Incorporating Rivalry in Reinforcement Learning for a Competitive Game

Pablo Barros, Ana Tanevska, Ozge Nilay Yalcin, Alessandra Sciutti*

Abstract

Recent advances in reinforcement learning with social agents have allowed us to achieve human-level performance on some interaction tasks. However, most interactive scenarios do not have performance alone as an end-goal; instead, the social impact of these agents when interacting with humans is as important and, in most cases, never explored properly. This preregistration study focuses on providing a novel learning mechanism based on a rivalry social impact. Our scenario explores different reinforcement learning-based agents playing a competitive card game against a human player. Our goal is to investigate if we can change the assessment of these agents from a human perspective by changing the agent's playing style based on a predicted and synthesized rivalry between agents and humans.

1 Introduction

The social aspects of interaction are usually dimmed when optimizing an artificial agent through reinforcement learning [1]. Most of the training loop must be done in an offline manner [1], or focuses on optimizing objective metrics that do not directly involve social aspects, such as planners [2] or human annotation feedback [3]. Most of the common success metrics in this regard are related to solving the task in fewer steps, reducing predicted values, or achieving some predefined intermediate objective goals. When the interaction with humans is the main goal, such agents are evaluated mostly based on their objective performance [2]. In the few examples where humans are present in the loop, the success measures are mostly related to the embodied interaction [4, 5], and not underlying decision-making process that these agents learned.

A scenario where these problems arise is in competitive interaction. In a competitive game scenario, an agent can learn through reinforcement learning how to adapt towards its opponents [6], even when these opponents are humans [7]. However, it is extremely difficult to measure the social aspects of this interaction, without relying on typical human-robot or human-computer interaction schemes [8]. Although providing important insight on some social aspects, these evaluations usually focus on controlled lab-scenarios [9], production of different robotic behavior [10], and dialogue [11]; this in turn neglects exploring how the agents' various learning strategies influence their explicit behaviour and interaction with humans [12], despite it being one of their most important characteristics. This can be evidenced even in the new area of explainable reinforcement learning [13, 14].

In this preregistration study, we address the problem of including social aspects in the learning strategies of artificial agents in a competitive scenario. We propose an objective human-centered metric, based on rivalry [15], to compose the reward function of the agents. Rivalry is a subjective social relationship arising between two actors based on the competitive characteristics of an individual, as well as the increasing stakes and psychological involvement in the situation [16]. We chose rivalry as it showcases the competing relation between individuals, which often affects their motivation and performance during game play [16, 17]. Our scenario will be evaluated using the multiplayer Chef's

---

*Pablo Barros, Ana Tanevska and Alessandra Sciutti are with the Italian Institute of Technology, Italy. Ozge Nilay Yalcin is with the Department of Computer Science, University of British Columbia, Canada.

Hat Card game and will include artificial agents based on Deep Q-Learning (DQL) and Proximal Policy Optimization (PPO). We separate our scenario into two steps: first we will map the original behavior of such algorithms when playing the game using the novel rivalry score. Second, we predict this metric by having agents identifying human rivals on inference-time, and use this information to update their behavior. To validate the behavior update, and to have a baseline for learning it with the agents, several humans will play the game against the agents and based on a series of explicit and implicit metrics, we will calculate the level of rivalry that each agent yields.

In this paper, we formalize these problems by proposing the following hypotheses:

- *Agents trained with different RL strategies yield distinct rivalry when playing the chef's hat card game against humans.*
- *Agents trained with a predicted rivalry as a part of their reward function can modulate human responses on a rivalry scale.*

In the remainder of this paper, we detail our motivation, methods, and our experimental setup, and present a short discussion on the possible impact that our research has on explainable reinforcement learning and related fields.

