Understanding Pedestrian Behavior in Shared Spaces: A Virtual Reality Study of Automated Shuttle Interactions

Danya Li[1], Yan Feng[2], and Rico Krueger[1]

[1]Department of Technology, Management, and Economics, Technical University of Denmark, Denmark
[2]Department of Transport and Planning, Delft University of Technology, Netherlands

Short summary

The integration of automated shuttles into shared urban spaces presents challenges due to the absence of traffic rules and the complex dynamics of pedestrian interactions. This paper presents a Virtual Reality (VR) study simulating diverse pedestrian-automated shuttle interactions in a shared space environment, including approaching from various angles and navigating continuous traffic. We focus on key behaviors, such as hesitation, deviation, gazing, and proxemics, that play a critical role in shaping safe and efficient interactions. Additionally, we propose a hybrid model that combines psychological principles with deep learning for trajectory and intention prediction for enhanced interpretability and accuracy. This approach aims to promote the harmonious integration of automated shuttles into urban environments, ensuring safer and more efficient shared spaces.

Keywords: pedestrian-automated vehicle interaction; shared space; Virtual Reality; neural stochastic differential equation.

1 Introduction

The emergence of automated shuttles offers promising solutions for last-mile transportation, yet their integration into urban environments introduces risks and uncertainties, especially for vulnerable road users like pedestrians. While autonomous vehicles (AVs) aim to enhance safety, their operation in shared spaces - urban environments without imposed traffic rules - poses unique challenges due to the diversity and complexity of interactions. Pedestrians in these shared spaces often exhibit significant behavioral changes when encountering small, automation-capable vehicles (Wang et al., 2022). Thus, understanding and predicting pedestrian behavior in these contexts is crucial for fostering a harmonious coexistence between pedestrians and automated shuttles.

AV-pedestrian interactions in shared spaces have been explored in real-world scenarios, but these efforts usually focus on pedestrian-vehicle interactions modeled by social force methods (Golchoubian et al., 2023). Existing datasets (Robicquet et al., 2016; Yang et al., 2019; Zhou et al., 2020) largely feature conventional cars, which differ significantly from automated shuttles in size, speed, and interaction dynamics. Besides, field experiments with automated shuttles (De Ceunynck et al., 2022) reported rare occurrences of critical situations due to the conservative settings and low speeds of automated shuttles. Moreover, these observational studies typically rely on vehicle-perspective data, leaving human-perspective data such as gaze behavior, underexplored.

Immersive technologies have also been employed in controlled studies to understand such an interaction in shared space. For instance, Woodman et al. (2019) examined pedestrian gap acceptance with a platoon of shuttles at varying speeds, gaps, and environments using virtual reality, while Feng et al. (2024) explored the influence of factors including social context, external human-machine interfaces (eHMIs), vehicle deceleration styles, and road conditions. Andrijanto et al. (2022) studied crossing and parallel interactions between pedestrians and automated shuttles using the cave automatic virtual environment. However, these studies generally focus on perpendicular crossing interactions, leaving other versatile interactions occurring in shared spaces unexplored. As a result, interaction models have been constrained to crossing-related variables and are often based on relatively simplistic statistical approaches.

To address these gaps, this paper presents a virtual reality study simulating interactions between pedestrians and automated shuttles in a shared space. More specially, we firstly explore previously unstudied scenarios involving shuttles approaching from different angles and/or navigating within continuous traffic. Secondly, we collect a diverse dataset incorporating human signals such as eye gaze to examine people's micro-level attention and investigate evasive and hesitation behaviors to enhance the understanding of safe and efficient pedestrian-automated shuttle interactions in shared spaces. Lastly, we will develop a hybrid model integrating cognitive theories with deep learning techniques for pedestrian trajectory prediction, aiming at achieving enhanced interpretability and accuracy.

