Leveraging User Preferences for Adaptive Decision-Making in Human-Agent Interaction

Anonymous Author(s)
Submission Id: «submission id»

ABSTRACT

Humans excel at rapidly modeling team members' latent factors, such as preferences and performance, even with limited interaction. For human-agent interactions, developing such models poses challenges due to the need for extensive prior knowledge or interaction data. To address this, we present a human-robot collaborative storytelling task where robot decision-making is formulated as a Markov Decision Process (MDP). We estimate human preferences and knowledge using a Large Language Model (LLM), then learn the transition probabilities of the MDP at every decision point from these estimates. Targeting to improve team performance and cooperation experiences, the robot then solves the MDP to maximize predicted rewards. We develop human preference and state estimation as a question-answering problem where the LLM offers probability distributions over feasible choices. Our method is evaluated in accordance with a baseline devoid of human knowledge and preferences. Results show that participant satisfaction, team cohesiveness, and narrative quality much better when the LLM-enhanced MDP used. By demonstrating the efficiency of including LLMs into adaptive robot decision-making processes, stressing their possibilities in improving real-time multi-agent team dynamics and opening new paths for more simple and effective human-robot interactions in different collaborative environments, this work advances the field of human-robot cooperation.

KEYWORDS

Legends, Myths, Folktales

ACM Reference Format:
Anonymous Author(s). 2025. Leveraging User Preferences for Adaptive Decision-Making in Human-Agent Interaction. In *Proc. of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2025), Detroit, Michigan, USA, May 19 – 23, 2025*, IFAAMAS, 11 pages.

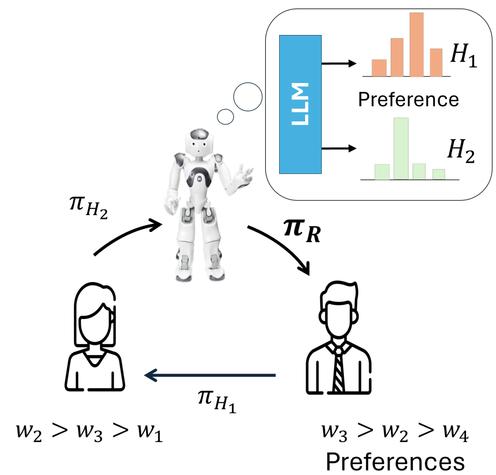

Figure 1: We consider a scenario where user preferences are modeled through an LLM in sequential human-agent collaborative tasks. We aim to leverage these preference distributions to guide the robot's adaptive decision-making, enhancing collaboration, task performance, and user experience.

1 INTRODUCTION

We have witnessed the proliferation of human-agent interaction in recent decades, particularly in the areas where they need to collaborate to achieve shared goals [20, 21, 65, 67]. To make informed decisions and adapt behavior in such collaboration scenarios, the agent is required to understand and model human preferences given the complex nature and frequency of the interaction [15, 66, 68]. This modification is essential for enhancing the overall task outcomes, user satisfaction, and team performance. With this in mind, research endeavors have been undertaken to model human preferences and utilize these models for the adaptive decision-making of agents in the development of sophisticated and effective human-agent collaboration systems [48, 67, 82]. Nevertheless, the modeling of human preferences in a human-robot team environment is a difficult task due to the dynamic nature of human preferences, which are influenced by real-time interaction.

This position is particularly well-suited for humans, as they have the capacity to interpret nonverbal signals, comprehend past experiences in analogous social contexts, and assess the behavioral history of others [27, 57]. This ability enables individuals to adjust their behavior in accordance with the knowledge, abilities, and preferences of others. However, server-related constraints make it challenging to replicate a comparable capability in autonomous agents [30, 54]. In contrast to humans, artificial agents are capable of intuitively understanding social dynamics and lack extensive prior knowledge. In real-world scenarios, certain methods that rely on hand-crafted techniques to model humans are not feasible and cannot be scaled [5, 13, 50, 84]. Conversely, data-driven methodologies employ deep learning methodologies to construct human models from interaction data [12, 51]. Nevertheless, these are constrained by the scarcity of interaction data, rendering them incapable of achieving excellence in novel or short-term collaboration scenarios. Additionally, the computational modeling of human preferences is challenging due to their complexity, dependence on context, and occasional inconsistency.

Recent research has demonstrated the potential of Large Language Models (LLMs) to comprehend and produce human-like text

Proc. of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2025), A. El Fallah Seghrouchni, Y. Vorobeychik, S. Das, A. Nowe (eds.), May 19 – 23, 2025, Detroit, Michigan, USA.

in a variety of application domains [49, 71, 81]. LLMs can be leveraged to estimate the values of latent factors that influence human decision-making and behavior. Research has been done to explore this capability of the LLMS to build human models with limited demonstration [2, 31, 53, 75]. Zhang et al. took this a step forward to utilize the LLM as a zero-shot human model in robot planning in interactive human-robot scenarios [80]. Such approaches usually formulate human preference estimation as a multiple-choice question-answering problem where a textual prompt is used to probe the LLM to generate a response within the option sets. Furthermore, LLMs can be implemented to modify an agent's perceptions of human collaborators, which leads to the acquisition of latent factor state distributions as the interaction progresses. These dynamic revisions of belief states may be incorporated into the agent's decision-making process to encourage more context-aware and adaptable behavior. Despite some attempts to model human preferences using LLM and apply this model to robot decision-making, the efficacy of this approach in a real-time human-robot team collaboration scenario has not yet been extensively investigated. Preference-aware decision-making in team settings added additional challenges as there can be conflict among the team members' preferences. How to balance among the preferences while making the adaptive decision is a challenging task. Furthermore, how these approaches impact participants' perception in interactive scenarios needs to be examined.

To address the aforementioned challenges, we extend the previous works to an interactive human-robot team setting where the preferences of the agents evolve based on the interaction dynamics. We consider a multi-agent sequential collaboration task that includes robotic and human agents, with each agent having its own preferences. The robotic agent is tasked with modeling the preferences and making adaptive decisions so that the collaborative effort of the other agents becomes intuitive, and, in turn, they end up with a coherent collaborative experience. The robotic agent's decision-making is referred to as a Markov Decision Process (MDP) in order to facilitate this. To facilitate this, the robotic agent's decision-making is referred to as a Markov Decision Process (MDP). The transition probabilities are dynamically determined by assessing the contributions of human gene preferences and their influence on the partnership. By resolving this MDP at each decision juncture, the robotic agent optimizes anticipated future benefits, which leads to enhanced team performance and more pleasurable collaborative interactions. To validate our approach, we developed a physical human-robot collaboration test bed where a team of one robot and human participants in a collaborative storytelling game. Here, each agent tries to maximize their utility by picking words from a storyboard and continuing the story narrative in their preferred theme. The robot agent leverages an LLM to extract a distribution of agents' preferences over the remaining words based on their predicted story theme and contribution level and utilizes this distribution in calculating the transition probability of MDP. Solving this MDP, the robotic agent learns a policy that maximizes the expected reward for the future word selection. The agent also probes the LLM to generate its response considering the estimated preferences of other agents.

To evaluate the effectiveness of our preference-aware adaptive strategy, we compare it against a myopic baseline that does not account for the preferences of other agents in a human-robot interaction study with 20 teams (n=40) in our collaborative storytelling test bed. The baseline strategy is predicated on the assumption that all agents will dependably contribute and optimize their immediate actions, without taking into account future impacts or team dynamics. In contrast, our preference-aware approach employs a large language model (LLM) to estimate human preferences and adjusts robot decisions to optimize long-term collaboration outcomes. According to our research, the preference-aware strategy significantly improved the satisfaction, narrative coherence, and team performance of participants in both subjective and objective metrics. The findings of this study suggest that the incorporation of human preferences into decision-making processes enhances the overall task experience, team cohesion, and collaboration fluency, thereby enabling more intuitive and effective human-robot interaction.

2 RELATED WORK

2.1 Human Modeling in HRI

Modeling humans has always remained an exciting field for the HRI research community as it allows robotic agents to make personalized and informed decisions [23, 47, 63]. Researchers have utilized methods, broadly categorizing them into two categories: theory of Mind (ToM) and data-driven approaches [9, 11, 77, 78]. The first category of works (e.g., Bayesian Theory of Mind) incorporates a set of assumptions about human mental processing and behavior and updates their beliefs based on these assumptions [5, 6, 32]. The second category of work requires a large amount of real human data to train a model to predict human behavior. However, these models suffer from scarcity of data, especially in short-length interactions [59]. Moreover, these models are computationally expensive and thus costly to train with a huge amount of data. In a multi-agent scenario, the problem becomes more challenging given the number of state spaces. Researchers also employ a hybrid mode in which the agent uses a pretrained model trained with RL and then fine-tunes it with a small amount of data to adapt to a particular individual[12].

