The Stereotype Content Model Applied to Human-Robot Interactions in Groups

Anonymous for Submission*

ABSTRACT

Stereotypes are a pivotal factor in determining our behavior towards robots. However, little is known about the underlining content of stereotypes towards these agents. Research has identified warmth and competence as two central dimensions of stereotype content around which people anchor their perception about others. Consequently, our goal with this study was to analyze in an entertainment group context, how autonomous robots, displaying high and low levels of competence and warmth, and different roles (opponent versus partner), are perceived in terms of warmth and competence and to what extent these variables affect users' emotional responses and can be leveraged to predict future intention to work with robots.

Additionally, we observed an effect of perceived warmth and competence in admiration, pity and contempt towards robots.

KEYWORDS

Stereotypes Content Model, Human-Robot Interaction, Groups, Entertainment

1 INTRODUCTION

Stereotypes play a pivotal role in the way we perceive and respond to other social agents [31]. In group contexts, stereotypes can affect the extent to which we trust [31] and identify with other people [36] and, thus, have an important impact on how we behave towards them [14, 34]. In Human-Human Interactions (henceforth, HHI), warmth and competence have been universally and repeatedly identified as two central categories of stereotype content, around which people organize and construct their perceptions about others [21, 22]. Likewise, stereotypes have also been found to hold a key role in Human-Robot Interactions (hereinafter, HRI). In particular, past research has hinted at an effect of stereotypes on several important dimensions of HRI, such as robot acceptance [55] and robot perception [29]. However, past research on HRI considering the role of stereotypes has focused predominantly in gender stereotypes and, thus far, has not payed much heed to the fundamental dimensions of stereotype content that can help predict the users' emotional and behavioral responses towards robots. In this study, our goal is to analyze how different levels of warmth and competence displayed by autonomous robotic agents can influence the users' perceptions, emotions and willingness to interact again with robots in the future. To achieve this purpose, we devised a scenario involving two humans and two autonomous robots displaying different levels of warmth and competence (high and low) interacting with each other either in the role of partners or opponents. The results of our study suggested that participants are able to create stereotypes about robots based on their level of warmth and competence and that different perceived combinations of these traits result in different emotional responses of users. In addition we also observed that the levels of warmth and competence can predict a users' intention to interact with robots in the future and that the participants' perceptions about robots (in terms of warmth and competence) present a certain degree of endurance.

*

Conference'17, July 2017, Washington, DC, USA
2020. ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

2 THE ROLE OF STEREOTYPES IN HUMAN-ROBOT INTERACTIONS

In recent years, the CASA (Computers Are Social Actors) paradigm has emerged as a stepping stone in HRI research by ushering in the idea that people often perceive and behave towards robotic and virtual agents in similar ways to those they perceive and behave towards other people [24, 44]. By suggesting that humans often apply social HHI constructs to HRI, this paradigm has placed the emphasis on creating increasingly life-like robots that can communicate with humans in social instinctive manners [24]. To do so, researchers in social robotics have begun to explore how the inclusion of certain human characteristics (such as personality) in robots can contribute to improve HRI, by being able, for example, to improve robots' acceptance or likability [26, 54].

In HHI, the stereotypes individuals develop and apply to make sense of the world have been thoroughly studied in regards to their effects on people perception. In this area, stereotypes are usually framed in terms of implicit personality theory and constitute social heuristics that due to their general acceptance, are considered as part of an ordinary process of social cognition [7]. These heuristics can be understood as a part of a larger set of social knowledge structures that guide the interpretation of other people' behavior and are, thus, relevant predictors of behavior [6, 7, 27]. Given the importance of stereotypes in HHI, researchers have begun to examine the role played by these social heuristics in HRI [56, 59, 60]. Previous studies have suggested that people tend to judge robots of the opposite gender more positively than same gender robots in dimensions like trustworthiness [53]. Gender stereotypes have also been found to have an effect in the perception of task suitability. For example, [55] found that, along with personality, gender is an important predictor of users' acceptance of robots in certain gender stereotyped roles. Robots displaying different genders also tend to be associated with different personality traits, with female robots being perceived as more communal whereas male robots tend to be perceived as more agentic [29]. The mapping of how these gender stereotypes influence HRI has been argued to help in the design of robots which can evoke increased user acceptance [29, 53]. However, an investigation in to the transversal and fundamental content of stereotypes is still lacking. The need to generate or test systematic dimensions of HRI has been outlined by some authors and its' usefulness stems from the fact that it allows the creation of robots that can evoke pre-determined responses. That, in turn, allows robots to engage in more successful interactions with users [38]. In this sense, we believe that a good starting point to achieve this end, can come from attempts to assess how models of stereotypes in psychology fare when predicting peoples' responses towards robots. These models can then be leveraged in HRI by proposing an association between specific traits and behavioral and emotional outcomes. However, if we are going to apply models of stereotypes to HRI, it is important to acknowledge some fundamental assumptions [4, 44].

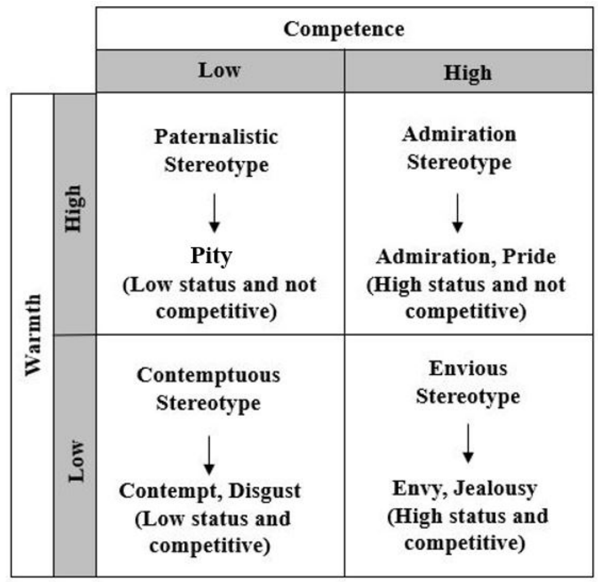
First, we assume that people tend to develop multiple stereotypes (i.e. shared images) that follow some form of social categorization, conveying both an ontological norm (*about what someone is*) and a deontological norm (*about what someone should be or behave*). Second, these stereotypes tend to follow a top-down logic (i.e. generalizations about the characteristics associated with a certain group of people tend to be extrapolated to the individuals that are members of those groups). Finally, these stereotypes can consistently affect peoples' behavior and that, robots playing social roles can evoke consistent stereotypes from users. By considering the above assumptions, it is important to examine how dimensions of stereotype content can be applied to HRI, what is their meaning and how they can be leveraged in these scenarios.

