



Unveiling the Dynamics of Human Decision-Making: From Strategies to False Beliefs in Collaborative Human-Robot Co-Learning Tasks

Rajul Kumar

Electrical and Computer Engineering
George Mason University
Fairfax, VA, USA
rkumar29@gmu.edu

Yidi Huang

Electrical and Computer Engineering
George Mason University
Fairfax, VA, USA
yhuang35@gmu.edu

Ningshi Yao

Electrical and Computer Engineering
George Mason University
Fairfax, VA, USA
nyao4@gmu.edu



Figure 1: Sidewalk shuffle between (a) Pedestrian-Pedestrian (b) Pedestrian-Delivery robot

ABSTRACT

As robots become more integrated into humans' daily activities, it is essential to understand how human decision varies during co-learning with robots in real-world scenarios. Despite great advances in developing humanoid robots, which aims to foster a seamless collaborative world where humans and robots coexist, a gap remains in the social bond between humans and robots, particularly in tasks demanding optimal teamwork. In alignment with current pioneering efforts in the human-robot collaboration field, this paper presents an experimental study leading to a rationale analysis and classification of human behavioral dynamics during a joint collaborative pick-and-place task with a robotic arm. Our post-experimental analysis categorized human behavioral dynamics into three distinct broad categories, which are "strategic explorers and decoders", "reactive navigators and dynamic responders", and "score maximizers and ideal collaborators". We provide in-depth analysis for each group, exploring potential reasons for their observed behavioral patterns and irrational decisions substantiated by intuitions from psychological and behavioral game theory, including concepts of false belief and strategy development.

CCS CONCEPTS

- Human-centered computing → Computer supported cooperative work.



This work is licensed under a Creative Commons Attribution International 4.0 License.

HRI '24 Companion, March 11–14, 2024, Boulder, CO, USA
 © 2024 Copyright held by the owner/author(s).
 ACM ISBN 979-8-4007-0323-2/24/03.
<https://doi.org/10.1145/3610978.3640743>

KEYWORDS

Human Decision Making, Co-learning, Joint Tasks, Human Robot collaboration, False belief

ACM Reference Format:

Rajul Kumar, Yidi Huang, and Ningshi Yao. 2024. Unveiling the Dynamics of Human Decision-Making: From Strategies to False Beliefs in Collaborative Human-Robot Co-Learning Tasks. In *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction (HRI '24 Companion)*, March 11–14, 2024, Boulder, CO, USA. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3610978.3640743>

1 INTRODUCTION

At least once in our lifetime, we all experience a moment walking through a narrow corridor or hallway, when another person approaches from the opposite direction, as illustrated in Figure 1 (a). Although we intend to avoid each other and proceed on our respective paths, we accidentally both move to the right, then to the left, and again to the right, in unintended synchronization. This undesirable and awkward situation is commonly known as sidewalk shuffle, hallway dancing, or footpath foxtrot. A similar situation can occur in robot-robot or human-robot navigation, as shown in Figure 1 (b). This can cause significant hindrance in smooth human-robot collaboration, especially now, when ground and aerial robots serve humans in various applications, including warehouse manipulation, delivery, rescue operations, and entertainment [5, 11, 17, 19, 22].

To avoid such an oscillation in human-robot collaboration, Young et al. developed a dual expert chatter algorithm (or DEA) and a human-aware dual expert algorithm (or HADEA), which they tested in simulations for a double feedback closed-loop human-robot co-learning system [27]. The authors refer to the term "chatter" to the instances where humans and robots make the same choices. They

used the famous Rescorla-Wagner model [18] in psychology to simulate human choices based on robot past choices [2, 15, 16, 20]. However, in joint human-robot real-world tasks, human decision-making behavior does not strictly obey the Rescorla-Wagner model [4, 6, 8, 10, 25]. It may be more rational, strategic, irrational, random, or a combination of all these [3, 7, 13, 21].