However, the performance of the mentioned models in modeling human latent states such as preferences and emotions is still unclear. Thus, further exploration of the model's performance for estimating human latent states, such as preferences in real-time interaction scenarios, is required.

2.2 LLM in Human Modeling

Large Language Models (LLMs) have become a groundbreaking technology in recent years, significantly impacting various fields, including Human-Robot Interaction (HRI). Their influence extends to improving communication [7, 10, 29], enhancing task planning capabilities [3, 19, 56, 83], and advancing decision-making processes [14, 35, 72, 76], as evidenced by numerous studies in these areas. Trained on massive datasets with billions of parameters, these models excel at producing human-like text and tackling diverse tasks [22, 28, 72, 75, 79, 80, 80]. Researchers have leveraged this ability to replicate human behavior and stimulate the emergence of the theory of mind [2, 31, 53, 80]. Recent research demonstrates LLMs' ability to serve as prior knowledge bases for inferring human intentions. By converting structured task information into natural language queries, these models can generate probability distributions reflecting user preferences. This functionality allows agents to dynamically interpret human behavior, even in novel situations lacking task-specific data, known as zero-shot scenarios [24, 80].

Although these recent works have demonstrated the potential of LLMs as a human model for HRI, the literature still needs more evidence, and there remains a gap in understanding how LLM-based human models excel in interaction human-robot scenarios. Moreover, how LLM-based human preference-aware robot strategies impact the performance and perception of human team members in human-robot collaborative tasks needs to be explored.

2.3 Adaptive Decision Making in HRI

In real-time, adaptive decision-making in HRI enables the robots to adjust their behavior based on human preferences, actions, and environmental context [26, 38, 39, 64]. Researchers explored various methods, including partially Observable Markov Decision Processes (POMDPs) and utility-based frameworks to empower robots to infer human intentions from the observations and update their strategies accordingly [13, 43–45]. In human-robot collaborative tasks, this capability allows robots to generate adaptive responses aligning with human team members' goals and preferences [18, 46]. However, most of these approaches rely on predefined models that need help to generalize to new scenarios.

Recent studies have utilized data-driven techniques, such as deep learning, which enables robots to anticipate human preferences from interaction data [16, 52, 55]. However, data-driven methods are computationally intensive, limiting their real-time application.

3 TECHNICAL APPROACH

3.1 Collaborative Task Model

We model the collaborative task involving N agents, $i = 1, 2, \ldots, N$, as a Markov Decision Process (MDP), which is defined by a set of states S, a set of actions A, a probabilistic transition function P, and a real-valued reward function r. The task unfolds sequentially in turns $t = 1, 2, \ldots, T$, where T is the maximum number of turns. The system's state evolves according to the transition function $P(s' \mid s, a^i)$, representing the probability of transitioning from state s to state s' when agent i (either human or robot) performs action a^i. Upon each transition, the agent receives a reward $r(s, a^i, s')$, which is designed to encourage behaviors that align with the task objectives. Additionally, at any given step t, we assume the agents' interaction history is accessible as $h_t = \left\{(s_k, \{a^i_k\}_i)\right\}_{k=1}^{t-1}$, where s_k is the state at step k and a^i_k denotes the action taken by agent i at step k.

3.2 User Preference Model

Since human preferences are not directly observable by an agent, we model them as latent factors inferred from the interaction history between the two humans and the robot. Let $z_t^{H_i} \in \mathcal{Z}^{H_i}$ represent the preference of human agent i at time t, where $\mathcal{Z}^{H_i}$ denotes the space of possible preferences for human i. Our objective is to estimate the joint distribution $p(z_t^{H_i} \mid s_t, h_t)$, where s_t is the current state, and h_t refers to the interaction history up to time t. To leverage large language models (LLMs) for human preference modeling, we define a textualization function f that transforms the state, history, and agent information into a natural language prompt: $x_t = f(s_t, h_t, q_t)$, where q_t represents the query posed to the LLM. We then use the LLM to estimate the human agents' preference distribution, formalized as $p_l(z_t^{H_i} \mid x_t) \approx p(z_t^{H_i} \mid s_t, h_t)$, with p_l representing the LLM-based joint probability distribution over possible human preferences.

The prompt x_t is structured to include a description of the task and environment, the interaction history h_t, the current state s_t, and a query q_t that probes each human's preference or likely action. This query formulation allows the LLM to infer human preferences in a zero-shot manner for multiple agents without requiring additional training data or fine-tuning. The LLM's pre-trained knowledge acts as a prior for human behavior, dynamically updated based on the observed interaction history. By framing human preference modeling as a question-answering task for the LLM, we can capture complex, context-dependent preferences and inter-human dynamics that may be challenging to model through traditional approaches.

3.3 Robot Agent's Decision Making

In the sequential task, each agent selects an action $a \in A$ with the aim of maximizing its reward. However, in creative tasks like collaborative storytelling, solely focusing on reward maximization can compromise task quality. Agents may need to prioritize utility over immediate rewards, balancing between optimal word selection and their ability to contribute meaningfully to the narrative. For instance, choosing a high-reward word might challenge the agent to maintain narrative coherence, thereby reducing its overall contribution to the story. To address this, we model human agents as Boltzmann rational actors, who seek to maximize their utilities—defined as functions of their expertise and preferences. Specifically, we represent the expertise and preference of an agent as$i$ at turn t as a latent factor, $z_t^{H_i}$ where $z \in \{p, e\}$, i.e., $p_t^{H_i}$ captures the agent's preference and $e_t^{H_i}$ denotes the agent's expertise at time step t. To infer human preferences and expertise, we use an LLM that provides a distribution over human action preferences and expertise given the current task state:

$$p_l(z_t^{H_i} | x_t) \approx p(z_t^{H_i} | s_t, h_t) \tag{1}$$

We assume that all actions $a \in A$ are unique. Specifically, if an agent κ selects action a_t^{κ} during turn t, that action becomes unavailable for subsequent agents, and the action space is updated as $A = A \setminus \{a_t^{\kappa}\}$. Consequently, the robot agent's action at turn t influences the available actions for the other agents in future turns. When selecting an action $a \in A$, the robot anticipates the future action choices of all other agents. It learns a policy by simulating or rolling out these decisions across the remaining turns $t+1, t+2, \ldots, T$, considering the estimated preferences of the agents at each step. While agent preferences may shift in future turns, the robot relies on the interaction history up to the current point to form a reasonable estimate of future preferences. Thus, the robot's decision-making process can be framed as an action distribution problem, aiming to allocate actions among all agents over the remaining turns, including the current one, to maximize

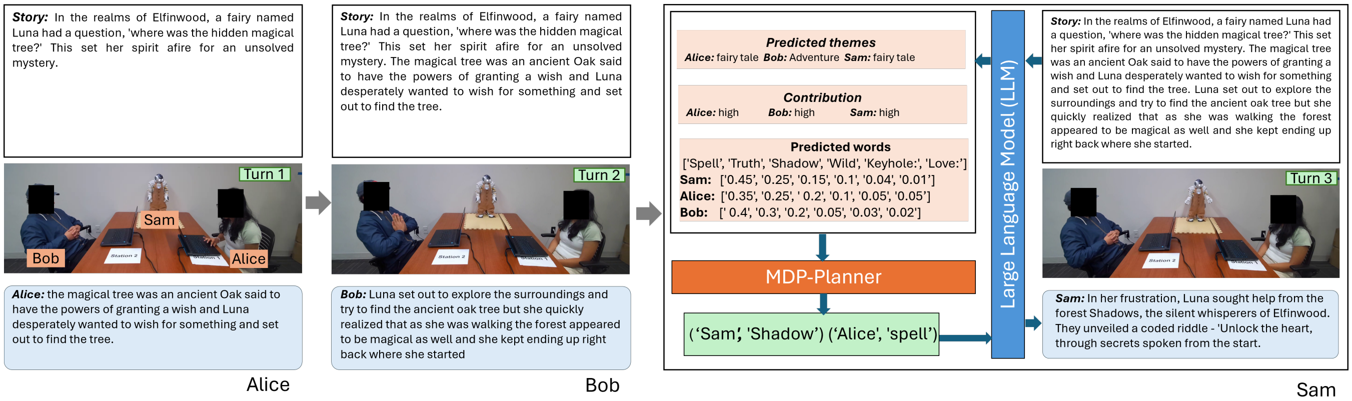

Figure 2: Collaborative storytelling testbed with two human participants and a robot agent taking turns to build a narrative from a shared storyboard. The robot uses a Markov Decision Process (MDP) with human preferences, estimated via a Large Language Model (LLM), to select actions that maximize rewards. The interface displays available words, turn order, and accumulated points, fostering teamwork and evaluating the robot's adaptive collaboration strategy.