3 THE STEREOTYPE CONTENT MODEL

As the process of stereotyping others is often described as automatic and mostly unavoidable [9] and given its' central role in determining the way we perceive and behave both towards other people [23, 30], stereotypes seem to be of central importance in HRI.

Generally, stereotypes have been defined as a form of gestalt view of personality perception, with some authors suggesting that certain traits hold a central role in the way we organize our perception of other agents [4, 5, 21]. Namely, in this regard, warmth and competence have been suggested to be important instances of stereotype content and are the two main dimensions of the Stereotype Content Model (SCM) [19, 21, 22]. This model argues that stereotypes have two main functional purposes: *to inform the person about a third party's intent (warmth)* and *about his/her ability to pursue that intent (competence)*. The usefulness of this model is that each of these different stereotypes has been steadily associated with certain emotional outcomes in HHI and has shown to be a good predictor of behavior in some situations (see fig. 1). For example, one area of HHI, which understanding has significantly benefited from input within the stereotypes' framework, has been the area of gender stereotypes (for a review, see [28]). These findings in HHI assume a double importance to better understand HRI. On one hand they constitute stable representations of the way in which people from different genders are perceived and can contribute to explain gender stereotypes in robots. On the other hand, they provide reliable general guidelines to orient a social design of robots that takes into account the role of stereotypes in shaping the users' behavior [19]. Thus, it is important to focus on stereotype content as it allows us to determine *"...who gets placed in what categories [of stereotypes] and why, and what attributes are associated with the various categories an why"* [61](p. 80).

Figure 1: Emotional responses and relational dynamics as a function of different combinations of stereotype dimensions, adapted from [21]).



4 TYPES OF STEREOTYPES AND THEIR EMOTIONAL AND BEHAVIORAL CONSEQUENCES

The SCM proposes the existence of qualitatively different dimensions of stereotype. These dimension are based on disparate perceptions of status and interdependence among individuals (collaboration and competition) [19]. More specifically, this model postulates that warmth and competence are the two central categories of stereotype content and that the conjugation of different levels (i.e. low and high) results in four distinct types of stereotypes. Firstly, contemptuous stereotypes are usually associated with groups that are perceived both as incompetent and cold (i.e. low warmth) [20–22]. Emotional responses associated to these groups include high feelings of contempt and disgust, often associated to dispositional attributions to explain another's negative outcomes, and therefore are perceived as being controllable by the individual (i.e. *He/She deserved it!*) [21]. Consequently, the behavioral responses to groups categorized in this quadrant include both active and passive harm (e.g: distancing or rejecting) [13, 48, 49]. Secondly, envy often occurs toward groups that are perceived as being highly competent but less warmth [21, 22, 30]. In these cases, envy is usually associated to a perception of unfair gains accompanied by other negative feelings (e.g. injustice, resentment) and positive feelings (e.g. respect, acknowledgement) [20–22]. These emotional ambivalence is hypothesized to be linked to a behavioral ambivalent response, which can include both passive facilitation (when the emotional response is mostly comprised of respect) or active harm (when the perception of injustice is predominant). Thirdly, paternalistic stereotypes are associated to groups who are perceived low on competence and high on warmth [21, 22, 30]. Individuals and groups stereotyped in a paternalistic way tend to be perceived as not harmful, with neither the intention to cause harm nor the ability to do so. Similar to contemptuous, these stereotypes are associated to ambivalent emotional and behavioral responses. Also, negative outcomes by target groups of this stereotype category tend to be perceived as uncontrollable[20–22]. These perceptions often lead to feelings of pity, involving sadness or compassion, which, in turn, may lead to active facilitation driven by empathy (e.g: provide help) or by passive harm, driven by sadness (e.g: denying help). Finally, admiration stereotypes are directed to groups whose achievements do not subtract or threaten the individuals' own goals. In terms of behavioral responses, target groups of this category generally trigger passive and active facilitation, because they are viewed by others with admiration and respect [21]. Feelings of admiration and pride tend to evoke positive reactions from other people and motivate positive contact and collaboration [1, 41].

In summary, groups perceived high in warmth will evoke active facilitation and prevent harmful behaviors, whereas high perceived competence will foster passive facilitation (e.g: collaboration) and prevent passive harm (viz. exclusion) [21].

5 THE SPECIFICITIES OF GROUP INTERACTIONS AND ENTERTAINMENT SCENARIOS IN HRI

A large bulk of research in HRI has been devoted to understand how multiple factors influence the establishment of HRI at an individual level. Although this question is, by itself, an interesting avenue of research, it contrasts with the natural human environments in which robots are being introduced [38, 45, 47]. More specifically, these real world environments can vary broadly in terms of structure and number of users [57] and thus, require robots that are designed with the specificities of group HRI in mind [38]. Group interactions are fundamentally different from interpersonal interactions because they add complexity to the analysis of the social exchanges [39]. The significance of exploring this type of setting comes from the idea that the sheer number of robots in an interaction can affect the relational dynamics among members of a group [32, 50]. Indeed, previous studies in HRI suggest the existence of different interaction patterns in group contexts, in comparison to those observed in one-on-one interaction settings [18, 32, 50].

Additionally, much of the research in HRI has been oriented towards identifying the factors that predict successful collaboration between humans and robots [10–12, 16]. However, research in HHI (and to a lesser extent, in HRI) has suggested that individuals and groups often establish different relationship dynamics that go beyond collaboration [11, 15, 25, 55] and that these different dynamics are associated with distinct behavioral patterns [15, 25, 46]. For example, in the context of gaming, research suggests that collaborative and competitive gaming are capable of evoking different behaviors from players. More specifically, whereas some studies suggest that competitive gaming increases aggression when compared to collaborative gaming [51], collaborative strategies have been associated to increased cohesion and positive socioemotional behaviors between members of a group [2]. This is relevant to our study because it relates to the perceived status and social goal orientations associated with groups displaying different stereotypes, and thus, it can have an effect on the development of stereotypes about robots.

6 GOALS AND HYPOTHESIS

The aim of this study is to analyze how different levels of warmth and competence, displayed by two autonomous robotic agents in a group entertainment scenario, can impact the users' feelings, perceptions and future intention to work with robots. In addition, we will analyze the role (partner or opponent) played by the robot on users' responses, which will include user's feelings of admiration, pity, envy and contempt directed at robots. Furthermore, we will also analyze the stability of the participants' recollections about their perception of robots, approximately one week after the initial interaction.

