record* 2672.37, pp. 153–163.

Mirzabagheri, Alireza, Ahmadi, Majid, Zhang, Ning, Alirezaee, Reza, Mozaffari, Saeed, and Alirezaee, Shahpour (2025). “Navigating Uncertainty: Advanced Techniques in Pedestrian Intention Prediction for Autonomous Vehicles—A Comprehensive Review”. In: *Vehicles* 7.2, p. 57.

Muktadir, Golam Md and Whitehead, Jim (2022). “Adversarial jaywalker modeling for simulation-based testing of autonomous vehicle systems”. In: *2022 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, pp. 1697–1702.

Muralidhar, P, Rani, C, Rajesh Kumar, M, et al. (2023). “Accident prevention for autonomous vehicle”. In: *2023 2nd International Conference on Vision Towards Emerging Trends in Communication and Networking Technologies (ViTECoN)*. IEEE, pp. 1–5.

Okazaki, Shigeyuki (1979). “A study of pedestrian movement in architectural space, part 1: Pedestrian movement by the application on of magnetic models”. In: *Trans. AIJ* 283, pp. 111–119.

Oxley, Jennifer A, Ihsen, Elfriede, Fildes, Brian N, Charlton, Judith L, and Day, Ross H (2005). “Crossing roads safely: an experimental study of age differences in gap selection by pedestrians”. In: *Accident Analysis & Prevention* 37.5, pp. 962–971.

Peesapati, Lakshmi N, Hunter, Michael P, and Rodgers, Michael O (2018). “Can post encroachment time substitute intersection characteristics in crash prediction models?” In: *Journal of safety research* 66, pp. 205–211.

Pekkanen, Jami, Giles, Oscar Terence, Lee, Yee Mun, Madigan, Ruth, Daimon, Tatsuru, Merat, Natasha, and Markkula, Gustav (2022). “Variable-drift diffusion models of pedestrian road-crossing decisions”. In: *Computational Brain & Behavior* 5.1, pp. 60–80.

Priisalu, Maria, Paduraru, Ciprian, and Smichisescu, Cristian (2023). "Varied realistic autonomous vehicle collision scenario generation". In: *Scandinavian Conference on Image Analysis*. Springer, pp. 354–372.

Rasouli, Amir and Tsotsos, John K (2019). "Autonomous vehicles that interact with pedestrians: A survey of theory and practice". In: *IEEE transactions on intelligent transportation systems* 21.3, pp. 900–918.

Reuters (2025). *Baidu's Apollo Go Robotaxi Service Expands Across Multiple Chinese Cities*. `https://www.reuters.com/business/autos-transportation/uber-partners-with-chinas-baidu-deploy-self-driving-taxis-international-markets-2025-07-15/`. Accessed: 2025-09-28.

Sawitzky, Tamara von, Grauschopf, Thomas, and Riener, Andreas (2023). "A flexible simulation environment for enhanced VRU research". In: *Adjunct Proceedings of the 15th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, pp. 283–286.

Shadab Siddiqui, MA, Rabbi, MS, Islam, Md Jobayer, and Ahmed, Radif Uddin (2025). "Comparison of Different Controller Architectures for Autonomous Driving and Recommendations for Robust and Safe Implementations". In: *Journal of Advanced Transportation* 2025.1, p. 9995539.

Song, Qunying, Tan, Kaige, Runeson, Per, and Persson, Stefan (2023). "Critical scenario identification for realistic testing of autonomous driving systems". In: *Software Quality Journal* 31.2, pp. 441–469.

Tian, Kai, Wei, Chongfeng, Lyu, Wei, Wang, Yueyang, Lee, Yee Mun, Merat, Natasha, Romano, Richard, and Markkula, Gustav (2025). "Interacting With Yielding Vehicles: A Perceptually Plausible Model for Pedestrian Road Crossing Decisions". In: *IEEE Transactions on Intelligent Transportation Systems*.

Wang, Yueyang, Srinivasan, Aravinda Ramakrishnan, Jokinen, Jussi PP, Oulasvirta, Antti, and Markkula, Gustav (2025a). "Pedestrian crossing decisions can be explained by bounded

optimal decision-making under noisy visual perception". In: *Transportation Research Part C: Emerging Technologies* 171, p. 104963.

Wang, Yueyang, Srinivasan, Aravinda Ramakrishnan, Lee, Yee Mun, and Markkula, Gustav (2025b). "Modeling Pedestrian Crossing Behavior: A Reinforcement Learning Approach With Sensory Motor Constraints". In: *IEEE Transactions on Intelligent Transportation Systems*.

Wang, Ziyu, Ma, Jing, and Lai, Edmund MK (2024). "A Survey of Scenario Generation for Automated Vehicle Testing and Validation". In: *Future Internet* 16.12, p. 480.

WebProNews (2025). *Waymo Launches Business Robotaxi Service With AI Analytics in Key Cities*. `https://www.webpronews.com/waymo-launches-business-robotaxi-service-with-ai-analytics-in-key-cities/`. Accessed: 2025-09-28.

World Health Organization (2013). *More than 270,000 pedestrians killed on roads each year*. `https://www.who.int/news/item/02-05-2013-more-than-270-000-pedestrians-killed-on-roads-each-year`. Accessed: August 5, 2025.

Yang, Yuhang, Kujanpää, Kalle, Babadi, I Amin, Pajarinen, Joni, and Ilin, Alexander (2023). "Suicidal pedestrian: Generation of safety-critical scenarios for autonomous vehicles". In: *2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, pp. 1983–1988.

Zhan, Guojian, Lyu, Yao, Li, Shengbo Eben, Jiang, Yuxuan, Zhang, Xiangteng, and Tao, Letian (2023). "Enhance Generality by Model-based Reinforcement Learning and Domain Randomization". In: *2023 7th CAA International Conference on Vehicular Control and Intelligence (CVCI)*. IEEE, pp. 1–6.

Chapter 6

Discussion and conclusions

6.1 Overview of the Main Findings

This thesis advances the modelling of pedestrian-vehicle interaction by integrating reinforcement learning with theories of human perception and motor control. The work builds progressively: first by incorporating perceptual and motor constraints, then by extending to multi-agent interaction, and finally by applying the pedestrian model derived from the COMMOTIONS framework to virtual testing of AVs. Across four studies (Chapters 2-5), the thesis addresses three research gaps identified in Chapter 1.3, illustrating the progression from theoretical development to practical application in AV evaluation.

Addressing Gap 1: Integrating mechanistic and machine learning approaches in the modelling of pedestrian-vehicle interaction

Research on pedestrian-vehicle interaction has largely developed along two independent methodological lines. Mechanistic models, including social force models (Helbing and Molnar, 1995; Yang et al., 2020), perception-based decision models (Tian et al., 2022; Pekkanen et al., 2022), and game-theoretic approaches (Fox et al., 2018), provide clear links to psychological and cognitive mechanisms. However, these models rely on hand-crafted behavioural rules and struggle to scale to continuous and high-dimensional control settings (Markkula et al., 2023).

Data-driven ML approaches, such as CNN, LSTM, and Transformer models (Deo and Trivedi, 2018; Yi et al., 2016; Alahi et al., 2016; Rasouli et al., 2019; Yang et al., 2025; Divya et al., 2025), stand in methodological contrast to mechanistic modelling. They can capture complex

correlations in trajectory data but depend heavily on large and representative training datasets, and they often generalise poorly when data are sparse or when scenarios deviate from the training distribution (Diaz-Ruiz et al., 2022; Roshdi et al., 2024).

This thesis addresses this methodological gap by embedding mechanistic principles directly within an RL environment. Although RL is itself a machine learning method, it is not used in this thesis as a purely data-driven predictive model. Instead, it serves as an optimisation method within a computational rationality framework. Specifically, the model is defined through assumptions about task structure, rewards, and human perceptual and motor constraints, and RL is then used to derive the boundedly optimal policy implied by these assumptions. Rather than specifying behavioural rules directly, this formulation allows behaviour to emerge from optimisation under human constraints. It therefore retains the theoretical grounding of cognitive models while gaining the scalability associated with ML approaches.

The resulting hybrid framework supports capabilities that neither approach achieves on its own. As shown in Chapter 4, it produces continuous, multidimensional human-like locomotion that is difficult to obtain from mechanistic models, while performing well in data-limited conditions that challenge data-driven ML approaches. This establishes a scalable framework for generating human-like behaviour without relying on inflexible rules or extensive training data.

Addressing Gap 2: Understanding pedestrian-vehicle interaction from the perspective of boundedly optimal behaviour under human constraints

This thesis advances the understanding of pedestrian-vehicle interaction by showing that many observed behaviours can be explained as boundedly optimal responses to human perceptual and motor constraints. Previous logit-based choice models and other cognitive models focus primarily on visual cues (Pekkanen et al., 2022; Tian et al., 2022), but overlook how the motor system and its effort costs influence behaviour. Yet human behaviour arise from both perception and action, and modelling only one of these mechanisms cannot capture their combined influence.

In parallel, existing social force and game-theoretic models describe interaction outcomes but rely on simplified assumptions about physical dynamics or assume perfect knowledge of the environment (Helbing and Molnar, 1995; Fox et al., 2018; Dang et al., 2025). Because road users in the real world act under uncertainty, these simplified or unrealistic assumptions prevent such models from explaining why road users behave as they do in dynamic and uncertain

environments.

Chapters 2-4 show that these behaviours can be understood instead as rational adaptations to perceptual and motor constraints. Chapter 2 demonstrates that speed-dependent gap acceptance emerges from increasing visual uncertainty at higher vehicle speeds, which was previously interpreted as the biases in the decision-making process (Petzoldt, 2014; Tian et al., 2022). Chapter 3 shows that adjustments in walking speed arise from an optimal trade-off between time pressure and biomechanical effort. Chapter 4 extends this perspective to interactive settings, illustrating that behaviours such as yielding and the spatio-temporal pattern of the interaction emerge from joint optimisation under perceptual and motor constraints of both agents.

Taken together, the results demonstrate that diverse behaviours in pedestrian crossing decisions, pedestrian locomotion, and pedestrian-vehicle interaction can be explained within a boundedly optimal framework.

Addressing Gap 3: Using human-like behavioural models for realistic and adversarial virtual testing of autonomous vehicles

Finally, this thesis bridges the gap between behavioural modelling and AV validation by integrating the COMMOTIONS pedestrian model into a closed-loop simulation environment with a full autonomous driving stack. Most existing testing pipelines rely on either simplistic rule-based agents (Papadimitriou et al., 2016; Rashid et al., 2024) or purely adversarial agents that prioritise collision generation over realism, such as the 'suicidal pedestrian' approach (Yang et al., 2023). Behavioural models developed for realism have rarely been evaluated against full-stack AV algorithms, leaving their practical relevance for AV testing unexplored.

Our application study (Chapter 5) shows that cognitively grounded pedestrian models produce more human-like interactions with the AV stack compared with simple rule-based pedestrian models. Moreover, by varying the parameters that govern human characteristics, the same framework can generate scenarios that are challenging yet behaviourally plausible, avoiding the unrealistic behaviour often seen in optimisation-based adversarial attacks (Dyro et al., 2024). Integrating the pedestrian model from COMMOTION framework into a closed-loop, full-stack AV simulation demonstrates how human-like modelling can support realistic and informative evaluation and optimisation of AV control. It enables testing conditions that better capture real-world safety challenges.

6.2 Contributions

6.2.1 Integrating perceptual and motor constraints with reinforcement learning

The first contribution of this thesis is the establishment of a hybrid modelling framework that combines the mechanistic structure of perceptual and motor constraints with the flexibility of reinforcement learning. As discussed in the previous section regarding addressing the research gaps, this formulation integrates visual uncertainty and biomechanical effort directly into the RL environment, allowing behaviour to emerge from boundedly optimal action rather than from hand-crafted rules or learning from large datasets.

This framework advances the modelling of pedestrian behaviour in several ways. Compared with earlier mechanistic models, which typically focus on discrete decisions such as gap acceptance (Tian et al., 2022), the proposed approach captures continuous locomotion shaped jointly by perception and motor execution. Compared with the more advanced mechanistic model, COMMOTIONS model, which was limited to one-dimensional movement (Markkula et al., 2023), this thesis modelled full two-dimensional walking trajectories and dynamic fixation control. By enabling richer movement patterns, these additions provide a stronger basis for modelling realistic road user behaviour in naturalistic settings.

Compared with data-driven ML approaches, the framework does not depend on large datasets and generalises more reliably in data-limited settings. As discussed earlier, behaviour here is generated through optimisation under perceptual and motor constraints rather than by replicating observed patterns, which makes the model suitable for exploring rare or safety-critical scenarios. At the same time, because behaviour emerges from explicit mechanisms, the framework can account for empirically observed patterns such as speed-dependent gap acceptance, hesitation, and adjustments in walking speed as rational adaptations to sensory uncertainty and motor effort (Wang et al., 2025; Wang et al., 2023). This provides a causal explanation that is not available in black-box learning models.

While many behavioural models are developed for online real-time prediction of surrounding road users, the models in this thesis are primarily intended for offline use as simulation agents in virtual testing environments. In online prediction, interpretability supports the diagnosis of model failures in safety-critical decision-making (Doshi-Velez and Kim, 2017; Rudin, 2019).

In offline simulation, interpretability also remains important for different reasons: transparent mechanisms allow researchers and developers to verify that simulated behaviour reflects plausible human perceptual and motor processes, to understand why specific risky interactions arise, and to tune parameters of agents to create realistic diverse scenarios (Roshdi et al., 2024). The framework developed in this thesis therefore provides not only behavioural realism but also mechanistic transparency, which enhances its suitability for simulation-based evaluation of AV.

Finally, this thesis also contributes a modular and generalisable modelling tool. Chapter 2 introduces the noisy perception system, and Chapter 3 adds a motor-effort model while retaining the same perceptual module. This incremental extension demonstrates that the framework can be expanded to capture additional behavioural mechanisms in a principled way. In addition, because the perceptual and motor components are implemented in a parametrised form, the framework can be adapted readily to different road user classes. Chapter 4 illustrates this by applying a similar structure to driver agents, enabling heterogeneous traffic simulations while maintaining consistency in the underlying behavioural principles.