the expected reward. This formulation treats the robot's decision-making at the current turn t as a Markov Decision Process (MDP) $M = (S, A, P, R, \gamma)$. Since it is the robot agent's turn, the action distribution always begins with the robot. The transition function P represents the current estimate of the system's dynamics and is defined as follows.

$$P(s_{t+1}|s_t, a_t^{\mathcal{K}}) \approx p_l(z_t^{\mathcal{K}}|x_t) \tag{2}$$

The robot's goal is to find an optimal policy $\pi^* : S \longrightarrow A^{\mathcal{K}}$ that maximizes the expected cumulative discounted reward:

$$\pi^* = \arg\max_{\pi} \mathbb{E}\left[\sum_{t=0}^{\infty} \gamma^t R(s_t, \pi(s_t))\right] \tag{3}$$

We summarize the agent's decision-making process in Algorithm 1. In its turn, the robotic agent uses the task state and history to generate a prompt for the LLM, which provides predictions guiding the decision. The robot then solves an MDP to select the optimal action, maximizing the expected reward based on the task context.

Algorithm 1: Robotic agent's Decision Making

Input: Current turn t, task state s_t', Action space A_t, History h_t, Maximum turns T, Reward function R
Output: Robot agent's action a^R
$x_t \leftarrow f(s_t', h_t)$ // textualize state and history to LLM prompt
$P_z \leftarrow LLM(x_t), z \in \{p, e\}$
$R_t \leftarrow \texttt{calculateReward}(R, p_e)$
$S_t \leftarrow \texttt{generateStates}(\Gamma, T)$
$\pi^R \leftarrow \texttt{solveMDP}(S_t, A_t, P_z, R_t, \gamma)$
$a^R \leftarrow \pi^R(s_t)$

4 EXPERIMENT DOMAIN

4.1 Collaborative Storytelling Game Domain

Researchers have explored collaborative storytelling to study how humans and agents are involved in a co-creative process and how

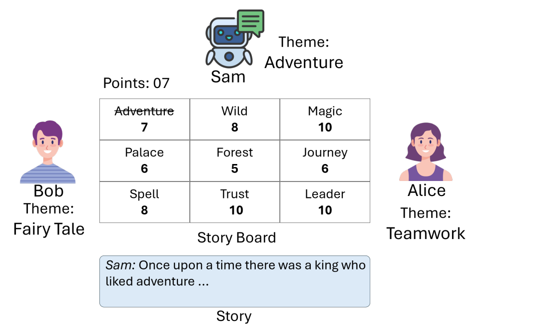

Figure 3: We consider a scenario where user preferences are modeled through an LLM in sequential human-agent collaborative tasks. We aim to leverage these preference distributions to guide the robot's adaptive decision-making, enhancing collaboration, task performance, and user experience.

the agent's strategy affects the participant's performance [40, 41, 60]. Collaborative storytelling creates an environment for expressing creativity and shared narrative building and also provides a means to study complex social interactions, teamwork, and communication dynamics [34, 73]. Researchers highlight that storytelling tasks push people to work together, helping them understand each other and share common goals [33, 74].

Recent research has extended storytelling into collaborative game environments where story narrative is embedded in interactive game mechanics [42]. Extending storytelling to a collaborative game provides the task with structured goals with a set of constraints that help adaptive and collaborative decision-making throughout the task [61]. In such an engaging environment, participants must balance between narrative creativity and task objectives. Moreover, the inclusion of game elements in storytelling encourages active engagement as the balance between story coherence and game objective demands continuous collaboration and shared problem-solving.

We select collaborative storytelling games as our experimental domain to study the effects of agents' adaptive and preference-aware collaboration strategies on teammates' perceptions and performances. In this game, a set, $\Gamma = \{R, H1, H2\}$ of three players, including a robot agent (R) and two human (H) participants, sequentially build a narrative by picking a word from a set of the storyboard words, W while each agent attempts to align the story with their preferred theme. We designed this collaborative game to serve as a structured yet creative environment to study the impact of preference-aware robot policies in human-agent collaboration. This domain allows us to explore how robots can model human preferences and adapt their actions in real time, balancing individual inputs to create a coherent narrative while enhancing team performance and satisfaction.

4.2 Robotic Agent's Decision Making Process

4.2.1 Deriving Optimal Policy. In the collaborative storytelling game, the robotic agent, at its turn, chooses a word $w \in W$ from the storyboard, makes the story narrative, and passes the turn to the next human agent. In the preference-aware strategy, the robotic agent considers human agents' word choice preferences and thus takes action that considers all estimated future word choices of all the agents. This problem can be considered as the word distribution problem among the agents sequentially according to their preferences, starting from the robotic agent. We formulate this as a Markov Decision Process (MDP) defined by the tuple $M = (S, A, P, R, \gamma)$, where:

- S: $\langle a_i, W_r \rangle$, where a_i is the current agent and W_r is the set of remaining words.
- A: Action space, $A(s) = \{w_1, w_2, \cdots, w_n\}$ consisting of selecting any word from the remaining words.
- P: $P(a_i, w_j)$, probability that agent a_i chooses word w_j, which is the known preference probability of each agent-word pair.
- R: $R(s, a)$, the reward for assigning word w_j to agent a_i determined by the word's weight multiplied by the agent's predicted contribution score.
- $\gamma \in [0, 1]$: Discount factor, reflecting the importance of future rewards.

Given a state $s = \langle a_i, W_r \rangle$ and an action w_j, the system transitions to a new state, $s' = \langle a_{i+1}, W_\mathrm{r} \setminus \{w_j\} \rangle$ with probability $P(a_i, w_j)$. The agent receives a real-valued reward $R(s_t, a_t) = C(w_j) \times S(a_i)$, where $C(w_j)$ is the predefined weight of the selected word, and $S(a_i)$ is the contribution score of the agent. The weight of a word is calculated as $C(w_j) = \alpha \times f_{len} + (1-\alpha) \times f_{feq}$, where f_len and f_freq are the Min-Max normalized length and frequency factor of the word, respectively. The frequency factor is determined from the number of occurrences of the word in the Gutenberg corpus, which is a widely used story corpus. We use the Gutenberg corpus because it provides a large collection of narrative texts, enabling an accurate estimation of word frequency and complexity in storytelling contexts. The Min-Max normalization ensures that the length and frequency factors are scaled between 0 and 1, maintaining consistency across the word selection process. Here, α is the relative weight between the length and frequency factors. Here, $S(a_i)$ is the agents contribution score that can be derives the interaction history.

The robot policy defines the strategy for word allocation. In each state $s = \langle a_i, W_r \rangle$, the policy $\pi(s)$ determines which word to assign to the agent. The goal is to find a policy π^* that maximizes the cumulative reward over the word distribution process. The robot aims to find an optimal policy π^* that maximizes the expected cumulative discounted reward:

$$\pi_R^* = \arg\max_\pi \mathbb{E}\left[\sum_{t=0}^{\infty} \gamma^t R(s_t, \pi(s_t))\right]. \tag{4}$$

By solving the MDP, the robot derives a policy π^R that maps states to robot actions, i.e., $a^R = \pi^R(s_t)$, This policy guides the robot's actions by accounting for the predicted preferences of human agents, making the robot an adaptive collaborator.

4.2.2 Preference Probability Estimation. In the collaborative storytelling game, each human participant chooses a theme from a set of themes, $\mathcal{T} = \{\text{adventure}, \text{fairy tale}, \text{teamwork}, \text{mystery}\}$. The robot agent also randomly selects a theme from the theme set. The human agent's theme choice is unknown to the robot, and the robot tries to estimate the theme based on the interaction history.

At the start of the game, the robot assumes a uniform distribution over the human participants' theme choice, i.e., $T^{H_i} \sim \text{Uniform}[\mathcal{T}]$. As the game progresses, we update the distribution over T^κ by leveraging a large language model, using the text prompt x_t, and calculating the theme distribution as $p(T_t^\kappa \mid x_t)$ For each agent $\kappa \in \Gamma$, we sample a theme, augment it with the prompt, and feed the prompt to a large language model (LLM) to obtain a distribution over the agent's word choice preferences:

$$p(P_t^\kappa \mid x_t) = \text{LLM}(x_t).$$

Additionally, we also estimate the agent's contribution level, which is modeled as a distribution over the set of levels, $L = \{\text{very poor}, \text{poor}, \text{moderate}, \text{good}, \text{very good}\}$. A mapping function $F : L \longrightarrow [0, 1]$ maps each contribution level to a real value.