2 Methodology

Modeling methods

Our approach is inspired by the drift-diffusion model (DDM) (Ratcliff, 1978), a psychological model of decision-making. The DDM describes evidence accumulation over time, starting from an *initial bias* and progressing with a *drift rate* until reaching a *decision boundary.* Mathematically, this process can be expressed as a stochastic differential equation (SDE):

$$dz = f(z,t)dt + g(z,t)dW \tag{1}$$

where z is the accumulated evidence, initialized as $z(0)$ (the starting bias); $f(z,t)$ is the drift rate, typically a linear function of relevant variables over time; $g(z,t)$ is the diffusion term capturing stochastic variability and dW is a Wiener process. Providing a cognitive perspective, it has been applied to model the first passage time in road-crossing contexts (Giles et al., 2019; Pekkanen et al., 2022).

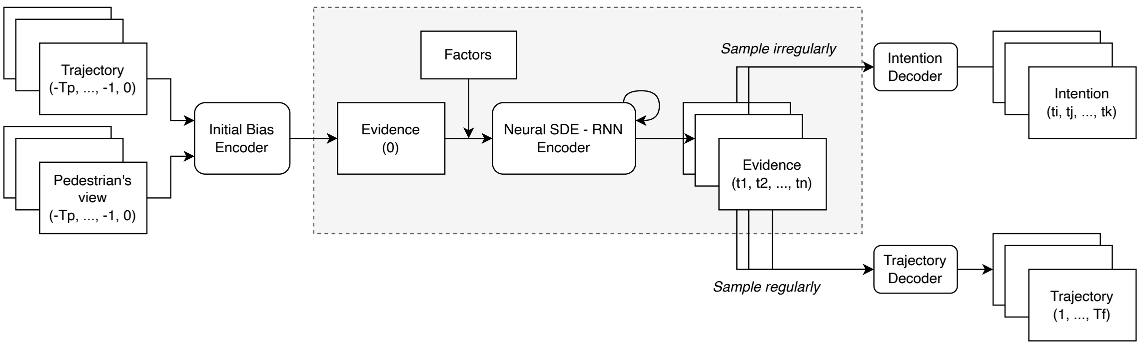


Figure 1: An overview of our proposed model.

Building on this foundation, we propose a neural extension of the DDM to predict pedestrian's decisions $X_{[1,\ldots,T_f]}$ based on past observations $X_{[-T_p,\ldots,0]}$. Our model, illustrated in Fig.1, consists of the following components:

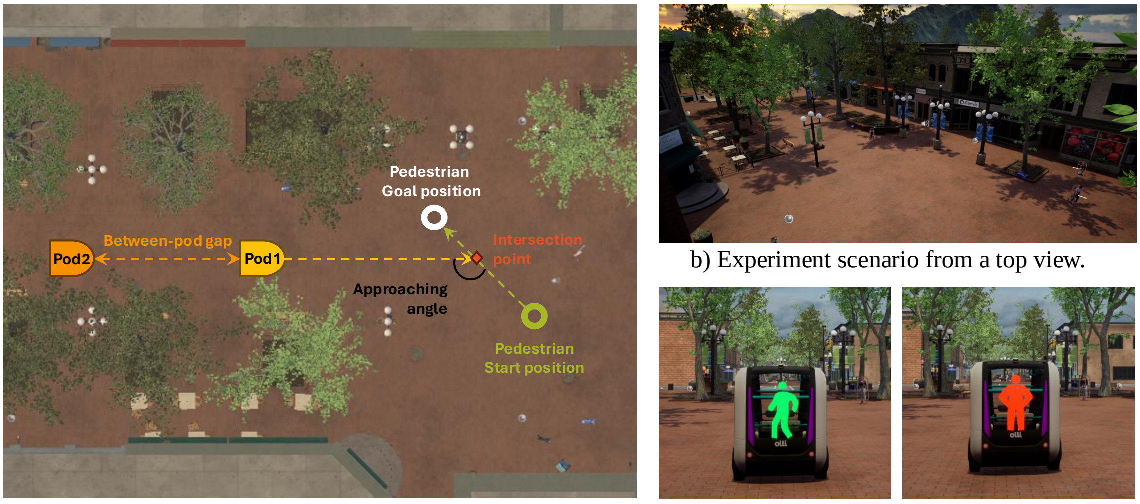
- *Initial bias encoder*: This module processes sequences of past trajectory data and visual inputs from the pedestrian's perspective to estimate the initial evidence (or starting bias).
- *Neural SDE encoder*: Extending the conventional SDE formulation, this module uses neural networks to parameterize the drift $f(x,t)$ term and/or diffusion term $g(x,t)$ (Tzen & Raginsky, 2019). We implement this in a recurrent neural network (RNN) structure to enable flexible, continuous-time latent evidence accumulation while incorporating various experimental factors as inputs.
- *Trajectory decoder*: This component translates the latent evidence into future trajectories, generating predictions at regular time intervals.
- *Intention decoder*: Using the same latent evidence, this module enables irregular sampling to model hesitation behaviors, which can be critical in ambiguous shared space scenarios.