6.2.2 Extending computational rationality to interactive settings

A second contribution of this thesis is the extension of computational rationality from individual behaviour to interactive pedestrian-vehicle interactions. Existing MARL approaches can reproduce adaptive traffic patterns (Schmidt et al., 2022; Konstantinidis et al., 2023), but they typically assume perfect state information and do not represent the perceptual or motor constraints that shape human behaviour. Conversely, computational rationality has been applied in human-computer interaction (Chen et al., 2015; Jokinen et al., 2020), but these models are restricted to single agents operating in static or non-interactive environments.

This thesis brings these two lines of research together by formulating pedestrian-vehicle interaction as a coupled optimisation problem in which both agents act under perceptual and motor constraints. Interaction patterns arise from this optimisation process, rather than from predefined rules or behavioural scripts. The resulting model reproduces features observed in naturalistic interaction, such as hesitation and the continuous adjustments pedestrians and drivers make in response to one another.

Taken together, these findings show that the framework of computational rationality provide a basis not only for modelling individual behaviour but also for capturing the adaptive dynamics

underlying pedestrian-vehicle interaction. To the best of our knowledge, this thesis presents the first multi-agent model grounded in computational rationality, demonstrating that this framework can be extended to interactive settings in which both agents learn to co-adapt under human-like perceptual and motor constraints.

The progression from Chapter 2 to Chapter 3 and Chapter 4 also highlights how RL algorithm choice depended on the structure of the modelling task. In Chapter 2, the task was relatively simple, involving a binary crossing decision of whether to cross before or after the vehicle passed. In that setting, the optimisation problem and action space were both limited, so RL algorithm choice was less central to the modelling contribution. By Chapter 3, however, the optimisation problem had become richer, and by Chapter 4 it had become substantially more complex, making the choice of RL algorithm more important for both optimisation and the resulting behaviour.

In Chapter 3, only the pedestrian policy was learned, while the vehicle followed predefined kinematic patterns and did not adapt to the pedestrian during the interaction. The pedestrian action space was also relatively simple, consisting of discrete speed choices executed at walking step intervals. In this context, PPO provided a suitable balance between optimisation stability, implementation simplicity, and sufficient policy flexibility to learn boundedly optimal crossing behaviour under different sensory and motor constraints. More complex alternatives, such as off-policy methods designed for richer continuous control, were not necessary for this setting and would have introduced additional algorithmic complexity without clear benefit for the modelling aims of that chapter.

In Chapter 4, by contrast, both pedestrian and vehicle were learning agents, the interaction was fully adaptive, and the action spaces were continuous. This created a more complex optimisation problem in which informative events could be relatively sparse and behavioural coordination had to emerge through repeated interaction. In this context, SAC offered two practical advantages. First, its off-policy formulation allowed past experience to be reused through a replay buffer, improving sample efficiency in a multi-agent interaction setting. Second, SAC learns a stochastic policy while explicitly encouraging exploration through entropy regularisation, which reduced the tendency to converge to overly rigid behavioural strategies. In preliminary experiments, PPO often led the agents to select actions near the boundaries of the action space in this setting, suggesting poorer optimisation behaviour and producing less behaviourally plausible

motion patterns. SAC was therefore adopted as a more suitable algorithm for capturing nuanced interaction patterns such as yielding, gradual speed adjustment, and behavioural variability across repeated interactions.

6.2.3 Improving autonomous vehicle evaluation and optimisation

A practical contribution of this thesis lies in the integration of human-like models into the evaluation and optimisation of AV control algorithm. Most existing adversarial generation methods prioritise danger over realism, often producing 'suicidal' agents that offer limited insight for real-world deployment (Yang et al., 2023; Dyro et al., 2024). In contrast, by embedding our pedestrian model into a closed-loop simulation with the full Autoware stack (Chapter 5), we generated scenarios that are challenging yet behaviourally plausible. The model exposed safety-critical interactions, such as low-PET near misses arising from human-like behavioural mechanisms rather than artificial aggression.

Furthermore, the thesis introduced a novel approach to AV controller tuning by framing evaluation as an optimisation problem within these realistic scenarios. Using the pedestrian model from the COMMOTIONS framework as a basis for adversarial generation, AV control algorithms were adjusted, in this case the braking distance parameter, to achieve a balance between safety and efficiency. This demonstrates that incorporating human-like behavioural realism can make AV testing more representative of real-world interactions while directly supporting the improvement of AV control performance.

6.2.4 Contributing to human-centred autonomous system design

Finally, the modelling approach developed in this thesis extend beyond road user interaction to the broader field of human-robot interaction (HRI). By modelling agents as computationally rational optimisers operating under perceptual and motor constraints, the framework can simulate or predict human adaptation across collaborative and competitive settings. This ability to anticipate human responses provides a foundation for using computational models to design autonomous systems that must coordinate their behaviour with human partners.

The developed framework illustrates that autonomous systems can benefit from learning not only what humans do, but also why they act in particular ways given their perceptual and motor limitations. Embedding such mechanisms into control policies can promote smoother

coordination, clearer communication of intent, and reduced behavioural uncertainty in shared environments (Ho and Griffiths, 2022; Fuchs et al., 2023; Gomaa and Feld, 2023). Ultimately, this line of research supports the development of autonomous systems that not only perform safely and efficiently, but also interact with humans in ways that are perceptually intuitive, behaviourally predictable, and socially appropriate.

6.3 Limitations

The findings of this thesis should be interpreted in light of several limitations related to the modelling assumptions, the representation of behavioural variation, the empirical scope of the studies, and the application context used for AV evaluation. While the thesis demonstrates that RL combined with perceptual and motor constraints can generate human-like behaviour in the scenarios studied, these results do not imply that the resulting models provide a complete account of pedestrian behaviour more generally. Computational rationality provides a useful framework for modelling behaviour in tasks that involve sequential decision-making under uncertainty and action constraints, but it is most naturally suited to behaviours that can be approximated as goal directed adaptation under such constraints. The real-world pedestrian behaviours may also reflect emotional responses, adherence to social conventions, or simple heuristic decision strategies, which may require richer or alternative modelling assumptions than those included in the present optimisation-based formulation. Computational rationality should therefore be understood here as a useful and informative modelling framework for the present research questions, rather than as a complete theory of pedestrian behaviour.

This limitation is also evident when computational rationality is considered alongside alternative modelling approaches. Heuristic and cue-based models can provide simpler descriptions of specific decisions, particularly when behaviour is strongly tied to a small number of perceptual variables (Fajen and Warren, 2003; Tian et al., 2022; Tian et al., 2023). Evidence accumulation models may offer more direct accounts of discrete timing decisions (Ratcliff et al., 2016; Giles et al., 2019; Pekkanen et al., 2022), while game theoretic models are well suited to representing strategic anticipation between agents (Fox et al., 2018; Camara et al., 2021; Kalantari et al., 2022; Dang et al., 2025). The contribution of the present thesis is therefore not to show that computational rationality replaces these alternatives, but to demonstrate that it offers a useful framework for integrating adaptive behaviour, perceptual uncertainty, and motor constraints

within a single learning-based formulation. Future work may benefit from hybrid approaches that combine this framework with heuristic components, such as structured priors, or action biases, to better capture behaviours that are not well described by bounded optimisation alone.

A second limitation of this work is the restricted treatment of heterogeneity. Although the models include variation through fitted non-policy parameters and stochasticity in perception and action, broader sources of individual difference are not explicitly represented. Factors such as age, familiarity with traffic environments, social motivations, cultural norms, and trust in vehicle behaviour may affect pedestrian decisions in practice. Similarly, the thesis does not address group behaviour or interactions shaped by the presence of multiple pedestrians. As a result, the behavioural diversity captured by the present models remains limited relative to the full heterogeneity observed in real populations.

A third limitation concerns the scope of the empirical and simulated contexts considered. The thesis focuses primarily on simplified pedestrian-vehicle interaction tasks, especially one-to-one interaction in relatively constrained traffic scenarios. Real-world environments involve richer scene structure, multiple competing goals, communication cues, occlusions, social context, and more varied traffic configurations. In addition, the experimental and trajectory data used for validation were obtained in relatively limited and task-specific settings. The results should therefore be interpreted as context-specific evidence that the proposed framework can reproduce important aspects of pedestrian-vehicle interaction, rather than as evidence of general validity across all pedestrian behaviours or traffic environments.

A fourth limitation concerns the calibration and validation approach used across the thesis. In several parts of the thesis, model calibration and evaluation were conducted within relatively specific experimental and simulation contexts, without relying on large independent training and test datasets or systematic cross-context out-of-sample validation of the kind often used in predictive ML approaches. This raises the possibility that some aspects of the models are partly over-fitted to the particular tasks, behavioural patterns, and environmental conditions represented in the available datasets. In Chapter 4, an additional three-fold training and validation analysis yielded the same relative ranking of model performance as the main analysis, which provides some reassurance that the superior performance of the VMC model was not simply a consequence of additional flexibility or a data-limited evaluation procedure. However, this does not fully resolve the broader question of generalisability. The results support the usefulness of

the proposed framework in the contexts studied here, but they do not by themselves establish transferability to other pedestrian-vehicle interaction settings. This point is also illustrated by the collision rates reported in Chapter 3. The overall collision rate observed in the experiment was 0.002, meaning that collisions were relatively uncommon within that specific experimental setting. However, this should not be interpreted as reflecting the rarity of real-world pedestrian crashes, which are far less frequent and arise under a much broader range of behavioural and environmental conditions. The collision rates of the SM, S, and M models, at 0.027, 0.044, and 0.040 respectively, therefore represent substantial increases relative to the experimental data rather than small deviations. These differences likely reflect the simplification of the present study and the stochasticity of the RL optimisation process, rather than any realistic representation of real-world crash occurrence. Applying the models in other contexts would likely require further calibration and validation using data from different road layouts, traffic conditions, pedestrian populations, and interaction types. In some cases, adaptation may also require modification of the task formulation, reward structure, observation design, or non-policy parameter ranges so that the model reflects the demands of the new setting.

A final limitation relates to the AV testing application in Chapter 5. The use of a human-like pedestrian model in closed-loop simulation demonstrates the potential value of behaviourally realistic agents for AV controller evaluation and optimisation, but this application remains a proof of concept. The study does not aim to reproduce real-world crash frequencies, nor does it establish that the framework provides a complete basis for safety validation across broader deployment conditions, traffic scenarios, or AV systems. The pedestrian model used in this part of the thesis was calibrated to human data but predominantly non-collision interactions, rather than to crash-involved behaviour. Consequently, the findings from Chapter 5 should be understood as showing that more realistic behavioural models can change the character of AV evaluation in ways that are likely to be more behaviourally realistic, rather than as demonstrating direct correspondence to real-world safety outcomes.

6.4 Future work

Building on the limitations discussed above, future work should expand the modelling scope and realism, strengthen empirical grounding, and tighten the link between these models and AV evaluation.

A first direction is to extend the current framework to richer and more complex interaction contexts. A related methodological question concerns the role of RL algorithm choice in these more complex settings. As Chapters 3 and 4 suggest, different optimisation algorithms may differ not only in training efficiency and stability, but also in the qualitative properties of the learned behaviour. RL algorithm choice should therefore not be treated as a purely technical detail, since it may influence which behavioural patterns are more easily learned and how realistically they are expressed in interaction. Future work should examine more systematically how algorithm choice affects behavioural realism, especially in interactive settings with continuous control and multiple learning agents. The environments studied in this thesis were limited to one-to-one pedestrian-vehicle interactions for simplicity. Scaling these models to more pedestrians and vehicles settings such as intersections, shared spaces, and dense pedestrian-vehicle environments would allow investigation of emergent coordination patterns, and implicit social conventions. Achieving this would require more than simply adding additional agents to the simulation. The perceptual models would need to accommodate increased surrounding road users. And because of the increased computational complexity, a more advanced algorithm and more computing power are required. Such extensions could bridge the gap between individual-level interaction modelling and macroscopic traffic behaviour, offering insight into how human decision-making mechanisms scale up to shape collective flow dynamics.

Another direction is to enrich the cognitive and perceptual representation of the agents. While the current framework already incorporates gaze-based perception, it treats gaze as a fixed sensory channel rather than a dynamically managed resource. Future work could introduce a more explicit model of attention, drawing on theories of selective attention and information sampling. For instance, representing how humans prioritise visual cues according to task relevance, perceived risk, or expected information gain would allow the model to capture how perceptual focus shifts in response to changing traffic conditions. Incorporating such mechanisms would make the model toward a more complete account of human information processing in interactive environments.

A third priority is empirical grounding and systematic validation. Although the models were validated against the real-world trajectory data, future research should leverage larger and more diverse datasets that combine motion capture and eye-tracking to capture both movement and visual attention. Such multimodal data would provide a more detailed basis for modelling hu-

man locomotion, enabling more detailed validation of how pedestrians move, look, and adapt in different conditions. With more realistic locomotion dynamics, potentially incorporating skeletal representations, simulated pedestrians could exhibit natural posture and motion patterns that convey decision intent, making their behaviour in simulation closer to that of real humans. Another related direction is to benchmark the framework against state-of-the-art supervised learning models. An open and important question is how expressive cognitive or RL-based models need to become, such as how many non-policy parameters or mechanisms are required, in order to approach or match the realism achieved by high-capacity supervised learning approaches.

Finally, future work could examine whether the same computational principles can be extended beyond pedestrian-vehicle interaction to broader forms of spatial and mobility behaviour. In such settings, the agent would still be modelled as pursuing a goal, such as reaching a destination or completing an activity, but the relevant constraints would extend beyond perceptual uncertainty and motor control in a crossing task. Individual characteristics such as income, health, physical ability, time pressure, or risk preference could instead be represented as constraints. Within this formulation, behaviour would emerge from how different individuals adapt their movement and choice strategies in order to achieve mobility goals under their own constraints. This could make the framework relevant not only to pedestrian crossing, but also to problems such as route choice, navigation, and other forms of human spatial decision-making.

6.5 Conclusion

This thesis has developed a computational rationality framework for modelling pedestrian-vehicle interaction, where human behaviour is represented as boundedly optimal decision-making under perceptual and motor constraints. Beyond capturing the interaction patterns, the framework also explains how adaptive and interactive crossing behaviour can emerge from agents optimising within human constraints.