A brain monitoring study revealed that a specific region of the brain linked to social cognition is more active in human-human interactions than in human-robot interactions [23], suggesting that the lack of social bonding in the latter could hinder effective collaboration in tasks requiring high cooperation from both sides, such as joint pick-and-place tasks. In line with these findings, it is essential to address the core challenge of accurately identifying human decision-making models during co-learning tasks with robots, to bridge the existing cognitive and social gaps between humans and robots [1, 9, 14, 24, 26]. Therefore, this paper presents an experimental study on a joint, concurrent pick-and-place task performed by a human and a robot, involving 11 human participants ($N = 11$). Additionally, we have classified the observed behaviors of the participants into three groups and provided a comprehensive reasoning for their behavioral choices and actions during the game.

2 EXPERIMENTAL FRAMEWORK AND PROCEDURE

Contrary to the hallway dancing problem, the term “chatter” in this paper for the pick-and-place task denotes the misalignment of choices between human and robot while simultaneously selecting between a red and blue marker, as shown in Figure 2. The co-learning experiment detailed in Section 2.2 was inclusive to all and did not target any specific demographic population, except age: all human subjects were university students possessing developed theory of mind [12].

2.1 Experimental Scenario and Setup

We leveraged a 6-DOF collaborative Ned2 robotic arm as a counterpart to human subjects. For six subjects, the robot operated using the DEA algorithm, while for the other five, the HADEA algorithm was used [27]. To further challenge the human decision-making process, we subjected some participants to eight initial random choices by robot, discussed in Section 4.

The participants were unaware that the robot followed these algorithms to avoid chatter and would try to be collaborative. Our intent was to observe whether or not they would make significant efforts to figure out if the robot was following a pre-set algorithm or predictable pattern. The observations of this experiment were discussed in Section 3. The implementation of chatter-avoiding DEA and HADEA algorithms in the context of this pick-and-place task is shown in Algorithm 1, where ' W_{Blue} ' and ' W_{Red} ' represent the weights for choosing red and blue markers, and ' R_t ' and ' H_t ' denote the robot's and the human's choice at each iteration.

2.2 Co-learning Experimental Workflow

At the beginning of the experiment, each participant was motivated to score as high as possible over 20 iterations of selecting a marker, either red or blue, to match the robot's choice every time. The step-by-step experimental procedure is described in the following.

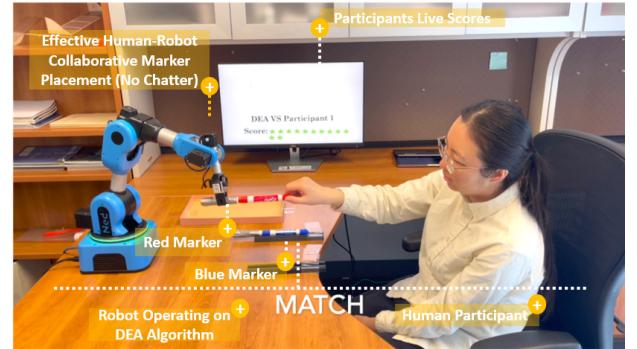


Figure 2: Experimental Setup showing No Chatter Instance

- Step 1: The participant was instructed to sit at a safe and specified distance from the robotic arm.
- Step 2: The participant was informed that they must choose a marker, red or blue, on the count of “1, 2, 3, go...”. And the robot would also make its selection simultaneously.
- Step 3: If the participant chooses the same marker as the robot, they would earn a 1 point and be required to place the marker in a nearby box, perceived as a robot completing the task collaboratively with sufficient human assistance, as shown in Figure 2. If the participant chooses a different marker, they would score 0 and would not perform any action, waiting instead for the next countdown to begin, while the robot attempted the task without assistance. And without human assistance, the robot struggled to put the marker in the box.
- Step 4: This selection game was repeated for 20 iterations. During the experiment, the participant’s score for each game iteration was displayed on a nearby screen.