In this context, we made the following predictions:

- **Warmth and Competence** We expect that participants would be able to distinguish the robots that display low and high attributes of warmth and competence. Thus, is is expected that the high competent robot will be perceived as being more competent than the low competent robot (H1), whereas the high warmth robot will be perceived as warmer than the low warmth robot (H2).
- **Collaborative and Competitive Interactions** We expect that participants' perceptions about robots to be more favorable (i.e., higher in warmth and competence) when robots play as partners than when they play as opponents (H3).
- **Emotional Responses** Based on the SCM, we expect that a similar pattern of emotional responses in HRI, as the ones that tend to occur in HHI, as displayed in figure 1. In particular, we expect both warmth and competence to have an effect in the participants' emotional responses towards robots (H4). More specifically, we expect to observe a higher degree of admiration felt towards robots displaying high levels of competence and warmth (H5) and a higher level of contempt towards robots displaying low levels of these traits (H6). Additionally, we expect to observe more feelings of pity directed at robots displaying high levels of warmth and low levels of competence (H7) and more feelings of envy directed at robots displaying the opposite pattern of these traits (H8).
- **Future Intention to Interact with Robots** We expect to observe a higher intention to interact with robots displaying high levels of competence, in the conditions in which warmth is maintained stable (low or high; H9) and a higher intention to interact again in the future with robots displaying high levels of warmth, when the level of competence is maintained stable (H10).
- **Recall of Perceptions about Robots** We expect to observe a certain degree of stability in the stereotypes developed about robots. More specifically, we expect participants to be able to correctly recall the stereotypes generated about robots, one week after initially interacting with them (H11).

7 METHOD

7.1 Sample

A convenience sample of 54 participants (37 men, 17 women) was recruited in the campus of a major technological institute. Participants were in average 24 years old (SD = 6 years) and approximately 82% reported having superior education.

7.2 Task and Robots

To test our hypotheses, an entertainment card-game scenario was devised. The task consisted of a traditional Portuguese card-game, called Sueca, involving 4 players (two robots and two humans). Players were teamed up in pairs and competed with the other pair of players, while collaborating with the other member of the pair they were assigned to. The rules and game events for this particular game can be consulted in [ANONYMIZED FOR REVIEW][1]. This game was chosen because it is a group scenario which involve both collaborative and competitive relational dynamics. To play the game we used two robots with similar embodiments (Emys heads) capable of both playing the game and autonomously interact with the other players. The robots were attached to a multi-touch screen table used as the game interface. In addition, we used a standard French deck of cards with imprinted fiducial markers that were recognized by the table, and two audio columns (placed by the side of each robot) to transmit its verbal interactions. Finally, two video cameras were used to record the interactions, but these records will not be addressed here.

7.3 Manipulation

To analyze the effects of warmth and competence on the relational dynamics during the HRI, we used two robots displaying different levels of these traits (high and low). Each of these variables was manipulated as follows:

- **Warmth:** Warmth was manipulated through the utterances displayed by the robot. The validation of these utterances can be found in section 8.1 and examples of the utterances can be consulted in table 1. A full list of utterances can also be consulted in [ANONYMIZED FOR REVIEW][2]
- **Competence:** Competence was manipulated through the algorithm to play the game. We implemented the Perfect Information Monte Carlo algorithm, proposed by [33] and successfully implemented for this card game in [17]. The manipulation is a consequence of changing the returning instruction to either maximize or minimize the average simulated reward, in order to produce a high or low competence level, respectively.
  To validate this manipulation we run an agent-based simulation of 500 independent games with the following configuration: a team of a low-competence agent and a rule-based agent, against a team of a high-competence agent and a rule-based agent. According to previous results by Correia et al., the rule-based agent achieved similar results to an average human player in their experiments, therefore it is expected that this setting will simulate the mix of two levels of competence in our experiment. In the agent-based simulations, the team with a high-competence agent won 96.4% of the games, revealing the effectiveness of our manipulation.

[1]See:ANONYMIZED LINK
[2]See:ANONYMIZED LINK

Each participant played as a partner of one of the two robots in each condition, and was randomly assigned to one of four conditions of the game in which the two robot players were programmed to display:

(a) distinct levels of warmth (high versus low), but similar low levels of competence; (b) distinct levels of warmth (high versus low), but similar high levels of competence; (c) distinct levels of competence (high versus low), but similar low levels of warmth; (d) distinct levels of competence (high versus low), but similar high levels of warmth.

More specifically, we alternated within and between subjects manipulation of high and low levels of competence. In the two first groups (a and b), competence was manipulated between groups, whereas warmth was manipulated between groups. Adversely, in groups (c) and (d) warmth was manipulated between groups, whereas competence was manipulated within groups. In each condition, participants played once with each robot (each displaying different levels of the variable manipulated within groups).

7.4 Measures and Procedure

A request detailing the procedure of the present study was submitted and accepted by the university' Ethical Committee. Participants were told that we were interested in analyzing how humans and robots interacted in a group entertainment scenario and informed consents were obtained from all participants before we begun the experiment. After this, we gave a brief explanation to participants of how to play the game using the touch table and the cards. Then participants began playing with the robots and were instructed to call the researcher (who was in an adjacent room) if any difficulty arose. At the end of a set of three games, participants evaluated both robots (partner and opponent) in terms of competence (4 items, e.g: *competent*) and warmth (4 items, e.g: *good-natured*), and an additional 4 filler items e.g: (*tolerant*) using a 7 point Likert-like scale (1-*"Does not describe well"* to 7-*"Describes it perfectly"*). These items were retrieved from the list of adjectives used in [19] and translated to the native language and were presented to the participants

Additionally, they indicated how much each feeling from a list of 4 feelings (pitty, envy, admiration and contempt) corresponded well or not to the way they felt about the robots that they had interacted with (both the robot that played as partner and the one that played as an opponent), using a similar scale to that described before.

After filling this initial questionnaire, participants were requested to switch partners and play an additional set of three Sueca games with the other robot. After ending this final round, participants were asked to evaluate the robots again using the same set of questions used in the first round.

In the end, we asked participants to fill a sociodemographic questionnaire for sample characterization. Participants were then thanked for their collaboration with a movie ticket voucher and requested to indicate their e-mail address so that they could respond to some follow-up questions. Overall, participants took approximately one hour and a half to complete the experiment.