The first three studies formalised how perceptual uncertainty, effort, and motor control shape the pedestrian-vehicle interaction, progressing from single-agent to multi-agent settings. The framework successfully reproduced a range of human phenomena, including gap acceptance, walking-speed adjustments, and yielding behaviour, demonstrating that realistic human-vehicle interactions can emerge from agents optimising within perceptual and motor constraints rather

than from scripted rules. These findings establish bounded optimality as a coherent principle connecting perception, action, and interaction in traffic behaviour.

Beyond modelling, the thesis also demonstrated the practical value of human-like behavioural models for AV testing. Integrating a behaviourally grounded pedestrian model within a closed-loop AV simulation showed that realistic interactions can generate more informative and realistic scenarios for controller evaluation and tuning. The results highlight that behavioural realism is not only critical for understanding human decision-making but also for ensuring the robustness and safety of autonomous systems.

In summary, this thesis represents the first application of computational rationality to pedestrian-vehicle interaction. It establishes a theoretical and computational foundation for modelling human-like decision and control in the constrained interaction settings studied here, while demonstrating the potential relevance of such models for the design and evaluation of human-centred autonomous systems.

References

Alahi, Alexandre, Goel, Kratarth, Ramanathan, Vignesh, Robicquet, Alexandre, Fei-Fei, Li, and Savarese, Silvio (2016). "Social lstm: Human trajectory prediction in crowded spaces". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 961–971.

Camara, Fanta, Dickinson, Patrick, and Fox, Charles (2021). "Evaluating pedestrian interaction preferences with a game theoretic autonomous vehicle in virtual reality". In: *Transportation research part F: traffic psychology and behaviour* 78, pp. 410–423.

Chen, Xiuli, Bailly, Gilles, Brumby, Duncan P, Oulasvirta, Antti, and Howes, Andrew (2015). "The emergence of interactive behavior: A model of rational menu search". In: *Proceedings of the 33rd annual ACM conference on human factors in computing systems*, pp. 4217–4226.

Dang, Meiting, Zhao, Dezong, Wang, Yafei, and Wei, Chongfeng (2025). "Dynamic Game-Theoretical Decision-Making Framework for Vehicle-Pedestrian Interaction With Human Bounded Rationality". In: *IEEE Transactions on Intelligent Transportation Systems*.

Deo, Nachiket and Trivedi, Mohan M (2018). "Convolutional social pooling for vehicle trajectory prediction". In: *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pp. 1468–1476.

Diaz-Ruiz, Carlos A, Xia, Youya, You, Yurong, Nino, Jose, Chen, Junan, Monica, Josephine, Chen, Xiangyu, Luo, Katie, Wang, Yan, Emond, Marc, et al. (2022). "Ithaca365: Dataset and Driving Perception Under Repeated and Challenging Weather Conditions". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 21383–21392.

Divya, J, Divyadharshini, M, and Divyapriya, V (2025). "Gaze-based Pedestrian Safety Prediction using Attention LSTM". In: *2025 3rd International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)*. IEEE, pp. 1838–1845.

Doshi-Velez, Finale and Kim, Been (2017). "Towards a rigorous science of interpretable machine learning". In: *arXiv preprint arXiv:1702.08608*.

Dyro, Robert, Foutter, Matthew, Li, Ruolin, Di Lillo, Luigi, Schmerling, Edward, Zhou, Xilin, and Pavone, Marco (2024). "Realistic extreme behavior generation for improved av testing". In: *arXiv preprint arXiv:2409.10669*.

Fajen, Brett R and Warren, William H (2003). "Behavioral dynamics of steering, obstable avoidance, and route selection." In: *Journal of Experimental Psychology: Human Perception and Performance* 29.2, p. 343.

Fox, Charles, Camara, Fanta, Markkula, Gustav, Romano, Richard, Madigan, Ruth, Merat, Natasha, et al. (2018). "When should the chicken cross the road?: Game theory for autonomous vehicle-human interactions". In.

Fuchs, Andrew, Passarella, Andrea, and Conti, Marco (2023). "Modeling, replicating, and predicting human behavior: A survey". In: *ACM Transactions on Autonomous and Adaptive Systems* 18.2, pp. 1–47.

Giles, Oscar, Markkula, Gustav, Pekkanen, Jami, Yokota, Naoki, Matsunaga, Naoto, Merat, Natasha, and Daimon, Tatsuru (2019). "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, pp. 366–372.

Gomaa, Amr and Feld, Michael (2023). "Towards adaptive user-centered neuro-symbolic learning for multimodal interaction with autonomous systems". In: *Proceedings of the 25th International Conference on Multimodal Interaction*, pp. 689–694.

Helbing, Dirk and Molnar, Peter (1995). "Social force model for pedestrian dynamics". In: *Physical review E* 51.5, p. 4282.

Ho, Mark K and Griffiths, Thomas L (2022). "Cognitive science as a source of forward and inverse models of human decisions for robotics and control". In: *Annual Review of Control, Robotics, and Autonomous Systems* 5.1, pp. 33–53.

Jokinen, Jussi PP, Wang, Zhenxin, Sarcar, Sayan, Oulasvirta, Antti, and Ren, Xiangshi (2020). "Adaptive feature guidance: Modelling visual search with graphical layouts". In: *International Journal of Human-Computer Studies* 136, p. 102376.

Kalantari, Amir Hossein, Markkula, Gustav, Uzondu, Chinebuli, Lyu, Wei, Pedro, Jorge Garcia de, Madigan, Ruth, Lee, Yee Mun, Holmes, Christopher, and Merat, Natasha (2022). *Vehicle-Pedestrian Interactions at Uncontrolled Locations: Leveraging Distributed Simulation to Support Game-Theoretic Modeling*. Tech. rep.

Konstantinidis, Fabian, Sackmann, Moritz, Hofmann, Ulrich, and Stiller, Christoph (2023). "Modeling interaction-aware driving behavior using graph-based representations and multi-agent reinforcement learning". In: *2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, pp. 1643–1650.

Markkula, Gustav, Lin, Yi-Shin, Srinivasan, Aravinda Ramakrishnan, Billington, Jac, Leonetti, Matteo, Kalantari, Amir Hossein, Yang, Yue, Lee, Yee Mun, Madigan, Ruth, and Merat, Natasha (2023). "Explaining human interactions on the road by large-scale integration of computational psychological theory". In: *PNAS nexus* 2.6, pgad163.

Papadimitriou, Eleonora, Auberlet, Jean-Michel, Yannis, George, and Lassarre, Sylvain (2016). "Simulation of Pedestrians and Motorised Traffic: existing research and future challenges". In: *Civil and Environmental Engineering: Concepts, Methodologies, Tools, and Applications*. IGI Global Scientific Publishing, pp. 1646–1662.

Pekkanen, Jami, Giles, Oscar Terence, Lee, Yee Mun, Madigan, Ruth, Daimon, Tatsuru, Merat, Natasha, and Markkula, Gustav (2022). "Variable-drift diffusion models of pedestrian road-crossing decisions". In: *Computational Brain & Behavior* 5.1, pp. 60–80.

Petzoldt, Tibor (2014). "On the relationship between pedestrian gap acceptance and time to arrival estimates". In: *Accident Analysis & Prevention* 72, pp. 127–133.

Rashid, Md Mobasshir, Seyedi, MohammadReza, and Jung, Sungmoon (2024). "Simulation of pedestrian interaction with autonomous vehicles via social force model". In: *Simulation modelling practice and theory* 132, p. 102901.

Rasouli, Amir, Kotseruba, Iuliia, Kunic, Toni, and Tsotsos, John K (2019). "Pie: A large-scale dataset and models for pedestrian intention estimation and trajectory prediction". In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 6262–6271.

Ratcliff, Roger, Smith, Philip L, Brown, Scott D, and McKoon, Gail (2016). "Diffusion decision model: Current issues and history". In: *Trends in cognitive sciences* 20.4, pp. 260–281.

Roshdi, Mohamed, Petzold, Julian, Wahby, Mostafa, Ebrahim, Hussein, Berekovic, Mladen, and Hamann, Heiko (2024). "On the Road to Clarity: Exploring Explainable AI for World Models in a Driver Assistance System". In: *2024 IEEE Conference on Artificial Intelligence (CAI)*. IEEE, pp. 1032–1039.

Rudin, Cynthia (2019). "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead". In: *Nature machine intelligence* 1.5, pp. 206–215.

Schmidt, Lukas M, Brosig, Johanna, Plinge, Axel, Eskofier, Bjoern M, and Mutschler, Christopher (2022). "An introduction to multi-agent reinforcement learning and review of its application to autonomous mobility". In: *2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, pp. 1342–1349.

Tian, Kai, Markkula, Gustav, Wei, Chongfeng, Lee, Yee Mun, Madigan, Ruth, Hirose, Toshiya, Merat, Natasha, and Romano, Richard (2023). "Deconstructing pedestrian crossing decisions in interactions with continuous traffic: An anthropomorphic model". In: *IEEE Transactions on Intelligent Transportation Systems* 25.3, pp. 2466–2478.

Tian, Kai, Markkula, Gustav, Wei, Chongfeng, Lee, Yee Mun, Madigan, Ruth, Merat, Natasha, and Romano, Richard (2022). "Explaining unsafe pedestrian road crossing behaviours using a psychophysics-based gap acceptance model". In: *Safety science* 154, p. 105837.

Wang, Yueyang, Srinivasan, Aravinda Ramakrishnan, Jokinen, Jussi PP, Oulasvirta, Antti, and Markkula, Gustav (2023). "Modeling human road crossing decisions as reward maximization with visual perception limitations". In: *2023 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, pp. 1–6.

– (2025). "Pedestrian crossing decisions can be explained by bounded optimal decision-making under noisy visual perception". In: *Transportation Research Part C: Emerging Technologies* 171, p. 104963.

Yang, Changzhi, Pan, Huihui, and Wang, Jue (2025). "Interact, Plan, and Go: Transformers With Social Intentions for Trajectory Prediction". In: *IEEE Transactions on Consumer Electronics*.

Yang, Dongfang, Özgüner, Ümit, and Redmill, Keith (2020). "A social force based pedestrian motion model considering multi-pedestrian interaction with a vehicle". In: *ACM Transactions on Spatial Algorithms and Systems (TSAS)* 6.2, pp. 1–27.

Yang, Yuhang, Kujanpää, Kalle, Babadi, I Amin, Pajarinen, Joni, and Ilin, Alexander (2023). "Suicidal pedestrian: Generation of safety-critical scenarios for autonomous vehicles". In: *2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, pp. 1983–1988.

Yi, Shuai, Li, Hongsheng, and Wang, Xiaogang (2016). "Pedestrian behavior understanding and prediction with deep neural networks". In: *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14*. Springer, pp. 263–279.

Appendix A

Supplementary material for Chapter 2

A.1 Reward function sensitivity analysis

This appendix examines how the reward parameters used in Chapter 2 affect behaviour and model comparison outcomes across the four main model variants: BM, LM, VM, and VLM. The baseline reward structure was designed to reflect the assumption that pedestrians aim to cross safely while avoiding unnecessary delay. In this formulation, successful crossing was rewarded positively, collisions were penalised symmetrically, and a time penalty discouraged prolonged waiting. In the LM and VLM variants, an additional looming-related term was included to represent aversion to crossing in front of a rapidly approaching vehicle. The baseline values used in the main analysis were selected initially through manual testing in the BM model to produce qualitatively reasonable crossing behaviour. The purpose of the present appendix is to examine this choice more systematically by testing how changes in the reward parameters affect model behaviour and model comparison outcomes.

A factorial sensitivity analysis was conducted for all four model variants: BM, LM, VM, and VLM. Two reward parameters were varied. The first was the crossing reward magnitude, with tested values $R \in \{4, 20, 100\}$. The second was the time-penalty coefficient, with tested values $tp \in \{0.002, 0.01, 0.05\}$. The setting $R = 20$ and $tp = 0.01$ corresponds to the baseline reward specification used in the main analysis, and the other tested values correspond to an increase or decrease of these baseline values by a factor of five. For each reward combination, the cor-

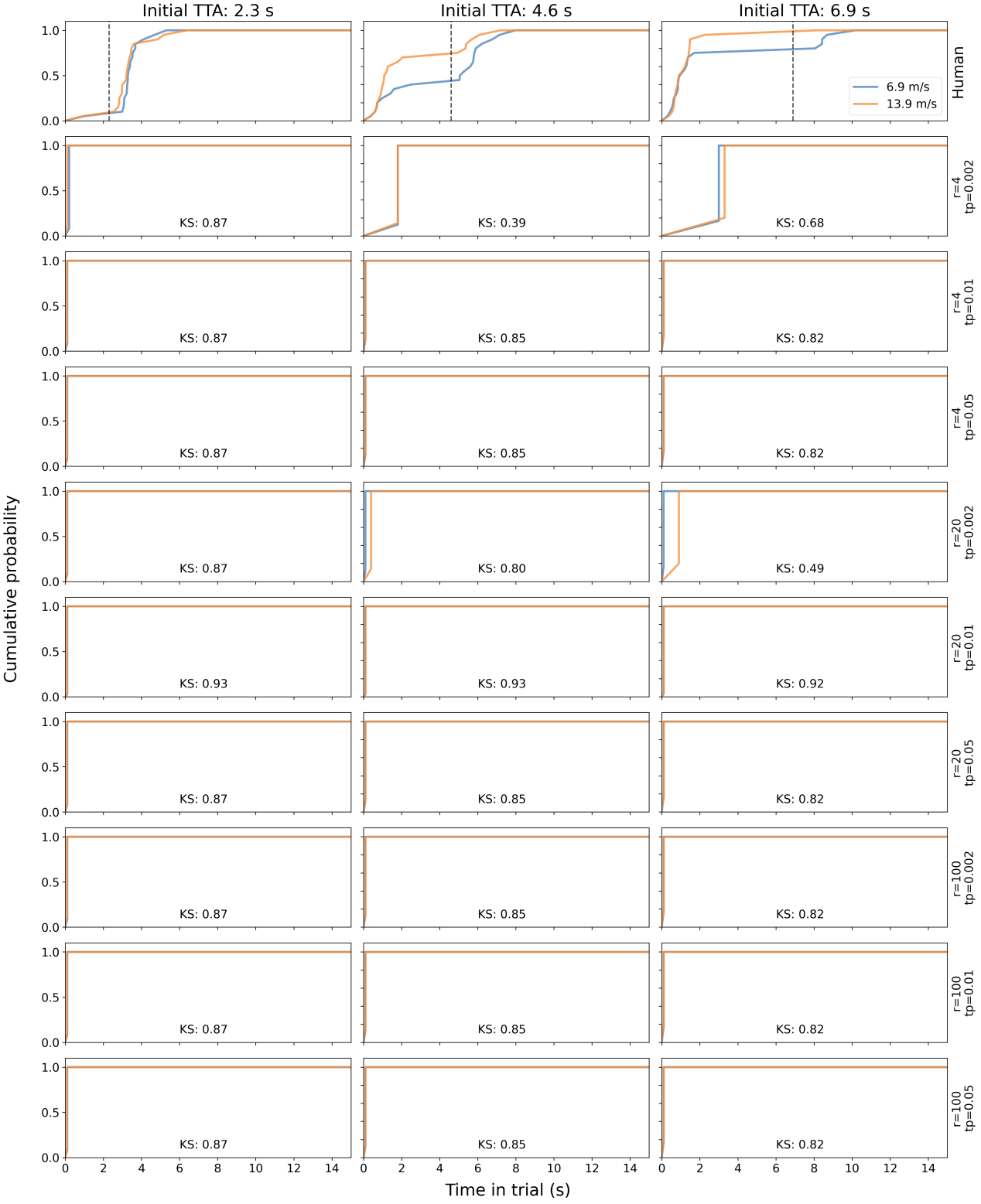

Figure A.1: Cumulative probability distributions of crossing initiation time (CIT) in constant-speed scenarios for the BM model under different reward settings. Columns correspond to the three initial TTA conditions. The top row shows the human data, and each subsequent row shows the BM model under a different combination of crossing reward magnitude r and time-penalty coefficient tp. Blue and orange lines correspond to the two vehicle-speed conditions. KS statistics are shown in each model panel.