In summary, our model combines the interpretability of the DDM with the flexibility of neural networks to jointly predict pedestrian intention and trajectory. By accounting for hesitation and integrating the pedestrian's visual perspective, it aligns closely with real-world behavior. Furthermore, the model's ability to capture complex, nonlinear dynamics enables it to extend beyond traditional psychological frameworks, offering a robust solution for understanding pedestrian decision-making in shared spaces.

Experiment design

We develop an immersive virtual reality experiment to examine how various factors influence pedestrian-automated shuttle interaction.

Virtual environment: The virtual environment of a shared space resembling Delft streets is selected in VR using Unreal Engine 5 (see Fig. 2 b), mimicking a shopping center. The environment has no elements that indicate the right of way, i.e., no traffic lights, no stop signs, no pedestrian zebra, and no lane specification for automated shuttles. An audio effect of a noisy shared space and automated shuttles and background pedestrians are added to enhance the realism of VR experiments.



a) An illustration of factor definition and experiment setup in the virtual environment.

b) Experiment scenario from a top view.

c) Pod's eHMI with pedestrian sign.
(Left) Yielding case; (Right) Non-yielding case.

Figure 2: An overview of experiment setup.

Factor design: Four within-subject variables were included, namely shuttle behavior (i.e., yielding, non-yielding), eHMI presence (i.e., with eHMI, without eHMI), approaching angle (i.e., 45, 90, 135 degrees), and the number of shuttles (i.e., single shuttle, two shuttles with a gap of 3 seconds, two shuttles with a gap of 5 seconds). An illustration of the last two variables is shown in Fig. 2 a), and a detailed explanation is shown below.

- *Shuttle behavior:* The shuttle can either stop or not stop for pedestrians in a shared space environment. For non-yielding cases, automated shuttles maintain the speed. For yielding cases, we choose the design of type I deceleration profile in Feng et al. (2024) and adapt it to detect pedestrians from 12 meters.
- *eHMI presence:* The eHMI is designed with a pedestrian sign displayed on the front window when it is activated. Depending on the shuttle's behavior, a red pedestrian sign means it cannot stop for the pedestrian while a shuttle signaling a green pedestrian sign means it can stop in front of the pedestrian. The displays are shown in Fig. 2 c).
- *Approaching angles:* The angle between the forward directions of the automated shuttle and the pedestrian forms the approaching angle. The 90-degree angle resembles the typical crossing scenario. Additionally, we consider 45 and 135 degrees to examine how visibility influences such an interaction.
- *Multiple shuttles:* We decide gap size of 3 and 5 seconds from our pilot study.