4.2.3 Generating Robot Narrative. The policy π_R^* provides the allocation of agent-word pairs based on the agents' preference distribution. We prompt the LLM with the robotic agent's currently selected word and the predicted word for the next agent, along with the current state of the story. The query in the prompt asks the LLM to generate the robotic agent's narrative using the robot's chosen word and theme while also considering the predicted words for the next agent. This ensures that the robot's narrative is intuitive for the next agent to build upon.

By generating the robot's narrative in alignment with human agents' predicted preferences, and as the allocation is derived from a policy learned by solving the MDP (which encodes all agents' preferences), the method ensures that the resulting story remains coherent with respect to all agents' themes. As the robotic agent continually updates its estimates of other agents' themes and word preferences, this approach guarantees adaptability in the robot's actions.

4.3 Prompt Design

To design effective prompts while minimizing invalid completions, we adhered to guidelines provided in prior works [80]. Our prompt design encompassed several key components: a detailed description of the collaborative storytelling game, including its rules and the features of the robot agent; the interaction history h_t and the current state s_t of the collaborative task; and a query q_t to guide the LLM's response. In line with [80], we structured the prompt to limit valid completions to elements of $z^{\mathcal{K}}$, assigning a single token label to each element of $z^{\mathcal{K}}$. For theme prediction, we employed a multiple-choice format, while for extracting agents' contribution levels, we used a 5-point Likert scale. Additionally, to extract the preference distribution over remaining words, the LLM was prompted to provide the probability of each word being associated with the predicted theme. To avoid invalid completions, we instructed the LLM to adhere to a strict format: "word: probability," and any deviations from this format were automatically detected as errors.

5 HUMAN-AGENT INTERACTION STUDY

5.1 Experimental Procedure

We evaluated our approach (preference-aware decision-making approach) against a myopic baseline in a real human-robot collaboration task conducted in our lab. The preference-aware approach leverages human preferences to optimize long-term team performance, following the policy outlined in Sect. **??**. In this approach, the agent infers human preferences using an LLM and formulates a policy by solving an MDP that accounts for the long-term expected rewards of all agents. In contrast, under the myopic condition, the agent disregards human preferences and follows a short-sighted policy, optimizing Equation 9 by assuming equal contributions from all agents. The experiment features a single factor (Robot Policy) with two levels. We designed a within-subject repeated measure study involving teams of two humans and a Nao humanoid robot. Each team participated in two tasks, each with a different robot policy. The tasks were counterbalanced, and each group was initially assigned the preference-aware or myopic condition as their starting scenario. We made the following hypotheses:

H1. Objective Measure of Task Performance: The task performance of the human-robot team will be better than the baseline condition, i.e., the preference-aware strategy will enable participants to produce high-quality stories with greater robot contribution compared to the baseline strategy.

H2. Subjective Performance Rating: Participants will rate the preference-aware strategy higher in terms of collaboration fluency, task experience, perceived robot competence, and interactivity compared to the baseline strategy.

Pre-Task Survey: Before starting the study, participants reviewed consent documents and task instructions. They then completed a pre-task survey to collect demographic information and details about their prior experience with robots. Participants were also asked to select a theme from a list of story themes and create a narrative using a given word based on the chosen theme. This narrative was later used to assess individual performance before the main task.

Human-Agent Collaborative Task Session: In this session, two human participants and the robot collaborated in a sequential storytelling game. A client-server system facilitated the coordination between the participants and the robot, while an interactive GUI displayed a storyboard containing the available words. Participants used the GUI to select words and continue the story in alignment with the chosen theme. The GUI also allowed participants to record their narratives, which were sent to the server, and view the current story with the narrator's name, previously selected words, and the team's accumulated points. The robot implemented either the preference-aware or myopic strategy during the game. Turns were taken clockwise, starting with the robot, with each participant given 30 seconds to think before continuing the narrative. The participant who selected the last word was responsible for concluding the story. At the start, the robot greeted the participants and explained the game rules. Upon task completion, the robot thanked participants for their participation and contributions.

In-Task Survey: After each task, participants completed a in-task survey to evaluate their experience. The survey included questions on various aspects, such as task engagement, collaboration fluency, and robot-specific assessments, using a 5-point Likert scale. Additionally, participants participated in a recall session, where they were asked to recall the story's main characters and key ideas.

Post-Session Survey: After completing both tasks, participants filled out a post-session survey. They were asked to choose the story they preferred from the two tasks and provide an explanation for their choice.

5.2 Participants

We conducted a human-agent collaborative storytelling study to evaluate our proposed approach. A total of 40 adults (2 people per team, resulting in a total of 20 teams) participated in the study, comprising 26 males (65%), 13 females (32.5%), and 1 individual (2.5%) who preferred not to disclose their gender. The mean age of the participants was 25 years (SD = 3.99). All participants were required to be fluent in English, available for in-person participation, and at least 18 years old. The sample consisted primarily of university students. Participants self-reported their experience with robots on a 5-point Likert scale (1: no experience, 5: expert-level experience), yielding an average score of 2.45. Each session lasted approximately one hour, and participants received a $15 gift card for their time. The study protocol was approved by the Institutional Review Board.

5.3 Measures

5.3.1 Objective Measures. To examine the impact of agent policies on participants' task performance and task quality within the framework of a storytelling task, we utilized six key metrics: story completion time, story length, robot narrative length, human participant narrative length, Narrative Structure Score (NSS), and story quality score. Metrics such as story completion time, story length, and narrative lengths were employed to assess participant productivity and engagement, while the NSS and story quality score were used to evaluate the coherence, creativity, and overall effectiveness of the produced stories. For NSS, we applied a coding scheme adapted from previous research on narrative structures [33, 36, 37]. For a given session s and participant i, the NSS was calculated as follows: $NSS_{s,i} = \frac{\text{Mentioned(CoreCharacters + CoreIdeas)}}{\text{All(CoreCharacters + CoreIdeas)}}$, which captures participants' ability to logically recount key plot elements. A perfect NSS score of 1.0 indicates that all core characters and ideas were mentioned. To assess story quality, we employed the average scores provided by three large language models (LLMs): Claude 3.5 Sonnet

[4] and Gemini 1.5 (Flash and Pro) [62]. Each model was given a standardized prompt to evaluate the stories generated under the two different robot policies and assign a score on a scale of 1 to 10. Notably, GPT-4 [1] was excluded from this evaluation to avoid bias, as parts of the robot's narrative incorporated into the stories were generated using this model.

5.3.2 Subjective Measures. To evaluate participants' perceptions of the agents, we employed a subset of items from a subjective fluency metric scale. This subset included a series of semantic differential scales designed to assess key dimensions such as *human-agent fluency*, *agent relative contribution*, *trust in agent*, *human-agent bond*, and *goal alignment*, all rated on a five-point Likert scale [25]. Additionally, we assessed participants' perceived task engagement, interactivity, and the agent's competence using a similar five-point Likert scale (1 = strongly disagree, 5 = strongly agree). These measures were adapted from established frameworks in the literature [69, 70]. Together, these subjective evaluations captured participants' attitudes towards the agents and their collaborative strategies during task performance.

6 RESULTS AND ANALYSIS

6.1 Task Performance

The paired-samples t-tests comparing the effects of the preference-aware strategy and the myopic baseline on collaborative task performance revealed mixed results. For story completion time, the preference-aware strategy (M = 542.01, SD = 185.75) took longer than the baseline (M = 494.30, SD = 115.26), with a mean difference of 47.72 (SD = 162.98), though this difference was not statistically significant, $t(19)$ = 1.309, p = .103, as shown in Fig. 4(a). In contrast, the preference-aware strategy produced significantly longer stories (M = 353.90, SD = 97.60) compared to baseline (M = 264.55, SD = 99.95), with a mean difference of 89.35 (SD = 108.71), $t(19)$ = 3.676, p < .001, as illustrated in Fig. 4(b). Human contribution was slightly higher in the preference-aware strategy (M = 238.00, SD = 92.86) than the baseline (M = 222.85, SD = 96.58), but this difference was not statistically significant, $t(19)$ = 0.716, p = .241, as shown in Fig. 4(c). However, robot contribution was significantly greater in the preference-aware strategy (M = 115.90, SD = 37.68) compared to baseline (M = 41.70, SD = 7.09), $t(19)$ = 8.851, p < .001, as shown in Fig. 4(d). Story quality also improved significantly under the preference-aware strategy (M = 7.10, SD = 0.29) compared to baseline (M = 6.52, SD = 0.71), $t(19)$ = 3.28, p = .002, as illustrated in Fig. 4(e). Finally, there was no significant difference in story recall scores (NSS) between the preference-aware (M = 0.62, SD = 0.22) and baseline strategies (M = 0.65, SD = 0.22), $t(39)$ = -0.731, p = .234, as seen in Fig. 4(f).

