Internal State Estimation in Crowds via Active Information Gathering

XUEBO JI, The University of Hong Kong, Hong Kong, Hong Kong
ZHERONG PAN and XIFENG GAO, Tencent America, Bellevue, Washington, USA
LEI YANG, Centre for Transformative Garment Production, The University of Hong Kong, Hong Kong, Hong Kong
XINXIN DU, Tsinghua University, Beijing, China
KAIYUN LI, University of Jinan, Jinan, China
YONGJIN LIU, Tsinghua University, Beijing, China
WENPING WANG, The University of Hong Kong, Hong Kong, Hong Kong
CHANGHE TU, Shandong University, Qingdao, China
JIA PAN, The University of Hong Kong, Hong Kong, Hong Kong

Accurately estimating human internal states, such as personality traits or behavioral patterns, is critical for enhancing the effectiveness of human–robot interaction, particularly in multi-agent settings. These insights are key in applications ranging from social navigation to autism diagnosis. However, prior methods are limited by scalability and passive observation, making real-time estimation in complex, multi-human settings difficult. In this work, we propose a practical method for active human personality estimation in crowds, with a focus on applications related to Autism Spectrum Disorder (ASD). Our method combines a personality-conditioned behavior model, based on the Eysenck 3-Factor theory, with an active robot information-gathering policy that triggers human behaviors through a receding-horizon planner. The robot's belief about human personality is then updated via Bayesian inference. We demonstrate the effectiveness of our approach through proof-of-concept studies in simulation, user studies with typical adults, and preliminary experiments involving participants with ASD. Our results show that our method can scale to tens of humans and reduce personality estimation error by 29.2% and uncertainty by 79.9% in simulation compared to the passive baseline. User studies with typical adults confirm the method's ability to generalize across complex personality distributions. Additionally, we explore its application in autism-related scenarios, demonstrating that the method can identify

This work was partially supported by the Natural Science Foundation of China (62461160309, U2336214, U23A20312), the NSFC-RGC Joint Research Scheme (N_HKU705/24), the Jiangsu-Hong Kong-Macau Project (BZ2024061, GRF 17201025, GRF 17200924), the Shandong Provincial Natural Science Foundation (No. ZR2025MS333), and the China Postdoctoral Science Foundation under the Grant (2024M751591).
Authors' Contact Information Xuebo Ji, The University of Hong Kong, Hong Kong, Hong Kong; e-mail: xbji@connect.hku.hk; Zherong Pan, Tencent America, Bellevue, Washington, USA; e-mail: zherong.pan.usa@gmail.com; Xifeng Gao, Tencent America, Bellevue, Washington, USA; e-mail: gxf.xisha@gmail.com; Lei Yang, Centre for Transformative Garment Production, The University of Hong Kong, Hong Kong, Hong Kong; e-mail: l.yang@transgp.hk; Xinxin Du, Tsinghua University, Beijing, China; e-mail: duxx@tsinghua.edu.cn; Kaiyun Li, University of Jinan, Jinan, China; e-mail: sep_liky@ujn.edu.cn; Yongjin Liu, Tsinghua University, Beijing, China; e-mail: liuyongjin@tsinghua.edu.cn; Wenping Wang, The University of Hong Kong, Hong Kong, Hong Kong; e-mail: wenping@cs.hku.hk; Changhe Tu, Shandong University, Qingdao, China; e-mail: chtu@sdu.edu.cn; Jia Pan (corresponding author), The University of Hong Kong, Hong Kong, Hong Kong; e-mail: jpan@cs.hku.hk.

ACM 2573-9522/2025/12-ART28
https://doi.org/10.1145/3772065

the difference between neurotypical and autistic behavior. The results suggest that our framework could serve as a foundation for future ASD-specific applications.

CCS Concepts: • **Computing methodologies** → **Artificial intelligence**; **Active learning settings**; Modeling and simulation; • **Applied computing** → *Psychology*; • **Computer systems organization** → *Robotics*;

Additional Key Words and Phrases: Internal State Estimation, Personality Estimation, Behavior Prediction, Autism Spectrum Disorder, Robot–Crowd Interaction, Active Learning

ACM Reference format:
Xuebo Ji, Zherong Pan, Xifeng Gao, Lei Yang, Xinxin Du, Kaiyun Li, Yongjin Liu, Wenping Wang, Changhe Tu, and Jia Pan. 2025. Internal State Estimation in Crowds via Active Information Gathering. *ACM Trans. Hum.-Robot Interact.* 15, 2, Article 28 (December 2025), 33 pages.
https://doi.org/10.1145/3772065

1 Introduction

Human internal state estimation has been an active interdisciplinary research area in the field of human–robot interaction over the past decades [5, 9, 55, 59]. Human behavior is generally influenced by internal states, which are not directly accessible to a robot during interaction [55]. For example, a driver's driving style—whether aggressive or defensive—can significantly impact their driving behavior in interactive traffic scenarios. Similarly, human navigation behavior is partially influenced by personality traits [9]; children with autism often demonstrate restricted and repetitive patterns of behavior during social interactions [53]. Accurately estimating these internal states is essential for the robot to obtain more information and sustain an effective interaction process. On one hand, valuable information can be extracted through behavioral data analysis, such as a robot identifying signs of autism based on observed actions. On the other hand, the estimation of the human state can enhance human–robot interaction, enabling the robot to achieve its objectives more efficiently and precisely. For example, with an accurate prediction of an autistic child's behavior, the robot could devise more effective intervention strategies. Recently, there has been a growing focus on methods where robots predict human behavior and employ the "coupled" method that incorporates a human model in the planning for joint robot–human actions [5, 8, 54].

However, the estimation of human internal states during interactions presents significant challenges. The robot must perform multiple layers of logical inference, specifically to predict human responses following its actions and evaluate whether these responses facilitate state estimation and interaction. In addition, interactions commonly occur in multi-agent settings, such as driving scenarios, social navigation, and group activities for children with autism in special education schools. The complex, nested interactions between multiple individuals in these multi-agent contexts pose significant challenges for accurate state estimation. In this work, the general goal is to enable the robot to actively estimate human internal states in crowds.

A row of theoretical frameworks has been proposed to estimate human states and behaviors. The **Partially Observable Markov Decision Process (POMDP)** lays the theoretical foundation for solving such problems [35]. POMDP models the robot–human interaction as a joint dynamic system conditioned on the human internal states, aka intentions or behavioral patterns. It then jointly updates the robot behaviors and estimates the internal states via Bayesian rules. This framework, however, is limited in its scalability. The human internal states have to be estimated in time in order for them to be effectively exploited [5], but directly solving POMDP is computationally expensive. In multi-agent settings, the robot simultaneously interacts with multiple people, where estimating the state for each individual via POMDP is intractable. To improve the computational efficiency of

POMDP, several approximations have been proposed for similar problem formulations [25, 34]. However, these methods only update the human internal states from "*passive*" *observations* [55], i.e., it cannot explore human behaviors that are important for interaction but are not exhibited without external triggers. There have been prior works that partially address the aforementioned problems, but a complete solution is yet to be presented.

In order to improve the efficacy of human internal state estimation, prior works have combined multi-modal information, including audio [28], vision [9, 43], gesture [44], and unintentional human behaviors [47]. Modern deep-learning approaches [1] can scale up to detect several human or vehicle trajectories at the same time, but they still rely on passive observations. One pioneering work on "active" information gathering is presented in [55]. It models the human and robot as a joint differential dynamic system and computes a robot trajectory that maximizes the information gain of human driving style via **Model Predictive Control (MPC)**. However, this method requires solving an expensive optimization problem, where both the human's and the robot's actions are optimized in a nested manner to minimize different objectives. Even for a single human, solving such a problem in real-time can be a major computational burden. Several follow-up works extend the idea of Sadigh et al. [55] to the application of grasping [10] and other cobot applications [14, 50], but they all suffer from limited scalability due to involving POMDP or independent optimization for each human.

To address these issues, we proposed a novel framework for estimating human internal state in multi-agent settings actively and efficiently. Our key insight is to employ active information gathering to acquire crucial information for more effective internal state estimation, while leveraging system-level inference to ensure the maintenance of both estimation quality and processing speed. We emphasize the importance of the concept "active gathering in an interactive context" for efficient internal state estimation and additionally consider the potential ASD-related applications. This work is conducted as a prospective study for assessing the technical feasibility of our insight, where we select *personality* as the internal state and *navigation behavior* as the action space to study the *scalable human personality estimation* problem. Proof-of-concept studies and user evaluations are performed in simulation and on typical human groups, respectively. Although forward-looking, we conducted additional evaluations on the *autistic* groups to explore the connections between the proposed method and the probabilistic estimation of autistic traits. Specifically, our *main contributions* are twofold.

Human Personality Estimation. We study the problem of personality estimation in typical human crowds. As numerous works have linked personality models with heterogeneous crowd behaviors that have obtained reliable results [2, 9, 29, 57], we follow the personality trait theory [22] as Guy et al. [29], which stresses that extensive changes in human behavior are mainly caused by personality traits. For example, when navigating through a crowd, aggressive agents tend to make the most progress along straighter paths, while shy agents are more easily diverted as they attempt to avoid others [29]. Our study primarily focuses on such personality-conditioned movement behaviors, particularly in response to the motions of nearby agents, such as when being approached. We use the established Eysenck 3-Factor personality model [22] based on a 3D behavior descriptor: **Psychoticism, Extraversion, and Neuroticism (PEN)**. We link human behavior in crowds with these three factors and focus on the observed navigation strategy. As a main point of departure from prior works, we propose an efficient and scalable framework to actively estimate personalities for tens of humans. Firstly, we train *a personality-conditioned human behavior network* via the trait theory [29], which could efficiently predict collision-free trajectories. As a result, the trajectories of multiple humans can be predicted recursively in a differentiable manner. Second, we propose a receding-horizon planner to compute robot trajectories that actively interact with crowds to trigger humans' navigation behavior, e.g., making the robot deliberately navigate through the crowd in ways that elicit reactions from nearby humans. Finally, the personalities of multiple observed humans are updated according to observed behaviors via Bayesian rules.

One core aspect of our method is enabling robots to perform active actions to collect more effective information. For example, through approaching actions, the robot may observe a person's response, e.g., a tendency to avoid, which can indicate a preference for limited social interaction, potentially reflecting traits such as shyness. Although passive observation can provide rich behavioral data, it often includes a high proportion of low-informative samples, particularly in sparse environments. For instance, no clear behavioral signals are available if a human is simply walking without engaging with others, and only natural human–human interactions offer meaningful information for personality estimation. Our method does not discard passive observations; instead, it complements them by actively generating informative interactions through intentional robot–human encounters. This form of active probing is especially beneficial in low-density interaction scenarios. However, balancing this active information gathering with *human comfort* during interaction is essential, posing a tradeoff. To address this, we reduce the aggressiveness of the robot's behavior by fine-tuning parameters and applying additional optimization, ensuring human comfort is prioritized throughout the interaction process. Human comfort-related evaluations are conducted through questionnaires and observations in our user studies.

Evaluation on Both Neurotypical and Autistic Groups. As a preliminary work to explore the technical feasibility of applying the proposed method in autism-related applications, we conduct a series of evaluations to validate the effectiveness of our framework and further examine its impact when applied to autistic participants. Specifically:

— We evaluate our method through proof-of-concept studies in *simulation*, to assess the performance of its various components and the advantages of active gathering over a passive approach. We show that our method can scale to tens of humans with unknown personalities in real time. Compared to the passive baseline, the active gathering converges faster on estimation and lowers the personality **Estimation Error (EE)** and uncertainty about personality by 29.2% and 79.9% on average, respectively.
— To validate the effectiveness of our method in estimating internal states of human participants beyond simulated agents, two user studies are conducted on *typical adults*, which show that our active method has effectiveness when applied on complex human personality distributions and can elicit human behaviors, which are conducive to personality estimation.
— To further investigate the applicability of the proposed method in ASD-related contexts, we implemented the proposed method in the form of a PC game and applied it to *participants with autism*. Our results demonstrate that the method can identify the difference between autistic and neurotypical participants based on the estimated internal states, which may be linked to the restricted and repetitive behavior patterns commonly observed in autism.

Overall, this work takes a step toward active internal state estimation in multi-agent settings, focusing on scalable personality estimation as a preliminary study for the ASD-specific applications. Although ASD-specific internal state and action space were not used to develop the model in this work, we still identified meaningful correlations between the proposed framework and ASD-related traits. These findings lay the groundwork for our future research, where we aim to use behavior data collected from ASD groups to refine the model and apply our framework to actively estimate internal states for autistic children, exploring its potential as a robot-assisted diagnosis strategy.

2 Related Work

Our method is closely related to recent progress in human behavior prediction, human personality modeling, multi-agent navigation, and human–robot interaction in autism. We briefly review prior works on each topic.

2.1 Human Behavior Prediction

Human behaviors take the form of navigation trajectories, gestures, languages, and so on, and early works like social **Long Short-Term Memory (LSTM)** [1] focus on detecting and predicting these external behaviors. The human behavior predictor can also be combined with robot planners [6] to improve the efficacy of human–robot collaboration. Some studies also enable robots to be aware of group dynamics within crowds and leverage group-based understanding and prediction to enhance the performance of robot navigation [36, 64]. However, without the explicit modeling of internal states, pure external behavior prediction would require more data for training and is weaker in generalization. This limit can be resolved via POMDP [5], where the robot trajectory computing and the human internal state estimating are conducted simultaneously.

All the aforementioned methods are passive predictions, where the predictor observes humans without interacting with them. The seminal work of Sadigh et al. [54] first notices the possibility for the robot to influence human behaviors, and their follow-up work [55] computes robot influencing trajectories to actively probe and detect a human's driving style. However, their method needs to solve a costly nested optimization to compute the active probing trajectory for a single human. Follow-up works [10, 49] allow the robot to influence humans not only for probing but also to better accomplish the task. However, the robot only interacts with a single human in these applications. Their poor scalability limits their application in multi-agent settings where the robot interacts with multiple humans simultaneously.