Sampling method: The combination of all variables yields 36 scenarios. However, to reduce participant fatigue, we determine 12 scenarios per participant based on a pilot study and use the Latin hypercube sampling (LHS) method (McKay et al., 1979) to sample from all combinations. Unlike random sampling, LHS stratifies the range of each factor into equal intervals and ensures that each interval is sampled exactly once, preventing overlap across the different factors. This approach preserves the Latin square property and thus provides more representative samples.

Experimental task: At the beginning of each scenario, there is a green and a white circle indicating the start and goal positions of the interaction task, respectively. Participants are required to step into the green circle, orient themselves toward the goal, and press the controller when they are ready to walk. Upon walking, the automated shuttle starts to approach the pedestrian from a distance of 17.3 meters at a speed of 15 km/h following the shared space speed limit in the Netherlands. Participants are instructed to reach the goal safely as they would do in daily life. The trial ends once participants reach the goal position, after which the next scenario is initiated.

Experiment procedure

The experiment procedure is summarized in Fig. 3. Upon arrival, participants are provided with written information about the experiment, including automated shuttle and eHMI design, procedure, and tasks. They then read and sign a consent form to confirm participation. Next, participants are equipped with a calibrated VR headset and start a familiarization phase consisting of two stages: 1) to practice moving freely in the virtual environment and 2) to familiarize the experiment task and the display and behavior of automated shuttles. The formal experiment begins after familiarization. Participants are randomly assigned to one of 12 sampled scenarios. After completing each scenario, they return to the predefined start position for the next trial. After finishing the VR experiment, participants remove the headset and fill in five questionnaires. Each participant receives 15 euros as compensation. Ethical approval was granted by the Human Research Ethics Committee of Delft University of Technology (Reference ID 4888).

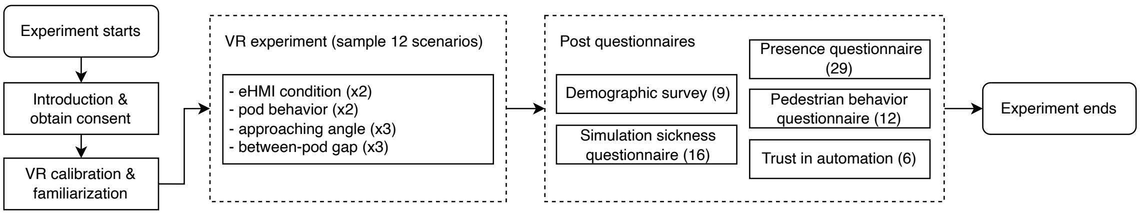


Figure 3: Experiment procedure. The numbers in the VR experiment block represent the levels of each variable, while the numbers in the post-questionnaire block indicate the number of items in each questionnaire.

Experiment apparatus

The VR experiment was conducted in a 10×5 meter room. HTC VIVE Pro Eye headset (resolution: 1440×1600 pixels per eye, 110 field-of-view, 90 Hz refresh rate) was used during the experiment. We also used the real walking locomotion style, allowing users to have continuous movements and rotations in the real world, which can be matched under a 1:1 scheme to the virtual environment.

Data Collection

Two types of data were collected during the experiment, including objective and subjective data.

Objective data: The whole interaction process in the virtual environment was recorded. The data recording started when participants left the start position (at the green circle) and ended when they reached the goal position (at the white circle). The collected data include: 1) timestamp, 2) participant's position, 3) head orientation, 4) eye gaze direction, fixation point, and fixation object. All data were recorded at 20 Hz.

Subjective data: The subjective part includes several questionnaires, namely 1) personal characteristics, 2) simulation sickness questionnaire, 3) presence questionnaire, 4) pedestrian behavior questionnaire, and 5) trust in automation. Personal characteristics include age, gender, nationality, achieved highest education level, dominant hand, walking frequency in daily life, familiarity with VR, familiarity with the concept of automated shuttles, and previous experience regarding interaction with automated shuttles. The simulation sickness questionnaire was adopted from (Kennedy et al., 1993) to measure the experienced simulation sickness of participants in the virtual environment. The presence questionnaire from (Witmer et al., 2005) measures the feeling of presence in the virtual environment. The pedestrian behavior questionnaire (Deb et al., 2017) measures three behavior scores that are relevant to people's attention, namely violation, lapse, and positive behaviors. Finally, trust in automation was adopted from (Payre et al., 2016).