2 Related Work

Reinforcement learning in competitive games. In the late 1990ies, several researchers tried to identify the impact of the Deep Blue artificial chess player [18] on the development of artificial intelligence [19, 20]. They all argue that beyond the technical challenge of beating a human, there is an underlying impact on how this agent affects the opponents' behavior during the entire interaction. Over time, these investigations were let aside by the mainstream community, which focused mostly on solving more complex problems. This vision is reflected by the recent development of deep reinforcement learning and the research on training artificial agents to play competitive games that have flourished since [21]. AlphaGo [22] demonstrated that these agents can play competitively against humans in very complex games. The recent development of agents that play the StarCraft computer game [7] pushes these boundaries even further. These agents learn how to adapt to dynamic environments, how to map hypercomplex states and actions, and how to learn new strategies [23]. Most of the studies, however, focus on the final goal of these agents: how to be competitive against humans. None of them focus on understanding the impact that such agents had on human opponents.
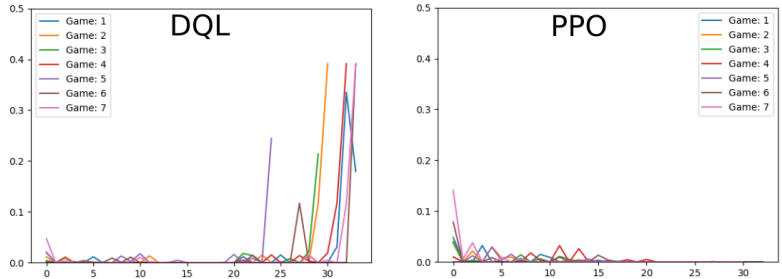

Figure 1: Illustration of the highest Q-Values of an agent trained with DQL and another trained with PPO, playing seven games of the Chef's Hat Card game.

DQL and PPO Playing Behavior on the Chef's Hat Card Game. In the same development wave, it was recently investigated the design and development of reinforcement learning agents to play the four players Chef's Hat competitive card game [24]. These agents were based on Deep Q-Learning (DQL) [25] and Proximal Policy Optimization (PPO) [26], and achieved success in learning how to win the game in different tasks: playing against random agents, self-play, and online adaptation towards the opponents. It was observed, however, that these agents present different behavior during game-play while maintaining a similar objective performance measured by overall wins over a series of games. By observing the Q-values of each agent during an entire match, it is possible to see in Figure 1 that the DQL agent usually presents higher Q-values at the end of a game, while the PPO

agent presents them at the beginning of the game, but with lower intensity. This led to a general understanding that, although both effective, these agents learned different strategies. What has not yet been done is to measure the social impact that such strategies can have when the agents play against humans.

Human-centric Analysis of RL. When analyzing the impact of artificial agents on humans, there are now decades of studies focused on Human-Robot [27] and Human-Computer [28] Interaction (HRI and HCI respectively). Most of these studies, however, focus on optimizing RL agents to solve a specific task, even social ones, without having much feedback on the social aspects of the task as part of their learning mechanism. Such agents are usually designed to learn an expected outcome, such as improving engagement [4], or to imitate humans [6]. None of the most recent studies focus on extracting the intrinsic behavior bias that different learning schemes apply to the final agent.

Rivalry in Competitive Games. One way to explain the behavior of an agent as a factor of its learning strategy is to measure its impact on humans. There exist several social metrics that take into consideration the interaction [29], but most of them focus on the subjective impressions humans have on the embodied interaction [30], the subjective quality of the interaction [31] and/or the efficiency of the interaction to solve a task [32]. In a competitive game, however, one of the most informative metrics is the rivalry [15] between the human and the agents. Rivalry is defined as a competitive relation between individuals or groups that is characterized by the subjective importance placed upon competitive outcomes (i.e., win or lose) independent of the objective characteristics of the situation (e.g., tangible stakes) [16, 17]. A proposed theoretical model of rivalry suggests the antecedents of rivalry as similarity factors, competitiveness, and relative performance of the agents [16]. The presence of a rivalry effect in turn affects the motivation of the individual and their performance [17]. We aim to evaluate how different agents affect the user's perception of the agents, and the user's performance due to the increased competitiveness and rivalry effects. Figure 2 shows the theoretical model we follow for estimating rivalry, and its effect on motivation and performance.
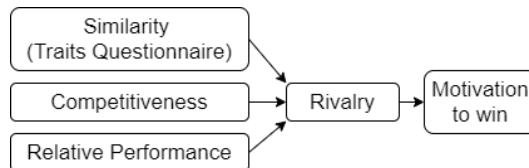

Figure 2: The theoretical model of rivalry used for our new rivalry metric proposition, based on the framework proposed by Kilduff et al. [16].