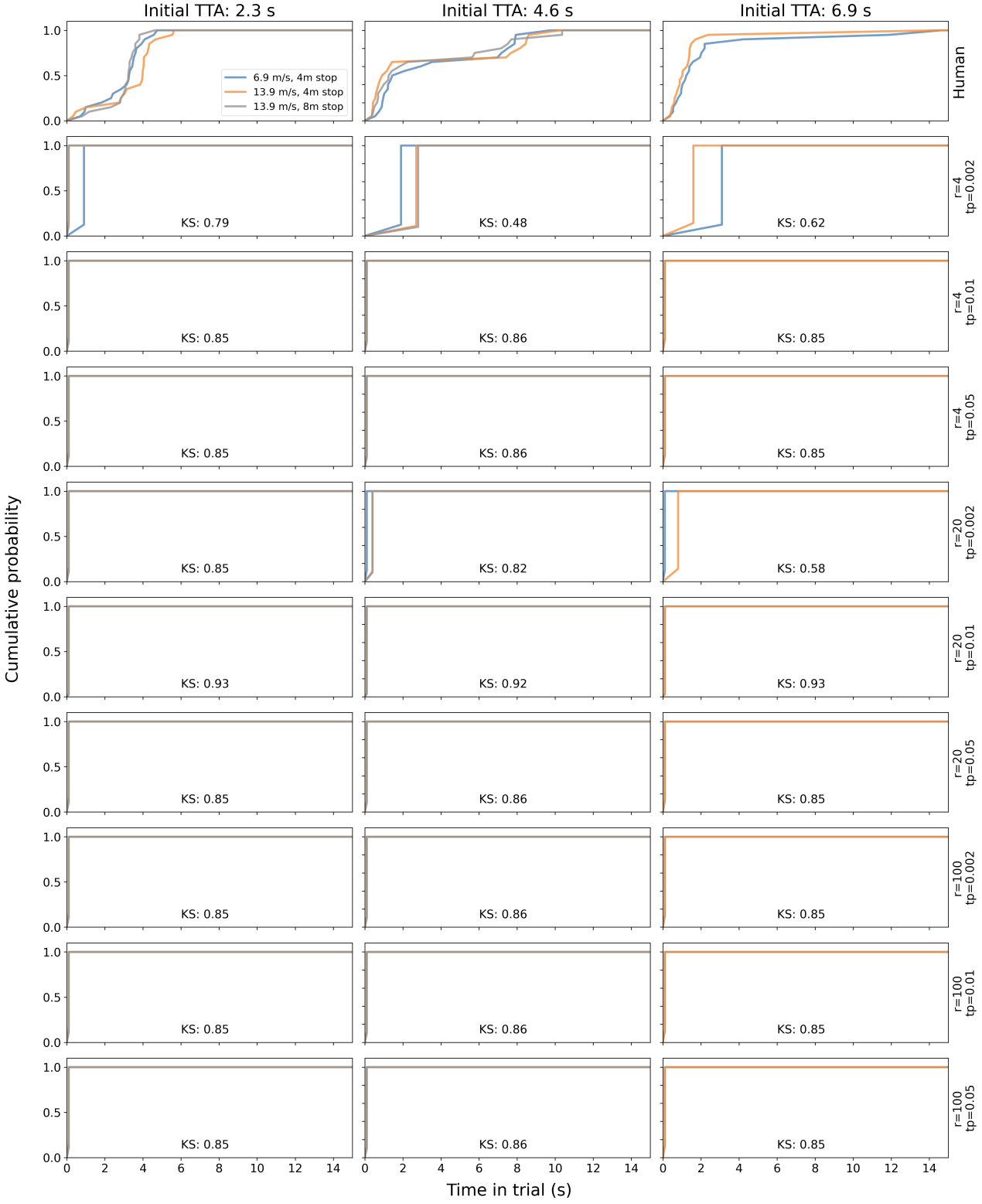

Figure A.2: Cumulative probability distributions of CIT in yielding scenarios for the BM model under different reward settings. Columns correspond to the three initial TTA conditions. The top row shows the human data, and each subsequent row shows the BM model under a different combination of crossing reward magnitude r and time-penalty coefficient tp. Blue, orange, and grey lines correspond to the three yielding conditions shown in the human data. KS statistics are shown in each model panel.

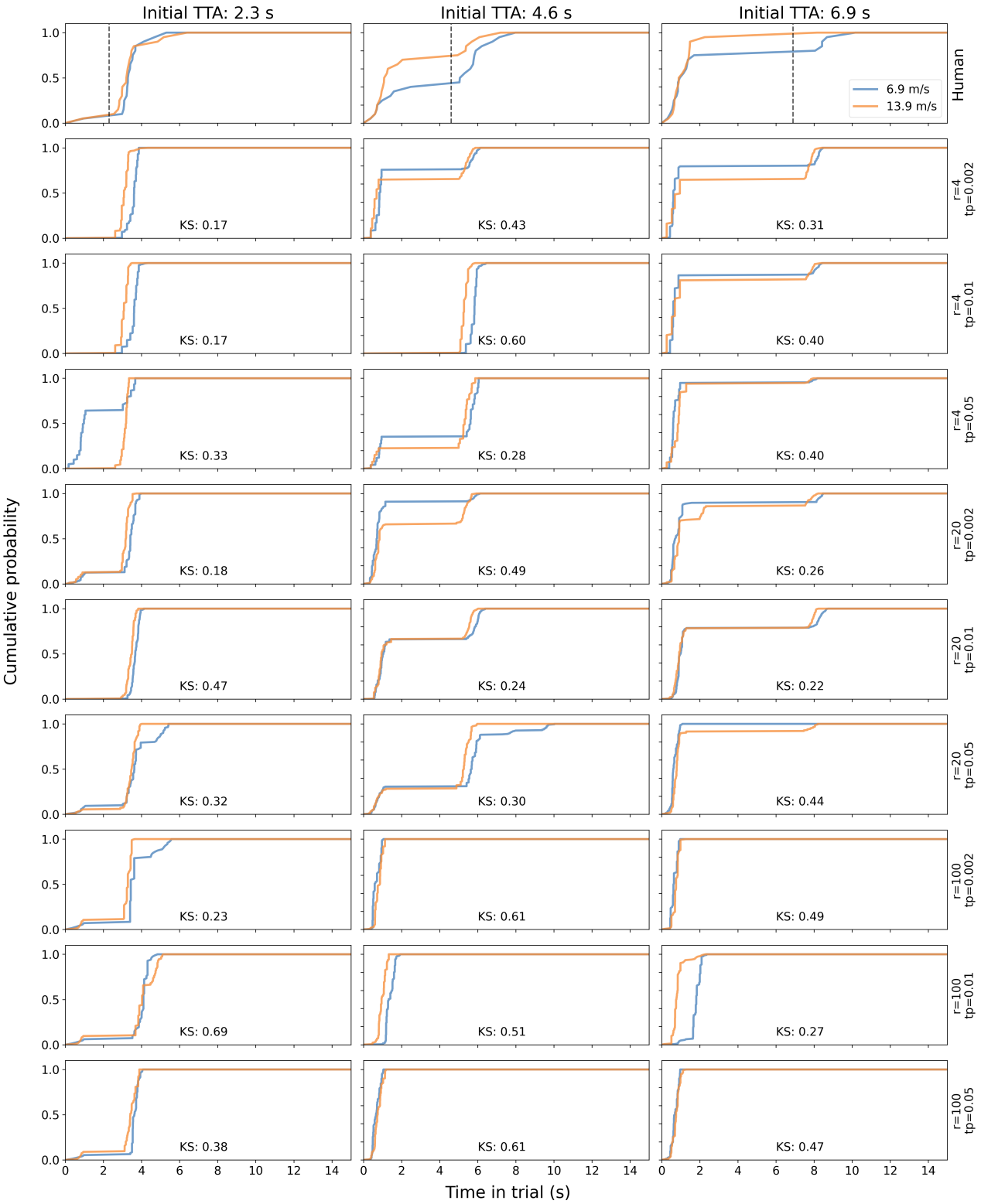

Figure A.3: Cumulative probability distributions of CIT in constant-speed scenarios for the LM model under different reward settings.

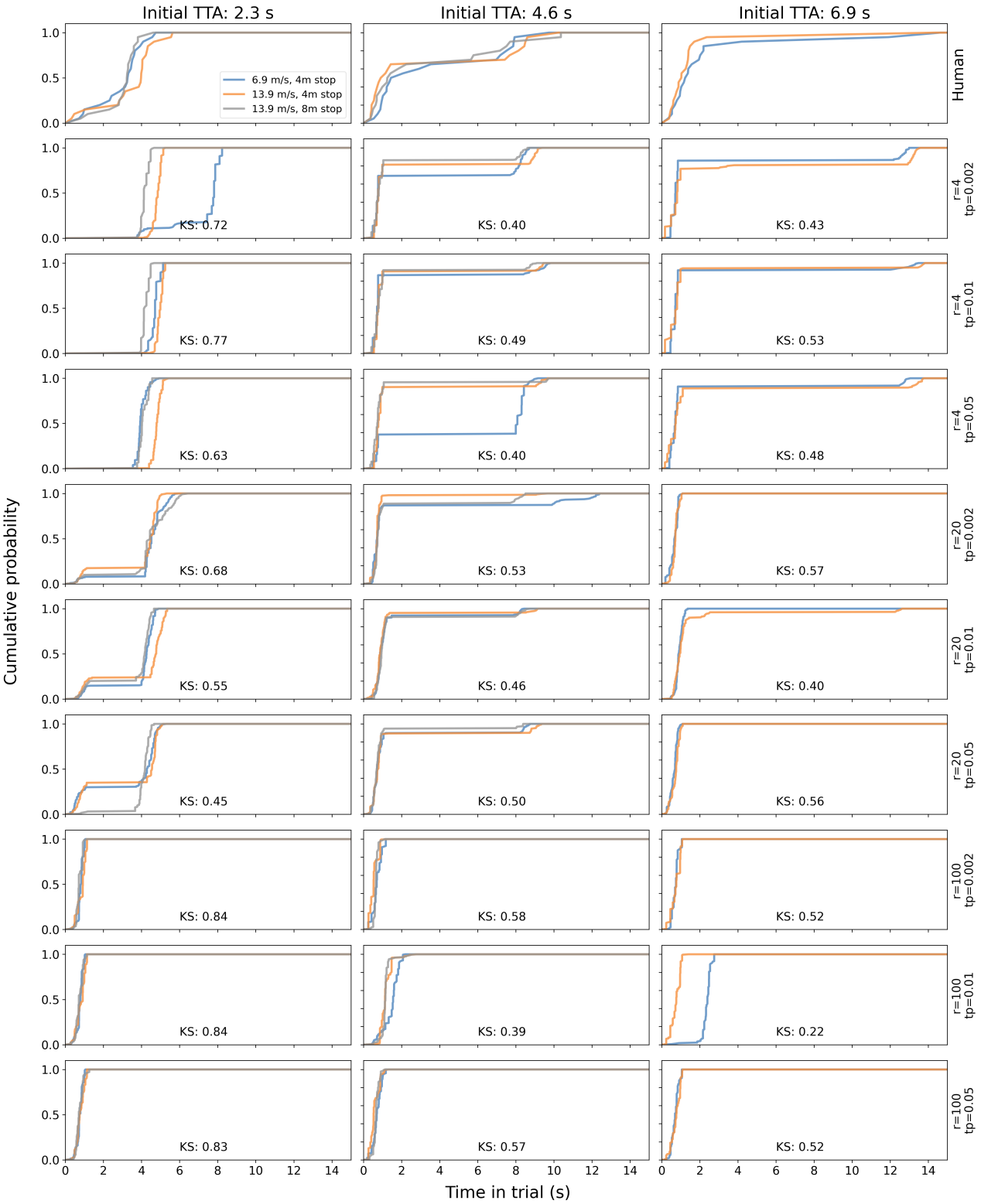

Figure A.4: Cumulative probability distributions of CIT in yielding scenarios for the LM model under different reward settings.