Participants

In total, 51 participants aged between 21 and 61 (M=26.62, SD=5.76) were recruited. All participants had normal vision or corrected vision and normal mobility. None of the participants dropped the experiment due to motion sickness. The participant characteristics are shown in Tab. 1.

Table 1: Demographic information of participants.

Description	Category	Number (Percentage)	Description	Category	Number (Percentage)
Gender	Male	27 (52.94%)	Dominant hand	Right hand	47 (92.16%)
	Female	24 (47.06%)		Left hand	4 (7.84%)
Highest education level	High school or eq.	2 (3.92%)	Previous experience with VR	Never	17 (33.33%)
	Associate degree or eq.	0 (0.00%)		Seldom	23 (45.10%)
	Bachelor's degree or eq.	19 (37.25%)		Sometimes	10 (19.61%)
	Master's degree or eq.	26 (50.98%)		Often	1 (1.96%)
	Doctoral degree or eq.	4 (7.94%)		Very often	0 (0.00%)
Familiarity with the concept of automated shuttles	Not at all	5 (9.80%)	Previous experience with automated shuttles	No experience	27 (52.94%)
	Slightly familiar	17 (33.33%)		Little experience	13 (25.49%)
	Moderate familiar	19 (37.25%)		Some experience	10 (19.61%)
	Very familiar	7 (13.73%)		A lot of experience	1 (1.96%)
	Extremely familiar	3 (5.88%)		Extensive experience	0 (0.00%)

3 Results

We expect to complete the inferential analysis on pedestrian trajectory and intention prediction before the conference. Here, we present some statistical results first. We derive the following metrics from the experimental data collected, categorized into five groups:

- Movement behaviors:
  - *Gap selection*: The gap participants choose to cross.
  - *Waiting time*: The time from the trial start to the participants' last movement initiation.
- Hesitation behaviors:
  - *Initiation count*: The number of crossing attempts, including the final successful attempt.
  - *Backward count*: The number of times participants stepped backward during the trial.
- Deviation behaviors:
  - *Mean deviation*: The average lateral offset between the participant's actual path and the ideal straight-line path.
  - *Max. deviation*: The maximal lateral offset between the participant's actual path and the ideal straight-line path.
- Gazing behaviors:
  - *Pre-crossing shuttle-gazing pct.*: The percentage of time that participants looked at the shuttle before starting their crossing.
  - *During-crossing shuttle-gazing pct.*: The percentage of time that participants focused on the shuttle while crossing.
- Proxemics:
  - *Lateral clearance*: The lateral distance between the participant and the shuttle as it passed.

Each participant completed 12 trials in total. The total number of trials amounts to 612. Removing those mistakenly triggered and bad signals, we obtained 580 trials. Table. 2 shows descriptive statistics for the five groups of indicators described above. Yielding shuttle behavior generally makes crossing easier and reduces hesitation behavior, as reflected by shorter waiting time, fewer initiation counts, smaller distance, and pre-crossing shuttle-gazing, although pedestrians spend more time observing the shuttle during the crossing. Interestingly, the presence of eHMI seems to increase hesitation behavior, possibly due to its timing being relatively late. However, overall, eHMI presence has minimal impact on deviation and gazing behaviors. Approaching angle significantly affects behavior, with acute angles leading to greater hesitation, larger deviations, and reduced lateral clearance. Poor visibility at acute angles (e.g., 45 degrees) also appears to result in reduced attention to traffic. Finally, scenarios involving two shuttles with 5-second gaps increase hesitation and reduce lateral clearance, highlighting the added complexity in decision-making under such conditions.