In competitive games, rivalry is a central concept that directly effects the opponent's behavior through their motivation in play. In human-to-human scenarios and economics, a healthy rivalry is considered to be an important factor that can positively affect the performance of opponents; while in other situations, it can also contribute to unnecessary risk-taking behavior. In human-in-the-loop online learning scenarios for competitive games, the absence of rivalry or competitiveness might result in the human opponent to lose motivation to play the game and show sub-optimal performance during game-play. The agents who are learning actively from the human, are bound to learn from this bad performance input which in turn would result in sub-optimal learning. By introducing the notion of rivalry, we aim to evaluate how different agents affect the user perception of the agents and in turn have an effect on user performance based on the increased competitiveness and rivalry effects.

3 Methodology and Experimental Protocol

Our evaluation setup includes two formal experiments to address each of our hypothesis: the first one, to identify the individual aspects of rivalry between a human and the agents, and to create the fundamental understanding of similarity that the agents will learn; and the second experiment that evaluates how the trained agents impact the rivalry assessment from the humans.

The Chef's Hat Card Game [33] introduces a competitive multi-player scenario designed to be used both in human-human and human-robot interaction (HRI).As a reinforcement learning task, it provides a controllable action-perception cycle, where each player can only perform a restricted set of actions. This, in turn, allows each player to behave as organically as possible and allows the

3

direct measure transfer from a real-world game into an OpenAI GYM-ready scenario without any functionality loss.

Chef's Hat is a 4-player round-based card game, where each person has a restaurant-context role (Chef, Sous-Chef, Waiter or Dishwasher). This role is updated after each game based on the order of finishing the previous match. At the beginning of the game, the players are dealt a full hand of cards (17 cards per player), and taking turns they need to dispose of their cards as quickly as possible. The cards represent ingredients for pizzas, and each round consists of the players making pizzas by discarding cards in a certain manner. Wins the match the first player that discards all its cards, receiving 3 points. The second player receives 2 points, the third receives 1 point and the last player receives no points. Wins the game the first player to reach 15 points. The details of the pizza-making rules and the role hierarchy are explained fully in the games' formal description [33].

Reinforcement Learning Agents. To simulate players in the game, we must train reinforcement learning agents. Following the original design proposed by the Chef's Hat authors [24], we will implement and train two types of agents: one based on Deep Q-Learning (DQL) and another one based on Proximal Policy Optimization (PPO). We need to implement these agents as their learned behavior will be our baseline for our final evaluation.

The Chef's Hat simulation environment defines the state space as an aggregation of the cards on the board and the current cards in the player's hand. The action space is composed of 199 allowed discard actions (which comprise each card and card combinations allowed by the game rules), and one pass action, totaling 200 actions. Each agent is trained using the same reward: 1 for an action that leads to a victory, and -0.01 for every other action. Using such ultimate reward guarantees that the agents will have to explore enough and find their strategies to win the game.

We will train the agents from scratch, first playing against agents that output random movements, and later on against themselves. As each game is composed of 4 players, we will mix different instances of each agent in different game sessions. The entire optimization of the agents, including topology definition, training parameters, and training strategies will be defined using standard optimization schemes, such as grid-search and TPE. As our goal is to obtain agents that learn how to win the game, the entire optimization for the first experimental task will be done based on a pure objective performance: the total number of victories in a set of games.