Algorithm 1 Dual Expert Algorithm

```

1: procedure ROBOT CHOICE( $H_t$ )
2:    $W_{Blue} = W_{Red} = 0.5$ 
3:   for  $i = 1$  to 20 do
4:      $R_t \leftarrow$  Blue
5:      $W_{R_t} \leftarrow W_{Blue}$ 
6:     if  $W_{Blue} < W_{Red}$  then
7:        $R_t \leftarrow$  Red
8:        $W_{R_t} \leftarrow W_{Red}$ 
9:     end if
10:    if  $R_t = H_t$  (No Chatter) then
11:       $W_{R_t} \leftarrow W_{R_t} < 0.5 ? W_{R_t} \times 2 : W_{R_t}$ 
12:    else (Chatter)
13:       $W_{R_t} \leftarrow W_{R_t} / 2$ 
14:    end if
15:  end for
16: end procedure

```

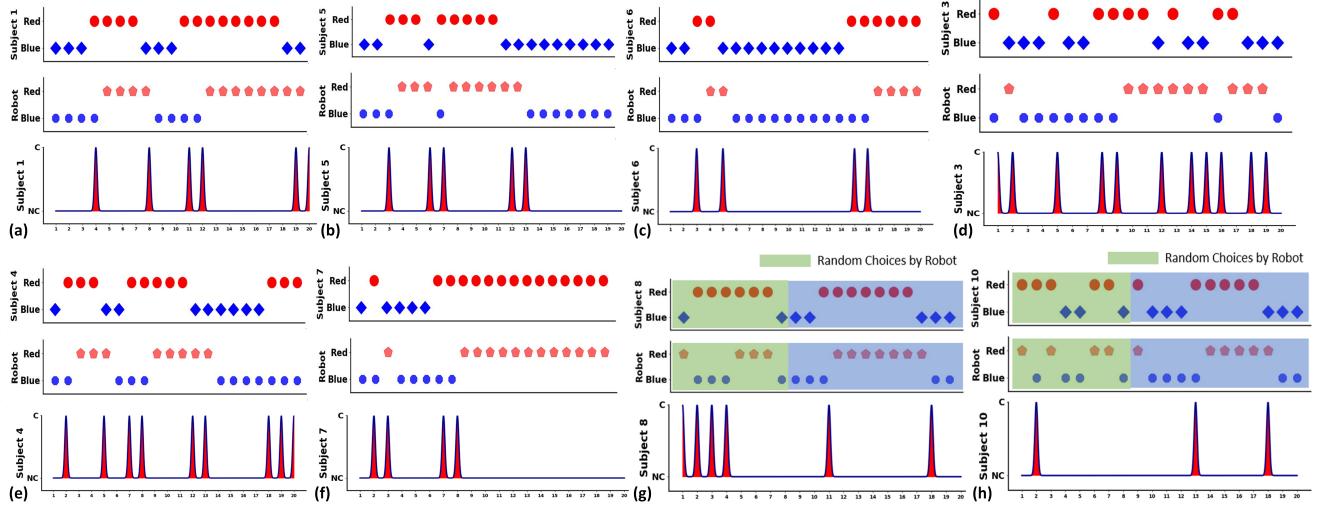


Figure 3: Experiment data: (a) (b) and (c) as Strategic Explorers; (d) and (e) as Dynamic Responders; (f) as Score Maximizers; and (g) and (h) with 8 random robot choices at the beginning.

3 BEHAVIOURAL CLASSIFICATION AND RATIONALE ANALYSIS

It is clear from our observations that human behavior exhibits considerable diversity, with decisions varying significantly from individual to individual. This variability presents a challenge to come up with a robust, accurate, and one-time model that encapsulates all aspects of human behavior. However, it seems practical and of real-world applicability to categorize behaviors into distinct groups, allowing the development of targeted behavioral models for each classified group. Subjects 1, 2, 5, and 6 were grouped as outlined in Section 3.1; Subjects 3 and 4 in Section 3.2; and Subject 7 in Section 3.3.