In summary, these shared-space interactions highlight several variations in hesitation, spatial negotiation, and trajectory deviations across different scenarios, revealing critical differences in how pedestrians respond to ambiguous situations. These results provide valuable empirical data to understand pedestrian behavior, which is essential for designing more effective predictive models.

Table 2: Descriptive statistics for movement, hesitation, deviation, gazing, and proxemics variables.

Measure	Shuttle behavior		eHMI presence		Approaching angle			Multiple shuttles		
	Non-yielding	Yielding	Absent	Present	45°	90°	135°	single shuttle	2 shuttles (gap=3s)	2 shuttles (gap=5s)
Gap selection [count]										
before the first shuttle	88	282	190	180	128	111	131	111	130	129
between two shuttles	118	4	62	60	28	54	40	84	2	36
after two shuttles	87	1	39	49	31	32	25	0	61	27
Waiting time [second]										
Mean	7.25	5.15	6.17	6.25	6.22	6.33	6.09	5.42	6.94	6.28
Standard deviation	4.64	3.57	4.26	4.30	4.30	4.14	4.41	3.28	4.77	4.52
Initiation count										
Mean	1.71	1.62	1.63	1.70	1.68	1.65	1.66	1.64	1.65	1.70
Standard deviation	0.67	0.69	0.65	0.71	0.72	0.59	0.73	0.66	0.64	0.75
Backward count										
Mean	0.19	0.07	0.12	0.15	0.18	0.07	0.15	0.10	0.12	0.18
Standard deviation	0.47	0.28	0.38	0.40	0.45	0.28	0.42	0.30	0.36	0.49
Mean deviation [cm]										
Mean	21.92	20.33	21.37	20.90	24.01	12.85	26.72	19.98	22.28	21.27
Standard deviation	21.99	19.48	21.25	20.34	25.33	7.71	22.42	16.97	22.90	22.07
Max. deviation [cm]										
Mean	46.84	44.04	45.96	44.94	48.92	28.34	59.35	41.45	48.82	46.14
Standard deviation	43.10	36.13	39.49	40.16	44.76	16.15	45.03	32.10	44.88	41.25
Pre-crossing shuttle-gazing pct.										
Mean	46.44%	43.80%	45.01%	45.26%	40.95%	46.43%	47.82%	45.32%	45.98%	44.10%
Standard deviation	0.30	0.31	0.31	0.30	0.31	0.28	0.32	0.29	0.32	0.32
During-crossing shuttle-gazing pct.										
Mean	16.15%	26.05%	21.97%	20.12%	14.00%	18.30%	30.53%	18.45%	22.20%	22.54%
Standard deviation	0.19	0.25	0.24	0.21	0.14	0.19	0.28	0.21	0.25	0.22
Lateral clearance [cm]										
Mean	110.45	128.63	110.64	111.10	109.14	120.87	99.25	115.58	108.55	106.95
Standard deviation	41.99	77.87	46.42	39.67	35.18	49.93	36.28	47.22	39.93	39.83

4 Conclusions

This study explores the intricate dynamics of pedestrian-automated shuttle interactions in shared spaces through a virtual reality experiment. By simulating diverse interaction scenarios, including varied approaching angles and continuous traffic patterns, we uncover critical behaviors such as hesitation, deviation, gazing, and proxemics, which are often overlooked yet essential for ensuring safety and efficiency in shared environments. Additionally, our hybrid model, integrating psychological principles with deep learning, could potentially enhance trajectory prediction accuracy while maintaining interpretability. Statistical analyses highlight the diversity of pedestrian behaviors observed in our experiment, providing a solid foundation for future model development and advancing our understanding of pedestrian-automated shuttle interactions.