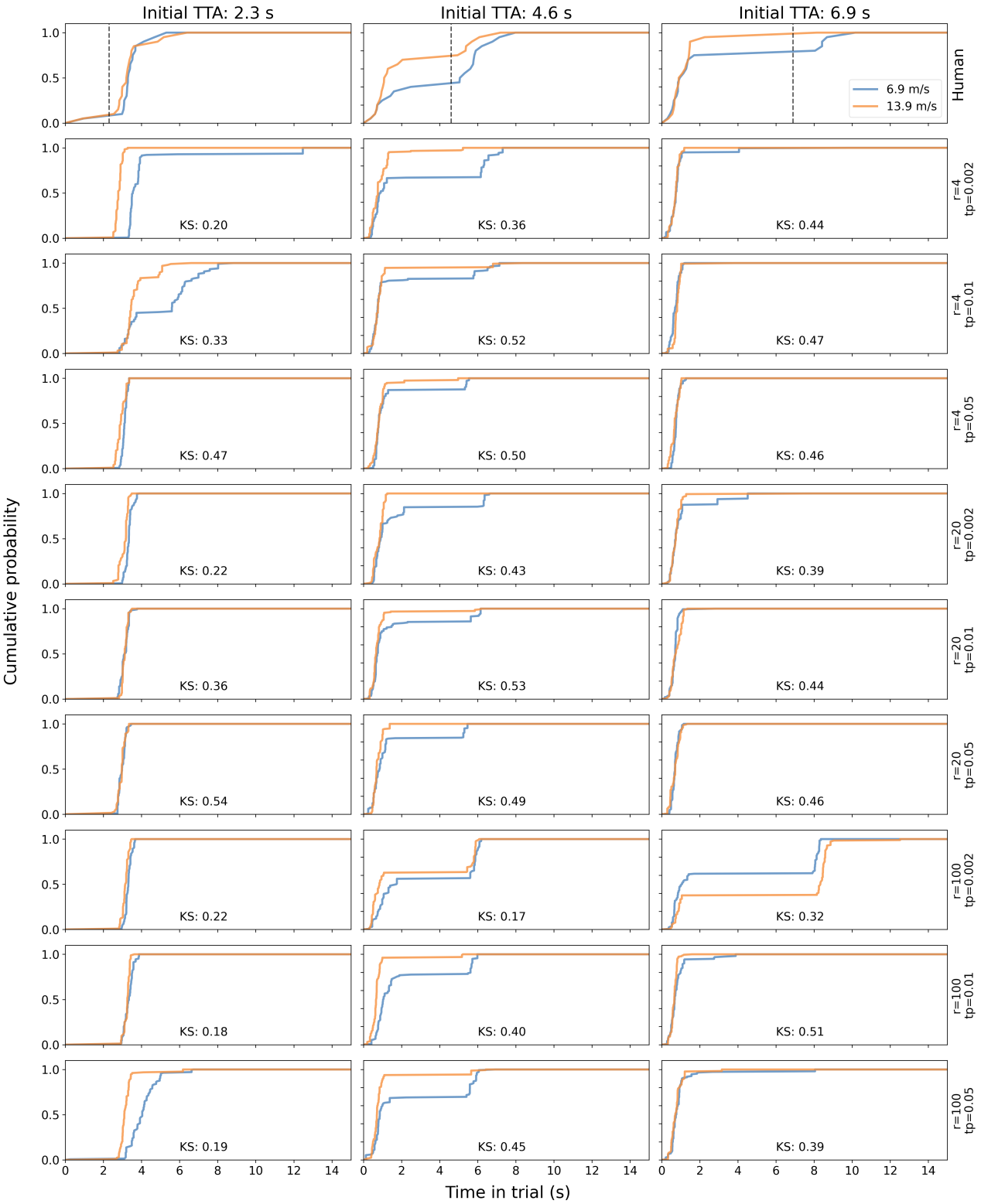

Figure A.5: Cumulative probability distributions of CIT in constant-speed scenarios for the VM model under different reward settings.

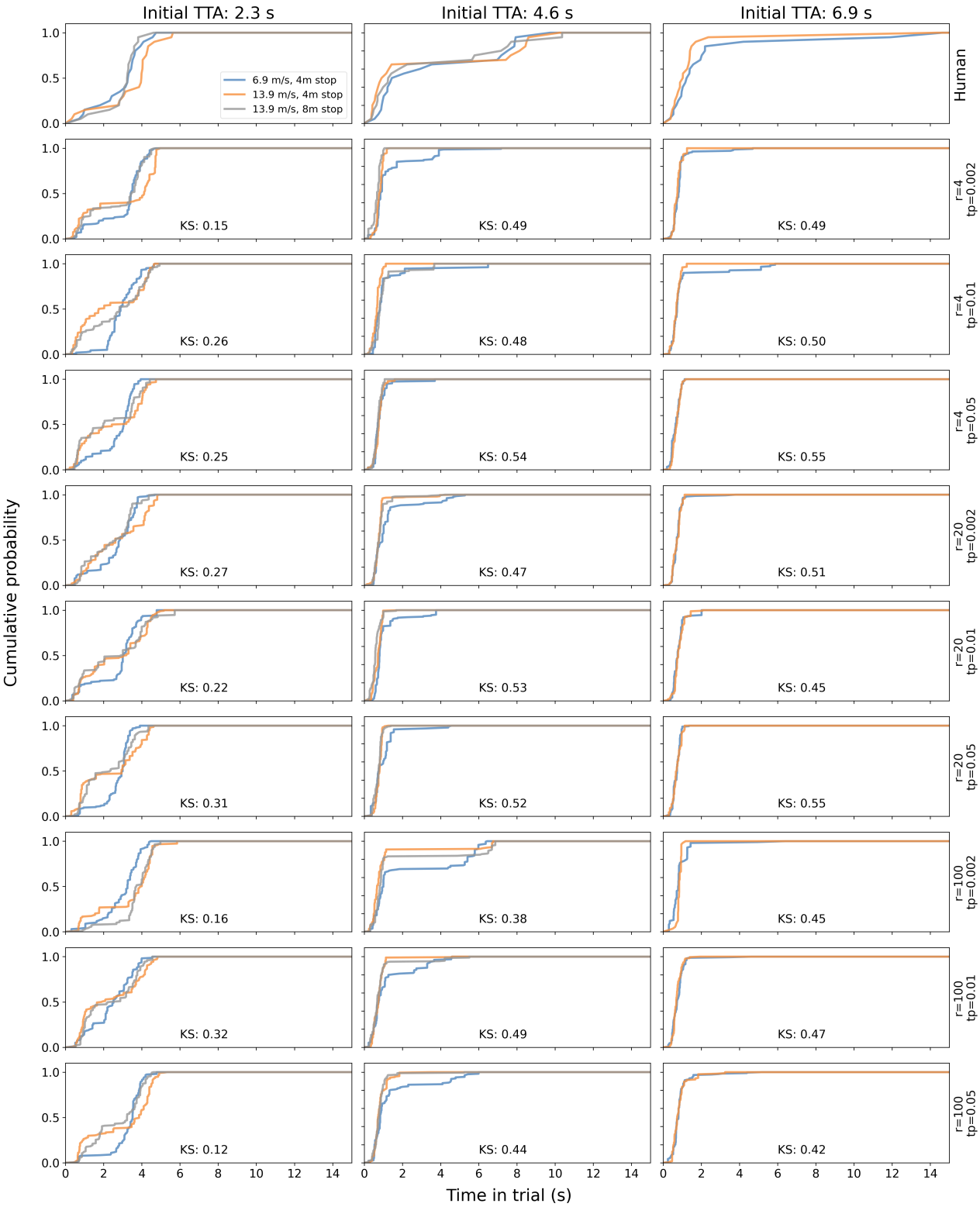

Figure A.6: Cumulative probability distributions of CIT in yielding scenarios for the VM model under different reward settings.

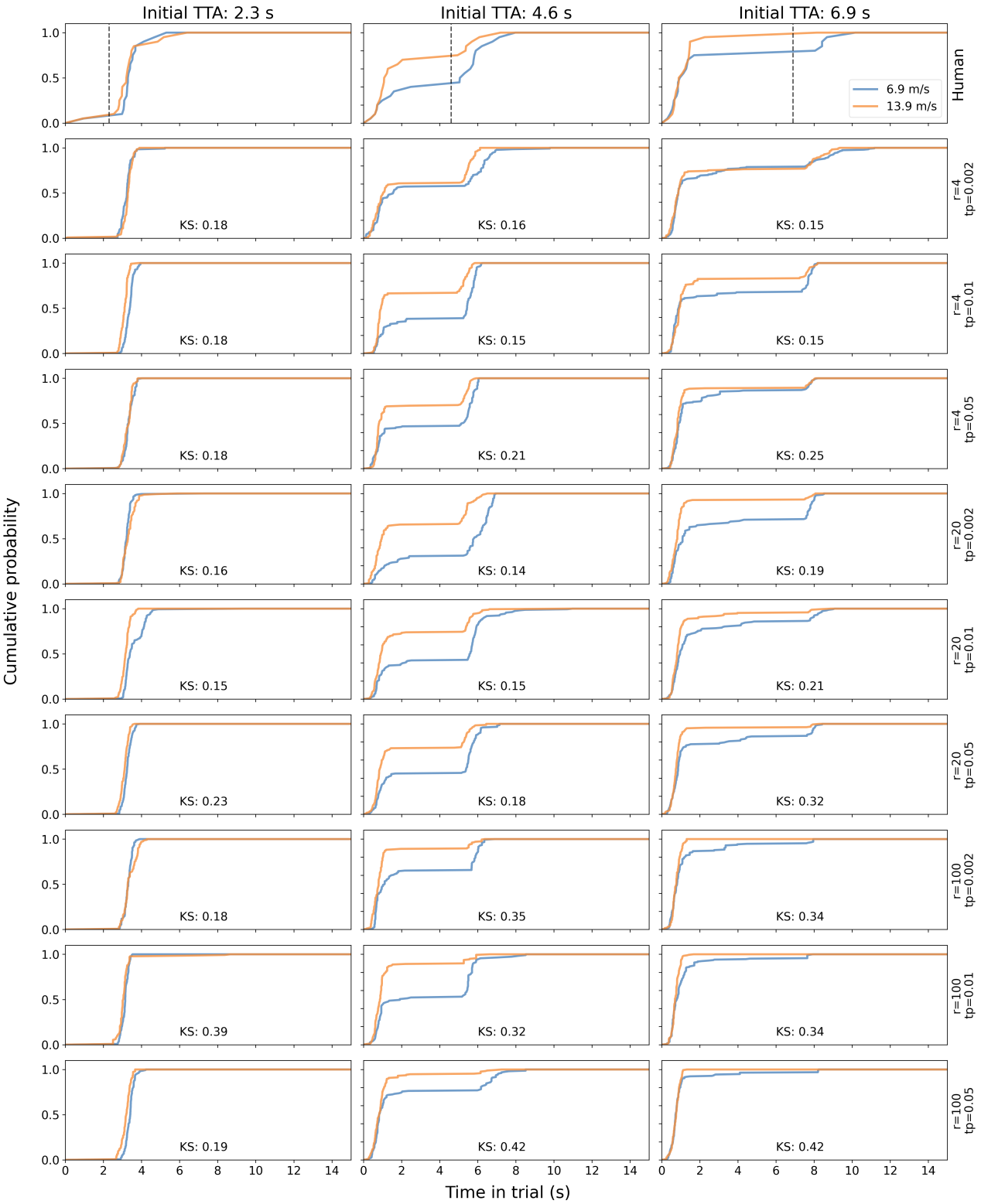

Figure A.7: Cumulative probability distributions of CIT in constant-speed scenarios for the VLM model under different reward settings.

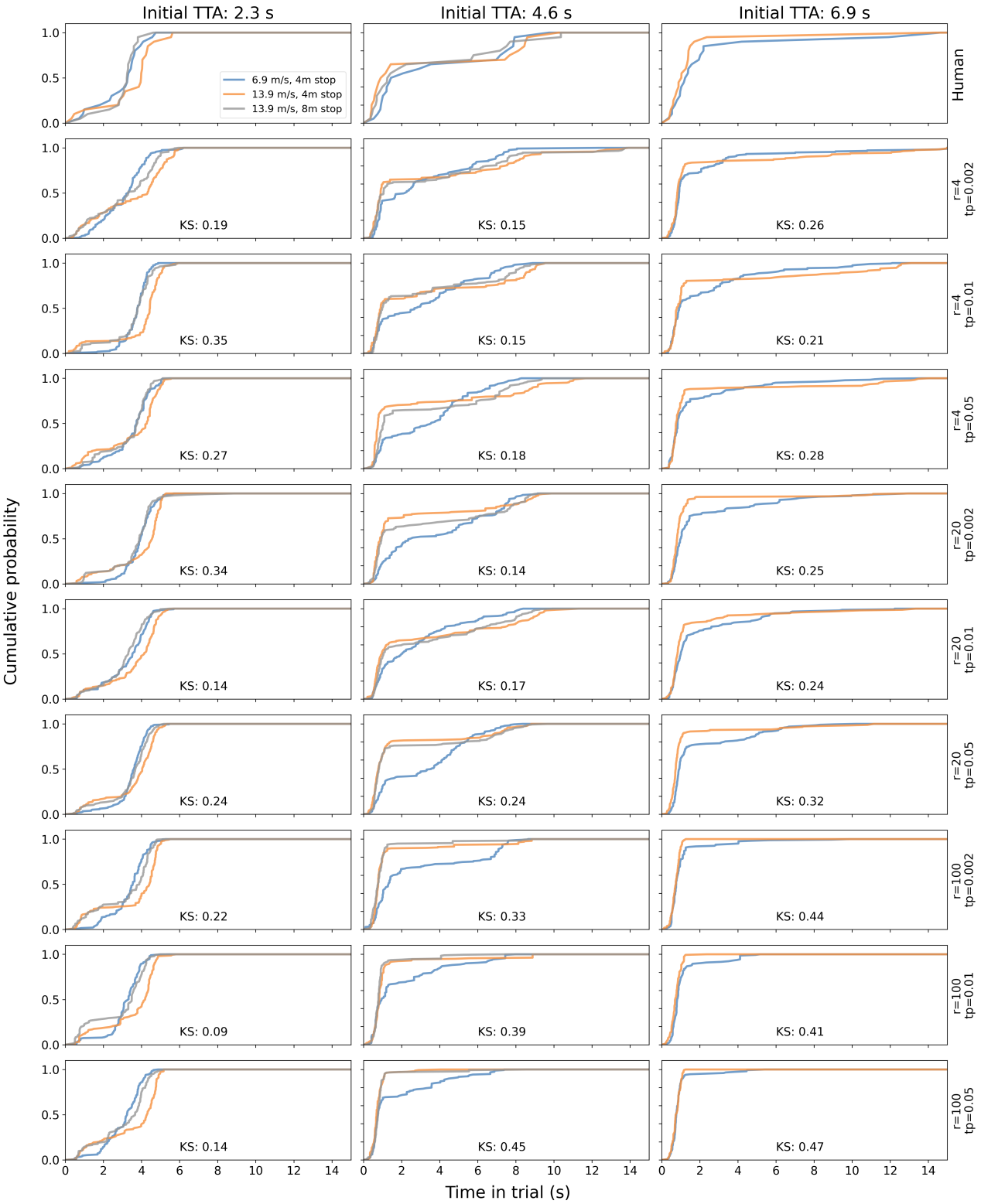

Figure A.8: Cumulative probability distributions of CIT in yielding scenarios for the VLM model under different reward settings.

responding model was trained independently. For model variants with fitted non-policy parameters, namely LM, VM, and VLM, the fitting procedure was also repeated for each reward setting so that the estimated parameters remained matched to the behaviour of the retrained model. Behavioural differences across reward settings were assessed using the KS statistic and by visually comparing the cumulative probability distributions of CIT. Figures A.1 to A.8 show the resulting cumulative probability distributions of CIT for all four model variants, separately for constant-speed and yielding scenarios.