One week after the study, participants were contacted via e-mail and asked to fill a short online questionnaire (via Qualtrics platform) about their interaction with the robots in this study. After receiving this e-mail, participants were given three week days to respond. In this questionnaire, participants were requested to indicate the final result of both of the sets of three games they played with each robot (a total of 6 games) and were given four options: (a) "Our team lost", (b) "Our team won", (c) "The game ended with a tie" and (c) "The game ended because one of the teams renounced".

Finally, we requested participants to indicate how much each of the robots was friendly and competent at playing Sueca. After responding to these follow-up questions, participants were then thanked again for their participation and given the researchers' e-mail in case they wanted additional information about the study.

8 RESULTS

8.1 Character Validation

A convenience sample of 13 participants (7 female and 5 female, average age=28, S.D.=5) was recruited and shown video-recorded interactions of people interacting with two robots in the context of a card-game scenario. Each one of the robots displayed low or high levels of warmth and each participant observed a total of 20 (10 per condition) interactions consisting of verbal utterances randomly selected from the verbal repertoire embedded in each of robots[3]. After observing each set of interactions, participants were asked to evaluate on a 7-point scale (1=*Does not describe well* to 7=*Describes it perfectly*) how much, each of the robots' behavior was accurately described by a list of 8 adjectives reflecting different traits associated with warmth and competence (e.g: competent and affectionate, respectively) adapted from [19]. A paired samples t-test of the results yielded significant differences across all dimensions. Considering previously established guidelines for the interpretation of the values for Cronbach' Alpha, all the scales used presented acceptable to excellent levels of reliability ranging from .56 to .91. As expected, the high warmth robot was perceived as being significantly warmer (M=6.17, S.D.=0.54) than the low warmth robot (t(11)=24.18, p<.001; M=2.56, S.D.=0.59), although it was also perceived as being less competent (M=2.44, S.D.=0.54) compared to the low warmth robot (M=4.18, S.D.=1.29; t(11)=-5.17, p<.001).

8.2 Data Analysis

Since this study involved analyses of HRI in a small group, the non-independence among participants within groups must be taken into account. Therefore, restricted maximum likelihood linear mixed-model (LMM) analyses were run on each of the seven evaluation parameters of the robots (perceived warmth, competence, pity, admiration, envy, and contempt), using the manipulated repeated conditions of Warmth (Low and High), Competence (Low and High) and Partnership (Partner and Opponent). All of the interaction terms between these three variables was estimated as fixed effects. Besides running the LMM to account for the fact that participants were nested within groups (i.e. each two participants belong to the same group), thereby controlling for their potential mutual influence within the group, we also took into account for the fact that the repeated evaluations were nested within participants (i.e. both robots were evaluated in each of the two sessions). A scaled identity covariance was used for both the random effects of the group and of the repeated measures, which is considered appropriate in studies using repeated measures of short duration [35]. Because the IBM SPSS statistical software (version 24) was used to run the LMM, the Satterthwaite correction for calculating the degrees of freedom will be reported. Results for the LMM 2 (robot's warmth: high, low) X 2 (robot's competence: high, low) X 2 (robot's role: partner, opponent) main effects and interactions on the outcomes are shown in Table 1.

[3]See:ANONYMIZED LINK

8.3 Perception of Warmth and Competence

8.3.1 Perceived Warmth. The LMM yielded a significant effect of warmth, F(1, 170)=224,99, p<.001). As expected, the low warmth robot was perceived as less warmer than the high warmth robot (M=3.63, S.E.=0.135 and M=5.86, S.E.=0.134, respectively). Interestingly, we also found a significant main effect of competence on the perceived levels of warmth of the robots, F(1, 205)=4.54, p=.034, indicating that the low competent robot was perceived as being warmer (M=4.89, S.E.=0.133) than the high competence robot (M=4.59, S.E.=0.130). Moreover, a significant interaction between warmth and the role displayed by the robot, F(1, 176)=5.025, p=.024; see table 1, indicated that the low warmth robot was perceived as being less warm when it played as a partner (M=3.50) than as an opponent (M=3.75). In contrast, the opposite pattern occurred for the high warmth robot, which was judged as being warmer when it played the role of partner (M 6.00) than when it played the role of opponent (M=5.73), F(1, 203)=111.73, p<.001 .

8.3.2 Perceived Competence. As expected, the MLM for competence yielded a significant main effect of the competence displayed by the robot (1). The robot that was programmed to display a high level of competence was evaluated by participants as being more competent than the robot displaying a low level of competence (M=5.43, S.E.=0.172 versus M=3.32, S.E.=0.175). In addition, we also found a main effect of warmth, F(1,196)=6.018, p=.015, indicating that the high warmth robot was also perceived as being more competent than the low warmth robot (M=4.59, S.E.=0.177 and M=4.15, S.E.=0.178, respectively). However, an interaction effect between warmth and competence also occurred, F(1, 179)=9.32, p=.003. Notably, in the condition in which both robots displayed a high level of warmth, the low competent robot was perceived as being less competent (M=2.90, S.E.=0.21) than in the condition in which both robots displayed a low level of competence (M=3.76, S.E.=0.21).

In the condition in which both robots displayed high levels of warmth, the same pattern was observed, with the low competence robot being perceived as less competent than the high competence robot (M=3.74, S.E.=0.21 and M=5.44, S.E.=0.20, respectively).

Finally, a significant two-way interaction between the robot's competence and role indicated that the low competent robot was perceived as being less competent when it played the role of partner (M=3.94) than when it played as an opponent (M=4.57; see table 1), whereas the high competent robot was perceived similarly in terms of competence regardless of the role it played.

Table 1: Results of the Multilevel moddeling on Warmth (W), Competence (C), and Role (Role) played by the robot in the participants' emotional responses and appraisals of warmth and competence (statistical significant level of p < 0.05 appear in bold).