Data Collection and Experimental Environment. Our experimental environment involves the agents playing the game against a human opponent. This will be done with a web-based interface for the Chef's Hat simulation environment, that will allow a human to play, in real-time, against three other artificial agents. The web-interface will be bounded to the same rules as the game's simulation, to maintain the same evaluation environment. Each human will play a full 15 points games against three agents, one based on DQL, one based on PPO, and one based on an agent that chooses random moves. A participant can finish the game by playing a minimum of 5 matches, each consisting of an average of 30 turns. This totals to at least 150 action/space pairs per person. We aim to have 60 participants, each playing one full game in both scenarios 1 and 2. This will generate 9000 space/action data points per scenario. For each match, we will collect all the objective and state information from all players, but most importantly the total game points (*points*) per match. We will also collect subjective information from the human players, to understand better their behavior and help us to formalize rivalry:

- Before the game starts, we will proceed with a simple questionnaire to self-assess the person's agency (ag_h), competence (ct_h), sense of communion (cm_h) [34] and competitiveness (C_h) [35] in a scale between 0 and 5.

- At the end of the game, the human player will assess each agent adversary based on their agency (ag_a), competence (co_a), and communion (cm_a) traits.

Proposing Rivalry. To optimally define the impact that each agent has on the players, we will use the aforementioned formalization of Rivalry [16]. Rivalry can be defined as a subjective social relationship arising between two actors based on the competitive characteristics of an individual, as well as the increasing stakes and psychological involvement in the situation. Thus, a proposed theoretical model of rivalry, illustrated in Figure 2, suggests that antecedents of rivalry are similarity

factors, competitiveness, and relative performance of the agents [16]. In our scenario, this translates into:

$$S_a = \sqrt{(ag_h - ag_a)^2 + (ct_h - ct_a)^2 + (cm_h - cm_a)^2} \tag{1}$$

$$P_a = (points_h - points_a)/15 \tag{2}$$

$$R_a = (S_a + C_h + P_a)/3 \tag{3}$$

where the index a defines one of the agents, the index h defines the human ratings, S_a is the similarity between the player and the agent, assessed by the player themselves; P_a is the relative performance for each agent, and R_a is the final rivalry rating.

Research in rivalry suggests that individuals tend to evaluate their abilities by comparing their performance to the persons who have similar characteristics to themselves. The similarity can be measured in demographics [16], gender [36], personality [16], perceived traits [37] and rank in competition [17] to assign social or behavioral attributes to the other. For this work, in addition to the ranking and performance similarity, we included trait similarity as a factor that includes competence, agency and communion traits as indicators of relevant behavioral attributes. These traits are frequently used in social stereotypes research (e.g., gender, nationality, age, status) [34], shown to affect liking [38] and social perception [39] was previously shown to affect competitive behavior [40]. Agency, competence and communion traits were also used in virtual agent research to examine users' judgements of expected agent behavior [41]. These traits were chosen as representative of traits due to their simple understanding, which facilitates the own and third-party assessment [42]. Also, due to its direct mapping between action and perception [43], which makes it ideal for a competitive card game analysis.

The presence of a rivalry effect affects the motivation of the individual and their performance. Thus, we expect that rivalry rates will change accordingly to the game development. As the users will evaluate the entire game behavior of an agent, and the only difference between players is the way they play the game, the measure of rivalry will reflect directly the user's perception of the outcome of the reinforcement learning strategies.

Predicting Rivalry. In our second experiment, each agent will calculate rivalry against the human player. To that, they follow the same rivalry calculation as humans. To predict similarity, tho, the agents will use a similarity predictor that will be trained based on the humans´ responses.

The agent-player similarity predictor ($pr_a(h)$) will match state+action chosen by humans (h) with the given competence, agency and communion traits (ag_a, ct_a and cm_a), and it follows the approach used in virtual agent research [41]. By collecting the data during the first experimental scenarios, we will form the dataset to train a similarity predictor, used then later in our second experimental scenario. Our similarity predictor will be built as an MLP that will map an interval of action/spaces with a set of similarity scores. We posit the 9000 action/space pairs collected in Scenario 1 to be sufficient to obtain an efficient similarity predictor. The similarity predictor of each agent will infer, during game-play, the traits of each human the agents play against. Each type of agent, DQL, and PPO will also have a single set of traits associated, given also by our human analysis of scenario one. Thus the similarity based on an agent´s (a) perspective (S_h) is given as:

$$S_h = \left(\sqrt{(pr_a(h) - (ag_a, ct_a, cm_a))^2} \right) \tag{4}$$

The performance measure of the agents will be given by their own assessment of their actions. To achieve this, each agent will compute the introspective confidence (ic_a) [44] of each action, which focuses on scaling the selected Q-value of an action towards the final goal using a logarithm transformation which computes the probability of success, in our case, of winning the game. The

introspective confidence gives us a self-assessment of the agent's actions, based on its own game experience. The agent´s perspective is given as:

$$C_a = \frac{\sum ic_a(act)}{totalActions} \qquad (5)$$

where *act*, is each of the actions the agent took during the game.

The relative performance (P_h) is calculated similarly to the human´s perspective, but taking the agent´s perspective into consideration. Thus, the predicted rivalry (R_h) is defined as the mean of ($S_h + C_a + P_h$).

To include the predicted rivalry in the reward function will impact on the agent behavior. As the rivalry prediction is a representation of subjective and objective human metrics,

First Experimental Scenario. Our first experimental scenario involves one human playing the game against three agents: two pre-trained agents, one using DQL and the other PPO, and an agent that only does random moves. Our goal with this experiment is to collect the human assessments about the agents, to provide an appropriate dataset for the rivalry prediction. Also, it will be important to identify the baseline assessment that each of the trained agents already carries, which will be completely related to their learning strategy during training.

Second Experimental Scenario. We will repeat the same setup as the first experimental scenario, except the following: each agent will be updated, during game-play, based on the novel reward function that includes a rivalry score. The rivalry score will be predicted by each agent in a process defined above. Each agent will include the rivalry against the human in its reward function, to change the human perception toward themselves. The agents will be trained once per finished match. For each agent, we will experiment with three conditions: increase rivalry, decrease rivalry, and maintain a rivalry.

The fine-tuning of the similarity model, parameter-wise and learning-wise, is part of the experimental methodology proposed in this experimental scenario. As the agents are already trained to learn how to play the game, we will investigate a weighted aggregation as a simple but effective manner to include the rivalry into the existing reward [45]. The precise definition of the reward calculation will come as a result of the analysis of the outcomes of Scenario 2.

Expected Contribution We believe that integrating the social representations given by our novel rivalry score will enrich current reinforcement learning strategies. We hope to pave the way to bridge the gap between pure reinforcement learning optimization and social learning applications.

References

[1] Zimo Liu, Jingya Wang, Shaogang Gong, Huchuan Lu, and Dacheng Tao. Deep reinforcement active learning for human-in-the-loop person re-identification. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 6122–6131, 2019.

[2] Hamidreza Modares, Isura Ranatunga, Frank L Lewis, and Dan O Popa. Optimized assistive human–robot interaction using reinforcement learning. *IEEE transactions on cybernetics*, 46(3):655–667, 2015.

[3] Nikhil Churamani, Matthias Kerzel, Erik Strahl, Pablo Barros, and Stefan Wermter. Teaching emotion expressions to a human companion robot using deep neural architectures. In *2017 International Joint Conference on Neural Networks (IJCNN)*, pages 627–634. IEEE, 2017.

[4] Ioannis Papaioannou and Oliver Lemon. Combining chat and task-based multimodal dialogue for more engaging hri: A scalable method using reinforcement learning. In *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, pages 365–366, 2017.

[5] Yuan Gao, Fangkai Yang, Martin Frisk, Daniel Hemandez, Christopher Peters, and Ginevra Castellano. Learning socially appropriate robot approaching behavior toward groups using deep reinforcement learning. In *2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, pages 1–8. IEEE, 2019.

[6] Stephanie Milani, Nicholay Topin, Brandon Houghton, William H Guss, Sharada P Mohanty, Oriol Vinyals, and Noboru Sean Kuno. The minerl competition on sample-efficient reinforcement learning using human priors: A retrospective. *arXiv preprint arXiv:2003.05012*, 2020.