Summary: The preference-aware strategy produced notably longer stories and significantly enhanced the robot's contribution. These findings suggest that the preference-aware approach fosters more engaging and elaborate narratives, with the robotic agent playing a more active role. However, this strategy didn't significantly affect task completion time, human contribution, or recall scores. Interestingly, while the robot's involvement increased substantially, human participation remained consistent, indicating a

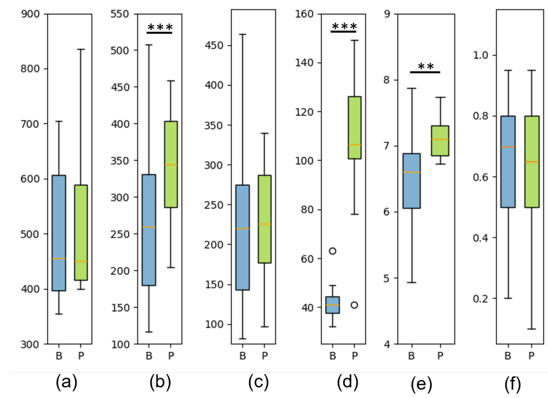

Figure 4: Box plot on participants task performance for the *Myopic baseline (B)* and *Preference-Aware (P)* strategies. Horizontal lines represent significant pairwise comparisons. Independent variables: (a) Story Time, (b) Story Length, (c) Human Contribution, (d) Robot Contribution, (e) Story Quality, (f) Recall Score. The significant values are shown in * ($* = p < .05, ** = p < .01, *** = p < .001$).

more balanced collaboration. This balance and improved story quality suggest that the preference-aware strategy effectively enhances the storytelling experience without overshadowing human input, potentially creating a more satisfying and productive creative partnership between humans and AI.

6.2 Perceived Collaboration Fluency

A paired-samples t-test was conducted to examine perceived collaboration fluency between the preference-aware and baseline strategies using several subjective fluency metrics: Human-Robot Fluency, Robot Relative Contribution, Trust in Robot, and the Bond and Goal subscales of the Working Alliance for H-R teams. The results indicate that the preference-aware strategy significantly improved fluency (M = 4.27, SD = 0.80) compared to the baseline (M = 3.55, SD = 1.09), $t(38)$ = 4.89, p < .001, as shown in Fig. 5(b). However, robot relative contribution was significantly lower for the preference-aware strategy (M = 3.71, SD = 1.02) than for the baseline (M = 4.05, SD = 0.97), $t(38)$ = -1.736, p = .045, as depicted in Fig. 5(e). Trust in the robot was significantly higher for the preference-aware strategy (M = 4.12, SD = 0.76) compared to baseline (M = 3.53, SD = 1.09), $t(38)$ = 4.50, p < .001, as seen in Fig. 5(g). For the Bond subscale, the preference-aware strategy (M = 2.756, SD = 0.67) indicate a significantly higher rating than the baseline (M = 2.55, SD = 0.59), $t(38)$ = 1.84, p = .037, as illustrated in Fig. 5(h). Lastly, performance ratings were also significantly higher for the preference-aware strategy (M = 3.35, SD = 0.53) than for baseline (M = 3.05, SD = 0.56), $t(38)$ = 2.62, p = .006 for task-goal, as seen in Fig. 5(c).

Summary: The preference-aware strategy significantly outperformed the baseline in several perceived collaboration fluency metrics. It demonstrated superior performance in overall fluency, trust, bonding, and performance measures, suggesting a more natural, reliable, and effective collaborative experience. Notably, the robot contribution scale, which measured agreement with "the human was the most important team member," indicate lower scores for the preference-aware strategy. This indicates that participants perceived the robot agent as having a more significant role in the collaboration when using this approach. This finding suggests that the preference-aware strategy successfully balanced the robotic agent's contributions with human input, creating a more equitable partnership where the human and the robotic agent were seen as essential to the team's success rather than the human dominating the interaction.

6.3 Perceived Task Experience

A paired-samples t-test was conducted to compare task experience ratings between the preference-aware strategy and the baseline. The results indicate that the mean task experience for the preference-aware strategy (M = 4.08, SD = 0.70) was significantly higher than for baseline (M = 3.82, SD = 0.78), with a mean difference of 0.26 (SD = 0.59) as illustrated in Fig. 5 (f). This difference was statistically significant, $t(38) = 2.73$, $p = .005$, indicating that the preference-aware strategy led to a significantly better task experience compared to the baseline strategy.

Summary: The analysis of perceived task experience demonstrates that the preference-aware strategy significantly enhanced participants' overall enjoyment and engagement during collaborative storytelling. With statistically considerably higher ratings and a moderate effect size, this strategy outperformed the baseline approach across various aspects of the experience, including enjoyment, desire for task continuation, willingness to recommend, relaxation, surprise, and overall engagement. These results suggest that by tailoring LLM responses to user preferences, the preference-aware strategy improved the quality of the stories produced and made the creative process more rewarding and immersive. This enhanced experience aligns with earlier findings of improved collaboration fluency and story quality. This indicates that the preference-aware approach creates a more satisfying, engaging, and user-centered environment for human-agent collaborative storytelling.

6.4 Perceived Robot Competence and Interactivity

A paired-samples t-test was conducted to compare perceived competence and interactivity scores between the preference-aware and baseline strategies. The results indicated that the mean competence score for the preference-aware strategy (M = 4.15, SD = 1.02) was significantly higher than for the baseline (M = 3.51, SD = 1.29), with a mean difference of 0.63 (SD = 1.20), $t(36)$ = 3.20, p = .001, as shown in Fig. 5(a). Similarly, the robot's interactivity score was significantly higher for the preference-aware strategy (M = 4.32, SD = 0.67) compared to the baseline (M = 3.73, SD = 1.15), with a mean difference of 0.59 (SD = 1.04), $t(36)$ = 3.45, p < .001, as illustrated in Fig. 5(d).

Summary: The analysis of perceived robot competence and interactivity demonstrates that the preference-aware strategy significantly enhanced participants' perceptions of the robot in both aspects. With statistically considerably higher ratings and large effect sizes, participants viewed the robot as more knowledgeable, competent, responsive, and interactive when it employed the preference-aware approach. This suggests that by tailoring its responses to

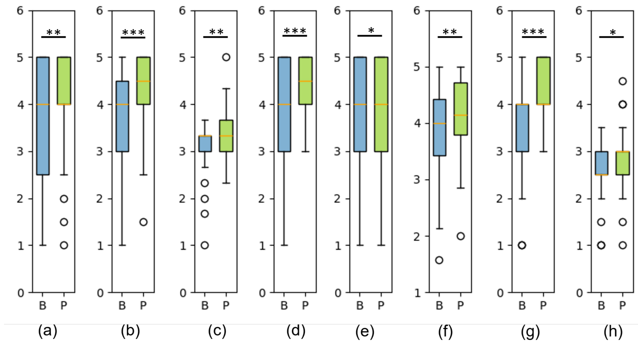

Figure 5: Box plot on participants perception for the *Myopic baseline (B)* and *Preference-Aware (P)* strategies. Horizontal lines represent significant pairwise comparisons. Independent variables: (a) Competence, (b) Fluency, (c) Task Goal, (d) Interactivity, (e) Robot Relative Contribution, (f) Task Experience, (g) Trust, (h) Bonding. The significant values are shown in * ($* = p < .05, ** = p < .01, *** = p < .001$).

user preferences, the robot appeared more capable of understanding and contributing meaningfully to the storytelling task while engaging more dynamically with its human partner. These findings align with earlier collaboration fluency and task experience results, reinforcing that the preference-aware strategy creates a more competent and engaging robotic collaborator. This improved perception of the robot's abilities and interactivity likely contributes to the higher trust, enjoyment, and overall satisfaction reported in the human-robot collaborative storytelling experience.

6.5 Qualitative Analysis

We collected qualitative data through open-ended questions in the post-task survey, capturing participants' evaluations of two stories created using different robot collaboration strategies. We performed open coding and thematic analysis on the responses [8, 17, 58]. Two team members independently reviewed the responses, allowing codes to naturally emerge. After multiple iterations, we clustered related codes into broader themes and organized the findings accordingly.