2.2 Human Personality Modeling

Personality is one of the most decisive factors in human behaviors [29]. Human personalities have been summarized by various dimensional models, which have been further incorporated into agent simulation systems. The dimensional models assume that the personality is affected by several factors. For example, the OCEAN model [66] identifies five factors, and Allbeck and Badler [2] and Durupinar et al. [20] relate them with low-level simulation parameters via a generative probabilistic model. The Myers-Briggs type indicator model [48] identifies four factors, and Salvit and Sklar [57] implement a swarm of agents covering a subset of the Myers-Briggs model. The personality trait theory [21] identifies three major factors, and Guy et al. [29] link them with the low-level simulator parameters of the **Reciprocal Velocity Obstacles (RVOs)** algorithm [62] using regressive analysis. Our method is based on [29], and we train a personality-conditioned navigation policy to predict the behavior of agents with different personalities.

2.3 Multi-Agent Navigation

The personality estimation in this work is related to social navigation. The classic centralized navigation algorithms [7, 58] assume an omniscient node coordinates and plans motions for a swarm of robots, but they suffer from a heavy communication burden and vulnerability to system error. Some rule-based decentralized microscopic approaches, e.g., the Social Force Model [32] and RVO [63], make humans navigate around by utilizing local information in a small vicinity, while several follow-up works utilize local collaboration to achieve better navigation efficacy or social norm compliance [26, 31]. However, it is difficult to translate task-level objectives, such as time optimality [24] and social awareness [16], into analytic rules. In view of this, learning-based algorithms like inverse **Reinforcement Learning (RL)** [39] and RL [24, 30] propose to learn low-level rules from high-level specifications. However, a common drawback of the above-mentioned techniques is that their behaviors are not personalized for each human. We resolve this problem by training a personality-conditioned navigation policy. Bera et al. [8] proposed a method for socially invisible robot navigation by reducing robot group entitativity through modulation of RVO parameters, using a data-driven mapping between these parameters and entitativity features. Although their approach involves psychology-related traits, it does not aim to infer human internal states from behavior, which is the core focus of our study.

2.4 Human–Robot Interaction in Autism

Traditional methods in ASD-related interventions mainly rely on behavioral therapy, which is time-consuming [40]. Recently, artificial intelligence has increasingly been applied in autism diagnosis, screening, and treatment [3], utilizing multimodal data like electroencephalography [12], eye-tracking [37], and facial and hand movement [18, 42]. Compared to solely relying on observation, interactive robots can better capture the attention of children with autism, improving data collection efficiency and offering more diverse intervention options [56]. Robot-based interventions are used to enhance the autistic children's social skills [60], emotional understanding [11], and imitation abilities [17]. While these approaches have shown some success, they mainly rely on manually designed interaction scenarios, which lack explicit modeling of a child's internal states. As a result, the interactions are limited in intelligence and adaptability, making it difficult to meet the diverse needs of individuals with autism. Although some studies use interaction data to model and estimate states [46], data collection remains "passive," which becomes challenging when children display low levels of expression or have limited behavioral abilities. In this work, we study active personality estimation as a preliminary approach for the ASD-specific applications. Our method shows the potential of using active information gathering and dynamic internal state estimation during interactions, paving the way for more flexible and effective robot-based diagnosis for autism in the future.

3 Problem Formulation and Assumptions

We study the problem of human personality estimation in crowds using active information gathering. We assume that there is a single robot R and N humans $H_{1\dots N}$ with personality descriptors $\phi_{1\dots N}$ navigating in a shared space. Both the robot and humans can observe the physical state of others within a certain range. According to [29], we further assume that the human H_i's navigation pattern is affected by the personality ϕ_i, which is unknown to the robot. Our goal is to estimate the personality of each person through the robot's active probe.

Joint Dynamic System. We represent the robot and humans as a joint dynamic system with state $s = \{s_R, s_{H_1}, \cdots, s_{H_N}\}$. At each timestep, the motion of the robot and all humans is characterized by f:

$$s^{t+1} = f(s^t, \mu_R^t, \mu_{H_{1\dots N}}^t),$$

where μ_R^t and $\mu_{H_{1\dots N}}^t$ are continuous control inputs for the robot and humans, respectively, and the dynamic system f encodes physical rules determining the complex interaction between the robot and humans, such as the collision avoidance. Following the personality trait theory [9, 29], we further assume human's personality ϕ_i is the only latent factor affecting its behavior, which is consistent and independent. Indeed, during the short period of robot–human interaction, the human emotion and goal-biased intention are also relatively stable.

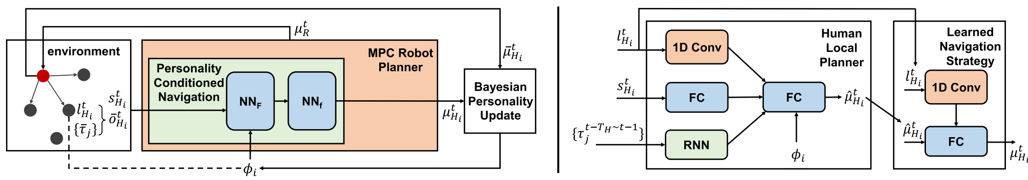


Fig. 1. The pipeline of our system (left) and the neural network structure of our human behavior model (right). In the pipeline, the red circle represents the robot, and the black circles represent humans. At each timestep, the robot observes the human's state $s^t_{H_i}$ and its surroundings $\bar{o}^t_{H_i}$, then uses a pre-trained personality-conditioned navigation policy to infer the human's expected action $\mu^t_{H_i}$ based on its current belief. By comparing $\mu^t_{H_i}$ with the actual observed action $\bar{\mu}^t_{H_i}$, the robot updates its belief over the human's personality ϕ_i using Bayesian inference. Finally, an MPC-based robot planner optimizes the robot control policy to actively probe human personalities. The personality-conditioned navigation policy includes a human local planner, NN_F, which takes the human's sensing data, local state, and neighbor trajectory history as input and outputs a preferred action conditioned on the human's personality. A learned navigation strategy NN_f refines this action to improve collision avoidance based on current sensing input.

Human Planner. We assume that each human H_i is a local planner that determines its behavior via a personality-conditioned function:

$$\mu^t_{H_i} = F(s^t_{H_i}, o^t_{H_i}, \phi_i),$$

where $s^t_{H_i}$ represents human H_i's physical state (e.g., current position p_{H_i}, velocity $\dot{p}_{H_i}$, goal) and $o^t_{H_i}$ is H_i's observation at the time instance t (e.g., states and historical trajectories of humans or robots in H_i's sensing range d_H). It is worth noting that since humans and robot have different goals, following previous work [55], we assume that humans will not make decisions to actively influence the robot.

Robot Planner. The main goal of our robot is estimating a human's personality. We assume that the robot does not know the number of humans in the environment and has only a limited sensing range d_R. To simplify the problem setting and focus on active personality inference, we further assume the robot has perfect sensing within the sensing range. We set up the robot to keep track of a belief $b(\phi_i)$ for each human it has observed and continually refine the belief as it collects more observations about this human. Notably, it does not mean we only consider each human's behavior individually: As humans are interactive individuals that always consider the surrounding environment with collision-avoidance constraints, different humans are interacting with each other and reacting based on their observations. Our method does take such interactions into account. Compared with observing and predicting passively in prior works [9, 43], we make the robot generate behaviors intelligently, which could have an active interaction with the crowds for drawing better inferences. In addition, another challenge of our problem lies in quickly and tractably estimating and updating $b(\phi_i)$ for dozens of humans, rather than dealing with only a single human in prior work [55].

4 Active Human Personality Estimation

In this section, we first describe the details of the dynamic system f and local planner models used in our framework and then introduce our personality estimation scheme. The full pipeline of our method is outlined in Figure 1.

4.1 Learning-Based Personality-Conditioned Navigation

Our method is based on the key observation drawn in prior research [52]: human behavior is influenced by several internal factors. Generally speaking, a human's overall behavior and moving patterns will be biased by their underlying personalities, even when they are pursuing the same task. Recent works [22] and [29] showed that for the particular task of navigation, the low-level simulator parameters of the RVO algorithm [62] and the Eysenck 3-factor personality model (PEN model) [22] have a strong correlation. Different from binary internal state features [55], the PEN model uses a 3D continuous indicator to model the three factors: PEN. Guy et al. [29] derive a mapping between the personality factors and RVO parameters using data-driven linear regression, which involves a wide range of participants, leading to personality-conditioned human behavior. However, their linear regression involves a pseudoinverse, i.e., the number of RVO parameters is larger than that of personality factors. As a result, using insufficient observation to estimate these parameters can lead to undetermined behaviors. Further, the RVO algorithm is non-differentiable, which limits the potential efficacy when applying model-predictive controls to estimate the trajectories for humans.

Human Local Planner. To resolve the above issues, we follow the idea of the personality-conditioned RVO behavior and propose a novel learning-based human local planner. As illustrated in Figure 1 (right), we use a feed-forward neural network to parameterize the human local planner F, which maps human H_i's physical state $s^t_{H_i}$ and its observation:

$$o^t_{H_i} = \left\{ l^t_{H_i}, \{\tau_j^{t-T_H\sim t-1} \mid \|p^t_{H_i} - p^t_j\| \leq d_H, j \in H_{-i} \cup \{R\}\} \right\}, \tag{1}$$

to the intermediate control action $\hat{\mu}^t_{H_i}$. Here, H_i's observation consists of its local sensing measurement $l^t_{H_i}$ and τ_j, which is a short (within time horizon T_H) historical trajectory of other humans H_{-i} or robots R within H_i's sensing range d_H. Following the setting in previous crowd simulation works [41], the local sensing measurement $l^t_{H_i}$ is implemented in a way similar to LiDAR, where m rays within 180° are used to simulate the perception from the front of a person. Finally, our network is also conditioned on H_i's personality ϕ_i. In summary, our local planner can be formulated as the following function:

$$\hat{\mu}^t_{H_i} = \mathrm{NN}_F(s^t_{H_i}, o^t_{H_i}, \phi_i). \tag{2}$$

The inputs to NN_F, $s^t_{H_i}$ and $o^t_{H_i} = \{l^t_{H_i}, \{\tau_j^{t-T_H\sim t-1}\}\}$, have different modalities, for which we use different feature encoders. The current physical state $s^t_{H_i}$ is encoded by additional fully connected layers. To encode the local sensing input $l^t_{H_i}$, we use an 1D convolutional network. To encode the neighbor historical trajectory data $\tau_j^{t-T_H\sim t-1}$, we use a **Recurrent Neural Network (RNN)**. All the encoded features along with the personality ϕ_i are then fed into the feed-forward network to yield the intermediate action $\hat{\mu}^t_{H_i}$.

Learned Navigation Strategy. The above local planner is for generating control actions for humans. However, simply executing such actions would lead to violations of the collision-free constraints among agents or against static obstacles. To mitigate collision avoidance, we follow [24] and train another neural network to learn the joint robot–human dynamic system function f in a decentralized manner. We introduce another feed-forward network NN_f that filters the intermediate human navigation action $\hat{\mu}^t_{H_i}$ to yield the ultimate collision avoidance action $\mu^t_{H_i}$. Our feed-forward network is also conditioned on human H_i's local sensing data:

$$\mu^t_{H_i} = \mathrm{NN}_f(l^t_{H_i}, \hat{\mu}^t_{H_i}). \tag{3}$$

H_i's final control action $\mu^t_{H_i}$ simply takes the form of a desired velocity, and we update the human's position by simply translating it by $\mu^t_{H_i}\Delta t$, where Δt is the timestep size.

Model and Policy Training. Unlike [24], which uses RL, we train both NN_f and NN_F using supervised learning. For training NN_F, we sample a large dataset of trajectories from the ground truth personality-conditioned RVO simulator [29] using randomized goal positions and personality parameters. For NN_f, we fix the physical radius of agents in the simulator and sample a dataset consisting of actions with different lengths and their corresponding RVO-handled collision-free actions. Both networks are trained with l_2-loss between the ground truth actions and the network predictions, with each network trained independently. Our key idea is to equip the robot with the learned personality-conditioned human local planner as a prior. This method significantly reduces the dimension of search space for the human personality. In addition, the learned neural networks offer a level of robustness and smoothness in interactions: the smoothing effect of neural networks helps ensure that crowd behavior does not exhibit drastic fluctuations in response to sudden changes in a surrounding agent's behavior [45]. This capability enables effective responses to abrupt changes in user input during the interaction process. Further, the learned policy allows the robot to recursively predict human behaviors over multiple future steps, without nested optimization as done in [55].

4.2 Personality Estimation

As mentioned before, a human's personality is assumed to be static and independent; thus, the robot could infer each human's personality separately. We have our robot maintaining a Bayesian belief $b(\phi_i)$ over the personality of each observed human H_i. At timestep t, the robot can observe the behavior of humans within its sensing range d_R, i.e., all H_i such that $\|p_{H_i} - p_R\| \leq d_R$. The robot's observation for human H_i is $z^t_{H_i} = \{s^t_{H_i}, \bar{o}^t_{H_i}\}$, which includes two parts: $s^t_{H_i}$ is the observed H_i's current state; $\bar{o}^t_{H_i}$ is the robot's estimation of what H_i should have locally observed. In particular, $\bar{o}^t_{H_i}$ is a subset of $o^t_{H_i}$ in Equation (1) where only H_i's neighbors that can also be observed by the robot are taken into consideration:

$$\bar{o}^t_{H_i} = \{l^t_{H_i}, \bar{\tau}_j^{t-T_H\sim t-1} | \|p^t_{H_i} - p^t_{H_j}\| \leq d_H \wedge \|p^t_R - p^t_{H_j}\| \leq d_R\}, \tag{4}$$

where $\bar{\tau}_j$ represents the historical trajectories of H_i's neighbors who are visible to the robot, which is a subset of τ_j in Equation (1).

Given the robot-observed human behavior $z^t_{H_i}$, together with the human's actual action $\overline{\mu}^t_{H_i}$ observed by the robot, we propose to update the robot's belief $b(\phi_i)$ using the Bayesian rule. In particular, the humans' actual action should follow the probabilistic model below:

$$p(\bar{\mu}^t_{H_i} \mid z^t_{H_i}, \phi_i) = \frac{1}{Z} \exp\left(-\alpha \mathcal{L}\left(\bar{\mu}^t_{H_i}, \mathrm{NN}_f\left(l^t_{H_i}, \mathrm{NN}_F(s^t_{H_i}, \bar{o}^t_{H_i}, \phi_i)\right)\right)\right), \tag{5}$$

where Z is a normalizing factor, α is a sharpness-controlling parameter, and $\mathcal{L}(\cdot,\cdot)$ is the loss function, which is the cosine-similarity function in our case.