	Predictors													
	Warmth (W)		Competence (C)		Role (R)		W*C		W*R		C*R		W*C*R	
Dependent Variable	F	Sig.	F	Sig.	F	Sig.	F	Sig.	F	Sig.	F	Sig.	F	Sig.
Warmth	**224.98**	**<.001**	**4.54**	**.034**	0.01	.935	2.45	.119	**5.20**	**.024**	1.03	.313	0.38	.540
Competence	**6.02**	**.015**	**169.85**	**<.001**	2.09	.150	**9.32**	**.003**	3.23	.074	**4.49**	**.036**	0.12	.734
Pity	**10.24**	**.002**	0.50	.479	1.95	.164	1.51	.220	0.65	.420	1.10	.297	3.27	.072
Admiration	**21.38**	**.000**	1.77	.185	0.53	.467	0.17	.677	**8.62**	**.004**	3.83	.052	**6.28**	**.013**
Envy	0.003	.955	3.49	.063	3.49	.063	2.98	.086	0.07	.797	0.83	.363	0.67	.415
Contempt	**21.69**	**<0.001**	0.15	.703	2.03	.156	0.01	.905	**8.97**	**.003**	1.48	.225	1.77	.186

8.4 Emotional Responses

8.4.1 Admiration. Overall, we found a main effect of warmth indicating that participants felt more admiration towards the high warmth robot than towards the low warmth robot (M=4.42, S.E.=0.24 and M=3.09, S.E.=0.25) (see table 1). Moreover, a significant interaction between warmth and competence was found, F(1, 180)=6.139, p=.014. In the condition in which robots displayed low levels of warmth, participants felt more admiration towards the high competent robot (M=3.87, S.E.=0.33) than towards the low competence robot (M=3.20, S.E.=0.34). In contrast, in the condition in which both robots displayed a high level of warmth, participants felt more admiration for the low competence robot (M=4.18, S.E.=0.33) than for the high competence robot (M=3.68, S.E.=0.33).

8.4.2 Pity. For Pity, the MLM yielded a significant main effect of warmth, indicating that the robot displaying a high level of warmth evoked more pity than the low warmth robot (M=3.18, S.E.=0.24 and M=2.29, S.E.=0.24, respectively). No other significant main or interaction effects were found for the participant's feelings of pity ((see table 1).

8.4.3 Contempt. The MLM for this dimension yielded a significant effect of the warmth displayed by the robot, F(1, 114)=4.22, p=.042. More specifically, robots displaying low levels of warmth evoked more contempt (M=3.22, S.E.=0.23) than robots displaying high levels of this trait (M=2.61, S.E.=0.23). The interaction between warmth and the role displayed by the robot also indicated that participants only felt more contempt for the low warmth robot (M=3.76) than for the high warmth (M=1.79) when the robot was their partner (see table 1).

8.4.4 Envy. The display of different levels of warmth and competence by the robots has not influenced participant's envy towards the robots (see table 1). Thus, both the high ((M=1.95, S.E.=0.16) and low competence (M=2.07, S.E.=0.16), as well as high warmth (M=1.99, S.E.=0.16) and low warmth (M=2.03, S.E.=0.17) attributes of the robots seemed to evoked similar low levels of envy in participants. Robots playing as partners (M=1.86, S.D.=.154) and as opponents (M=2.16, S.E.=0.15) also evoked similar low levels of envy, F(1, 182)=2.851, p=.093. No interactions were found between the independent variables.

8.5 Future Intention to Work with Robots

To analyze participants' future intention to work with robots, we conducted a MLM in which we compared the four conditions (correspondent to each combination of high and low levels of warmth and competence). This analysis yielded yielded significant differences between all conditions (F(3, 203)=36.63, p<.001). In particular, participants displayed a higher willingness to interact again with the robot displaying high levels of warmth and competence (M=5.82, S.E.=0.33), than with the robots displaying high warmth and low competence (M=4.51, S.E.=0.33). Furthermore, participants also preferred to interact again with the latter (i.e. high warmth and low competence robot) than with the robot displaying low warmth and high competence (M=3.50, S.E.=0.33). Consequently, the lowest intention to interact again in the future, was directed at the robots displaying low warmth and low competence (M=2.12, S.E.=0.34).

8.6 Recall of Perceptions about Robots

Twenty-nine participants responded to a brief online questionnaire about the experiment, approximately one week and a half after their collaboration. Overall, 54,6% were able to correctly remember the score of each game. Paired t-test were used to compared participants' recall of the competence and warmth displayed by each robot, in each role. In the two conditions in which participants played with both the high and the low warmth robots, no differences were found for perceived warmth, both when the robot played as a partner, t(11)=-1.24, p=.241, and when it played as an opponent, t(11)=-.76, p=.463. Furthermore, for participants' recall of competence, the high competent robot was reported as being more competent than the low competent robot, in both role conditions: as a partner t(11)=-2.46, p=.032, M=4.58, S.D.=0.99 and M=3.08, SD=1.50, respectively; and as opponent, t(11)=-6.69, p<.001, M=5.08, S.D.=1.16 and M=2.25, S.D.=86, respectively.

Moreover, when comparing participants' recall of the warmth and competence displayed by the robots in the conditions in which the competence was maintained stable (high or low) within groups, we observed a difference between the two robots, both when they played as partners (t(16)=-3.85, p=.001) and as opponents (t(16)=-6.93, p<.001). More specifically, the high warmth robot was remembered as being warmer (M=5.71, S.D.=1.26 and M=5.47, SD=1.37, respectively) than the low warmth robot (M=2.71, S.D.=1.21 and

M=2.53, S.D.=1.18, respectively). Additionally, we also found differences on participants' recall of the levels of competence displayed by the robots, both when they played as partners, t(16)=-3.85, p=.001, and when they played as opponents, t(16)=-5.63, p<.001. More specifically, participants reported higher levels of competence for the high warmth robot (Partner: M=4.76, S.D.=1.52; Opponent: M=5.71, S.D.=1.26), in comparison to the low warmth robot (Partner: M=3,0, S.D.=1.17; Opponent: M=2.53, S.D.=1.18).

9 DISCUSSION

In this paper, we examined whether stereotypes affect the way we perceive and respond to robots in a social group situation involving more than one robot and more than one person. Our goal was to analyze how different levels of warmth and competence displayed by the robotic agents can intertwine to create an holistic stereotype of the robot, and also evoke different emotional responses from users. Furthermore, we also analyzed the role of the warmth and competence displayed by the robot in the participants' future intention to work with robots, as well as their recollection of the levels of warmth and competence displayed by the robots. Overall, our results seem to support several of our hypotheses. More specifically:

9.1 Robot's Role and Display of Warmth and Competence

9.1.1 H1 and H2. In the current study, participants were able to recognize robots displaying high and low levels of warmth through the verbal utterances. However, our results only partially verified our hypotheses and it was interesting to find an an effect of the level of competence on the perceived level of warmth of the robot. The fact that the low competent robot was perceived as being warmer than the high competent robot might have been due to a compensation effect on the participants' appraisal of the robot (viz. groups perceived low on one dimension, are often evaluated high on the other, due to the participants' assumption that the target of the stereotype must have some redeeming qualities, similar to what was observed by [37]). However, when comparing the low and high competence robot in terms of perceived warmth, the low competence robot was perceived more favorably. This is in line with the interpretation of social goals [21], the low competence of the robots might have been interpreted as incapable to pursue harmful intents, and thus judged more positively in terms of warmth than the high competent robot. Furthermore, this finding hints at the existence of an halo effect in robots' perception, in which high warmth robots are associated with other positive traits, in this case, competence.