Coherence and Engagement in Storytelling: Participants highlighted the importance of coherence and meaningful story flow in their evaluations. For instance, several participants appreciated when the robot aligned with the overall narrative, such as P14, who mentioned, *"The flow of the story for the second one was better and ended up creating a meaningful story."* Others noted how the robot's coherence made the task easier. P27 said, *"The robot followed my chain of thought and developed the plot."* In contrast, some participants criticized instances where the robot failed to maintain coherence, with P11 noting, *"The second story seemed to ignore the context we gave it."* This feedback suggests that participants value a collaborative storytelling approach, with coherent storylines enhancing engagement and ease of task completion.

Adaptability and Context Awareness: The robot's ability to adapt to the context and direction set by human collaborators was highly valued. Participants appreciated when the robot could understand and work within the established narrative framework. P34 expressed this by saying, *"It followed my intuition and story goal of mine."* P26 highlighted the importance of context awareness, noting, *"As it used the same context I was intending to continue the narrative."* These examples demonstrate the significance of the robot's ability to align with and support the human collaborators' narrative intentions.

Detailed and Supportive Contributions: Participants often preferred more substantial and detailed contributions from the robot, especially when these inputs helped set up the next part of the story. P22 noted, *"The robot's contributions were more detailed and overall had more interesting content. This allowed me to be more creative."* P18 appreciated the robot's supportive role, saying, *"The robot gave longer sentences and actively tried to help us complete the story by giving sentences that related to the unused words."* These comments highlight the value of robot contributions that not only add depth to the story but also facilitate human creativity.

7 DISCUSSION

In this work, we proposed a method to incorporate human team preferences in robotic agents' decision-making in collaboration tasks. We also examined the method's effectiveness in a real human-robot collaborative storytelling task. Analyzing both task performance and subjective user experiences, we uncovered several key insights regarding the interaction between adaptive decision-making, task engagement, and collaboration fluency in human-robot teams.

Preference-aware strategies improve task performance and story quality: Our findings show that using a preference-aware strategy significantly improved the quality of the collaboratively created stories and enhanced overall task performance. Participants produced longer and more coherent narratives when the robot adapted its decisions based on the estimated preferences of human collaborators. This supports our hypothesis H1, as the preference-aware strategy enhanced the robot's contribution and led to better overall team performance and higher narrative quality compared to the baseline strategy. The improvement in story quality underscores the importance of adaptive decision-making, especially in tasks where creative contributions need to align with human expectations.

Balancing human and robot contributions enhances collaboration fluency: Our results indicate that balancing contributions between humans and robots is crucial for successful collaboration. In the preference-aware condition, the robot contributed significantly to the story without overshadowing the human participants, leading to higher ratings of collaboration fluency. This finding supports hypothesis H2, which posits that the preference-aware strategy would foster a more natural and balanced interaction between humans and robots, improving team dynamics. By making contributions aligned with human preferences, the robot helped create an environment of shared responsibility, enhancing engagement and fluency in the collaborative process.

Task engagement and perceived relevance of robot contributions: Interestingly, while the preference-aware strategy improved task performance, it also significantly enhanced participants' subjective experience. Participants rated the robot's contributions as more relevant and reported greater engagement when the robot considered their preferences. This finding supports both H1 and H2, as it demonstrates that the preference-aware strategy improves objective outcomes (H1) and enhances the subjective experience of collaboration (H2). The robot's ability to align its actions with the participants' narrative goals made the interaction feel more intuitive and satisfying, highlighting the importance of considering both performance and user experience in human-robot interaction design.

Importance of Adaptability and Context Awareness in Collaborative Tasks: The robot's ability to adapt to the evolving narrative context and align its actions with human collaborators' intentions was critical to the success of the storytelling task. Participants responded positively when the robot demonstrated an understanding of the direction they intended for the story and contributed in a way that supported the narrative flow. This adaptability allowed the robot to maintain coherence in the story, enhancing the overall quality and making the collaboration feel more natural. The robot's context awareness was highly valued, as it helped the team stay aligned on a common storyline, fostering a sense of partnership between humans and robots. This aligns with our hypothesis H2, emphasizing the importance of fluency and interaction quality in human-robot collaboration. The ability of the robot to interpret and extend the context set by human collaborators also contributed to higher engagement and task satisfaction. These findings highlight the importance of designing robotic agents that can dynamically respond to the task environment and human input, particularly in creative and interactive tasks where maintaining coherence is essential for team success.

8 CONCLUSION

Integrating human preferences into robot decision-making can significantly enhance collaboration quality and task performance in human-robot teams. In this work, we demonstrated that a preference-aware strategy, which adapts the robot's actions based on participants' preferences, improves story coherence, participant engagement, and overall collaboration fluency in a storytelling task. The robot's ability to align with human intentions created a more balanced and satisfying collaborative experience. In contrast to baseline approaches, this method promotes a deeper partnership between humans and robots by dynamically adjusting contributions to support human creativity and task goals. Our findings suggest that adaptive strategies considering human input can improve both the objective outcomes and the subjective collaboration experience. Moving forward, the qualitative insights from this study will inform the design of more refined prompts for future robot interactions, ensuring a deeper understanding of human preferences. Additionally, the stories generated during the study provide valuable data for training models to further personalize robot behaviors in real-time. Future work will focus on extending these preference-aware strategies to other collaborative tasks and exploring more sophisticated methods for learning and predicting human preferences from interaction data to enhance human-robot collaboration across various contexts.

REFERENCES

[1] Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, et al. 2023. Gpt-4 technical report. *arXiv preprint arXiv:2303.08774* (2023).

[2] Gati Aher, Rosa I Arriaga, and Adam Tauman Kalai. 2022. Using large language models to simulate multiple humans and replicate human subject studies.(2022). *DOI: https://doi. org/10.48550/ARXIV* 2208 (2022).

[3] Michael Ahn, Anthony Brohan, Noah Brown, Yevgen Chebotar, Omar Cortes, Byron David, Chelsea Finn, Chuyuan Fu, Keerthana Gopalakrishnan, Karol Hausman, et al. 2022. Do as i can, not as i say: Grounding language in robotic affordances. *arXiv preprint arXiv:2204.01691* (2022).

[4] AI Anthropic. 2024. Claude 3.5 sonnet model card addendum. *Claude-3.5 Model Card* (2024).

[5] Chris L Baker, Julian Jara-Ettinger, Rebecca Saxe, and Joshua B Tenenbaum. 2017. Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nature Human Behaviour* 1, 4 (2017), 0064.

[6] Chris L Baker and Joshua B Tenenbaum. 2014. Modeling human plan recognition using Bayesian theory of mind. *Plan, activity, and intent recognition: Theory and practice* 7, 177-204 (2014), 86.

[7] Erik Billing, Julia Rosén, and Maurice Lamb. 2023. Language models for human-robot interaction. In *Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*. 905–906.

[8] R Boyatzis. 1998. *Transforming qualitative information: Thematic analysis and code development*. Sage.

[9] Lindsey J Byom and Bilge Mutlu. 2013. Theory of mind: Mechanisms, methods, and new directions. *Frontiers in human neuroscience* 7 (2013), 413.

[10] Szeyi Chan, Shihan Fu, Jiachen Li, Bingsheng Yao, Smit Desai, Mirjana Prpa, and Dakuo Wang. 2024. Human and LLM-Based Voice Assistant Interaction: An Analytical Framework for User Verbal and Nonverbal Behaviors. *arXiv preprint arXiv:2408.16465* (2024).

[11] Boyuan Chen, Carl Vondrick, and Hod Lipson. 2021. Visual behavior modelling for robotic theory of mind. *Scientific Reports* 11, 1 (2021), 424.

[12] Kaiqi Chen, Jeffrey Fong, and Harold Soh. 2022. Mirror: Differentiable deep social projection for assistive human-robot communication. *arXiv preprint arXiv:2203.02877* (2022).

[13] Min Chen, Stefanos Nikolaidis, Harold Soh, David Hsu, and Siddhartha Srinivasa. 2018. Planning with trust for human-robot collaboration. In *Proceedings of the 2018 ACM/IEEE international conference on human-robot interaction*. 307–315.

[14] Yaran Chen, Wenbo Cui, Yuanwen Chen, Mining Tan, Xinyao Zhang, Dongbin Zhao, and He Wang. 2023. Robogpt: an intelligent agent of making embodied long-term decisions for daily instruction tasks. *arXiv preprint arXiv:2311.15649* (2023).

[15] Rohan Choudhury, Gokul Swamy, Dylan Hadfield-Menell, and Anca D Dragan. 2019. On the utility of model learning in hri. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 317–325.