Then, the Bayesian rule for updating the human personality takes the following form:

$$b^{t+1}(\phi_i) = \frac{p(\bar{\mu}^t_{H_i}|z^t_{H_i}, \phi_i) b^t(\phi_i)}{\int p(\bar{\mu}^t_{H_i}|z^t_{H_i}, \phi_i) b^t(\phi_i) d\phi_i}. \tag{6}$$

In practice, we find the above per-human update rule is still too costly to compute. Since the personality feature space is only 3D and the personality follows a specific distribution (determined by the mapping between RVO parameters and personality factors), we propose to discretize it into a set of M representative personalities $\phi^{1\cdots M}$ via appropriate quantization. Then, we conduct the discrete Bayesian update as:

$$b^{t+1}(\phi_i^j) = \frac{p(\bar{\mu}_{H_i}^t \mid z_{H_i}^t, \phi_i^j) b^t(\phi_i^j)}{\sum_{j=1}^{M} p(\bar{\mu}_{H_i}^t \mid z_{H_i}^t, \phi_i^j) b^t(\phi_i^j)}. \tag{7}$$

4.3 Active Personality Probing

At every timestep t, the robot only considers the observed part of the crowds and their corresponding personality beliefs, but for simplicity, our following description slightly relaxes the notation as if all humans were observed by the robot. Equipped with the learned navigation policy and using the partial observations, the robot could recursively predict the future trajectories of the joint system. With the belief over personality, however, the joint dynamic system becomes stochastic, and estimating the exact distribution of trajectories is intractable. To alleviate the computational cost of prediction, we use the expected personality $\hat{\phi}_i = \sum_{j=1}^{M} \phi_i^j b(\phi_i^j)$ to predict the trajectories.

Based on the trajectory prediction, we formulate robot planning as a non-linear optimization problem in continuous action space and solve it via **Model Predictive Path Integral (MPPI)** Control. Our method is inspired by [55] but extends it to the prediction problem of crowds with complex personality representation, which brings additional challenges. Following the concept of traditional MPC, we design a cost function for the robot to compromise between various objectives. Our robot controller is expected to accomplish three goals simultaneously: exploring the environment, avoiding collisions, and most importantly, collecting information actively to help the robot better estimate humans' personalities.

Exploring the Environment. We use a global planner to guide the robot to explore the environment so that it has more chances to encounter humans. The global planner provides the robot with a loopy path consisting of several goal positions, and the robot would traverse to each goal position cyclically. We define the first term of the robot cost function as follows:

$$C_{\text{goal}} = \sum_{\Delta t=1}^{T_R} \exp\left(\alpha_{\text{goal}} \|p_R^{t+\Delta t} - g^{t+\Delta t}\|^2\right), \tag{8}$$

where T_R is the robot's control horizon, α_{goal} is the hyper-parameter, and $g^{t+\Delta t}$ is the next goal position at timestep $t + \Delta t$ given the robot's current location.

Collision Avoidance. We use another cost function to guide the robot to cautiously avoid collisions. To this end, we introduce the following cost:

$$C_{\text{coll}} = \sum_{i=1}^{N} \sum_{\Delta t=1}^{T_R} \exp\left(-\alpha_{\text{coll}} \min(d_{\text{safe}}, \|p_R^{t+\Delta t} - p_{H_i}^{t+\Delta t}\|^2)\right), \tag{9}$$

where d_{safe} is a user-defined safe distance threshold and α_{coll} is the hyper-parameter.

Active Information Gathering. The main target of the robot is generating useful actions, which could be beneficial for personality estimation. As mentioned before, the robot could infer the future system states when following a preplanned trajectory, so it could naturally obtain the tentative future belief $(b^{t+1}(\phi_i), \ldots, b^{t+T_R}(\phi_i))$ for all observed humans via Bayesian update. Thus, we defined the information gain over humans' beliefs as the final term of the robot cost function:

$$C_{\text{active}} = \sum_{i=1}^{N} H\left(b^{t+T_R}(\phi_i)\right) - H\left(b^t(\phi_i)\right), \quad H\left(b(\phi_i)\right) = -\sum_{j=1}^{M} b(\phi_i^j) \log\left(b(\phi_i^j)\right). \tag{10}$$

The above cost term sums up the negative information gain of the belief over each observed human's personality. Optimizing this cost is equivalent to finding actions that could produce more valuable information, leading to a better personality estimation. Note that we will simultaneously optimize

Algorithm 1: MPPI-Based Robot Planner

Input: Number of samples K, control horizon T_R, current observed state s^t, current belief $b^t(\phi_{1\cdots N})$.
Output: Optimal action at current timestep μ_R^t

Sample K Bezier curves $\mathcal{S}_{0\cdots K}$
for each robot control signal trajectory $\mathcal{S}_i$ **do**
    Sample on $\mathcal{S}_i$ to get robot control signals $\mu_R^{t\cdots t+T_R}$
    Predict joint state trajectory $\{s^{t+1}, \cdots, s^{t+T_R}\}$
    Bayesian update to get tentative beliefs $b^{t+T_R}(\phi_{1\cdots N})$
    Compute cost $C(\mathcal{S}_i)$
end for
$\mathcal{S}^* \leftarrow \arg\min_{\mathcal{S}\in\mathcal{S}_{0\cdots K}} C(\mathcal{S})$
$\hat{\mu}_R^t \leftarrow$first timestep of $\mathcal{S}^*$
Return $\mathrm{NN}_f(l_R^t, \hat{\mu}_R^t)$

the information gain of all observed humans, which means we do not explicitly tell the robot which human needs more attention, and the robot shall make a tradeoff automatically when interacting with the crowds.

MPPI-Based Controller. With the above cost terms, we formulate the robot control problem as a non-linear optimization:

$$\arg\min_{\hat{\mu}_R^t} C = w_{\text{goal}}C_{\text{goal}} + w_{\text{coll}}C_{\text{coll}} + w_{\text{active}}C_{\text{active}}, \tag{11}$$

where w are different weight coefficients. We use an MPPI-based controller to find the optimal action sequence for our robot, which could minimize the gross cost. In order to reduce the dimension of sampling space and improve the smoothness of trajectory, we sample K Bezier curves $\mathcal{S}_{0\ldots K}$ as the potential robot action trajectories. For each $\mathcal{S}_i$, we predict the corresponding human trajectories to estimate the cost $C(\mathcal{S}_i)$. We choose $\mathcal{S}^*$ with the lowest cost, and $\mathcal{S}^*$'s first timestep action $\hat{\mu}_R^t$ is then fed to the local collision avoidance network NN_f proposed in Section 4.1 to generate the final action $\mu_R^t = \mathrm{NN}_f(l_R^t, \hat{\mu}_R^t)$. It should be mentioned that our neural network-based human behavior inference could use GPUs for parallelization and acceleration, which enables our optimization to perform nearly in real-time. Algorithm 1 outlines the MPPI-based optimization procedure.

We test this planner and our personality estimator through proof-of-concept studies in simulation and conduct user evaluations with typical adults and autistic participants separately, which could serve as a foundation for future active-robot-based ASD-specific applications.

5 Proof-of-Concept Studies in Simulation

In this section, we show the performance of each part of our approach through proof-of-concept studies in simulation settings. A continuous 2D simulator is set up where the robot and humans are represented as circular holonomic agents. We use ChAOS [13] for 3D simulation, which is an open source crowd animation software. We consider the scenario with only one robot and a crowd of different sizes.

5.1 Implementation Details

All our experiments are carried out on a PC with an 8-core, 2.30 GHz Intel Core i7-11800H CPU and an Nvidia GeForce RTX 3060 GPU with 6 GB memory. Our method is implemented using Python, and we use PyTorch for the training of neural networks and parallel trajectory prediction. In the following discussion, all dimensions mentioned in the simulations are expressed in pixel units; for instance, a radius of 5 indicates that an agent's radius occupies 5 pixels in the simulated environment.

Mapping between RVO Parameters, PEN Factors, and Adjectives. Our learning-based human local planner leverages the personality-conditioned RVO behavior, as introduced in [29], as the ground truth, where the mapping between the RVO parameters and the PEN personality factors was summarized as:

$$\begin{pmatrix} Psychoticism \\ Extraversion \\ Neuroticism \end{pmatrix} = A_{\text{pen}} \begin{pmatrix} \frac{1}{13.5}(NeighborDist - 15) \\ \frac{1}{49.5}(Max.Neighbors - 10) \\ \frac{1}{14.5}(PlanningHoriz. - 30) \\ \frac{1}{0.85}(Radius - 0.8) \\ \frac{1}{0.5}(Pref.Speed - 1.4) \end{pmatrix},$$
$$A_{\text{pen}} = \begin{pmatrix} 0.00 & 0.08 & 0.08 & -0.32 & 0.63 \\ -0.02 & 0.13 & 0.08 & -0.22 & 1.06 \\ 0.03 & -0.01 & -0.03 & 0.39 & -0.37 \end{pmatrix}. \tag{12}$$

By sampling RVO parameters (e.g., Neighbor Dist), we can generate specific agent behaviors along with the corresponding PEN values computed from this established mapping. In addition, to facilitate the presentation of specific evaluation results, we introduced a binary classification scheme based on the PEN factor profiles. Adopting the personality adjectives "Aggressive" and "Shy" from prior work linking them to the PEN framework [29], we defined these traits operationally within the simulation context: "Shyness" was characterized by heightened fear, inhibition, and avoidance in (or anticipation of) social situations [15], while "Aggressiveness" referred to actions intended to inflict harm upon others [4]. To establish the classification, we generated a diverse set of RVO agent behaviors within the simulator. A psychological expert then annotated each agent's behavior as either "Aggressive" or "Shy" based on these definitions. The results were fitted by a linear Support Vector Machine, deriving an approximate classification rule: an agent is classified as "Aggressive" if the arithmetic mean of its "P" (Psychoticism) and "E" (Extraversion) factor scores exceeds its "N" (Neuroticism) score; conversely, it is classified as "Shy" if its "N" score is comparatively higher. Meanwhile, we observed that in the mapping calculated in [29], the values of "P" and "E" exhibit a negative correlation with "N," which aligns with the classification method we obtained. Consequently, we adopt this classification method for subsequent evaluations requiring categorical information.

Dataset Preparation, Neural Network Architectures, and Training. We prepare separate training datasets for NN_F and NN_f. A total of 2,000 scenarios are generated first, each containing a random number (between 1 and 8) of convex obstacles with 4–7 edges. The shapes and positions of the obstacles are sampled uniformly. To train NN_F, we simulate 60 agents in each scenario across five different cases, where agents are assigned uniformly sampled RVO parameters. Each agent has a sensing range of $d_H = 100$ and is simulated over a 150-step trajectory. To reduce the computational burden during training, only 10 down-sampled steps are recorded per trajectory. At each recorded step, we store the agent's sensing input, state, the historical trajectories of neighboring agents, personality vector, and executed action. Since the number of observable neighbors can vary, we consider only the three closest neighbors within the sensing range. If fewer than three neighbors are detected, the missing trajectory slots are filled with a placeholder vector of −1. For each neighbor, a historical trajectory of length $T_H = 5$ is recorded. For training NN_f, we reuse the previously generated scenarios and maintain the same agent sensing range. 100 agents are simulated per scenario with fixed parameters where the physical agent radius is set to 5. This higher agent density increases the likelihood of inter-agent contact, which is beneficial for learning collision avoidance rules. In each scenario, we only conduct one-step simulation, where preferred actions are sampled in random directions, with magnitudes uniformly drawn from the range [1, 10]. For each agent, we record the sensing input, the preferred action, and the corresponding collision-free action computed by the RVO algorithm. This dataset totally contains 2,000,000 data samples, with scenarios reused evenly throughout the data collection process.

In NN_F, the physical state $s^t_{H_i}$ is encoded using a two-layer **Multilayer Perceptron (MLP)**, resulting in a 32-D feature. The sensing input $l^t_{H_i}$ is processed through two 1D convolutional layers with ReLU activations and max pooling, followed by a fully connected layer that projects the output into a 64-D feature space. The trajectory data $\{\tau_j^{t-T_H\sim t-1}\}$ is encoded using a two-layer RNN, where the final hidden state at the last timestep is flattened and passed through a fully connected layer to obtain a 64-D feature. These resulting features are concatenated and passed through a two-layer MLP, where the output is then concatenated with the personality vector ϕ_i and further processed by a three-layer MLP to obtain the intermediate action. In NN_f, the sensing input is encoded using the same network architecture as in NN_F. The resulting sensing feature is directly concatenated with the intermediate action $\hat{\mu}^t_{H_i}$ and passed through a three-layer MLP to produce the adjusted action. The tanh activation is applied to the final outputs of both NN_F and NN_f for normalization.

The two models are trained independently. Both are optimized using an l_2 loss between the ground truth actions and the network predictions, with a learning rate of 0.001 and a batch size of 256. Network parameters are updated using the Adam optimizer [38]. Given that the scenarios and agent configurations in the dataset are randomly sampled and that the models are designed to plan actions based on local observations, the resulting models can be interpreted as generalizable personality-conditioned local behavior models, enabling them to generalize across different environmental settings (e.g., scenarios with different obstacle distributions) without the need for retraining.

Details of Personality Estimation and Probing. In our Bayesian update-based personality update framework, the continuous personality space is discretized to ensure computational tractability. The numerical range of the three PEN factors is defined as $[-10, 10]$. A discretization step size of 2 is used. However, as indicated in A_{pen}, the PEN factors are not fully independent, where one factor can be expressed as a linear combination of the other two. We uniformly discretize the values of P and E, and compute N according to A_{pen}. Any resulting triplet that falls outside the defined boundaries is discarded, resulting in $M = 68$ representative personality vectors. The belief of each human is initialized as a uniform distribution. During the personality estimation and probing phase, the robot's planner operates with the following default parameters: $\alpha = 0.1, Z = 1.0, d_R = 200, T_R = 10, K = 30, \alpha_{\text{goal}} = 0.003, \alpha_{\text{coll}} = 0.1, d_{\text{safe}} = 15, w_{\text{goal}} = 0.0005, w_{\text{coll}} = 0.04, w_{\text{active}} = 1.0$. GPU-based acceleration is employed throughout the entire optimization pipeline, including belief updates, human behavior inference, and optimal action computation.