Despite conceptual independence, appraisals of warmth and competence have been found to interact in previous studies, with groups that are described as being high in one dimension (e.g: warmth), also being increasingly better evaluated in the alternative dimension (competence). Overall, these results suggest some ambivalence in stereotypes when comparing group's levels of warmth and competence [37].

9.1.2 H3. Furthermore, and in line with our hypotheses, the role played by the robot (partner vs. opponent) seems to have affected the judgments of warmth. When both robots display high levels of warmth, the robot playing as a partner is perceived more favorably in terms of warmth in comparison to the robot playing as an opponent, suggesting an positive ingroup bias. However, when comparing robots displaying low levels of warmth, this relationship is inverted, thus suggesting a form of a black-sheep effect, in which members of the ingroup (in this case, a partner) that do not exhibit a social expected trait (high warmth) are judged more severely than members of the outgroup [42].

Additionally, when considering competence judgments, we observed that when both robots display low levels of competence, the robot playing as partner was judged to be less competent than the robot playing as opponent. This might have occurred by an increase in the salience of the trait of competence when both the individual and the robot were trying to collaboratively achieve the same goal or by the cumulative effect of the low (and in this context, undesirable) levels of warmth and competence.

9.2 Emotional Responses

9.2.1 H4. In the condition in which both robots displayed high warmth, participants admired more the robot with a low level of competence in comparison with the robot displaying high levels of this trait. This is incongruent with the predictions of the SCM, which state that admiration is linked to groups perceived as high in both competence and warmth. A possible explanation for this result might have resulted from positive assessments associated with incompetent agents. Previous studies have, to some extent, suggesting that incompetent-like behavior (e.g: cheating [52] or making mistakes [43]) by robotic agents can have a positive effect in robot evaluation as it increases likability (see Pratfall Effect [3]).

Nonetheless, when both robots displayed low levels of warmth, higher levels of admiration were directed at the high competent robot. Suggesting that the level of competence only has an effect in predicting admiration towards robots, when both robots display low levels of warmth.

9.2.2 H5. Regarding the feeling of pity, our hypothesis was partially verified. In particular, participants reported higher levels of pity towards robots displaying high levels of warmth (as expected). However, we found no difference in this variable considering the level of competence displayed by the robot.

9.2.3 H6. For contempt we were able to find differences in the judgments of the robot playing the role of a partner, indicating that contempt was felt stronger towards the low warmth robot that towards the high warmth robot. This is partially congruent with predictions from the SCM given that it predicts that contempt is associated with groups perceived as low in warmth. However, no differences were observed between the low and high competence robot. This result suggests that in HRI, the feeling of contempt is primarily associated with the level of warmth displayed by the robot and not to its' level of competence.

9.2.4 H7. Regarding the envy felt towards the robots, no effects of warmth, competence or the role played by the robot were found. As far as the authors are concerned, no previous studies have considered the specificities of feelings of envy directed at robots, nor what are its' social determinants. In our case, we hypothesize that the lack of effects found on this variable might be due to a floor effect, given that the levels of this feeling reported by participants were very low.

Overall, the warmth displayed by the robot seemed to be a stronger predictor of the participant's emotional responses towards robots than robot's competence (specially for pitty and contempt; emotions associated to low levels of competence in the original model), which is interesting considering that the context was competitive, a situational factor that could have contributed to make the competence stereotypes more relevant for participants' goal achievement (win the game). However, when considering feelings of admiration (associated to high levels of competence in the original model), participants' responses were determined by an interaction between warmth and competence.

9.3 Future Intention to Interact and Recall of Perceptions About Robots

9.3.1 H8 and H9. These hypotheses were verified. Regardless of the level of warmth, participants preferred to interact again with the high competence robot and when both robots displayed similar levels of competence, participants showed a preference, in terms of future intention to interact, for the high warmth robot. Additionally, we also observed that participants show a higher intention to interact again with the robot displaying high warmth (and low competence) than with the robot displaying high competence (and low warmth), suggesting that warmth plays a more pivotal role than competence in predicting future intention to interact with robots.

9.3.2 H10. Finally, this hypothesis was also verified. In our study, participants' recollections about the levels of warmth and competence displayed by robots were congruent with the levels of these traits displayed by the robots. This suggests a degree of stability in the stereotypical perceptions about robots that is maintained even after the interaction.

9.4 Limitations

Past research on HRI has suggested that cultural congruence between the robot and the user can improve some relational outcomes, such as likability and request compliance [58]. As such, the fact that Sueca is a traditional Portuguese card-game and as a result holds a cultural load, might have had a positive influence in the overall robots' perception. Unencumbered, we opted for this task because it provides an interesting scenario in which we can analyze small group interactions between humans and robots and because the predictability of game events (and the emotional and behavioral responses associated with it) allow us to develop autonomous robotic agents that can interact in these contexts. In this context, it is also important to consider a possible novelty effect, which over a short period of time, can cause atypical patterns of interaction [8]. Moreover, we used a single-item scale, however the reliability of our results could have benefited from the use of multiple items.

Despite these limitations, we contribute to the literature in HRI by establishing an initial understanding of how the content of stereotypes is related to the emotional responses and future intention to interact again with robots in a group context. Previous research has demonstrated the pivotal role of stereotypes in group interactions, as well as its' instrumental role in predicting behavioral and emotional responses to other individuals. In this paper, we propose the usefulness of these constructs in predicting emotional responses towards robots and we demonstrate its' effects in a group social interaction HRI setting.

10 FUTURE ENDEAVORS

Social robots are expected to keep growing, not only multiplying in numbers, but also acquiring increasingly more developed social skills and taking on more complex social roles [40]. As such, it is important to build a firm understanding of how humans and robots interact in social groups [38]. Following this line of thought, the content of stereotypes that people have towards robots seems to be a determinant factor to take into consideration. As such, future studies aimed at studying the role of stereotypes in HRI should include efforts directed at verifying the generalization of our findings across different contexts and analyzing the behavioral outcomes associated with each quadrant. This is of a particular importance given that is a well established fact that responses to robots are prone to multiple context related factors (e.g: culture) [38]. Furthermore, we would like to call for more studies analyzing the specific context of small groups overall, and in specific the role of different relational dynamics (e.g: collaboration and competition). We believe that this approach has the potential to significantly contribute towards the increased successful development of robots that can act and collaborate based on the assumption that work is generally performed by a multitude of individuals, and thus "*... open [the] possibility for robots to better support teamwork through more social roles*" [38] (p.4).