[7] Oriol Vinyals, Igor Babuschkin, Wojciech M Czarnecki, Michaël Mathieu, Andrew Dudzik, Junyoung Chung, David H Choi, Richard Powell, Timo Ewalds, Petko Georgiev, et al. Grandmaster level in starcraft ii using multi-agent reinforcement learning. *Nature*, 575(7782):350–354, 2019.

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

[9] Mehdi Khamassi, George Velentzas, Theodore Tsitsimis, and Costas Tzafestas. Robot fast adaptation to changes in human engagement during simulated dynamic social interaction with active exploration in parameterized reinforcement learning. *IEEE Transactions on Cognitive and Developmental Systems*, 10(4):881–893, 2018.

[10] Hannes Ritschel and Elisabeth André. Real-time robot personality adaptation based on reinforcement learning and social signals. In *Proceedings of the companion of the 2017 acm/ieee international conference on human-robot interaction*, pages 265–266, 2017.

[11] Heriberto Cuayáhuitl, Donghyeon Lee, Seonghan Ryu, Yongjin Cho, Sungja Choi, Satish Indurthi, Seunghak Yu, Hyungtak Choi, Inchul Hwang, and Jihie Kim. Ensemble-based deep reinforcement learning for chatbots. *Neurocomputing*, 366:118–130, 2019.

[12] Aaquib Tabrez and Bradley Hayes. Improving human-robot interaction through explainable reinforcement learning. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 751–753. IEEE, 2019.

[13] Prashan Madumal, Tim Miller, Liz Sonenberg, and Frank Vetere. Explainable reinforcement learning through a causal lens. *arXiv preprint arXiv:1905.10958*, 2019.

[14] Pedro Sequeira and Melinda Gervasio. Interestingness elements for explainable reinforcement learning: Understanding agents' capabilities and limitations. *Artificial Intelligence*, 288:103367, 2020.

[15] Cody T Havard. What is rivalry and where we go from here. In *Rivalry in Sport*, pages 9–35. Springer, 2020.

[16] Gavin J Kilduff, Hillary Anger Elfenbein, and Barry M Staw. The psychology of rivalry: A relationally dependent analysis of competition. *Academy of Management Journal*, 53(5):943–969, 2010.

[17] Gavin J Kilduff. Driven to win: Rivalry, motivation, and performance. *Social Psychological and Personality Science*, 5(8):944–952, 2014.

[18] Murray Campbell, A Joseph Hoane Jr, and Feng-hsiung Hsu. Deep blue. *Artificial intelligence*, 134(1-2):57–83, 2002.

[19] Dennis DeCoste. The future of chess-playing technologies and the significance of kasparov versus deep blue. In *Deep Blue Versus Kasparov: The Significance for Artificial Intelligence*, pages 9–13, 1997.

[20] Dennis DeCoste. The significance of kasparov versus deep blue and the future of computer chess. *ICGA Journal*, 21(1):33–43, 1998.

[21] Kun Shao, Zhentao Tang, Yuanheng Zhu, Nannan Li, and Dongbin Zhao. A survey of deep reinforcement learning in video games. *arXiv preprint arXiv:1912.10944*, 2019.

[22] David Silver, Aja Huang, Chris J Maddison, Arthur Guez, Laurent Sifre, George Van Den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, et al. Mastering the game of go with deep neural networks and tree search. *nature*, 529(7587):484–489, 2016.

[23] Kun Shao, Yuanheng Zhu, and Dongbin Zhao. Cooperative reinforcement learning for multiple units combat in starcraft. In *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 1–6. IEEE, 2017.

[24] Pablo Barros, Ana Tanevska, and Alessandra Sciutti. Learning from learners: Adapting reinforcement learning agents to be competitive in a card game. *arXiv preprint arXiv:2004.04000*, 2020.

[25] Hado Van Hasselt, Arthur Guez, and David Silver. Deep reinforcement learning with double q-learning. In *Thirtieth AAAI conference on artificial intelligence*, 2016.