Environments Used in Evaluation. Figure 2 shows our 2D and 3D simulation scenarios. We use two closed environments, which enable sustainable interactions: a large *empty room* where humans move randomly and a *ring-like* environment where humans keep moving around the circular space in the same direction. In both environments, human velocities are assigned based on their personalities, sampled within the range $[1, 10]$, while the robot's velocity is fixed at 8. The physical radius of both human agents and the robot is set to 5. Each environment measures $1{,}200 \times 1{,}200$ pixels. The robot's and humans' positions are initialized randomly, where we use repeated sampling to ensure that each agent is at least 50 pixels away from other agents or obstacles. In the empty room environment, both human and robot goals are sampled randomly and updated iteratively, where we ensure that the distance between two adjacent goals is larger than half the size of the environment. In the ring-like environment, human goals are sampled iteratively in free space along the counterclockwise direction of movement, maintaining a consistent flow. The robot, moving in the same direction as the humans, is assigned a fixed set of 8 evenly distributed target points

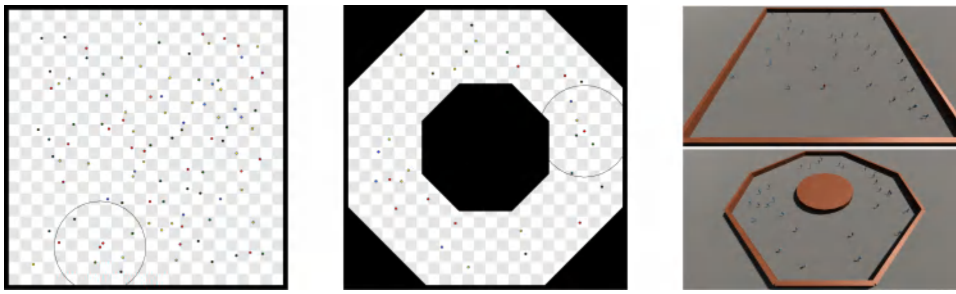


Fig. 2. Two environments used in experiments. Left: An empty room; Middle: A ring-like environment. Right: The corresponding environment in 3D simulation. The red disk-shaped agent with a circular sensor range in the 2D simulation and the red human in the 3D simulation represent the robot.

located at the center of the ring. Upon reaching a goal, whether randomly sampled or predefined, an agent immediately transitions to its next goal.

Baseline. For the baseline, however, all other works for active internal state estimation are only considering a single human and one robot, to the best of our knowledge. These techniques cannot be trivially extended to multiple humans because of the computational cost and change of formulation. While [8] shares some similarities with our work, it is not a suitable baseline, as it does not explicitly estimate human internal states based on observed interaction information. As a result, we mainly consider two algorithms with different cost settings, leading to completely different performances. The first is called an *inactive controller*, where we remove the "active" term C_{active} from our cost function and the robot fully relies on passive observations. We denote it as our baseline, and we compare it with the full version of our approach: the *active controller* that uses active probing actions. Both methods follow the same default settings, except for the inclusion of the C_{active} term in one of them.

5.2 Performance of Human Behavior Model

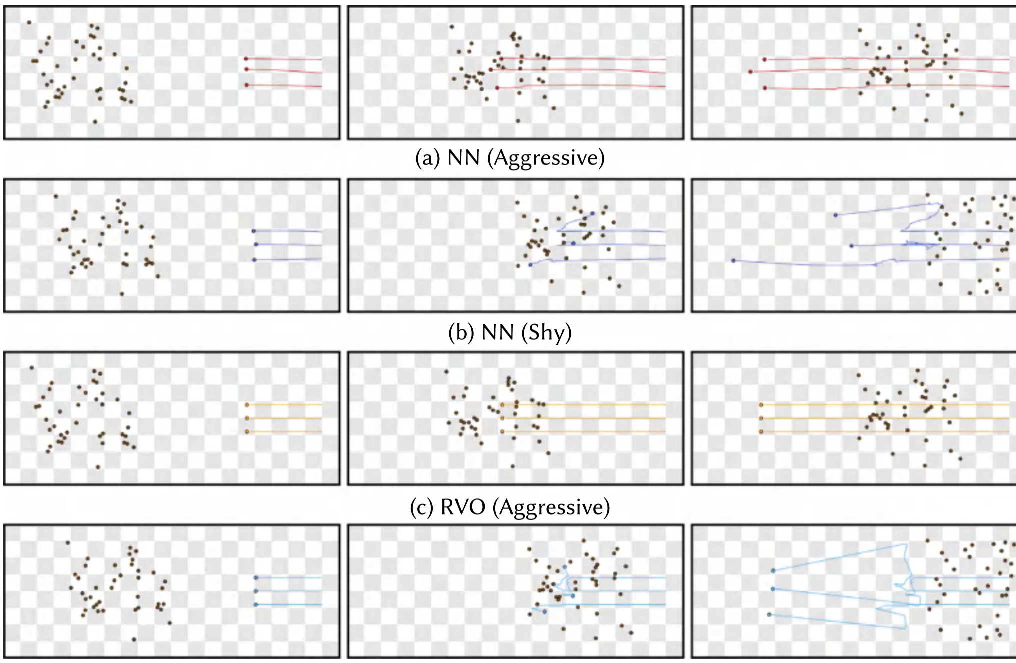
Our learned human behavior could be understood as a soft, differentiable version of RVO. Figure 3 illustrates three humans' behaviors generated by RVO and our learning-based method when facing the same crowd situation (the crowd is moving from left to right). We experiment with two personality instances, sampled from the "Aggressive" and "Shy" categories, respectively, based on the classification method described before. The three humans share the same personality instance in each case. As shown in Figure 3, aggressive humans controlled by both NN and RVO show strong aggressiveness, which has less avoidance behaviors when facing the crowd. Conversely, shy people tend to avoid contact with others and try to dodge or wait. The humans controlled by NN exhibit a similar tendency, which implies our neural-network-based human behavior model could provide explainable personality-conditioned behavior.

5.3 Passive versus Active Robot Behaviors

In this part, we will demonstrate the performance of our active controller by comparing it with the inactive baseline. We will demonstrate two main aspects:

— Our method results in reasonable interaction behaviors;
— Our method results in more interactions, which leads to more accurate personality estimation.

Interactive Behavior. Figure 4(a) plots the trajectories generated by the inactive controller (left) and our active controller (right), respectively, when one human and the robot try to cross an



(a) NN (Aggressive)

(b) NN (Shy)

(c) RVO (Aggressive)

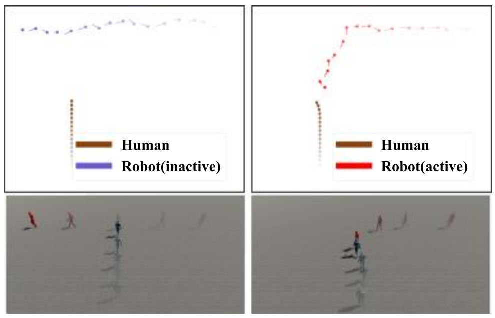
(d) RVO (Shy)

Fig. 3. Human behaviors generated using different models (either NN or RVO) with different personality types (Aggressive or Shy). In each case, a crowd (brown circles) with the same control parameters starts on the left side and moves toward the right. Meanwhile, three humans with specific personality traits (whose trajectories are shown) begin on the right and move toward the left. For each case, we illustrate three stages of the experiment: before the interaction (left column), during the interaction (middle column), and after the interaction (right column) between the three humans and the crowd.

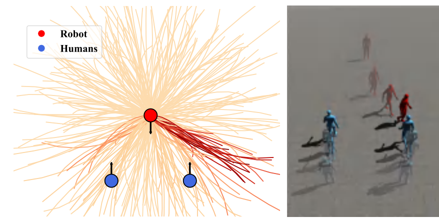
intersection in different directions. Both the human and the robot are plotted as solid circles. Compared with the inactive robot moving straightforward, the active robot changes its moving direction in advance and tries to interact with the human actively, which triggers more human reaction to serve as the basis for personality estimation.

We further find that the active controller can automatically guide the robot to focus on humans “stranger” to the robot, i.e., of whom the entropy of belief is high. If the belief in one human’s personality is quite low, the robot will find that the benefits from interacting with that human are also low, because interacting with him or her will bring little entropy change. We illustrate this process in Figure 4(b). The robot (red) is moving downward and observes two humans (blue) moving in the opposite direction. The human on the left is someone it is “familiar” with, i.e., the entropy of belief is low, and the right human is someone the robot knows nothing about. Our active controller will then sample several trajectories and compute the information gain brought by each trajectory. Figure 4(b) illustrates the sampled trajectories and the corresponding information gain, i.e., C_{active}, in which darker color represents lower cost. It is clear that the robot is more inclined to interact with the unfamiliar human on the right, which could be seen as an implicit greedy exploration strategy.

Number of Interactions. We visualize the number of interactions that the robot observed within 20,000 simulation time steps under the active controller and inactive controller, respectively. In a



(a) Robot behaviors generated by the inactive controller (left) and the active controller (right), respectively.



(b) **Left**: The sampled trajectories of the robot (red) when facing two humans (blue), one is familiar to it (bottom left) and another is a stranger (bottom right). Warmer color corresponds to lower C_{active}. The robot tends to turn right and interact with the stranger. **Right**: The same case in 3-D simulation.

Fig. 4. Interactive behavior of the robot.

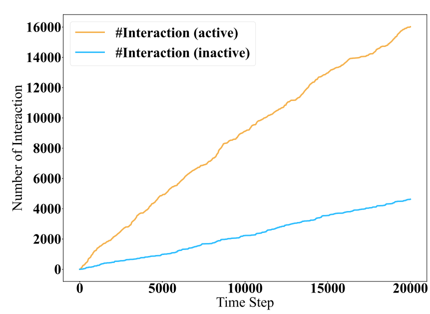


Fig. 5. The number of observed interactions changing history under active and inactive controllers.

single simulation timestep, an interaction is defined as occurring when the distance between two agents is less than $0.3d_H$. The experiment is carried out in the room environment with 30 humans. Figure 5 shows how the number of interactions changes over time for both algorithms. It could be clearly seen that the active controller brings more robot–human interactions, which we believe are more likely to contain useful information for better human personality estimation. It is worth noting that, while our active controller leads to much more interactions, it still well guarantees collision avoidance. Specifically, we count the collision times when executing using both active and inactive controllers, where we judge a collision to happen if the distance between the robot and a human is less than or equal to twice the agent's physical radius. We have observed 0 collisions in both two methods, which verifies the robust collision handling of our system. In practice, we can balance the navigation purpose and information gathering by adjusting the weights of the optimization objectives. For instance, increasing the weights of "collision avoidance" and "moving to goal" can prioritize general navigation over active interaction behaviors. This allows the robot to simulate the effect of "passing by" humans, ensuring *human comfort* is considered during interactions.

Performance of Personality Estimation. We use two metrics to quantitatively evaluate these two methods. We define EE as the **Mean Squared Error (MSE)** between the expected personality $\hat{\phi}_i$ and ground truth personality $\phi_i^\star$. We further define the **Belief Entropy (BE)** as the entropy of the updated belief, which could be represented as $H\left(b(\phi_i)\right)$. The definition of the number of interactions is the same as previously described. We run experiments with 1 robot and 1, 5, 10, 30, 100 humans

Table 1. Results in the Ring-Like Scenario

#Human	1	5	10	30	100
EE active	1.674	2.195	1.882	1.997	1.847
BE active	0.051	0.003	0.064	0.334	0.588
#Itr active	2,029	3,655	11,804	19,024	96,275
EE inactive	3.456	2.834	2.724	2.799	2.130
BE inactive	4.077	2.931	3.189	2.405	1.191
#Itr inactive	82	794	1,690	12,137	103,012

Table 2. Results in the Square Scenario

#Human	1	5	10	30	100
EE active	1.241	2.301	1.985	1.965	2.097
BE active	0.418	0.560	0.838	0.625	0.583
#Itr active	791	2,802	5,909	16,026	70,643
EE inactive	2.591	3.006	2.993	2.825	2.224
BE inactive	4.190	3.392	3.647	2.337	0.998
#Itr inactive	26	273	714	4,628	53,590

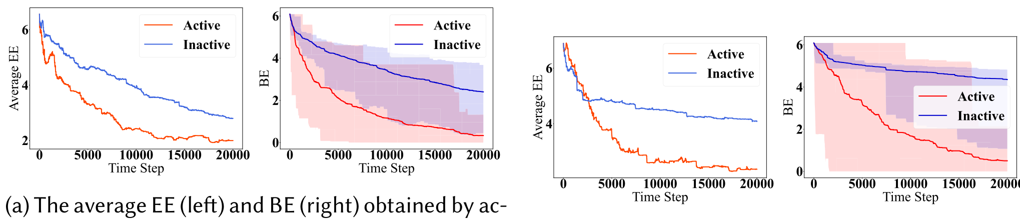
for both environments. The ground truth personalities of humans are generated by uniformly sampling from the RVO parameter space and mapping the samples to the PEN space based on Equation (12).

All experimental results are summarized in Tables 1 and 2. It is evident that the active controller brings more robot–human interactions ("Itr" is an abbreviation for "Interaction"). We also observe that the number of interactions initiated by the inactive controller is slightly higher than those by the active controller in the ring-like scenario with 100 humans, likely due to random variations. This suggests that the environment with 100 humans is quite crowded, resulting in numerous natural interactions. Meanwhile, the active controller always performs better than the inactive one in both EE and BE. The performance of the active controller does not show significant correlation with crowd density. By comparison, the EE and BE obtained by the inactive controller reduce with the increasing number of humans, which is marginally higher than the active method when the number of humans is 100. This implies that our active controller could be more effective in sparse environments, i.e., large scenarios with few humans. The reason behind this is also easy to understand: Sparse environments can induce less passive interaction between humans, which requires more active behaviors of the robot for information gathering. It should also be mentioned that, in practice, we inject Gaussian noise $\mathcal{N}(0, \sigma^2 = (\pi/6)^2)$ to the output of NN_F to mimic the system error of the personality-driven navigation model, which is not injected in the trajectory estimation procedure in Algorithm 1. The above results indicate that our method is robust to this noise.