REFERENCES

[1] Michele G Alexander, Marilynn B Brewer, and Richard K Hermann. 1999. Images and affect: A functional analysis of out-group stereotypes. *Journal of Personality and Social Psychology* 77, 1 (1999), 78.

[2] Craig A Anderson and Melissa Morrow. 1995. Competitive aggression without interaction: Effects of competitive versus cooperative instructions on aggressive behavior in video games. *Personality and Social Psychology Bulletin* 21, 10 (1995), 1020–1030.

[3] Elliot Aronson, Ben Willerman, and Joanne Floyd. 1966. The effect of a pratfall on increasing interpersonal attractiveness. *Psychonomic Science* 4, 6 (1966), 227–228.

[4] Solomon E Asch. 1946. Forming impressions of personality. *The Journal of Abnormal and Social Psychology* 41, 3 (1946), 258.

[5] Solomon E Asch and Henri Zukier. 1984. Thinking about persons. *Journal of Personality and Social Psychology* 46, 6 (1984), 1230.

[6] Richard D Ashmore and Frances K Del Boca. 1979. Sex stereotypes and implicit personality theory: Toward a cognitive—Social psychological conceptualization. *Sex roles* 5, 2 (1979), 219–248.

[7] Richard D Ashmore and Frances K Del Boca. 1981. Conceptual approaches to stereotypes and stereotyping. *Cognitive processes in stereotyping and intergroup behavior* 1 (1981), 35.

[8] Ruth Aylett. 2016. Games Robots Play: Once More, with Feeling. In *Emotion in Games*. Springer, 289–302.

[9] John A Bargh. 1999. The cognitive monster: The case against the controllability of automatic stereotype effects. (1999).

[10] Andrea Bauer, Dirk Wollherr, and Martin Buss. 2008. Human–robot collaboration: a survey. *International Journal of Humanoid Robotics* 5, 01 (2008), 47–66.

[11] Cynthia Breazeal, Andrew G Brooks, David Chilongo, Jesse Gray, Guy Hoffman, Cory D Kidd, Hans Lee, Jeff Lieberman, and Andrea Lockerd. 2004. Working collaboratively with humanoid robots.. In *Humanoids*, Vol. 4. 253–272.

[12] Cynthia Breazeal, Guy Hoffman, and Andrea Lockerd. 2004. Teaching and working with robots as a collaboration. In *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems-Volume 3*. IEEE Computer Society, 1030–1037.

[13] MARILYNN B Brewer and MICHELE G Alexander. 2002. Intergroup emotions and images. *From prejudice to intergroup emotions: Differentiated reactions to social groups* (2002), 209–225.

[14] Noel T Brewer, Neil D Weinstein, Cara L Cuite, and James E Herrington. 2004. Risk perceptions and their relation to risk behavior. *Annals of behavioral medicine* 27, 2 (2004), 125–130.

[15] Terence Burnham, Kevin McCabe, and Vernon L Smith. 2000. Friend-or-foe intentionality priming in an extensive form trust game. *Journal of Economic Behavior & Organization* 43, 1 (2000), 57–73.

[16] J Edward Colgate, J Edward, Michael A Peshkin, and Witaya Wannasuphoprasit. 1996. Cobots: Robots for collaboration with human operators. (1996).

[17] Filipa Correia, Patrícia Alves-Oliveira, Tiago Ribeiro, Francisco S Melo, and Ana Paiva. 2017. A social robot as a card game player. In *13th AAAI conference on artificial intelligence and interactive digital entertainment*.

[18] Filipa Correia, Samuel Mascarenhas, Rui Prada, Francisco S Melo, and Ana Paiva. 2018. Group-based emotions in teams of humans and robots. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 261–269.

[19] Amy JC Cuddy, Susan T Fiske, and Peter Glick. 2004. When professionals become mothers, warmth doesn't cut the ice. *Journal of Social issues* 60, 4 (2004), 701–718.

[20] Amy JC Cuddy, Susan T Fiske, and Peter Glick. 2007. The BIAS map: behaviors from intergroup affect and stereotypes. *Journal of personality and social psychology* 92, 4 (2007), 631.

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

[22] Amy JC Cuddy, Susan T Fiske, Virginia SY Kwan, Peter Glick, Stephanie Demoulin, Jacques-Philippe Leyens, Michael Harris Bond, Jean-Claude Croizet, Naomi Ellemers, Ed Sleebos, et al. 2009. Stereotype content model across cultures: Towards universal similarities and some differences. *British Journal of Social Psychology* 48, 1 (2009), 1–33.

[23] Alexander M Czopp, Aaron C Kay, and Sapna Cheryan. 2015. Positive stereotypes are pervasive and powerful. *Perspectives on Psychological Science* 10, 4 (2015), 451–463.

[24] Kerstin Dautenhahn. 1999. Embodiment and interaction in socially intelligent life-like agents. In *Computation for metaphors, analogy, and agents*. Springer, 102–141.

[25] Morton Deutsch. 1949. A theory of co-operation and competition. *Human relations* 2, 2 (1949), 129–152.

[26] Brian R Duffy. 2003. Anthropomorphism and the social robot. *Robotics and autonomous systems* 42, 3-4 (2003), 177–190.

[27] Alice H Eagly, Richard D Ashmore, Mona G Makhijani, and Laura C Longo. 1991. What is beautiful is good, but…: A meta-analytic review of research on the physical attractiveness stereotype. *Psychological bulletin* 110, 1 (1991), 109.

[28] Naomi Ellemers. 2018. Gender stereotypes. *Annual review of psychology* 69 (2018).

[29] Friederike Eyssel and Frank Hegel. 2012. (s) he's got the look: Gender stereotyping of robots 1. *Journal of Applied Social Psychology* 42, 9 (2012), 2213–2230.

[30] Susan T Fiske, Juan Xu, Amy C Cuddy, and Peter Glick. 1999. (Dis) respecting versus (dis) liking: Status and interdependence predict ambivalent stereotypes of competence and warmth. *Journal of social issues* 55, 3 (1999), 473–489.

[31] Margaret Foddy, Michael J Platow, and Toshio Yamagishi. 2009. Group-based trust in strangers: The role of stereotypes and expectations. *Psychological Science* 20, 4 (2009), 419–422.

[32] Marlena R Fraune, Steven Sherrin, Selma Sabanović, and Eliot R Smith. 2015. Rabble of robots effects: Number and type of robots modulates attitudes, emotions, and stereotypes. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 109–116.