3.1 Strategic Explorers and Decoders

In this group, participants showed a common trend of sequences, including initial exploration and pattern decoding, followed by strategy adjustment and one or two instances of false belief towards the end. The behavior of subjects 1, 2, 5 and 6 closely mirrors with slight variations in the occurrence of these events shown in Figure 3 (a) (b) and (c).

Figure 4 explains the behavior logic and the response rationales of Subject 1. From the human choices depicted in Figure 3 (a) of subject 1, we can see that subject started by selecting blue and succeeded in the first 3 trials. However, after 3 trials, the subject switched to red possibly for two primary reasons: first, there was a greater probability that the subject wanted to test the robot, especially since the game had just started. The participant might have been curious about how the robot would react to a change in their choice and if the robot's actions were genuinely influenced by their decisions. Second, there was a lesser chance that the subject had a false belief about the robot's choice. Given it was early in the game and the robot hadn't switched colors before, the probability of the participant assuming the robot would choose red was relatively low. However, it is aberrant that choosing red has a higher probability.

At this point, the subject also began making efforts to decode the robot's pattern, and it seems the subject perceived the robot has a pre-defined pattern, of changing its choice from red to blue (or vice versa) after every 3 or 4 iterations. With this perceived logic, the subject adjusted their choice from blue to red in the 11th iteration. At this point, it became evident to the subject that the robot was adapting to their choices, suggesting that the human was leading the robot in decision making.

As time passed, the subject became more motivated to achieve a higher score, prompting a significant change in their previously developed strategy. After three consecutive red choices, the subject decided to test their previous strategy in the 14th iteration sticking to red. Towards the end, it appears that the subject switched to blue, primarily driven by a false belief about robot action. The subject might have assumed that since they had previously chosen blue, and given that several iterations had passed without selecting blue, the robot might anticipate a blue choice. This led to the subject choosing blue in the 19th iteration. However, the robot, operating on a chatter-avoiding DEA algorithm, took a few iterations to avoid chatter and reduce weight for red choice, resulting in the robot choosing red for the last two iterations.

3.2 Reactive Navigators and Dynamic Responders

From the chatter frequency plots in Figure 3 (d) and (e) of subjects 3 and 4, it is evident that they exhibit the highest chatter frequencies among all participants, with 11 and 9 number of chatter spikes, respectively. Consequently, they achieved the lowest scores in the game, of 9 and 11 points, respectively, out of 20. These plots suggest that the behavior of these subjects is a mixture of reactive and predictive decision-making. They often provide an immediate counter-response to the robot's selections, through out the experiment, indicating an inclination to be more on outsmarting the robot on a move-by-move basis instead of sticking to their previously

developed strategy or attempting to decode a pattern in the robot's decisions.

Unlike the other participants, these participants did not maintain a consistent color choice for extended periods to discover the robot's reactions; instead, their decisions were noticeably more dynamic. Modeling this small subset of individuals poses a significant challenge due to the aforementioned uncertainties and frequent irrationalities in the behaviors.

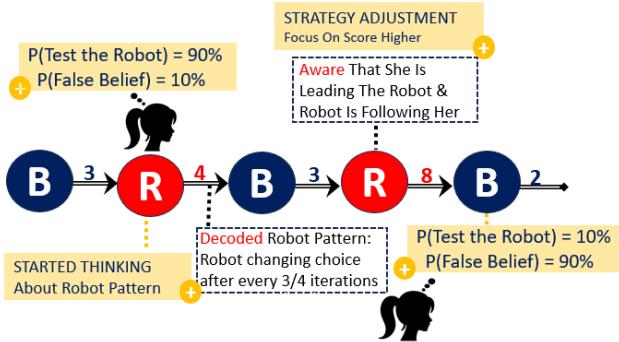


Figure 4: Rationale analysis of subject 1's decision behaviour

3.3 Score Maximizers and Ideal Collaborators

In this group, we classify participants who are intensely focused on achieving a high score and who aim to collaborate effectively with the robot to complete the task as soon as possible without testing or tricking the robot much. Figure 5 explains Subject 7's behavioral logic and response rationales.