To clearly visualize the above experimental results, we take the ring-like scenario with 30 humans as an example and plot the change of EE and BE. We set the total number of steps to 20,000 to allow the robot sufficient time to explore the environment and accumulate observations of the 30 individuals while maintaining an appropriate rate for belief updates. As shown in Figure 6(a), our active controller exhibits faster convergence and lower final value than the inactive controller in terms of EE. In addition, the BE of the active method also drops faster, which means the additional interactions do provide more useful information for reducing estimation uncertainty.

As mentioned before, our active controller could be more effective in sparse environments. We perform an additional experiment in the same scenario as Figure 6(a) but enlarge the environment by three times, of which results are illustrated in Figure 6(b). Our active controller leads to a consistently fast decrease of EE and BE, but the inactive method has a worse performance due to inefficient interaction.

Separating the Influence of Perception and Probing. We investigate the impact of the robot's probing behaviors while controlling for potential biases arising from differences in perception between the active and inactive controllers. The active controller naturally collects more data than the inactive controller by actively generating informative interactions, while due to the limited sensing range, these extra data may include both observing more humans and gathering information from probing actions. To isolate the effect of probing, we conducted additional experiments in which both the



(a) The average EE (left) and BE (right) obtained by active and inactive controllers, respectively. The shaded part represents the BE range of all human individuals.

(b) Same as (a), but enlarge the scenario by 3 times. The active method converges well, but the performance of the inactive method becomes worse.

Fig. 6. The changes of average EE and BE in ring-like cases.

Table 3. Results in the Ring-Like Scenario with Infinite Robot Sensing Range

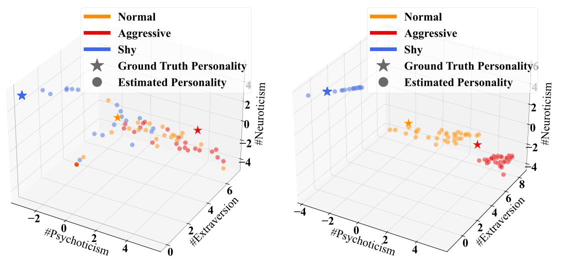
#Human	1	5	10	30	100
EE active	1.466	2.194	1.879	1.738	1.811
BE active	0.836	0.001	0.004	0.109	0.109
AMD active	464.391	191.401	74.026	32.937	29.811
EE inactive	3.050	3.415	2.351	1.910	1.878
BE inactive	2.475	1.063	1.014	0.524	0.153
AMD inactive	544.498	219.549	130.708	62.253	35.500

active and inactive robots were provided with an unlimited sensing range, allowing the robot to continuously observe all human behaviors and removing perception biases. To further explore the origins of the performance differences, we introduced a new metric: the **Average Minimum Distance (AMD)**. AMD is calculated by measuring the closest distance between the robot and each human at every timestep and averaging this value across all time steps.

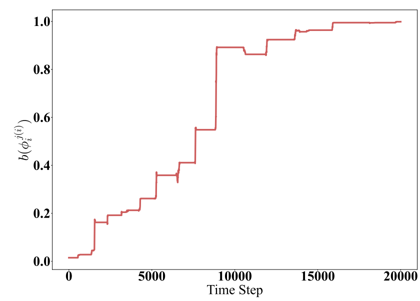
We conducted experiments in the ring-like scenario, with results shown in Table 3. The AMD for the active controller is lower than that of the inactive controller, indicating the impact of the robot's probing behavior. Additionally, the active controller consistently outperforms the inactive controller in both EE and BE. This demonstrates that even when both controllers have similar perception abilities, the active controller's probing behavior still contributes to improved performance. Moreover, compared to the limited sensing range condition, the performance of the active controller remains similar, while the performance of the inactive controller improves, especially in denser environments such as those with 30 or 100 humans. This improvement occurs because, with an unlimited sensing range, the robot can continuously observe the behavior of all humans, allowing it to gather significantly more interaction data in dense environments. This finding also supports our claim that when there is less passive interaction between humans in the environment (such as in sparse environments), it will require more active behaviors of the robot for information gathering.

5.4 Running Time

We have calculated the running time of our algorithm when generating a trajectory for the robot as well as optimizing personality estimation for multiple humans. Compared with the prior work [55] that requires approximately 0.3 s to solve a 5-horizon-length optimization step for 1 human, our algorithm uses a longer horizon length of $T_R = 10$ and solves each timestep of the optimization in 0.075 s, 0.192 s, and 0.463 s for 1, 10, and 30 humans, respectively. This achieves real-time performance suitable for practical deployment, as the estimation is completed before the robot finishes its action.



(a) The estimation effect at the early stage of our algorithm (left) and after many interactions (right). The estimated personalities of 3 kinds of humans are plotted using scatters and the ground truth is represented as stars. In the beginning, the robot can not distinguish humans with different personalities; but in the later stage, the estimation shows a clear boundary.



(b) Changing process of the belief of human H_i's personality. The probability of correct personality estimation continues to rise and becomes close to 1 after 15000 time steps.

Fig. 7. Visualizations of the personality estimation process.

Due to the neural network's parallelism advantage, the running time of our method grows with a slope significantly less than 1 with respect to the number of humans, indicating improved efficiency when handling multi-agent settings.

5.5 Visualizations of Personality Estimation

To further analyze the influence of our active controller on each human, we select one human and illustrate how the belief of his/her personality changes during the experiment. Let human H_i's ground truth personality be $\phi_i^\star$ and the closest discretized personality be $\phi_i^{j(i)}$, we plot the changing process of $b\left(\phi_i^{j(i)}\right)$ in Figure 7(b). It could be seen that the belief keeps increasing, which means the robot's estimation of $\phi_i^\star$ is becoming more and more accurate.

To better visualize the overall estimation result, we perform an additional experiment that uses the empty scenario with 100 humans. We use three personality instances; two are sampled as aggressive and shy cases based on the classification described earlier, and a third is selected at the classification boundary, which we refer to as the normal case. These personalities are assigned to the 100 simulated humans, with 33 labeled as aggressive, 34 as shy, and 33 as normal. We use our active method for robot control. Figure 7(a) illustrates the estimation results at different stages. At the beginning (left), the expected personalities of most humans are inaccurate. Over 20,000 simulation time steps (right), the estimation gets very close to the ground truth.

5.6 Privacy Security

Although we update beliefs for each human in the estimation process, we can choose to only retain the personality distribution over the entire crowd, securing the sensitive human's personality information. Following the setting that involves 100 humans with three kinds of personalities, for example, we assume that the robot has a priori knowledge of and carries out inference only on the three personality instances and classify each human's personality. In each timestep, we classify humans by calculating the distance between their expected personalities and the personality category label, then assign them to the category with the closest match. Then, we compile statistics on the proportion of each category in crowds at different stages and discard other information. As shown in Figure 8, the robot knows little about the personality distribution of the crowd initially. With the increase of exploration, the estimation of the proportions gradually becomes accurate,

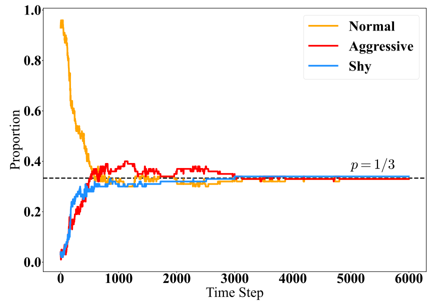


Fig. 8. Proportion changing history of each personality. The dotted line represents the real situation, i.e., each kind of personality accounts for one-third.

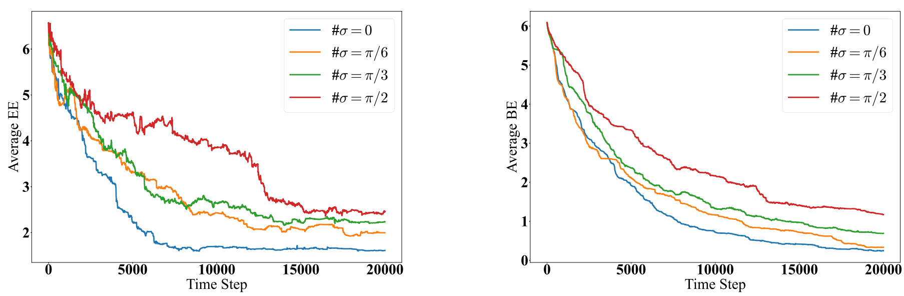


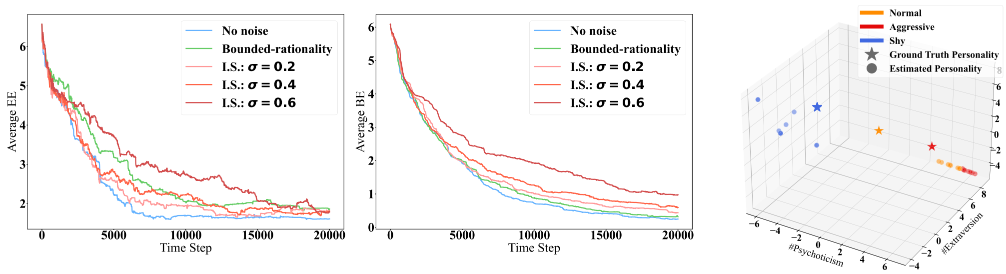
Fig. 9. Average EE (left) and BE (right) when the robot sensing is under noise of different levels.

i.e., about one third of each category. As a result, we can still obtain valuable crowd personality distribution without disclosing human identity. Note that personality distributions other than this example can also be handled, since the classification is based on the estimated personality values, and our method has generally low EE, as reported in Section 5.3.

5.7 Robustness to Model Corruption

In practice, we inject noise into the learned navigation simulator. Such noise is injected only into the simulator, but not the trajectory estimation procedure. As a result, we could mimic the sensing error, which brings additional difficulties to the personality estimation. To explore the effect brought by the noise, we firstly set up four noise levels by changing the value of the standard deviation: $\sigma = 0, \pi/6, \pi/3, \pi/2$, where $\sigma = 0$ means there is no noise, and when $\sigma > \pi/2$, the original behavior of the human will be difficult to identify. We then test the average EE of our method, as illustrated in Figure 9 (left). While the convergence speed becomes slower with the increase of noise scale, our method can provide estimation results with a small gap of 0.85 in EE between the lowest and highest noise levels evaluated in the experiment. We further investigate the changes in BE under these levels of noise. Figure 9 (right) shows that BE increases as noise levels rise. This indicates that, when the robot interacts with simple and unchanging human behavior, its estimation uncertainty decreases rapidly and converges to a low value. In contrast, when confronted with more complex and flexible human behaviors, the uncertainty remains relatively high. We further verify this finding through evaluations related to autism, which is discussed in later sections.

We further investigate two additional noise-adding methods to model human decision-making variability. The first method is based on a bounded-rationality-inspired model, where humans occasionally make suboptimal decisions based on a probability derived from a softmax decision



(a) Average estimation error (left) and belief entropy (right) when humans are simulated with bounded-rational behaviors or when internal states (denoted as I.S. in the figure) are affected by noise at varying levels.

(b) Estimation result when simulating with RVO.

Fig. 10. Performance of the model to various noises and model inconsistency (non-learning-based RVO model).

model. Specifically, for each human action μ_H, we uniformly sample seven candidate actions μ_{cand_i} within a 30° range around μ_H, spaced 10° apart. We compute the cosine similarity between each candidate μ_{cand_i} and μ_H to define the utility function $\mathcal{U}_i$, and apply the softmax function to sample from these candidates:

$$P(\mu_{\text{cand}_i}) = \frac{e^{\beta \mathcal{U}_i}}{\sum_j e^{\beta \mathcal{U}_j}}, \tag{13}$$

where the rationality parameter β is set to 6. The second method introduces Gaussian noise to the internal state ϕ with a mean of 0 and standard deviations (σ) of 0.2, 0.4, and 0.6, respectively. Figure 10(a) (left) presents the robot's EE across these conditions. As observed, all cases successfully converge, although the convergence speed may be influenced by the standard deviation of the noise (e.g., the case with $\sigma = 0.6$ converges more slowly). Figure 10(a) (right) illustrates the changes in BE. The bounded-rationality model achieves better convergence, and the noise added to the internal state follows a similar trend as before, where the BE increases as the noise level rises.

Then we investigate the case where the human model used for trajectory prediction is inconsistent with the actual human behavior model. We replace the human model in the simulator with the RVO [29] and still use the learned navigation policy for prediction and MPPI for robot control. The experiment is carried out in a ring-like scenario with 30 humans, with the same three uniformly distributed personality instance settings as before. Figure 10(b) illustrates that our method could still distinguish humans with different personalities, and there are visible boundaries between different categories. However, humans with personality "normal" are inferred to be overly aggressive. We think this is the main impact of model inconsistency—robot's decision boundary for human personalities becomes blurred, i.e., it becomes more difficult to distinguish two types of humans with insignificant behavioral differences. Despite that, our method can still classify humans by their personalities, although the exact type of personality might be inaccurate.

6 User Evaluation with Neurotypical Adults

In the previous part, we have demonstrated that our methods can generate interactive robot behaviors that are more efficient and accurate in detecting the overall personality distribution of a crowd compared to the passive baseline in the simulator. In this section, we show the results of two follow-up user studies, which evaluate our method's personality estimation ability for human participants and the effectiveness of active interaction.

6.1 User Personality Estimation

A user study was performed to validate our method's performance in estimating human participant's personality, where we set up a 2D game in the simulator in which participants could control one human's movement in a crowd, and we involved an active robot to estimate the participant's personality. We recruited 20 participants from a university campus with a female proportion of 40% and an age range of 22–31. The participants were asked to sign an informed consent form first and would get an equivalent coupon based on the measurement of 25.6 USD per hour after the experiment was completed. This user study took approximately 20 minutes for each participant.

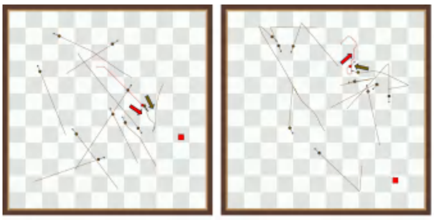
Study Goals. The main aim of this study is to explore our method's performance in human personality estimation. We further aim to investigate users' attitudes towards the robot's active interaction behavior.

Experimental Design. This study consisted of a personality test section and two parts of games. In the first stage, we used a version of the Eysenck Personality Questionnaire, EPQ-RS [23], which contains 48 yes or no questions and is aligned with the personality model we used, to test the participant's personality from the perspective of psychology, of which the result was treated as a trustworthy personality ground truth. Afterwards, the participant will enter the game section.