References

Andrijanto, A., Chen, Z., Kodama, T., Yano, H., & Itoh, M. (2022). Application of LargeSpace for Investigating Pedestrians' Behaviors when Interacting with Autonomous Vehicles in Shared Spaces. In *Proceedings - 2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops, VRW 2022.*

Deb, S., Strawderman, L., Carruth, D. W., DuBien, J., Smith, B., & Garrison, T. M. (2017, November). Development and validation of a questionnaire to assess pedestrian receptivity toward fully autonomous vehicles. *Transportation Research Part C: Emerging Technologies*.

De Ceunynck, T., Pelssers, B., Bjørnskau, T., Aasvik, O., Fyhri, A., Laureshyn, A., … Martensen, H. (2022, August). Interact or counteract? Behavioural observation of interactions between vulnerable road users and autonomous shuttles in Oslo, Norway. *Traffic Safety Research*.

Feng, Y., Xu, Z., Farah, H., & van Arem, B. (2024). Does Another Pedestrian Matter? A Virtual Reality Study on the Interaction Between Multiple Pedestrians and Autonomous Vehicles in Shared Space. *IEEE Transactions on Intelligent Transportation Systems*.

Giles, O. T., Markkula, G., Pekkanen, J., Yokota, N., Matsunaga, N., Merat, N., & Daimon, T. (2019, July). *At the Zebra Crossing: Modelling Complex Decision Processes with Variable-Drift Diffusion Models* (preprint).

Golchoubian, M., Ghafurian, M., Dautenhahn, K., & Azad, N. L. (2023). Pedestrian Trajectory Prediction in Pedestrian-Vehicle Mixed Environments: A Systematic Review. *IEEE Transactions on Intelligent Transportation Systems*.

Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993, July). Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. *The International Journal of Aviation Psychology*.

McKay, M. D., Beckman, R. J., & Conover, W. J. (1979). A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. *Technometrics*.

Payre, W., Cestac, J., & Delhomme, P. (2016, March). Fully Automated Driving: Impact of Trust and Practice on Manual Control Recovery. *Human Factors*.

Pekkanen, J., Giles, O. T., Lee, Y. M., Madigan, R., Daimon, T., Merat, N., & Markkula, G. (2022, March). Variable-Drift Diffusion Models of Pedestrian Road-Crossing Decisions. *Computational Brain & Behavior*.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological review*.

Robicquet, A., Sadeghian, A., Alahi, A., & Savarese, S. (2016). Learning Social Etiquette: Human Trajectory Understanding In Crowded Scenes. In *Computer Vision – ECCV 2016.*

Tzen, B., & Raginsky, M. (2019, October). *Neural Stochastic Differential Equations: Deep Latent Gaussian Models in the Diffusion Limit.*

Wang, Y., Hespanhol, L., Worrall, S., & Tomitsch, M. (2022, September). Pedestrian-Vehicle Interaction in Shared Space: Insights for Autonomous Vehicles. In *Proceedings of the 14th International Conference on Automotive User Interfaces and Interactive Vehicular Applications.*

Witmer, B. G., Jerome, C. J., & Singer, M. J. (2005, June). The Factor Structure of the Presence Questionnaire. *Presence: Teleoperators and Virtual Environments*.

Woodman, R., Lu, K., Higgins, M. D., Brewerton, S., Jennings, P. A., & Birrell, S. (2019, November). Gap acceptance study of pedestrians crossing between platooning autonomous vehicles in a virtual environment. *Transportation Research Part F: Traffic Psychology and Behaviour*.

Yang, D., Li, L., Redmill, K., & Ozguner, U. (2019, June). Top-view Trajectories: A Pedestrian Dataset of Vehicle-Crowd Interaction from Controlled Experiments and Crowded Campus. In *2019 IEEE Intelligent Vehicles Symposium (IV).*

Zhou, W., Berrio, J. S., De Alvis, C., Shan, M., Worrall, S., Ward, J., & Nebot, E. (2020). Developing and Testing Robust Autonomy: The University of Sydney Campus Data Set. *IEEE Intelligent Transportation Systems Magazine*.