From the decision plot shown in Figure 3 (f), we observe that Subject 7 started with a blue choice and, perhaps out of curiosity to test the robot, quickly switched to red to see the robot's reaction. Subsequently, the subject reverted to the initial blue choice. This rapid alternation between blue and red at intervals of one trial suggests a score-centric behavior, indicating that the subject was not keen on spending time testing or outsmarting the robot. Instead, their main focus was on achieving a high score.

Subject maintains this blue choice until the 6th iteration with an impressive score of 4 out of 6. However, at this point, the subject developed the false belief that the robot would now select red. This assumption was possibly based on the human's switch to red from blue during the 2nd iteration, after which the robot, operating on the DEA algorithm, also changed to red in the 3rd iteration to avoid chatter with human's choice. This sudden false realization excites the subject to shift from blue to red in the 6th iteration. After recognizing that subject is leading and guiding the robot's decisions, the human confidently pursued their ultimate goal of maximizing their score and consistently chose red until the end.

4 NON-DETERMINISTIC ROBOT CHOICES

To increase the cognitive load on human decision making, we make the initial 8 choices of robot random. From 9th iteration, the robot started following the DEA algorithm. Figures 3 (g) and (h) show two notable observations; subject 10 achieved the highest score; luckily, during their initial 8 trials, there was only 1 instance of

chatter. On the contrary, Subject 8 encountered chatter in first 4 of the initial eight trials. This analysis suggests that, despite having a logical and sequential model for human behavior, unpredictable and contrasting outcomes can occur at any time.

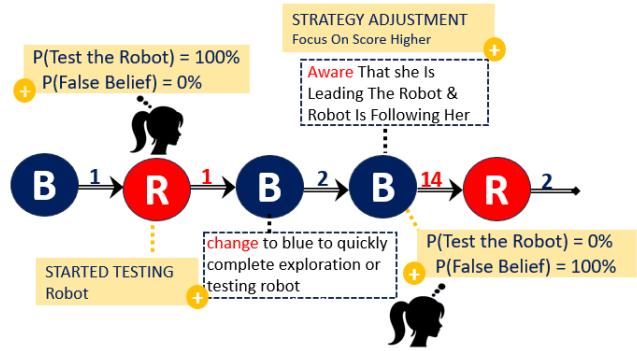


Figure 5: Rationale analysis of subject 7's decision behaviour

5 CONCLUSION AND FUTURE WORK

In this paper, we have explored the chaos of human decision-making process during a joint co-learning pick-and-place task between a human and a collaborative robotic arm. While the arm operates on a Dual-Expert Algorithm and a Human-Aware Dual-Expert Algorithm to avoid chatter and maintain adaptiveness, it quickly adjusts to match human decisions post-chatter occurrence. Through the execution of human-robot co-learning experimental tasks, we have collected decision choices (red/blue) for both human participants and robot. And we tried to infer common trends and outliers in human decisions. We have classified the human behavior into three broad, distinct categories: firstly, those subjects who believed the robot followed a pre-defined pattern so they attempted to decode it initially; secondly, subjects who exhibited more antagonistic and reactive behaviors in response to the robot's choices; and thirdly, those who focused on maximizing their performance by aiming to score highly during the game. Across all groups, we have observed a sudden episode of false belief in participants about the robot's next action, leading to a deviation from their current choice or strategy.

Acknowledging the imperative need to bridge the cognitive and social gaps that currently exist, thus making collaborative tasks more efficient and effective, our work is expected to serve as a foundational framework for developing robust human decision-making mathematical models. Ultimately, these models could assist in developing more robust chatter-free collaborative expert algorithms for robots in various real world applications, such as warehouse robots, ground and aerial delivery robots that experience human-robot or robot-robot adversarial problems.