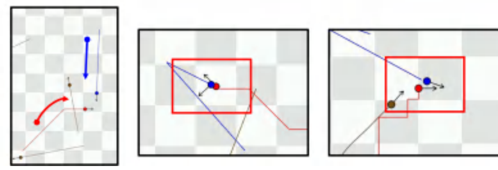
The game was built in a 2D simulator with several agents (representing humans) moving around in a square area similar to Figure 2. Agents were plotted as solid circles and could be divided into three categories: a user-controlled agent, an active robot, and several RVO agents with random parameters. The robot's parameters were adjusted slightly to reduce some interaction enthusiasm while maintaining the inference ability, where the robot's velocity was reduced from 8 to 6 and w_{active} was reduced from 1.0 to 0.1. The agent under the user's control was highlighted with red color, and all other agents were plotted in brown, including our active robot. The participant could use the arrow keys on the keyboard to control the movement of the highlighted agent. Participants were instructed to immerse himself/herself in the game context and imagine it as a real crowd scenario while without being informed of the possible presence of the robot. They needed to complete two parts of games (with a randomly disrupted order) for different research purposes. Before starting the game, the participants will first enter a 160-second trial phase to familiarize themselves with the operations.

Game Part I: Moving as Daily Life. This part of the game was intended to test the participant's original personality using our active robot. The participant was randomly given a target block representing an interested area, which will be automatically changed when he/she reached. The participant's task was moving towards the generated target area while handling the interactions with others, where both of them are equally important with no priority. The participant was encouraged to act as consistently as possible with his/her real daily life, and the behavior was not restricted. To improve the participant's engagement level [51, 61], the participant was instructed that there is a reward for moving towards the target and collisions might be punished where the relative reward/punishment value is measured by the participants themselves. Note that as various psychological studies have shown that there exists a close relationship between personality and reward setting, participants can be more immersed in incorporating their own personalities into the game when given some reward and punishment settings [19, 27]. This part of the game lasted for 160 seconds.

Game Part II: Probing one Human's Personality. This part was intended to explore human participants' behaviors when they are trying to estimate another human's personality. The participant was asked to probe another highlighted agent's personality while physically moving without being given any target areas. The participant's behaviors were recorded for analysis. This part lasted for 50 seconds.



(a) Behaviors of two participants in game part I. The red solid circle represents the agent which is under the user's control, where the red square is his/her interested area. The bold arrows show the local moving directions of the user (red) and his/her closest neighbor (brown). Left: A user disregards the collisions. Right: A user flees to the distance to avoid interactions.



(b) Trajectories of three participants when aiming to probe one human's personality. The red agent is controlled by a user and the blue one is the target agent that the user aims to probe. Left: An approaching behavior. Middle: A user collides with the target agent. Right: A user misses the opportunity to interact with the target agent.

Fig. 11. User behaviors observed in the study.

At the end of the user study, the participant was asked whether he/she felt any other agents exhibiting unacceptable behavior during the whole game, which would be used to measure human attitudes towards active interactions.

Hypotheses. We investigated the following hypotheses:

- $\mathcal{H}_0$: In game part I, the user personality estimated by the proposed method shows a correlation with the ground truth, which indicates that our method has effectiveness when applied on unknown and complex human personality distributions.
- $\mathcal{H}_1$: In game part II, the user's personality-probing behavior is similar to our active robot.

Result Analysis. To better demonstrate the personality distribution and estimation result, we conducted classification based on the PEN personality values as mentioned before, where the personality is divided into Aggressive (referred to as c_{agg}) and Shy (referred to as c_{shy}). The results of EPQ-RS show that the proportion of participants with c_{agg} is 55.0%.

Result Analysis for Part I. We first validate the correlation between the 3D user personality values estimated by our active robot and ground truth. We calculate an overall Spearman correlation coefficient with the associated p-value, where the alternative hypothesis is set to be a positive correlation. The result shows that the personality estimations have a moderate correlation with ground truth at a statistically significant rate ($rs = 0.4654, \mathrm{p} < 0.001$), which supports the hypothesis $\mathcal{H}_0$. Then, we classify the personalities as mentioned before. The overall classification accuracy of the proposed method is 70.0% and the weighted F1-score is 0.694. We observed rich user behaviors in this study, like emergency collision avoidance, avoiding interaction with crowds in advance, or indifference to collisions. Figure 11(a) shows part of the participants' behaviors. The left figure demonstrates an aggressive behavior where the user disregards the collision with the neighbors and moves directly to the target; the right figure shows an avoiding moving pattern where the user moves away from the target area to avoid interactions. The above results indicate that our method can roughly infer a human participant's personality.

Result Analysis for Part II. We recorded each participant's probing behavior and observed that all the participants demonstrated the behavior of approaching the target agent, which is similar to our active robot. The left figure in Figure 11(b) shows one participant's approaching behavior, which aims to obtain more observations by influencing the target agent's movement. It shows that our active robot's behavior is intelligent and partially human-like, which supports $\mathcal{H}_1$. However, the participants' ability in active interaction is weaker than our active robot. On the one hand, a user's performance in collision avoidance is less than satisfactory. As shown in Figure 11(b) (middle), the user collides with the target agent accidently when probing its personality. Our statistics show that 40% of the participants caused more than three times of collisions during the probing process. On the other hand, user's inaccurate prediction of the target agent's future trajectory leads to a low interaction efficiency. Figure 11(b) (right) shows an example. Due to the user's inaccurate prediction of the other agent's future trajectory, the user finds that the target agent has already moved to another location when he/she reaches the expected interaction position. By comparison, our active robot demonstrates robust collision avoidance and human behavior prediction abilities, which have been discussed in Section 5.

Finally, we summarized the participants' attitudes toward other agents' behaviors. 80% of the participants stated that they did not experience unacceptable behavior during the whole gaming process. It suggests that the proposed method can partially maintain human comfort while ensuring **Estimation Accuracy (EA)**.

6.2 Behavior Perception Questionnaire

To validate that the proposed active method brings more crucial information than the passive baseline, we conducted an additional user study where we generated some simulation videos with different settings (with active/passive robot) and asked users to label some humans' personalities from the video. The study was taken by the 20 participants who had been recruited in the previous game. The informed consent form was signed first, and a coupon based on the measurement of 25.6 USD per hour was sent after the participant finishing the study. This user study was designed in the form of a questionnaire, which took approximately 10 minutes to finish.

Study Goals. This study aimed to compare the user's perception of personality from the given videos under two cases: scenarios with our active robot or with a passive robot.

Experimental Design. The basic task for the participant is using the given personality adjectives to classify the behavior of the specified agent in the crowd simulation. We generated crowd behaviors with different settings in the simulator. The environment was set to be a square room. The number of agents had three categories: 5, 10, and 30, where 5 agents represented a sparse crowd, which contains little natural interaction, and 30 agents is a dense case in which the interaction between agents was quite frequent. Agents that represent humans were controlled by RVO with two kinds of personality settings, aggressive and shy. We involved a robot in the crowd, which is controlled either by our active method or passive baseline. For each category of crowd density and each robot type, we generate six cases randomly and get a video pool, which contains totally 36 videos. All agents' initial positions were randomly sampled, and each human agent's parameter was sampled from the corresponding personality parameter distribution.

In each video, all agents were plotted as solid circles. Two of the human agents were highlighted to attract the user's attention, while the other agents looked the same. Participants were asked to view the generated videos and answer the corresponding questions. We sampled from the video pool randomly to select an instance and showed it to the participant. Each participant needed to watch eight videos, and the videos could be rewatched. In each instance, the participant was instructed to carefully observe the highlighted human's behavior and classify the highlighted human's personality given the adjective candidates: *Aggressive*, *Shy*, and *Hard to decide*. The participant would further be asked the degree of difficulty he/she feels in evaluating each highlighted human's personality (ranged from 0 to 100), which could be understood as the evidence or confidence for supporting the evaluation. The participant was advised to carefully observe the highlighted human's trajectory and

Table 4. The Summary of Results in the Behavior Perception Study

Conditions	Overall		5 agents		10 agents		30 agents	
	EA (%)***	AD***	EA (%)***	AD***	EA (%)	AD***	EA (%)	AD
Active	81.5	20.1	78.8	21.0	83.3	19.1	81.8	20.7
Passive	63.9	42.9	40.9	56.8	77.3	40.5	83.3	26.2

Asterisks indicate statistical significant (*$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$).

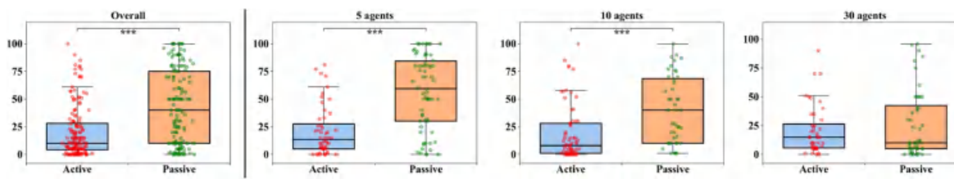


Fig. 12. Results of the evaluation difficulty in behavior perception questionnaire. The left figure shows the overall result, the right three figures show the results under different crowd densities. The vertical axis of each figure represents the evaluation difficulty score. Asterisks indicate statistical significant (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$).

its interactions with the others and mainly concerned about its navigation and collision avoidance pattern.

Hypotheses. We investigated the following hypotheses:

$\mathcal{H}_2$: The participants show higher personality EA and lower estimation difficulty in crowd simulations with our active robot than the passive robot.

$\mathcal{H}_3$: The participants show higher personality EA and lower estimation difficulty in dense crowds than sparse crowds.

Result Analysis. Referring to conditions containing active and passive robot as C_{active} and C_{passive} respectively, we summarized the behavior perception results under C_{active} and C_{passive} separately in Table 4. The EA was calculated to measure the proportion of agents whose personality is successfully estimated, while the **Average Difficulty (AD)** was used to measure the averaged difficulty for participants to evaluate the agent's personality.

Overall Results. The first column in Table 4 shows the overall results. Compared with the EA under C_{passive} (63.9%), the situation under C_{active} is obviously better (81.5%). A Pearson's Chi-square test is performed, which indicates that there is a significant difference in the user estimation under the two conditions ($\chi^2 = 12.4557, p < 0.001$). Meanwhile, the average evaluation difficulty under C_{active} is 20.1, which is less than half of the value under C_{passive} (42.9). We perform a one-tailed Mann-Whitney U-test, and it shows that our active method performs better than the passive baseline in evaluation difficulty at a statistically significant rate ($U = 7{,}776.0, p < 0.001$). The evaluation difficulty values of each participant are visualized in Figure 12 (left), which also demonstrates that the difficulty for the user's evaluation is relatively lower under the condition of C_{active}. The above results support our hypothesis $\mathcal{H}_2$.

Results under Different Crowd Densities. We further break down the results to analyze the changing of evaluation results with different numbers of agents, where 5 agents represent a sparse crowd and 30 agents is a quite dense case. As shown in Table 4, with the growing of agent numbers, both *EA* and *AD* under C_{active} have no obvious changes, while the indicators under C_{passive} show a significant improvement. The *EA* and *AD* under C_{active} significantly outperform the values under C_{passive} in the case of five agents, and the gap becomes smaller as the crowd becomes more dense. The statistical test results show a similar pattern. The evaluation difficulty values with different numbers of agents are shown in Figure 12. It also demonstrates that the level of evaluation difficulty under C_{active} stays stable with the changing of agent densities, while it shows a clear downward trend under C_{passive}. The above results support our hypothesis $\mathcal{H}_3$, which is consistent with our previous claim in Section 5.3: Sparse environment contains less natural interaction between humans, which requires more robot's active information-gathering behaviors.

7 User Evaluation with Autistic and Neurotypical Participants

In the previous sections, we validated the effectiveness of our approach to the personality estimation problem through experiments conducted both in simulation and with typical adults. These experiments demonstrate the technical feasibility of our method for internal state estimation. As mentioned before, our approach is potentially to be applied in the context of ASD. Prior to this next step, in this work we conducted studies in autism using our current framework, where we fine-tune the algorithm's settings and apply a PC game-based user experiment with both autistic and neurotypical participants to explore meaningful preliminary results. Our studies were approved by Tsinghua University's **Institutional Review Board (IRB)** (Project No: 20220110).

A comparative user study was conducted with autistic and neurotypical participants to validate whether our method could identify the difference between the two groups based on their estimated internal states. We partially followed the game setup described in the previous section, where the participant could control one human's movement in a crowd in the game, and we involved an active robot to estimate the participant's internal state. Besides, the experiment's target and settings were adjusted to fit the specific needs of the groups in this study. We recruited 7 participants with autism, aged 9 to 20, from a special education school, 6 of whom were male. All participants with ASD included in this study have received a formal diagnosis from a licensed clinician. In addition, 8 neurotypical participants, aged 10 to 16, were recruited from regular high schools, with 6 of them being female. Participants, or their family members/guardians, were asked to sign an informed consent form first and would get a participation fee based on the measurement of 14.1 USD per hour upon completing the experiment. The study was conducted offline; each participant was required to take part in the experimental PC game for three consecutive days, completing a total of six sessions, with two sessions per day.

Study Goals. The primary goal of this study is to investigate whether our method can detect differences in the estimated internal states between autistic and neurotypical participants. If such differences are identified, we aim to further explore their meaning and significance.

Experimental Design. The experiment consists of two parts: an intelligence test and a PC game. We first assessed the participants' intelligence quotients (IQs) using the **Combined Raven's Test (CRT)** [65], a widely used non-verbal intelligence test that evaluates abstract reasoning and problem-solving abilities. The test consists of 72 questions and is suitable for a wide age range and individuals with varying cognitive abilities. Then, the participants were asked to play the game for six times in different time slots.