[16] Paul F Christiano, Jan Leike, Tom Brown, Miljan Martic, Shane Legg, and Dario Amodei. 2017. Deep reinforcement learning from human preferences. *Advances in neural information processing systems* 30 (2017).

[17] Victoria Clarke and Virginia Braun. 2017. Thematic analysis. *The journal of positive psychology* 12, 3 (2017), 297–298.

[18] Melina Daniilidis, Nathaniel Steele Dennler, Maja Matarić, and Stefanos Nikolaidis. 2024. Adapting Task Difficulty in a Cup-Stacking Rehabilitative Task. In *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*. 374–378.

[19] Danny Driess, Fei Xia, Mehdi SM Sajjadi, Corey Lynch, Aakanksha Chowdhery, Brian Ichter, Ayzaan Wahid, Jonathan Tompson, Quan Vuong, Tianhe Yu, et al. 2023. Palm-e: An embodied multimodal language model. *arXiv preprint arXiv:2303.03378* (2023).

[20] Yiming Gao, Feiyu Liu, Liang Wang, Zhenjie Lian, Dehua Zheng, Weixuan Wang, Wenjin Yang, Siqin Li, Xianliang Wang, Wenhui Chen, et al. 2024. Enhancing Human Experience in Human-Agent Collaboration: A Human-Centered Modeling Approach Based on Positive Human Gain. *arXiv preprint arXiv:2401.16444* (2024).

[21] Sarah Gillet, Marynel Vázquez, Sean Andrist, Iolanda Leite, and Sarah Sebo. 2024. Interaction-shaping robotics: Robots that influence interactions between other agents. *ACM Transactions on Human-Robot Interaction* 13, 1 (2024), 1–23.

[22] Catalina Gomez Caballero, Junjie Yin, Chien-Ming Huang, and Mathias Unberath. [n.d.]. How Large Language Model-Powered Conversational Agents Influence Decision Making in Domestic Medical Triage Contexts. *Frontiers in Computer Science* 6 ([n. d.]), 1427463.

[23] Michael A Goodrich, Alan C Schultz, et al. 2008. Human–robot interaction: a survey. *Foundations and Trends® in Human–Computer Interaction* 1, 3 (2008), 203–275.

[24] Moritz A Graule and Volkan Isler. 2024. Gg-llm: Geometrically grounding large language models for zero-shot human activity forecasting in human-aware task planning. In *2024 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 568–574.

[25] Guy Hoffman. 2019. Evaluating fluency in human–robot collaboration. *IEEE Transactions on Human-Machine Systems* 49, 3 (2019), 209–218.

[26] Heye Huang, Jinxin Liu, Guanya Shi, Shiyue Zhao, Boqi Li, and Jianqiang Wang. 2024. Adaptive Decision-Making for Autonomous Vehicles: A Learning-Enhanced Game-Theoretic Approach in Interactive Environments. *arXiv preprint arXiv:2402.11467* (2024).

[27] Md Mofijul Islam, Alexi Gladstone, and Tariq Iqbal. 2023. Patron: perspective-aware multitask model for referring expression grounding using embodied multimodal cues. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 37. 971–979.

[28] Shyam Sundar Kannan, Vishnunandan LN Venkatesh, and Byung-Cheol Min. 2023. Smart-llm: Smart multi-agent robot task planning using large language models. *arXiv preprint arXiv:2309.10062* (2023).

[29] Callie Y Kim, Christine P Lee, and Bilge Mutlu. 2024. Understanding large-language model (llm)-powered human-robot interaction. In *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*. 371–380.

[30] Hema S Koppula, Ashesh Jain, and Ashutosh Saxena. 2016. Anticipatory planning for human-robot teams. In *Experimental Robotics: The 14th International Symposium on Experimental Robotics*. Springer, 453–470.

[31] Michal Kosinski. 2023. Theory of mind might have spontaneously emerged in large language models. *arXiv preprint arXiv:2302.02083* (2023).

[32] Jin Joo Lee, Fei Sha, and Cynthia Breazeal. 2019. A Bayesian theory of mind approach to nonverbal communication. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 487–496.

[33] Iolanda Leite, Marissa McCoy, Monika Lohani, Daniel Ullman, Nicole Salomons, Charlene Stokes, Susan Rivers, and Brian Scassellati. 2015. Emotional storytelling in the classroom: Individual versus group interaction between children and robots. In *Proceedings of the tenth annual ACM/IEEE international conference on human-robot interaction*. 75–82.

[34] Mike EU Ligthart, Mark A Neerincx, and Koen V Hindriks. 2020. Design patterns for an interactive storytelling robot to support children's engagement and agency. In *Proceedings of the 2020 ACM/IEEE international conference on human-robot interaction*. 409–418.

[35] Jacob P Macdonald, Rohit Mallick, Allan B Wollaber, Jaime D Peña, Nathan McNeese, and Ho Chit Siu. 2024. Language, Camera, Autonomy! Prompt-engineered Robot Control for Rapidly Evolving Deployment. In *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*. 717–721.

[36] Kathleen A McCartney and Katherine Nelson. 1981. Children's use of scripts in story recall. *Discourse Processes* 4, 1 (1981), 59–70.

[37] Fiona McGuigan and Karen Salmon. 2006. The influence of talking on showing and telling: Adult-child talk and children's verbal and nonverbal event recall. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition* 20, 3 (2006), 365–381.

[38] Austin Narcomey, Nathan Tsoi, Ruta Desai, and Marynel Vázquez. 2024. Learning Human Preferences Over Robot Behavior as Soft Planning Constraints. *arXiv preprint arXiv:2403.19795* (2024).

[39] Heramb Nemlekar, Neel Dhanaraj, Angelos Guan, Satyandra K Gupta, and Stefanos Nikolaidis. 2023. Transfer learning of human preferences for proactive robot assistance in assembly tasks. In *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*. 575–583.

[40] Eric Nichols, Leo Gao, and Randy Gomez. 2020. Collaborative storytelling with large-scale neural language models. In *Proceedings of the 13th acm siggraph conference on motion, interaction and games*. 1–10.

[41] Eric Nichols, Leo Gao, Yurii Vasylkiv, and Randy Gomez. 2021. Collaborative storytelling with social robots. In *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 1903–1910.

[42] Eric Nichols, Leo Gao, Yurii Vasylkiv, and Randy Gomez. 2021. Design and analysis of a collaborative story generation game for social robots. *Frontiers in Computer Science* 3 (2021), 674333.

[43] Stefanos Nikolaidis, Jodi Forlizzi, David Hsu, Julie Shah, and Siddhartha Srinivasa. 2017. Mathematical models of adaptation in human-robot collaboration. *arXiv preprint arXiv:1707.02586* (2017).

[44] Stefanos Nikolaidis, Anton Kuznetsov, David Hsu, and Siddhartha Srinivasa. 2016. Formalizing human-robot mutual adaptation: A bounded memory model. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 75–82.

[45] Stefanos Nikolaidis, Swaprava Nath, Ariel D Procaccia, and Siddhartha Srinivasa. 2017. Game-theoretic modeling of human adaptation in human-robot collaboration. In *Proceedings of the 2017 ACM/IEEE international conference on human-robot interaction*. 323–331.

[46] Stefanos Nikolaidis, Yu Xiang Zhu, David Hsu, and Siddhartha Srinivasa. 2017. Human-robot mutual adaptation in shared autonomy. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*. 294–302.

[47] Olivia Nocentini, Laura Fiorini, Giorgia Acerbi, Alessandra Sorrentino, Gianmaria Mancioppi, and Filippo Cavallo. 2019. A survey of behavioral models for social robots. *Robotics* 8, 3 (2019), 54.

[48] Liubove Orlov-Savko, Abhinav Jain, Gregory M Gremillion, Catherine E Neubauer, Jonroy D Canady, and Vaibhav Unhelkar. 2022. Factorial Agent Markov Model: Modeling Other Agents' Behavior in presence of Dynamic Latent Decision Factors. In *Proceedings of the 21st International Conference on Autonomous Agents and Multiagent Systems*. 982–990.

[49] Allen Z Ren, Anushri Dixit, Alexandra Bodrova, Sumeet Singh, Stephen Tu, Noah Brown, Peng Xu, Leila Takayama, Fei Xia, Jake Varley, et al. 2023. Robots that ask for help: Uncertainty alignment for large language model planners. *arXiv preprint arXiv:2307.01928* (2023).

[50] Dorsa Sadigh, Shankar Sastry, Sanjit A Seshia, and Anca D Dragan. 2016. Planning for autonomous cars that leverage effects on human actions.. In *Robotics: Science and systems*, Vol. 2. Ann Arbor, MI, USA, 1–9.