REFERENCES

- [1] Dario Antonelli and Dorota Stadnicka. 2019. Predicting and preventing mistakes in human-robot collaborative assembly. *IFAC-PapersOnLine* 52, 13 (2019), 743–748. <https://doi.org/10.1016/j.ifacol.2019.11.204>
- [2] Anja Austermann and Seiji Yamada. 2009. Learning to understand parameterized commands through a human-robot training task. In *The 18th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2009)*. Toyama, Japan, 757–762. <https://doi.org/10.1109/ROMAN.2009.5326220>

- [3] Giulia Calabretta, Gerda Gemser, and Nachoem M. Wijnberg. 2017. The Interplay between Intuition and Rationality in Strategic Decision Making: A Paradox Perspective. *Organization Studies* 38, 3-4 (2017), 365–401. <https://doi.org/10.1177/0170840616655483>
- [4] Gianluca Calcagni, Ernesto Caballero-Garrido, and Ricardo Pellón. 2020. Behavior Stability and Individual Differences in Pavlovian Extended Conditioning. *Frontiers in Psychology* 11 (2020), 612. <https://doi.org/10.3389/fpsyg.2020.00612>
- [5] Li-Keng Cheng. 2023. Effects of service robots' anthropomorphism on consumers' attribution toward and forgiveness of service failure. *Journal of Consumer Behaviour* 22, 1 (2023), 67–81. <https://doi.org/10.1002/cb.2112>
- [6] Andrew Farley, Jie Wang, and Joshua A. Marshall. 2022. How to pick a mobile robot simulator: A quantitative comparison of CoppeliaSim, Gazebo, MORSE and Webots with a focus on accuracy of motion. *Simulation Modelling Practice and Theory* 120 (2022), 102629. <https://doi.org/10.1016/j.simpat.2022.102629>
- [7] Piotr Fratczak, Yee Mey Goh, Peter Kinnell, Andrea Soltoggio, and Laura Justham. 2019. Understanding Human Behaviour in Industrial Human-Robot Interaction by Means of Virtual Reality. In *Proceedings of the Halfway to the Future Symposium 2019*, 1–7. <https://doi.org/10.1145/3363384.3363403>
- [8] Samuel J. Gershman. 2015. A unifying probabilistic view of associative learning. *PLoS Computational Biology* 11, 11 (2015), e1004567. <https://doi.org/10.1371/journal.pcbi.1004567>
- [9] Adrien Gregorj, Zeynep Yücel, Francesco Zanlungo, Claudio Feliciani, and Takayuki Kanda. 2023. Social aspects of collision avoidance: a detailed analysis of two-person groups and individual pedestrians. *Scientific Reports* 13 (2023), 5756. <https://doi.org/10.1038/s41598-023-32883-z>
- [10] Erin Hedlund-Botti and Matthew C. Gombolay. 2023. Investigating Learning from Demonstration in Imperfect and Real World Scenarios. In *Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction (HRI '23)*, 769–771. <https://doi.org/10.1145/3568294.3579980>
- [11] Shanie Honig and Tal Oron-Gilad. 2018. Understanding and Resolving Failures in Human-Robot Interaction: Literature Review and Model Development. *Frontiers in Psychology* 9 (2018). <https://doi.org/10.3389/fpsyg.2018.00861>
- [12] Baris Korkmaz. 2011. Theory of Mind and Neurodevelopmental Disorders of Childhood. *Pediatric Research* 69 (2011), 101–108. <https://doi.org/10.1203/PDR.0b013e318212c177>
- [13] Benedikt Leichtmann, Verena Nitsch, and Martina Mara. 2022. Crisis Ahead? Why Human-Robot Interaction User Studies May Have Replicability Problems and Directions for Improvement. *Frontiers in Robotics and AI* 9 (2022). <https://doi.org/10.3389/frobt.2022.838116>