The game framework is similar to the one described in Section 6.1. It was built with a square area, which contains a user-controlled agent (operated via keyboard), an active robot, and several RVO agents. The user-controlled agent was highlighted in red, while all other agents, including the active robot, were shown in brown. The participant's task was using the arrow keys to control the highlighted agent to move toward some randomly generated target areas iteratively while avoiding collisions with other agents. To accommodate the specific groups in this study, we made several modifications, including the following:

(1) To ensure the participants' comfort during the interaction, we first adjusted the initial weights of the robot's optimization objectives as before. Additionally, we implemented a dynamic parameter adjustment procedure where the weight of "active gathering" was adjusted adaptively throughout the interaction to reduce the robot's enthusiasm. Specifically, we introduced two new parameters, m_r and $\tilde{m}$. The parameter m_r measures the current adequacy of the robot's interaction, defined as follows:

$$m_r = \begin{cases} m_r + 2, & \text{if } \|p_r - p_{\text{user}}\| < 0.45 d_H \\ \max(0, m_r - 0.3), & \text{otherwise} \end{cases},$$

where $\|p_r - p_{\text{user}}\|$ denotes the distance between the robot and the user. The threshold $\tilde{m}$ represents a margin value for interaction adequacy. The weight of active information gathering, w_{active}, is adjusted adaptively based on m_r and $\tilde{m}$:

$$w_{\text{active}} = \max\left(0, 1.0 - \frac{m_r}{\tilde{m}}\right).$$

This adjustment implies that as interactions between the robot and user increase, w_{active} decreases rapidly, reducing the robot's activeness and shifting its focus more toward navigation. When the robot distances itself from the user, m_r decreases slowly, allowing the robot's activeness to gradually return to its normal level. These adjustments prioritized general navigation over active interaction to create a more comfortable experience.

(2) We modified the main objective and context of the game. Rather than asking participants to imagine themselves acting in the real world, which can be particularly difficult for individuals with autism, we reframed the experiment as a "pure game scenario". Participants were instructed that it was a PC game in which they needed to reach the given target areas while avoiding collisions with others. To reinforce the game concept, we added collision effects and set up a scoring system: 30 points were deducted for collisions, and 100 points were awarded for reaching the target.
In this context, the estimated internal state no longer reflects the participant's "real personality". Instead, we refer to it as an "internal state that influences the participant's behavior in the game context," and we investigate the features of this internal state.

Both autistic and neurotypical participants were recruited. Before starting the experiment, autistic participants completed a trial phase lasting at least 30 minutes over 3 days to familiarize themselves with the game content and controls. Neurotypical participants were allowed unlimited familiarization time until they were satisfied, with the longest lasting no more than 5 minutes. This process ensured that all participants understood the game content, rules, and could operate effectively. Then, each participant played the PC game for a total of six sessions across three different time slots (two sessions per slot), with each session lasting 6–7 minutes.

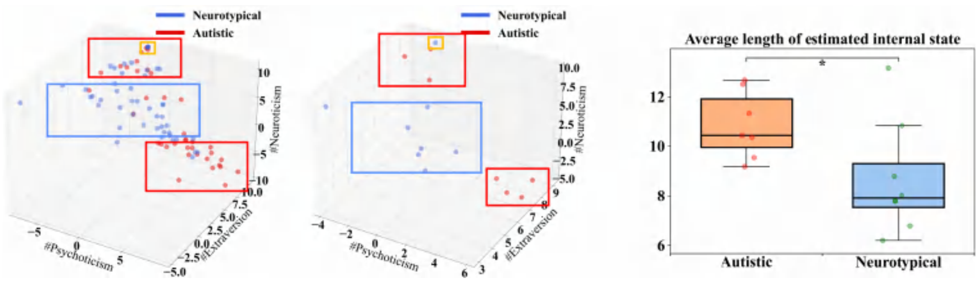
Hypotheses. We investigated the following hypotheses:

$\mathcal{H}_4$: The estimated internal states of autistic and neurotypical participants exhibit different distributions.

$\mathcal{H}_5$: The estimated internal states of autistic and neurotypical participants demonstrate different patterns both within a single game and across multiple game sessions.

Result Analysis. Through the experiments, we obtained six pieces of data for each participant, including both their in-game behaviors and their estimated 3D internal states. Based on this data, we conducted the following analyses:

Analysis of Internal State Distributions. We plotted each participant's estimated internal states in 3D space. Figure 13(a) (left) displays all session estimations for each participant, while Figure 13(a)



(a) Overall (left) and average (right) internal state estimation for autistic (red) and neurotypical (blue) participants.

(b) The average length of each participant's estimated internal states.

Fig. 13. Visualization for the estimated internal states.

(right) shows the average estimation for each participant. Red and blue dots represent autistic and neurotypical participants, respectively. We found that the estimations of neurotypical participants generally cluster near the origin point (within the blue frame), while autistic participants' estimations are closer to the boundary (within the red frame). This suggests that the robot's estimations for autistic participants are more extreme, with higher estimated values. To validate this, we calculated the average length of each participant's estimated internal states. The results, shown in Figure 13(b), confirm that autistic participants have generally higher length values than neurotypical participants. A one-tailed Mann-Whitney U-test supports this trend at a statistically significant rate with $p < 0.05$. Typically, simple and unchanging behavior patterns, such as consistently aggressive or tense behavior throughout the game, can drive the robot's estimations toward extreme values, as the robot's estimations will continuously shift in the same direction. We believe this result is mainly caused by the restricted and repetitive behavior patterns characteristic of autism. In the game, autistic participants tended to maintain a consistent movement pattern when facing various different situations due to these behaviors. For instance, some consistently rushed toward targets without regard for collisions, while others persistently avoided other agents even when the target was nearby. In contrast, neurotypical participants exhibited more complex and adaptive behaviors, adjusting flexibly to different situations. The above findings support our hypothesis $\mathcal{H}_4$.

In addition, we observed one outlier: a neurotypical participant with extreme estimated values (within the yellow frame in Figure 13(a) and at the top right in Figure 13(b)). Notably, this participant had a considerably lower intelligence score than the other neurotypical participants (85 compared to an average of 122.8). This suggests that extreme estimated values in this experiment may correlate with intelligence quotients, as participants with autism often experience intellectual disabilities, generally scoring lower than neurotypical participants. This relationship will be explored further in our future work. Anonymized experimental information for participants is provided in Table 5.

BE Change within Single Case. We further investigated the behavioral characteristics of autistic and neurotypical participants within a single game session. As our active robot continuously updated its belief of each participant's internal state during the game, we selected one game session to plot each participant's BE changes over time, which represent the robot's estimation uncertainty. As shown in Figure 14 (left), the BEs for autistic participants (red) show a greater reduction and converge to a lower value, while the curves for the neurotypical group (blue) are relatively higher. A similar trend was observed in a previous simulation experiment in Section 5.7, where we found that humans with higher behavioral noise exhibited slower decreases in estimation uncertainty, which then converged to a higher value. It indicates that the behavior of autistic participants tends to be more consistent and simple within a single game session, allowing the

Table 5. Anonymized Experimental Information for Participants

	Autistic				Neurotypical		
Id	Gender	Age (year)	Intelligence	Id	Gender	Age (year)	Intelligence
1	M	16	70	1	F	10	130
2	M	20	60	2	F	15	120
3	F	10	70	3	F	16	125
4	M	20	60	4	F	15	130
5	M	14	90	5	F	15	105
6	M	12	–	6	M	15	120
7	M	9	85	7	M	15	85
				8	F	15	130

For gender: F, female; M, male.

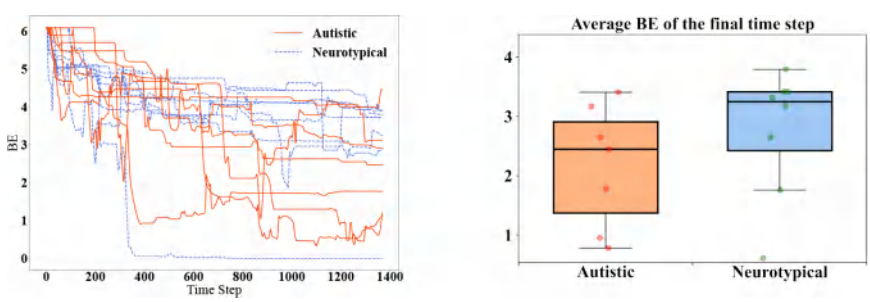


Fig. 14. Left: The BE curve of each participant in one game session. Right: The average BE in the last game timestep of each participant.

robot's estimations to converge quickly. In contrast, neurotypical participants demonstrate more complex behaviors and adjust their behavior pattern according to different situations, which slows the decrease in estimation uncertainty. Note that the lowest blue curve in Figure 14 (left) represents the low-intelligence outlier case mentioned before. We further calculated the average BE at the final timestep of each game for each participant, shown in Figure 14 (right), which confirms that the values for autistic participants are generally lower than those for neurotypical participants.

Estimation Consistency across Multiple Games. Except for single-game behavior, we also investigated the consistency of estimated internal states across multiple games for both autistic and neurotypical groups. With six estimations per participant, we used the **Intraclass Correlation Coefficient (ICC)** to analyze consistency within each group separately. The results, shown in Table 6, indicate that the autistic group exhibits higher consistency across games compared to the neurotypical group. This suggests that autistic participants are more likely to maintain consistent behavioral patterns over extended time frames and are less inclined to adapt their behavior to changing conditions. These findings further support that the characteristics observed in our estimated internal state distributions and trends may relate to the restricted and repetitive behavioral patterns typical of participants with autism. The above results support our hypothesis $\mathcal{H}_5$.

Discussions. We discuss supplementary observations gathered during the experimental process and additionally discuss the overall advantages of the proposed method over conventional approaches.

Observed User Behaviors. During the experiment, we observed that participants' in-game behaviors often mirrored aspects of their real-life behavior. For instance, two autistic participants rarely

Table 6. The ICC of Both Autistic and Neurotypical Groups

Group	ICC	F	p-value	CI95%
Autistic	0.827	29.649	<0.001	[0.71, 0.91]
Neurotypical	0.561	8.674	<0.001	[0.39, 0.74]

demonstrated collision avoidance when encountering other agents in the game. Upon investigation and interviews, their guardians confirmed that these participants also rarely avoid people in real life. For example, if two people are walking and talking side by side on the road, they will directly pass between them. Additionally, one autistic participant consistently moved along the edges of the game interface—a behavior aligned with his tendency to walk along walls in real life. These observations suggest that our experiment successfully captured meaningful behavioral information. Furthermore, no autistic participants experienced emotional outbursts or impatience during the game, indicating that their experience was relatively comfortable and that the active robot did not cause significant emotional fluctuations.

Advantages over Conventional Approaches. The proposed game-based interactive approach holds promise for enhancing future ASD diagnostic processes, with several potential advantages over conventional methods such as screening forms or clinician observation. It does not rely on verbal or cognitive skills, making it accessible to individuals with intellectual or literacy challenges. The interactive nature of the game helps elicit spontaneous behaviors in a dynamic and engaging environment, reducing the likelihood of response bias or masking. Additionally, the method typically requires around 15 minutes and supports concurrent testing, offering potential for reducing diagnostic burden. While existing diagnostic tools are primarily designed for children [33], our approach is not age-specific and may be applicable to older adolescents or adults, as suggested by the inclusion of participants around 20 years old in our study.

In summary, our PC-game-based evaluation provided valuable insights into user behaviors, identifying differences in the estimated internal states of autistic and neurotypical participants and highlighting their connection to the restricted and repetitive behavior patterns typical of autistic participants. These findings lay a foundation for our future work, which will involve using behavior data collected from ASD participants to refine the model.

8 Conclusion

In this work, we introduce a scalable framework for active internal state estimation in human–robot interactions, focusing on personality estimation in crowds. Our approach uses the Eysenck 3-Factor model to link observed behaviors with personality traits, enabling robots to predict human behavior in real time and to gather information actively while considering human comfort. We demonstrate the efficacy of our method through evaluations on both typical and autistic groups. Specifically, our results show that active information gathering outperforms passive observation, reducing both personality EEs and uncertainty. The evaluations on typical adults demonstrate the method's adaptability to diverse personality distributions. The preliminary study with autistic participants highlights its potential to identify behavior patterns associated with ASD, paving the way for future ASD-specific applications.

Our work has several limitations. First, we only evaluated the performance of our framework in a game-based setting, including both simulations and real-world studies using a PC game. However, the accuracy of internal state estimation is affected by discrepancies between the human behavior model and actual human behavior, where unpredictable situations might arise when deploying the system on a real-world robot. In our future work related to autism, we plan to establish a more realistic human behavior model using real-world interaction data. Second, we consider personality as the only internal state for estimation, while human behavior may also be influenced by other internal state factors, such as context, pedestrian type, intention, and emotion. Exploring the effect between multiple internal states is crucial in future work. Third, the robot's interactive behavior is limited to navigation policy in this work. We plan to explore more diverse interaction strategies in the future. Further, while our model accounts for human comfort, fine-tuning to achieve optimal interaction dynamics may require further adjustment across varied contexts. Finally, it would be interesting to explore multiple robots for more efficient cooperative internal state estimation in crowds.

We are excited to take a step toward active internal state estimation in multi-agent settings and to explore the scalable personality estimation, which lays the foundation for our future work in the ASD-specific applications.

References

[1] Alexandre Alahi, Kratarth Goel, Vignesh Ramanathan, Alexandre Robicquet, Li Fei-Fei, and Silvio Savarese. 2016. Social LSTM: Human trajectory prediction in crowded spaces. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 961–971.

[2] Jan M. Allbeck and Norman I. Badler. 2008. Creating crowd variation with the ocean personality model. *AAMAS'08* 3, 209 (2008), 1217–1220.

[3] Panagiota Anagnostopoulou, Vasiliki Alexandropoulou, Georgia Lorentzou, Andriana Lykothanasi, Polyxeni Ntaountaki, and Athanasios Drigas. 2020. Artificial intelligence in autism assessment. *International Journal of Emerging Technologies in Learning* 15, 6 (2020), 95–107.

[4] Craig A. Anderson and Brad J. Bushman. 2002. Human aggression. *Annual Review of Psychology* 53, 1 (2002), 27–51.

[5] Haoyu Bai, Shaojun Cai, Nan Ye, David Hsu, and Wee Sun Lee. 2015. Intention-aware online POMDP planning for autonomous driving in a crowd. In *Proceedings of the 2015 IEEE International Conference on Robotics and Automation (ICRA)*, 454–460. DOI: https://doi.org/10.1109/ICRA.2015.7139219

[6] Tirthankar Bandyopadhyay, Kok Sung Won, Emilio Frazzoli, David Hsu, Wee Sun Lee, and Daniela Rus. 2013. Intention-aware motion planning. In *Algorithmic Foundations of Robotics X*. Emilio Frazzoli, Tomas Lozano-Perez, Nicholas Roy, and Daniela Rus (Eds.), Springer, 475–491.