For the BM model, Figures A.1 and A.2 show that reward manipulation could alter crossing behaviour, with some settings producing undesired delayed crossing. The LM, VM, and VLM models show more different behavioural patterns across reward settings, as shown in Figures A.3 to A.8. Across the tested settings with $R \in \{4, 20\}$, the relative ranking of the model variants was preserved, with VLM performing best, followed by LM, VM, and BM. This indicates that the main qualitative conclusion of the model comparison was robust to moderate changes in reward setting. By contrast, some settings with $R = 100$ led to degraded fit and more aggressive crossing behaviour, especially in the simpler model variants. To be specific, a larger reward magnitude increased the incentive to cross and therefore shifted the learned policies towards earlier crossing decisions. The time penalty had a smaller overall effect, although very small values such as $tp = 0.002$ could lead to later crossing in some settings by reducing the cost of waiting.

These results provide justification for the baseline reward parameters used in the main analysis. The baseline setting $R = 20$ and $tp = 0.01$ lies within a stable and behaviourally reasonable region. More generally, the analysis suggests that moderate reward magnitudes, such as $R \in [4, 20]$, are acceptable in the present setting, whereas substantially larger values such as $R = 100$ can produce overly aggressive crossing behaviour and degrade model fit. The time penalty is less influential overall, but values that are too small may weaken the tendency to avoid unnecessary delay. This also clarifies the implications of selecting different reward values: increasing the crossing reward tends to promote earlier and more aggressive crossing decisions, whereas reducing the time penalty can make the model more willing to wait. The reward structure therefore does not require precise tuning, but its qualitative balance remains behaviourally important.

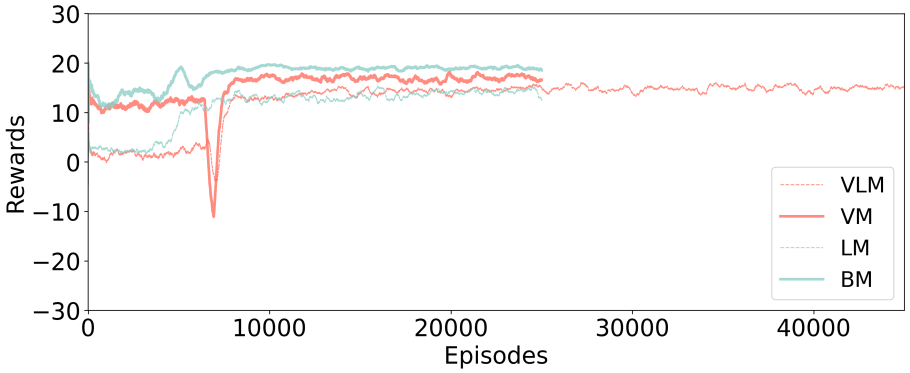

Figure A.9: Convergence plot for different models.

A.2 Model convergence

Figure A.9 illustrates the average rewards of four main model variants recorded for each of the 500 episodes, spanning the entirety of the training phase. From Figure A.9, it can be noted that the reward for the models, including the looming aversion assumption (LM and VLM), achieve lower rewards overall due to the change in the reward structure introduced by this assumption. Additionally, adding visual constraints to the baseline model (going from BM to VM) led to a decline in the average rewards. This is due to the VM agent's inability to know the precise position of the approaching vehicle. As a result, the agent tends to make more cautious decisions, leading to increased time penalties. However, there was no noticeable difference in the achievable rewards for LM and VLM. A possible reason could be that for LM, the looming aversion penalty is directly proportional to the inverse τ, while in VLM, the looming aversion penalty corresponds to the estimated inverse τ. From the reward maximisation perspective, both model variants aim to minimise this looming penalty within their respective environmental contexts, potentially explaining the similar overall rewards.

A.3 Additional Model Variants

Beyond the core model variants outlined in Section 2.2.2, we explored four additional variations to deepen our understanding of the pedestrian decision-making process:

- **VLDM:** The model variant considers motor delay an additional human-like constraint, reflecting the time lag between decision-making and action execution. This variant expands upon the VLM model variant, incorporating motor delay into the model.
- **VNM:** A variation of the looming aversion model that tests an alternative representation of risk aversion. Instead of an inverse τ, we evaluated a model where aversion is formulated

against near-collision situations, providing an alternative perspective on risk assessment in pedestrian behaviour.

- **VLM (E):** A version of the VLM model that employs a single set of non-policy parameters across the entire human dataset instead of fitting them to each participant. This approach aimed to assess the efficacy of a more generalised model application.
- **VLM (S):** This variation of the VLM model involves training individual networks for each unique set of parameters.

These models were designed to assess the influence of additional factors and assumptions on pedestrian crossing behaviour. While these models provided valuable insights, they did not show an overall improvement in performance compared to our primary model, VLM. Figure A.10 and A.11 and Table A.1 show the results obtained for these models.

The VLDM model integrates motor delay (m) as an additional non-policy parameter, alongside visual constraints (σ_{v}) and looming aversion (c). VLDM showed similar performance to the VLM model regarding log-likelihood and MAD. However, it obtained a higher AIC score due to the increased model complexity.

The VNM model explores the concept of pedestrian aversion to near-collision situations. This model variant provides an alternative way of modelling pedestrian aversion to being in safety-critical situations as an alternative to the formulation based on looming in the LM/VLM models. This model assumes pedestrians inherently aim to avoid situations where a vehicle invades their personal space, leading to an additional penalty in the model when this occurs. While VNM exhibited more human-like CITs than the VM model, it still lagged behind the comprehensive performance of the VLM model regarding AIC. This disparity in model performance could imply that the integration of temporal-spatial clues (e.g., looming) may play a more important role in pedestrian decision-making than spatial factors alone (e.g., proximity to a near-collision zone).

Table A.1: Quantitative assessment of additional model performance

Model type	**Log Lik.**	**Params.**	**AIC**	**MAD(s)**
VLDM	−536	60	1192	0.35
VNM	−564	40	1208	0.99
VLM(E)	−547	2	1098	0.36
VLM(S)	−572	40	1225	0.39

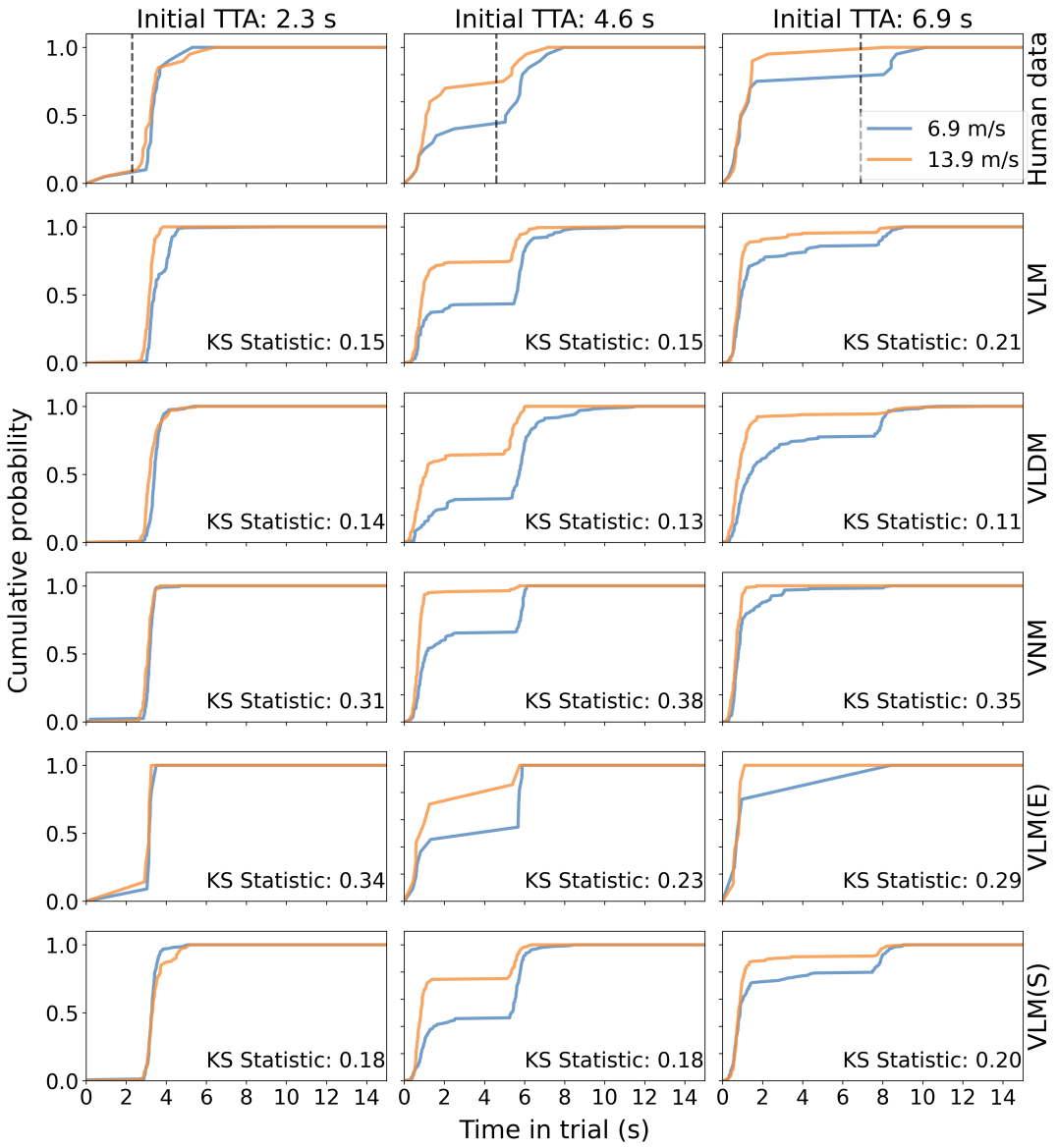

Figure A.10: Cumulative probability for Crossing Initiation Time (CIT) in constant speed scenarios for human data and additional models. Black dashed vertical lines in Human data indicate the times vehicles passed pedestrians. To avoid repetition, these black lines are not included in the model results. Vehicles passed pedestrians at the same time in the trial for each initial TTA condition. The VLM model is included in this figure for comparison.

The VLM (E) model, which uses a single set of non-policy parameters for the entire dataset, performed better in terms of AIC value than VLM, reflecting a simpler model structure. However, it did not match the log-likelihood performance of VLM. This finding suggests that individualising non-policy parameters may be beneficial for accurately capturing individual pedestrian behaviour.

In the VLM (S) model, instead of conditioning a single RL policy on the values of the non-policy parameters, we trained separate RL policy networks for each of the 10x10=100 considered

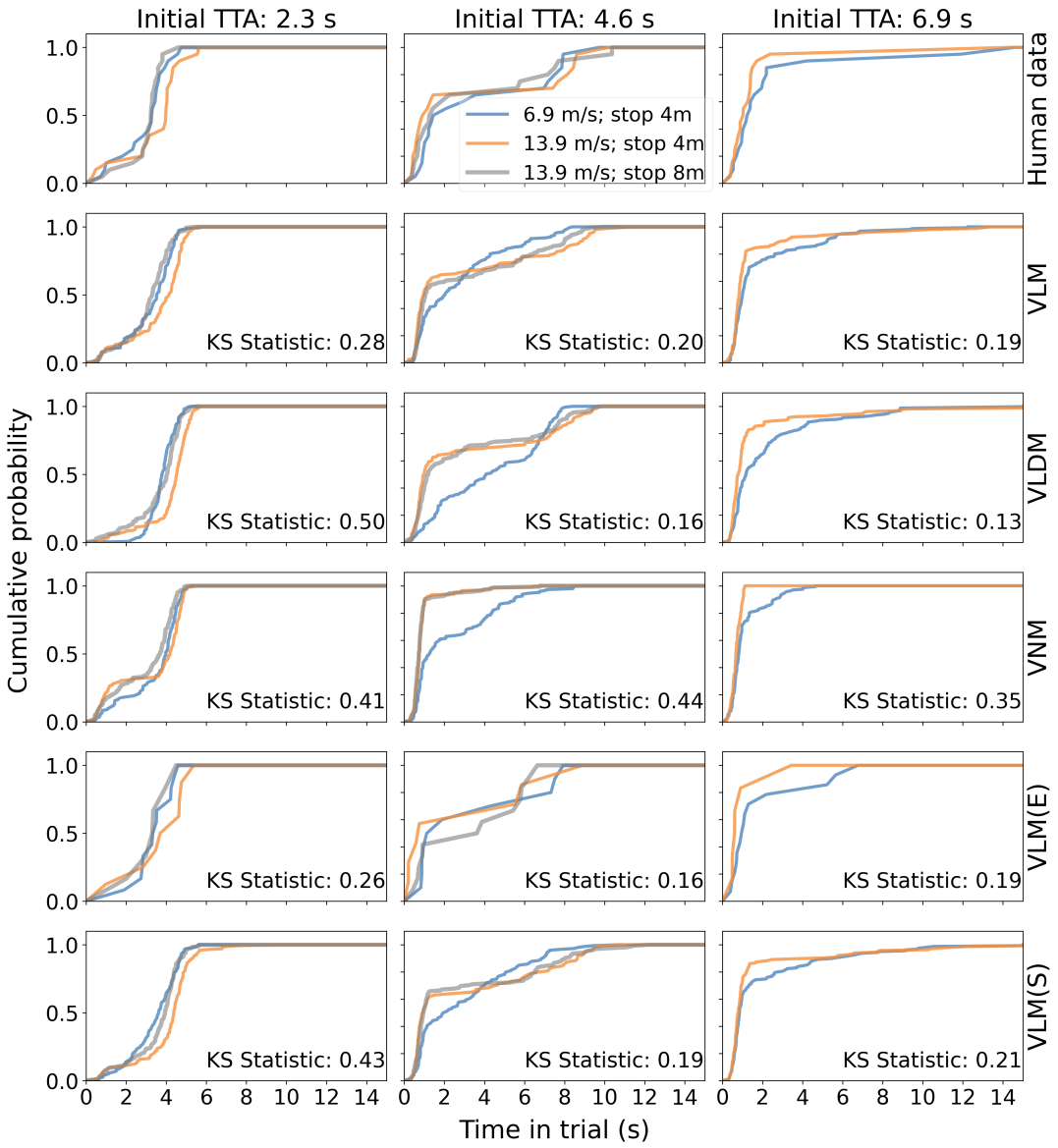

Figure A.11: Cumulative probability for CIT in yielding scenarios for human data and additional models. The VLM model is included in this figure for comparison.

values of the non-policy parameters. Despite the increased computational effort, VLM (S) did not achieve better performance metrics than the original VLM model, as shown in Table A.1. This result highlights the efficiency of our chosen parameter-conditioning in the VLM model.