- [14] Dylan P. Losey and Dorsa Sadigh. 2019. Robots that Take Advantage of Human Trust. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. Macau, China, 7001–7008. <https://doi.org/10.1109/IROS40897.2019.8968564>
- [15] Robert Lowe, Alexander Almér, and Christian Balkenius. 2019. Bridging Connectionism and Relational Cognition through Bi-directional Affective-Associative Processing. *Open Information Science* 3, 1 (2019), 235–260. <https://doi.org/10.1515/opis-2019-0017>
- [16] André Lizardo, Eduardo Alonso, and Esther Mondragón. 2017. A Rescorla-Wagner Drift-Diffusion Model of Conditioning and Timing. *PLoS Computational Biology* 13, 11 (2017). <https://doi.org/10.1371/journal.pcbi.1005796>
- [17] Christoforos Mavrogiannis, Francesca Baldini, Allan Wang, Dapeng Zhao, Pete Trautman, Aaron Steinfeld, and Jeann Oh. 2023. Core Challenges of Social Robot Navigation: A Survey. *Journal of Human-Robot Interaction* 12, 3 (2023), Article 36. <https://doi.org/10.1145/3583741>
- [18] Ralph R. Miller, Robert C. Barnet, r, and Nicholas J. Grahame. 1995. Assessment of the Rescorla-Wagner model. *Psychological Bulletin* 117, 3 (1995), 363–386. <https://doi.org/10.1037/0033-2909.117.3.363>
- [19] Debasmita Mukherjee, Kashish Gupta, and Homayoun Najjaran. 2022. A Critical Analysis of Industrial Human-Robot Communication and Its Quest for Naturalness Through the Lens of Complexity Theory. *Frontiers in Robotics and AI* 9 (2022). <https://doi.org/10.3389/frobt.2022.870477>
- [20] Robert A. Recorla and Allan R. Wagner. 1972. *A Theory of Pavlovian Conditioning: Variations in the Effectiveness of Reinforcement and Nonreinforcement*. Appleton-Century-Crofts, New York, 64–99 pages.
- [21] Herbert A. Simon. 1993. Decision Making: Rational, Nonrational, and Irrational. *Educational Administration Quarterly* 29, 3 (1993), 392–411. <https://doi.org/10.1177/0013161X93029003009>
- [22] Christina Soyoung Song and Youn-Kyung Kim. 2022. The role of the human-robot interaction in consumers' acceptance of humanoid retail service robots. *Journal of Business Research* 146 (2022), 489–503. <https://doi.org/10.1016/j.jbusres.2022.03.087>
- [23] Nicolas Spatola and Thierry Chaminade. 2022. Precuneus brain response changes differently during human–robot and human–human dyadic social interaction. *Scientific Reports* 12 (2022), 14794. <https://doi.org/10.1038/s41598-022-14207-9>
- [24] Pete Trautman. 2017. Breaking the Human-Robot Deadlock: Surpassing Shared Control Performance Limits with Sparse Human-Robot Interaction. <https://doi.org/10.15607/RSS.2017.XIII.005>
- [25] Joanna Oi-Yue Yau and Gavan P. McNally. 2023. The Rescorla-Wagner model, prediction error, and fear learning. *Neurobiology of Learning and Memory* 203 (2023), 107799. <https://doi.org/10.1016/j.nlm.2023.107799>
- [26] Wenfeng Yi, Wenhan Wu, Xiaolu Wang, and Xiaoping Zheng. 2023. Modeling the Mutual Anticipation in Human Crowds With Attention Distractions. *IEEE Transactions on Intelligent Transportation Systems* 24, 9 (2023), 10108–10117. <https://doi.org/10.1109/TITS.2023.3268315>
- [27] Carol Young, Ningshi Yao, and Fumin Zhang. 2019. Avoiding Chatter in an Online Co-Learning Algorithm Predicting Human Intention. In *2019 IEEE 58th Conference on Decision and Control (CDC)*. Nice, France, 6504–6509. <https://doi.org/10.1109/CDC40024.2019.9030038>