[33] Matthew L. Ginsberg. 2001. GIB: Imperfect information in a computationally challenging game. *Journal of Artificial Intelligence Research* (2001), 303–358.

[34] David L Hamilton. 2015. *Cognitive processes in stereotyping and intergroup behavior*. Psychology Press.

[35] Ronald H Heck, Lynn Tabata, and Scott L Thomas. 2013. *Multilevel and longitudinal modeling with IBM SPSS*. Routledge.

[36] Aline Hitti and Melanie Killen. 2015. Expectations about ethnic peer group inclusivity: The role of shared interests, group norms, and stereotypes. *Child Development* 86, 5 (2015), 1522–1537.

[37] Charles M Judd, Laurie James-Hawkins, Vincent Yzerbyt, and Yoshihisa Kashima. 2005. Fundamental dimensions of social judgment: understanding the relations between judgments of competence and warmth. *Journal of personality and social psychology* 89, 6 (2005), 899.

[38] Malte Jung and Pamela Hinds. 2018. Robots in the Wild: A Time for More Robust Theories of Human-Robot Interaction. *ACM Transactions on Human-Robot Interaction (THRI)* 7, 1 (2018), 2.

[39] David A Kenny, Lucia Mannetti, Antonio Pierro, Stefano Livi, and Deborah A Kashy. 2002. The statistical analysis of data from small groups. *Journal of personality and social psychology* 83, 1 (2002), 126.

[40] KPMG. 2016. MS Windows NT Social Robots. https://assets.kpmg.com/content/dam/kpmg/pdf/2016/06/social-robots.pdf

[41] Diane M Mackie and Eliot R Smith. 2016. *From prejudice to intergroup emotions: Differentiated reactions to social groups*. Psychology Press.

[42] José M Marques, Vincent Y Yzerbyt, and Jacques-Philippe Leyens. 1988. The "black sheep effect": Extremity of judgments towards ingroup members as a function of group identification. *European Journal of Social Psychology* 18, 1 (1988), 1–16.

[43] Nicole Mirnig, Gerald Stollnberger, Markus Miksch, Susanne Stadler, Manuel Giuliani, and Manfred Tscheligi. 2017. To err is robot: How humans assess and act toward an erroneous social robot. *Frontiers in Robotics and AI* 4 (2017), 21.

[44] Clifford Nass, Jonathan Steuer, and Ellen R Tauber. 1994. Computers are social actors. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 72–78.

[45] Illah R Nourbakhsh, Clayton Kunz, and Thomas Willeke. 2003. The mobot museum robot installations: A five year experiment. In *Intelligent Robots and Systems, 2003.(IROS 2003). Proceedings. 2003 IEEE/RSJ International Conference on*, Vol. 4. IEEE, 3636–3641.

[46] Raquel Oliveira, Patrícia Arriaga, Patrícia Alves-Oliveira, Filipa Correia, Sofia Petisca, and Ana Paiva. 2018. Friends or Foes?: Socioemotional Support and Gaze Behaviors in Mixed Groups of Humans and Robots. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 279–288.

[47] Mark Roberts, Thomas Apker, Benjamin Johnston, Bryan Auslander, Briana Wellman, and David W Aha. 2015. Coordinating Robot Teams for Disaster Relief.. In *FLAIRS Conference*. 366–371.

[48] Ira J Roseman. 2001. A model of appraisal in the emotion system. *Appraisal processes in emotion: Theory, methods, research* (2001), 68–91.

[49] Paul Rozin, Jonathan Haidt, and Clark R McCauley. 1999. Disgust: The body and soul emotion. *Handbook of cognition and emotion* (1999), 429–445.

[50] Nicole Salomons and Brian Scassellati. 2018. Trust and Conformity when Interacting with a Group of Robots. In *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 315–316.

[51] Brad E Sheese and William G Graziano. 2005. Deciding to defect: The effects of video-game violence on cooperative behavior. *Psychological science* 16, 5 (2005), 354–357.

[52] Elaine Short, Justin Hart, Michelle Vu, and Brian Scassellati. 2010. No fair!! an interaction with a cheating robot. In *Human-Robot Interaction (HRI), 2010 5th ACM/IEEE International Conference on*. IEEE, 219–226.

[53] Mikey Siegel, Cynthia Breazeal, and Michael I Norton. 2009. Persuasive robotics: The influence of robot gender on human behavior. In *Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on*. IEEE, 2563–2568.

[54] Adriana Tapus, Cristian Ţăpuş, and Maja J Matarić. 2008. User—robot personality matching and assistive robot behavior adaptation for post-stroke rehabilitation therapy. *Intelligent Service Robotics* 1, 2 (2008), 169.

[55] Benedict Tay, Younbo Jung, and Taezoon Park. 2014. When stereotypes meet robots: the double-edge sword of robot gender and personality in human–robot interaction. *Computers in Human Behavior* 38 (2014), 75–84.

[56] Benedict Tiong Chee Tay, Taezoon Park, Younbo Jung, Yeow Kee Tan, and Alvin Hong Yee Wong. 2013. When stereotypes meet robots: The effect of gender stereotypes on people's acceptance of a security robot. In *International Conference on Engineering Psychology and Cognitive Ergonomics*. Springer, 261–270.

[57] Rudolph Triebel, Kai Arras, Rachid Alami, Lucas Beyer, Stefan Breuers, Raja Chatila, Mohamed Chetouani, Daniel Cremers, Vanessa Evers, Michelangelo Fiore, et al. 2016. Spencer: A socially aware service robot for passenger guidance and help in busy airports. In *Field and service robotics*. Springer, 607–622.

[58] Lin Wang, Pei-Luen Patrick Rau, Vanessa Evers, Benjamin Krisper Robinson, and Pamela Hinds. 2010. When in Rome: the role of culture & context in adherence to robot recommendations. In *Proceedings of the 5th ACM/IEEE international conference on Human-robot interaction*. IEEE Press, 359–366.

[59] Samantha F Warta. 2015. If a robot did "the robot," would it still be called "the robot" or just dancing? Perceptual and social factors in human-robot interactions. In *Proceedings of the human factors and ergonomics society annual meeting*, Vol. 59. SAGE Publications Sage CA: Los Angeles, CA, 796–800.

[60] Astrid Weiss, Jenay Beer, Takanori Shibata, and Markus Vincze. 2014. Socially assistive robots for the aging population: Are we trapped in stereotypes?. In *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*. ACM, 497–498.

[61] Leslie A Zebrowitz. 1996. Physical appearance as a basis of stereotyping. *Stereotypes and stereotyping* (1996), 79–120.