[51] Edward Schmerling, Karen Leung, Wolf Vollprecht, and Marco Pavone. 2018. Multimodal probabilistic model-based planning for human-robot interaction. In *2018 IEEE international conference on robotics and automation (ICRA)*. IEEE, 3399–3406.

[52] Pratyusha Sharma, Balakumar Sundaralingam, Valts Blukis, Chris Paxton, Tucker Hermans, Antonio Torralba, Jacob Andreas, and Dieter Fox. 2022. Correcting robot plans with natural language feedback. *arXiv preprint arXiv:2204.05186* (2022).

[53] Haojun Shi, Suyu Ye, Xinyu Fang, Chuanyang Jin, Layla Isik, Yen-Ling Kuo, and Tianmin Shu. 2024. MuMA-ToM: Multi-modal Multi-Agent Theory of Mind. *arXiv preprint arXiv:2408.12574* (2024).

[54] Lei Shi, Paul-Christian Bürkner, and Andreas Bulling. 2023. Inferring Human Intentions from Predicted Action Probabilities. *arXiv preprint arXiv:2308.12194* (2023).

[55] Lucy Xiaoyang Shi, Zheyuan Hu, Tony Z. Zhao, Archit Sharma, Karl Pertsch, Jianlan Luo, Sergey Levine, and Chelsea Finn. 2024. Yell At Your Robot: Improving On-the-Fly from Language Corrections. *arXiv preprint arXiv: 2403.12910* (2024).

[56] Ishika Singh, Valts Blukis, Arsalan Mousavian, Ankit Goyal, Danfei Xu, Jonathan Tremblay, Dieter Fox, Jesse Thomason, and Animesh Garg. 2023. Progprompt: Generating situated robot task plans using large language models. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 11523–11530.

[57] Ronal Singh, Tim Miller, Joshua Newn, Eduardo Velloso, Frank Vetere, and Liz Sonenberg. 2020. Combining gaze and AI planning for online human intention recognition. *Artificial Intelligence* 284 (2020), 103275.

[58] Anselm Strauss, Juliet Corbin, et al. 1990. *Basics of qualitative research*. Vol. 15. sage Newbury Park, CA.

[59] DJ Strouse, Kevin McKee, Matt Botvinick, Edward Hughes, and Richard Everett. 2021. Collaborating with humans without human data. *Advances in Neural Information Processing Systems* 34 (2021), 14502–14515.

[60] Ming Sun, Iolanda Leite, Jill Fain Lehman, and Boyang Li. 2017. Collaborative storytelling between robot and child: A feasibility study. In *Proceedings of the 2017 Conference on Interaction Design and Children*. 205–214.

[61] Yuqian Sun, Zhouyi Li, Ke Fang, Chang Hee Lee, and Ali Asadipour. 2023. Language as reality: a co-creative storytelling game experience in 1001 nights using generative AI. In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, Vol. 19. 425–434.

[62] Gemini Team, M Reid, N Savinov, D Teplyashin, Lepikhin Dmitry, T Lillicrap, JB Alayrac, R Soricut, A Lazaridou, O Firat, et al. 2024. Gemini 1.5: Unlocking multimodal understanding across millions of tokens of context. In arXiv [cs. CL]. arXiv.

[63] Andrea Thomaz, Guy Hoffman, Maya Cakmak, et al. 2016. Computational human-robot interaction. *Foundations and Trends® in Robotics* 4, 2-3 (2016), 105–223.

[64] Alessandro Umbrico, Riccardo De Benedictis, Francesca Fracasso, Amedeo Cesta, Andrea Orlandini, and Gabriella Cortellessa. 2023. A mind-inspired architecture for adaptive hri. *International Journal of Social Robotics* 15, 3 (2023), 371–391.

[65] Vaibhav V Unhelkar, Przemyslaw A Lasota, Quirin Tyroller, Rares-Darius Buhai, Laurie Marceau, Barbara Deml, and Julie A Shah. 2018. Human-aware robotic assistant for collaborative assembly: Integrating human motion prediction with planning in time. *IEEE Robotics and Automation Letters* 3, 3 (2018), 2394–2401.

[66] Vaibhav V Unhelkar, Shen Li, and Julie A Shah. 2020. Decision-making for bidirectional communication in sequential human-robot collaborative tasks. In *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*. 329–341.

[67] Vaibhav V Unhelkar, Shen Li, and Julie A Shah. 2020. Semi-supervised learning of decision-making models for human-robot collaboration. In *Conference on Robot Learning*. PMLR, 192–203.

[68] Vaibhav V Unhelkar and Julie A Shah. 2019. Learning models of sequential decision-making with partial specification of agent behavior. In *Proceedings of the AAAI conference on artificial intelligence*, Vol. 33. 2522–2530.

[69] Marynel Vázquez, Elizabeth J Carter, Braden McDorman, Jodi Forlizzi, Aaron Steinfeld, and Scott E Hudson. 2017. Towards robot autonomy in group conversations: Understanding the effects of body orientation and gaze. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*. 42–52.

[70] Marynel Vázquez, Elizabeth J Carter, Jo Ana Vaz, Jodi Forlizzi, Aaron Steinfeld, and Scott E Hudson. 2015. Social group interactions in a role-playing game. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction Extended Abstracts*. 9–10.

[71] Sai Vemprala, Rogerio Bonatti, Arthur Bucker, and Ashish Kapoor. 2023. Chatgpt for robotics: Design principles and model abilities. 2023. *Published by Microsoft* (2023).

[72] Sai H Vemprala, Rogerio Bonatti, Arthur Bucker, and Ashish Kapoor. 2024. Chatgpt for robotics: Design principles and model abilities. *IEEE Access* (2024).

[73] Philipp Wicke and Tony Veale. 2018. Interview with the Robot: Question-Guided Collaboration in a Storytelling System.. In *ICCC*. 56–63.

[74] Clarice Jiaying Wong, Yong Ling Tay, Ruohan Wang, and Yan Wu. 2016. Human-robot partnership: A study on collaborative storytelling. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 535–536.

[75] Jimmy Wu, Rika Antonova, Adam Kan, Marion Lepert, Andy Zeng, Shuran Song, Jeannette Bohg, Szymon Rusinkiewicz, and Thomas Funkhouser. 2023. Tidybot: Personalized robot assistance with large language models. *Autonomous Robots* 47, 8 (2023), 1087–1102.

[76] Ziyi Yang, Shreyas S Raman, Ankit Shah, and Stefanie Tellex. 2024. Plug in the safety chip: Enforcing constraints for llm-driven robot agents. In *2024 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 14435–14442.

[77] Chuang Yu, Baris Serhan, and Angelo Cangelosi. 2024. ToP-ToM: Trust-aware Robot Policy with Theory of Mind. In *2024 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 7888–7894.

[78] Chuang Yu, Baris Serhan, Marta Romeo, and Angelo Cangelosi. 2023. Robot theory of mind with reverse psychology. In *Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*. 545–547.

[79] Andy Zeng, Maria Attarian, Brian Ichter, Krzysztof Choromanski, Adrian Wong, Stefan Welker, Federico Tombari, Aveek Purohit, Michael Ryoo, Vikas Sindhwani, et al. 2022. Socratic models: Composing zero-shot multimodal reasoning with language. *arXiv preprint arXiv:2204.00598* (2022).

[80] Bowen Zhang and Harold Soh. 2023. Large language models as zero-shot human models for human-robot interaction. In *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 7961–7968.

[81] Hongxin Zhang, Weihua Du, Jiaming Shan, Qinhong Zhou, Yilun Du, Joshua B Tenenbaum, Tianmin Shu, and Chuang Gan. 2023. Building cooperative embodied agents modularly with large language models. *arXiv preprint arXiv:2307.02485* (2023).

[82] Michelle D Zhao, Reid Simmons, and Henny Admoni. 2023. Learning Human Contribution Preferences in Collaborative Human-Robot Tasks. In *Conference on Robot Learning*. PMLR, 3597–3618.

[83] Zhehua Zhou, Jiayang Song, Kunpeng Yao, Zhan Shu, and Lei Ma. 2024. ISR-LLM: Iterative Self-Refined Large Language Model for Long-Horizon Sequential Task Planning. In *2024 IEEE International Conference on Robotics and Automation (ICRA)*. 2081–2088. https://doi.org/10.1109/ICRA57147.2024.10610065

[84] Brian D Ziebart, Nathan Ratliff, Garratt Gallagher, Christoph Mertz, Kevin Peterson, J Andrew Bagnell, Martial Hebert, Anind K Dey, and Siddhartha Srinivasa. 2009. Planning-based prediction for pedestrians. In *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 3931–3936.