[26] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017.

[27] Silvia Rossi, Francois Ferland, and Adriana Tapus. User profiling and behavioral adaptation for hri: a survey. *Pattern Recognition Letters*, 99:3–12, 2017.

[28] Christian Arzate Cruz and Takeo Igarashi. A survey on interactive reinforcement learning: Design principles and open challenges. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference*, pages 1195–1209, 2020.

[29] Robin R Murphy and Debra Schreckenghost. Survey of metrics for human-robot interaction. In *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 197–198. IEEE, 2013.

[30] Lola Cañamero. Embodied affect for real-world human-robot interaction. In *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, pages 459–460, 2020.

[31] Julia Peltason, Nina Riether, Britta Wrede, and Ingo Lütkebohle. Talking with robots about objects: a system-level evaluation in hri. In *2012 7th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 479–486. IEEE, 2012.

[32] Momotaz Begum, Richard W Serna, David Kontak, Jordan Allspaw, James Kuczynski, Holly A Yanco, and Jacob Suarez. Measuring the efficacy of robots in autism therapy: How informative are standard hri metrics'. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, pages 335–342, 2015.

[33] Pablo Barros, Alessandra Sciutti, Inge M Hootsmans, Lena M Opheij, Romain HA Toebosch, and Emilia Barakova. It's food fight! introducing the chef's hat card game for affective-aware hri. *arXiv preprint arXiv:2002.11458*, 2020.

[34] Alice H. Eagly, Christa Nater, David I. Miller, Michele Kaufmann, and Sabine Sczesny. Gender stereotypes have changed: A cross-temporal meta-analysis of u.s. public opinion polls from 1946 to 2018. *American Psychologist*, 75(3):301–315, Apr 2020.

[35] Robert D Smither and John M Houston. The nature of competitiveness: The development and validation of the competitiveness index. *Educational and Psychological Measurement*, 52(2):407–418, 1992.

[36] Carol T Miller. Self-schemas, gender, and social comparison: A clarification of the related attributes hypothesis. *Journal of Personality and Social Psychology*, 46(6):1222, 1984.

[37] John M Tauer and Judith M Harackiewicz. Winning isn't everything: Competition, achievement orientation, and intrinsic motivation. *Journal of Experimental Social Psychology*, 35(3):209–238, 1999.

[38] Bogdan Wojciszke, Andrea E Abele, and Wiesław Baryla. Two dimensions of interpersonal attitudes: Liking depends on communion, respect depends on agency. *European Journal of Social Psychology*, 39(6):973–990, 2009.

[39] Amy JC Cuddy, Susan T Fiske, and Peter Glick. Warmth and competence as universal dimensions of social perception: The stereotype content model and the bias map. *Advances in experimental social psychology*, 40:61–149, 2008.

[40] Susan T Fiske, Amy JC Cuddy, Peter Glick, and Jun Xu. A model of (often mixed) stereotype content: competence and warmth respectively follow from perceived status and competition. *Journal of personality and social psychology*, 82(6):878, 2002.

[41] Procheta Nag and Özge Nilay Yalçın. Gender stereotypes in virtual agents. In *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents*, pages 1–8, 2020.

[42] Andrea E Abele and Bogdan Wojciszke. Agency and communion from the perspective of self versus others. *Journal of personality and social psychology*, 93(5):751, 2007.

[43] Antonin Carrier, Eva Louvet, Bruno Chauvin, and Odile Rohmer. The primacy of agency over competence in status perception. *Social Psychology*, 2014.

[44] Francisco Cruz, Richard Dazeley, and Peter Vamplew. Explainable robotic systems: Understanding goal-driven actions in a reinforcement learning scenario. *arXiv preprint arXiv:2006.13615*, 2020.

[45] Kristof Van Moffaert, Madalina M Drugan, and Ann Nowé. Scalarized multi-objective reinforcement learning: Novel design techniques. In *2013 IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning (ADPRL)*, pages 191–199. IEEE, 2013.