[7] Michael A. Bender, Antonio Fernández, Dana Ron, Amit Sahai, and Salil Vadhan. 1998. The power of a pebble: Exploring and mapping directed graphs. In *Proceedings of the 30th Annual ACM Symposium on Theory of Computing*, 269–278.

[8] Aniket Bera, Tanmay Randhavane, Emily Kubin, Austin Wang, Kurt Gray, and Dinesh Manocha. 2018. The socially invisible robot navigation in the social world using robot entitativity. In *Proceedings of the 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '18)*. IEEE, 4468–4475.

[9] Aniket Bera, Tanmay Randhavane, and Dinesh Manocha. 2017. Aggressive, tense or shy? Identifying personality traits from crowd videos. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI)*, 112–118.

[10] Aaron Bestick, Ruzena Bajcsy, and Anca D. Dragan. 2017. Implicitly assisting humans to choose good grasps in robot to human handovers. In *Proceedings of the International Symposium on Experimental Robotics*. Springer, 341–354.

[11] Rachael Bevill, Paul Azzi, Matthew Spadafora, Chung Hyuk Park, Hyung Jung Kim, JongWon Lee, Kazi Raihan, Myounghoon Jeon, and Ayanna M. Howard. 2016. Multisensory robotic therapy to promote natural emotional interaction for children with ASD. In *Proceedings of the 2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 571–571.

[12] William J. Bosl, Helen Tager-Flusberg, and Charles A. Nelson. 2018. EEG analytics for early detection of autism spectrum disorder: A data-driven approach. *Scientific Reports* 8, 1 (2018), 6828.

[13] Julien Bruneau and Julien Pettré. 2019. ChAOS: Crowd Animation Open Source. Retrieved from https://project.inria.fr/crowdscience/project/ocsr/chaos/

[14] Yuhang Che, Allison M. Okamura, and Dorsa Sadigh. 2020. Efficient and trustworthy social navigation via explicit and implicit robot–human communication. *IEEE Transactions on Robotics* 36, 3 (2020), 692–707.

[15] Jonathan M. Cheek and Arnold H. Buss. 1981. Shyness and sociability. *Journal of Personality and Social Psychology* 41, 2 (1981), 330.

[16] Yu Fan Chen, Michael Everett, Miao Liu, and Jonathan P. How. 2017. Socially aware motion planning with deep reinforcement learning. In *Proceedings of the 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 1343–1350.

[17] Pauline Chevalier, Gennaro Raiola, Jean-Claude Martin, Brice Isableu, Christophe Bazile, and Adriana Tapus. 2017. Do sensory preferences of children with autism impact an imitation task with a robot? In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, 177–186.

[18] Marco Del Coco, Marco Leo, Pierluigi Carcagni, Paolo Spagnolo, Pier Luigi Mazzeo, Massimo Bernava, Flavia Marino, Giovanni Pioggia, and Cosimo Distante. 2017. A computer vision based approach for understanding emotional involvements in children with autism spectrum disorders. In *Proceedings of the IEEE International Conference on Computer Vision Workshops*, 1401–1407.

[19] Colin G. DeYoung, Jacob B. Hirsh, Matthew S. Shane, Xenophon Papademetris, Nallakkandi Rajeevan, and Jeremy R. Gray. 2010. Testing predictions from personality neuroscience: Brain structure and the big five. *Psychological Science* 21, 6 (2010), 820–828.

[20] Funda Durupinar, Nuria Pelechano, Jan Allbeck, Uğur Güdükbay, and Norman I. Badler. 2009. How the ocean personality model affects the perception of crowds. *IEEE Computer Graphics and Applications* 31, 3 (2009), 22–31.

[21] Hans J. Eysenck and Michael W. Eysenck. 1985. Personality and individual differences: A natural science approach. Plenum Press, New York.

[22] Hans J. Eysenck and Michael W. Eysenck, et al. 1987. *Personality and Individual Differences*. Plenum, New York, NY.

[23] Sybil B. G. Eysenck, Hans J. Eysenck, and Paul Barrett. 1985. A revised version of the psychoticism scale. *Personality and Individual Differences* 6, 1 (1985), 21–29.

[24] Tingxiang Fan, Pinxin Long, Wenxi Liu, and Jia Pan. 2020. Distributed multi-robot collision avoidance via deep reinforcement learning for navigation in complex scenarios. *The International Journal of Robotics Research* 39, 7 (2020), 856–892.

[25] Alan Fern, Sriraam Natarajan, Kshitij Judah, and Prasad Tadepalli. 2014. A decision-theoretic model of assistance. *Journal of Artificial Intelligence Research* 50 (2014), 71–104.

[26] Julio Erasmo Godoy, Ioannis Karamouzas, Stephen J. Guy, and Maria Gini. 2016. Implicit coordination in crowded multi-agent navigation. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence*.

[27] Jeffrey A. Gray. 1970. The psychophysiological basis of introversion-extraversion. *Behaviour Research and Therapy* 8, 3 (1970), 249–266.

[28] Yue Gu, Xinyu Li, Shuhong Chen, Jianyu Zhang, and Ivan Marsic. 2017. Speech intention classification with multimodal deep learning. In *Proceedings of the Canadian Conference on Artificial Intelligence*. Springer, 260–271.

[29] Stephen J. Guy, Sujeong Kim, Ming C. Lin, and Dinesh Manocha. 2011. Simulating heterogeneous crowd behaviors using personality trait theory. In *Proceedings of the 2011 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, 43–52.

[30] Ruihua Han, Shengduo Chen, Shuaijun Wang, Zeqing Zhang, Rui Gao, Qi Hao, and Jia Pan. 2022. Reinforcement learned distributed multi-robot navigation with reciprocal velocity obstacle shaped rewards. *IEEE Robotics and Automation Letters* 7, 3 (2022), 5896–5903. DOI: https://doi.org/10.1109/LRA.2022.3161699

[31] Liang He, Jia Pan, Sahil Narang, Wenping Wang, and Dinesh Manocha. 2016. Dynamic group behaviors for interactive crowd simulation. arXiv:1602.03623. Retrieved from https://arxiv.org/abs/1602.03623

[32] Dirk Helbing and Peter Molnar. 1995. Social force model for pedestrian dynamics. *Physical Review E* 51, 5 (1995), 4282.

[33] Tomoya Hirota and Bryan H. King. 2023. Autism spectrum disorder: A review. *JAMA* 329, 2 (2023), 157–168.

[34] Shervin Javdani, Siddhartha S. Srinivasa, and J. Andrew Bagnell. 2015. Shared autonomy via hindsight optimization. *Robotics Science and Systems: Online Proceedings* 2015 (2015), 10-15607.

[35] Leslie Pack Kaelbling, Michael L. Littman, and Anthony R. Cassandra. 1998. Planning and acting in partially observable stochastic domains. *Artificial Intelligence* 101, 1–2 (1998), 99–134.

[36] Kapil Katyal, Yuxiang Gao, Jared Markowitz, Sara Pohland, Corban Rivera, I-Jeng Wang, and Chien-Ming Huang. 2022. Learning a group-aware policy for robot navigation. In *Proceedings of the 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 11328–11335.

[37] Jae Han Kim, JaeSeong Hong, Hangnyoung Choi, Hyun Goo Kang, Sangchul Yoon, Jung Yeon Hwang, Yu Rang Park, and Keun-Ah Cheon. 2023. Development of deep ensembles to screen for autism and symptom severity using retinal photographs. *JAMA Network Open* 6, 12 (2023), e2347692–e2347692.

[38] Diederik P. Kingma and Jimmy Ba. 2014. Adam: A method for stochastic optimization. arXiv:1412.6980. Retrieved from https://arxiv.org/abs/1412.6980

[39] Markus Kuderer, Henrik Kretzschmar, Christoph Sprunk, and Wolfram Burgard. 2012. Feature-based prediction of trajectories for socially compliant navigation. *Robotics: Science and Systems* VIII, 8 (2012), 193–200.

[40] Justin B. Leaf, Joseph H. Cihon, Ronald Leaf, John McEachin, Nicholas Liu, Noah Russell, Lorri Unumb, Sydney Shapiro, and Dara Khosrowshahi. 2022. Concerns about ABA-based intervention: An evaluation and recommendations. *Journal of Autism and Developmental Disorders* 52, 6 (2022), 2838–2853.

[41] Jaedong Lee, Jungdam Won, and Jehee Lee. 2018. Crowd simulation by deep reinforcement learning. In *Proceedings of the 11th ACM SIGGRAPH Conference on Motion, Interaction and Games*, 1–7.

[42] Baihua Li, Arjun Sharma, James Meng, Senthil Purushwalkam, and Emma Gowen. 2017. Applying machine learning to identify autistic adults using imitation: An exploratory study. *PloS One* 12, 8 (2017), e0182652.
[43] Shengchao Li, Lin Zhang, and Xiumin Diao. 2020. Deep-learning-based human intention prediction using RGB images and optical flow. *Journal of Intelligent & Robotic Systems* 97, 1 (2020), 95–107.
[44] Manja Lohse, Reinier Rothuis, Jorge Gallego-Pérez, Daphne E. Karreman, and Vanessa Evers. 2014. Robot gestures make difficult tasks easier: The impact of gestures on perceived workload and task performance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1459–1466.
[45] Pinxin Long, Wenxi Liu, and Jia Pan. 2017. Deep-learned collision avoidance policy for distributed multiagent navigation. *IEEE Robotics and Automation Letters* 2, 2 (2017), 656–663.
[46] Elisabeta Marinoiu, Mihai Zanfir, Vlad Olaru, and Cristian Sminchisescu. 2018. 3D human sensing, action and emotion recognition in robot assisted therapy of children with autism. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2158–2167.
[47] Bilge Mutlu, Fumitaka Yamaoka, Takayuki Kanda, Hiroshi Ishiguro, and Norihiro Hagita. 2009. Nonverbal leakage in robots: Communication of intentions through seemingly unintentional behavior. In *Proceedings of the 4th ACM/IEEE International Conference on Human Robot Interaction*, 69–76.
[48] Isabel Briggs Myers and Peter B. Myers. 2010. *Gifts Differing: Understanding Personality Type*. Nicholas Brealey.
[49] Stefanos Nikolaidis, David Hsu, and Siddhartha Srinivasa. 2017. Human-robot mutual adaptation in collaborative tasks: Models and experiments. *The International Journal of Robotics Research* 36, 5–7 (2017), 618–634.
[50] Stefanos Nikolaidis, Minae Kwon, Jodi Forlizzi, and Siddhartha Srinivasa. 2018. Planning with verbal communication for human-robot collaboration. *ACM Transactions on Human-Robot Interaction* 7, 3 (2018), 1–21.
[51] Giuseppe Pagnoni, Caroline F. Zink, P. Read Montague, and Gregory S. Berns. 2002. Activity in human ventral striatum locked to errors of reward prediction. *Nature Neuroscience* 5, 2 (2002), 97–98.
[52] Lawrence A. Pervin. 2003. *The Science of Personality*. Oxford University Press.
[53] Darrel A. Regier, Emily A. Kuhl, and David J. Kupfer. 2013. The DSM-5: Classification and criteria changes. *World Psychiatry* 12, 2 (2013), 92–98.
[54] Dorsa Sadigh, Shankar Sastry, Sanjit A. Seshia, and Anca D. Dragan. 2016. Planning for autonomous cars that leverage effects on human actions. In *Proceedings of the Robotics: Science and Systems*, Vol. 2. Ann Arbor, MI, 1–9.
[55] Dorsa Sadigh, S. Shankar Sastry, Sanjit A. Seshia, and Anca Dragan. 2016. Information gathering actions over human internal state. In *Proceedings of the 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 66–73.
[56] Mohammed A. Saleh, Fazah Akhtar Hanapiah, and Habibah Hashim. 2021. Robot applications for autism: A comprehensive review. *Disability and Rehabilitation. Assistive Technology* 16, 6 (2021), 580–602.
[57] Jordan Salvit and Elizabeth Sklar. 2011. Toward a Myers-Briggs type indicator model of agent behavior in multiagent teams. In *Proceedings of the International Workshop on Multi-Agent Systems and Agent-Based Simulation*. Springer, 28–43.
[58] Guni Sharon, Roni Stern, Ariel Felner, and Nathan R. Sturtevant. 2015. Conflict-based search for optimal multi-agent pathfinding. *Artificial Intelligence* 219 (2015), 40–66.
[59] Zhihao Shen, Armagan Elibol, and Nak Young Chong. 2019. Inferring human personality traits in human-robot social interaction. In *Proceedings of the 2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 578–579.
[60] Ryo Suzuki and Jaeryoung Lee. 2016. Robot-play therapy for improving prosocial behaviours in children with autism spectrum disorders. In *Proceedings of the 2016 International Symposium on Micro-NanoMechatronics and Human Science (MHS)*. IEEE, 1–5.
[61] Michael T. Treadway, Joshua W. Buckholtz, Ashley N. Schwartzman, Warren E. Lambert, and David H. Zald. 2009. Worth the 'EEfRT'? The effort expenditure for rewards task as an objective measure of motivation and anhedonia. *PloS One* 4, 8 (2009), e6598.
[62] Jur Van Den Berg, Stephen J. Guy, Ming Lin, and Dinesh Manocha. 2011. Reciprocal n-body collision avoidance. In *Proceedings of the 14th International Symposium on Robotics Research (ISRR)*. Springer, 3–19.
[63] Jur Van den Berg, Ming Lin, and Dinesh Manocha. 2008. Reciprocal velocity obstacles for real-time multi-agent navigation. In *Proceedings of the 2008 IEEE International Conference on Robotics and Automation*. IEEE, 1928–1935.
[64] Allan Wang, Christoforos Mavrogiannis, and Aaron Steinfeld. 2022. Group-based motion prediction for navigation in crowded environments. In *Conference on Robot Learning*. PMLR, 871–882.
[65] Dong Wang, Ming Di, and Ming Qian. 2007. A report on the third revision of Combined Raven's Test (CRT-C3) for children in China. *Chinese Journal of Clinical Psychology* 15, 6 (2007), 559–561.
[66] Jerry S. Wiggins. 1996. *The Five-Factor Model of Personality: Theoretical Perspectives*. Guilford Press.

Received 2 February 2024; revised 26 June 2025; accepted 15 September 2025