Regarding VLDM, with the motor delay, m, as an additional input parameter alongside σ_v and c, from Table A.1, we can find that this model matched the VLM's performance in terms of both log-likelihood and MAD. However, its additional parameters result in a higher AIC compared to VLM.

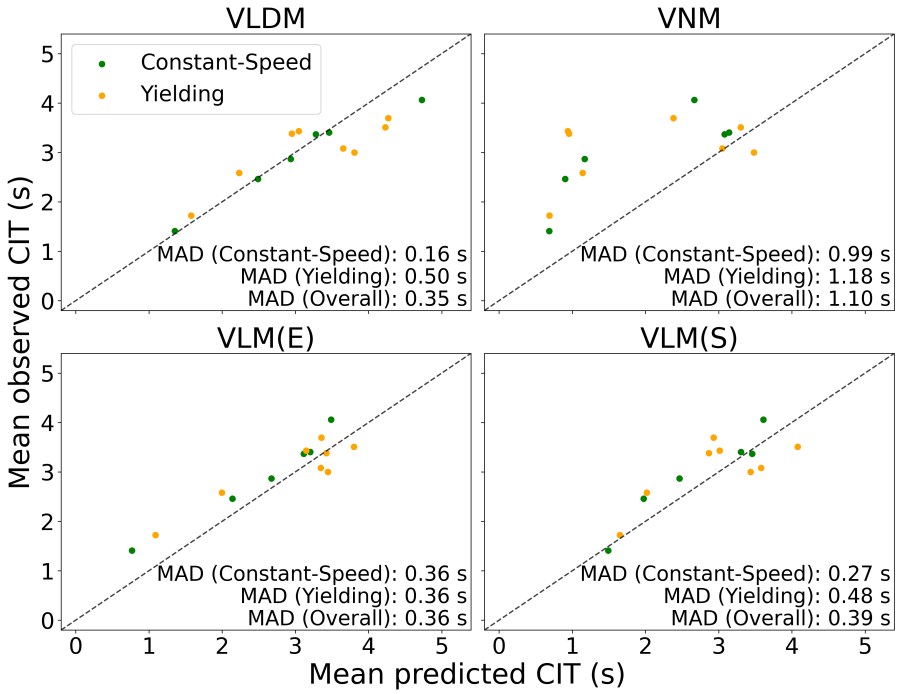

Figure A.12: Predicted vs observed mean CIT across the different scenarios.

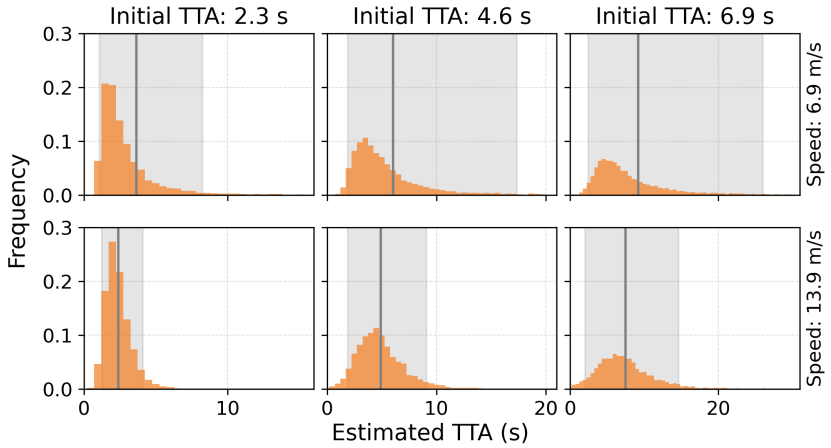

Figure A.13: Distributions of initial estimated TTA from the output of the visual perception model, across the different scenarios in the experiment. The grey vertical line is the mean value, and the shading area shows the 5th and 95th percentiles of the estimated TTA.

=

Appendix B

Supplementary material for Chapter 3

B.1 Experiment description

The experiment is described in full in (Lee et al., 2026), but a summary is provided here because the empirical task structure is central to the modelling developed in this chapter. The data came from a virtual reality pedestrian experiment conducted in the University of Leeds Highly Immersive Kinematic Experimental Research (HIKER) laboratory, a cave-based simulator with a 9 m by 4 m walking area and immersive projected virtual environment. Ethical approval was obtained from the University of Leeds Research Committee (Ref: 0536). In each trial, participants stood at the kerb and decided whether to cross the road between two approaching vehicles coming from their right, crossing naturally when they felt comfortable to do so. The experimental design used five within-participant factors: vehicle approach speed (25/30 mph), time gap between the two vehicles (3/5 s), behaviour of the second vehicle (yielding via deceleration versus non-yielding), presence or absence of eHMI in yielding trials, and time of day (daytime/nighttime), together with one between-participant factor, age group (younger/older).

Each participant completed both a daytime block and a nighttime block. Within each block, the decelerating trials with eHMI present and absent were each repeated across the four kinematic combinations of speed and time gap, and the non-decelerating trials were repeated more often so that decelerating and non-decelerating trials were balanced overall. This resulted in 32 experimental trials per block, presented in a randomised order, with day and night block order

counterbalanced across participants using a Latin square approach. In addition, a final failure trial was included at the end of each block, in which the vehicle displayed the eHMI but did not decelerate, giving a total of 33 trials per block.

The eHMI used in the experiment was a cyan slow pulsing light band placed around the windscreen and roofline of the vehicle. In the yielding trials, the second vehicle began to decelerate when it was 30 m from the pedestrian and came to a stop 2.5 m before the crossing path; when the eHMI was present, it was activated at the same moment that deceleration began. Prior to the experimental trials, participants completed a practice block in both daytime and nighttime environments to familiarise them with both the virtual environment and the task. The practice and experimental blocks used the same written task instructions. Participants were explicitly told that the cyan eHMI signalled the message 'I am yielding', that is, that the approaching vehicle intended to give way. This induction is important for interpreting the present modelling results, because the empirical data reflect behaviour after participants had been introduced to the intended meaning of the eHMI, rather than responses to a completely unfamiliar signal.

Appendix C

Supplementary material for Chapter 4

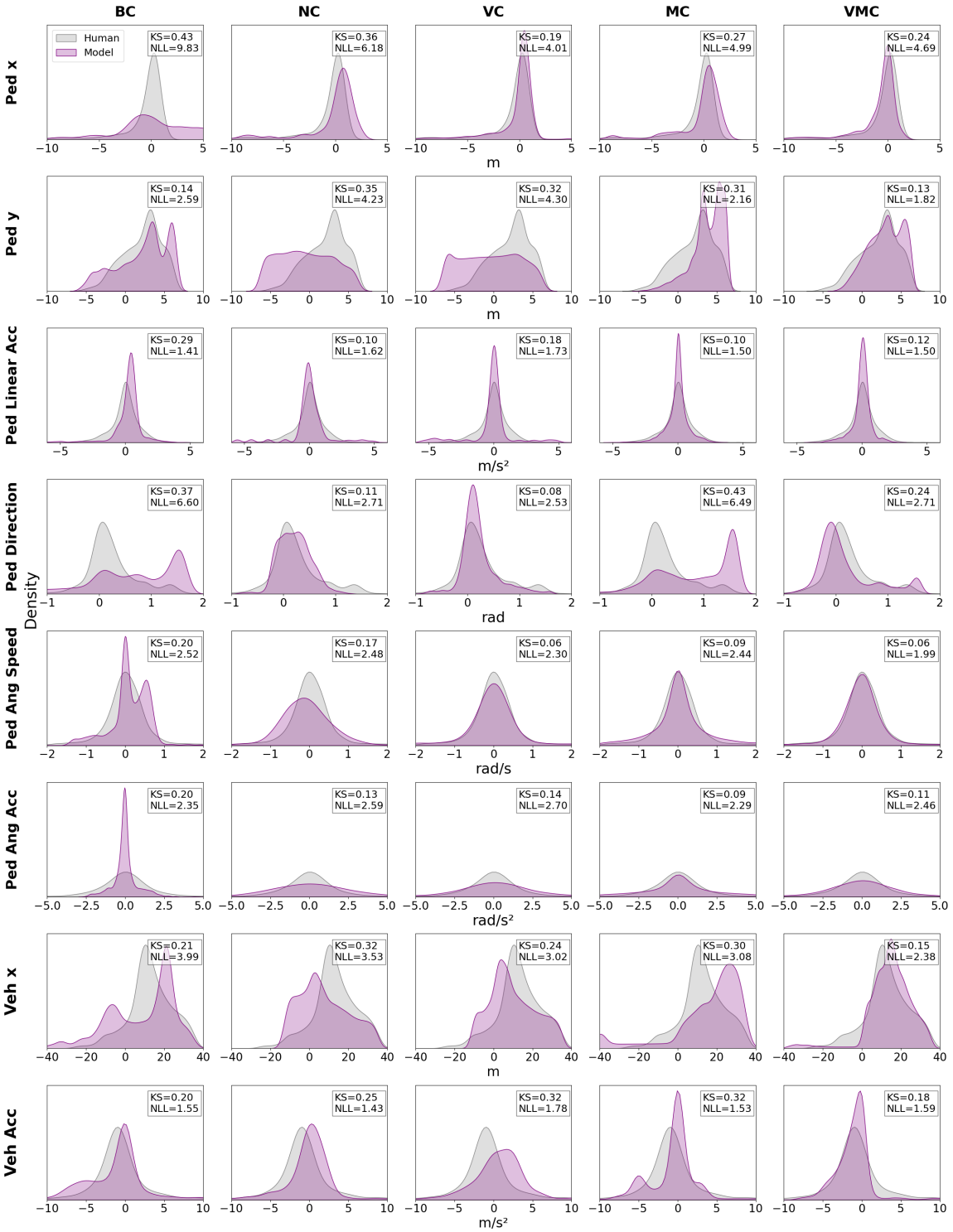

Figure C.1: Comparison of model and real-world behavioural metric distributions across all features used in fitting (see Section 2.2.5), in addition to the four key metrics shown in Figure 4.6. Each column corresponds to one model variant, and each row to one behavioural metric.

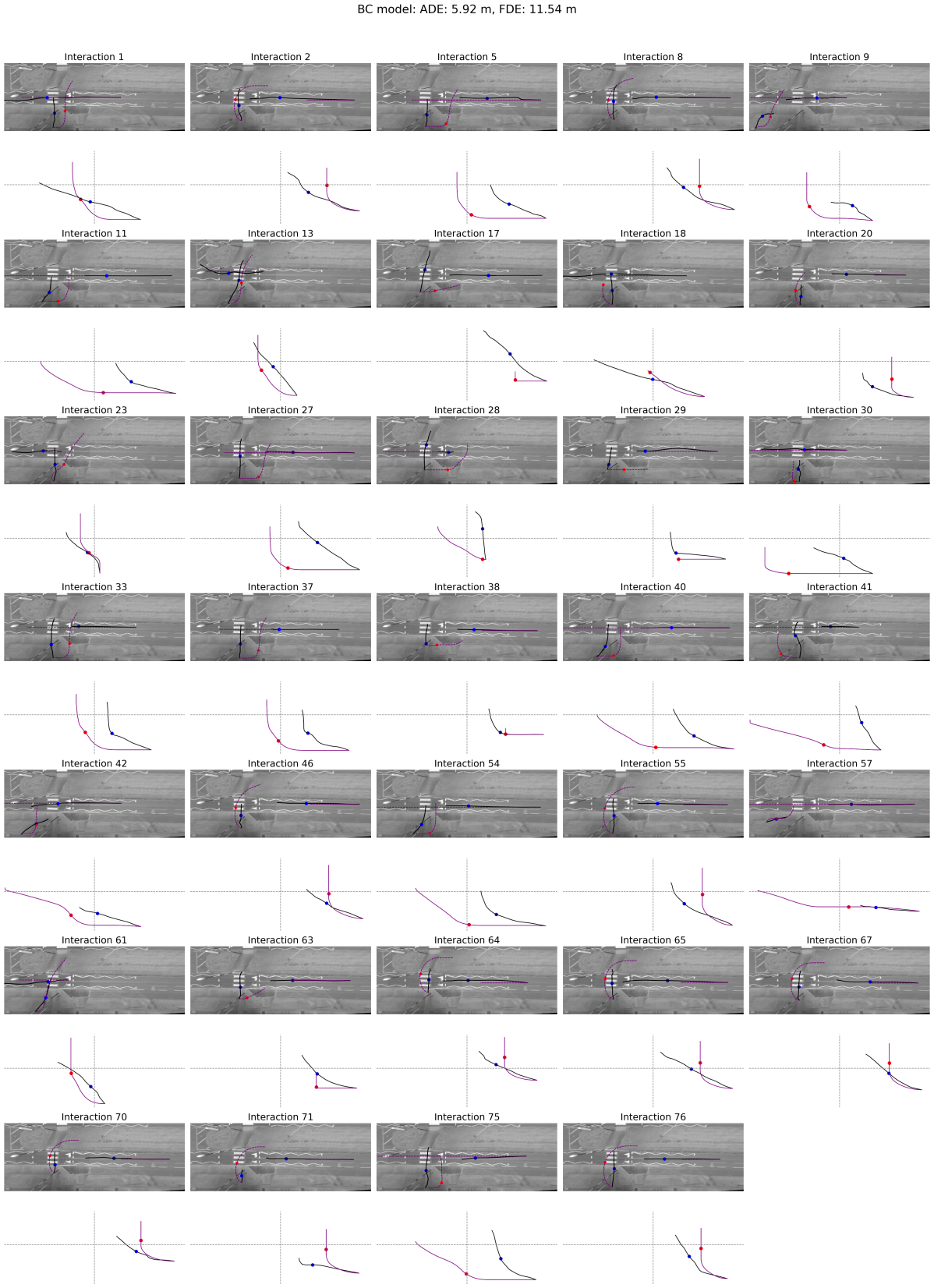

Figure C.2: Model and human trajectories for the Behavioural Cloning (BC) model across all pedestrian-vehicle interactions. For each interaction, the top panel shows the trajectories overlaid on the real-world road layout, while the bottom panel depicts the relative spatial positions of the vehicle (x-axis: longitudinal distance to crossing point) and pedestrian (y-axis: lateral distance to crossing point) at each time step. Human trajectories are shown in black, and model trajectories are shown in purple. Red and black dots indicate the 3-second positions of the model and human agents, respectively.

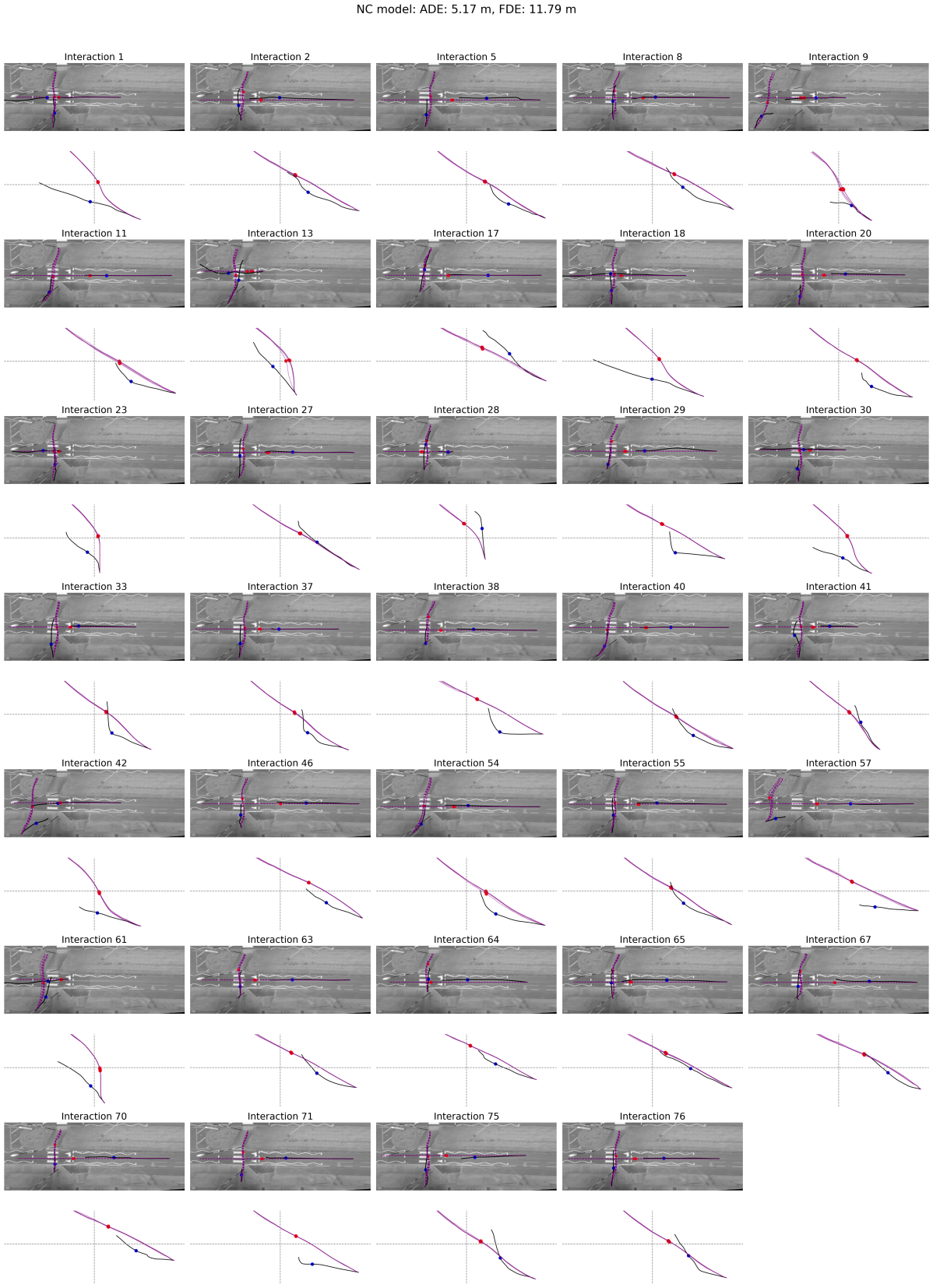

Figure C.3: Model and human trajectories for the No-Constraint (NC) model across all pedestrian-vehicle interactions.

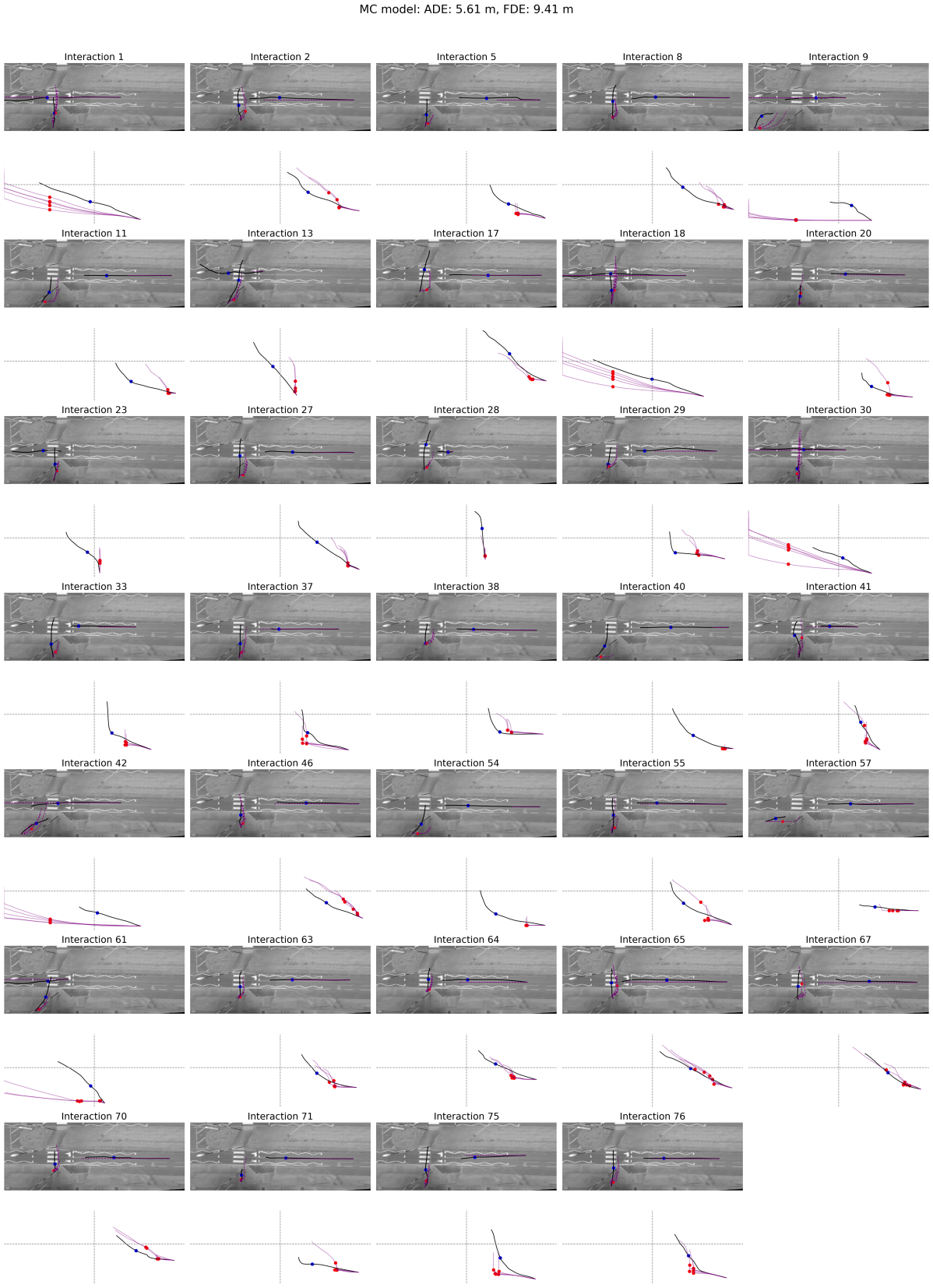

Figure C.4: Model and human trajectories for the Motor-Constraint (MC) model across all pedestrian-vehicle interactions.

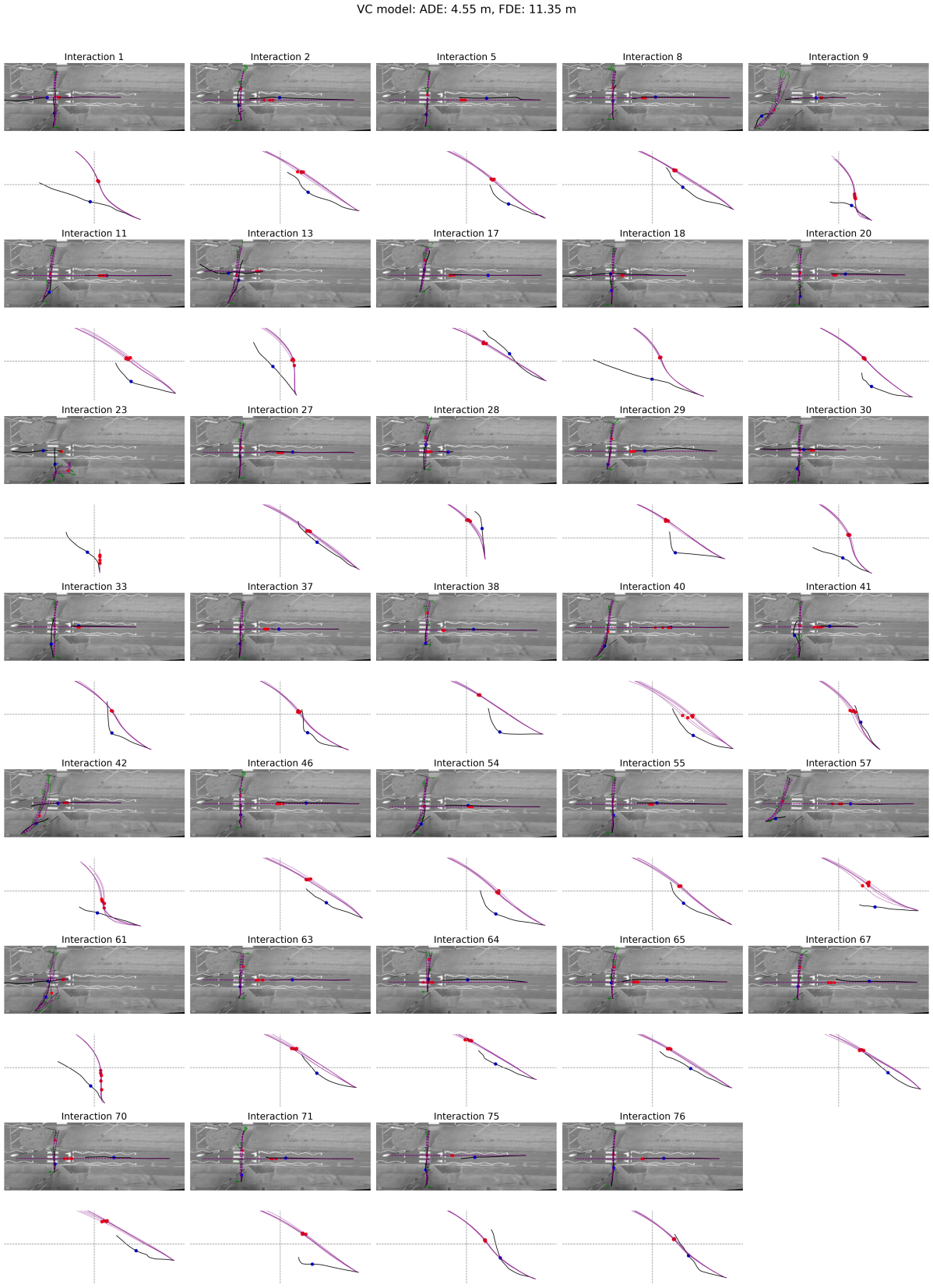

Figure C.5: Model and human trajectories for the Visual-Constraint (VC) model across all pedestrian-vehicle interactions.

Figure C.6: Model and human trajectories for the Visual and Motor-Constraint (VMC) model across all pedestrian-vehicle interactions.