Gender Effects in Perceptions of Robots and Humans with Varying Emotional Intelligence

1st Meia Chita-Tegmark
Human-Robot Interaction Laboratory
Tufts University
Medford, MA
mihaela.chita-tegmark@tufts.edu

2nd Monika Lohani
Affective and Cognitive Regulation Lab.
University of Utah
Salt Lake City, UT
monika.lohani@utah.edu

3rd Matthias Scheutz
Human-Robot Interaction Laboratory
Tufts University
Medford, MA
matthias.scheutz@tufts.edu

***Abstract*—Robots are machines and as such do not have gender. However, many of the gender-related perceptions and expectations formed in human-human interactions may be inadvertently and unreasonably transferred to interactions with social robots. In this paper, we investigate how gender effects in people's perception of robots and humans depend on their emotional intelligence (EI), a crucial component of successful human social interactions. Our results show that participants perceive different levels of EI in robots just as they do in humans. Also, their EI perceptions are affected by gender-related expectations both when judging humans and when judging robots with minimal gender markers, such as voice or even just a name. We discuss the implications for human-robot interactions (HRI) and propose further explorations of EI for future HRI studies.**

***Index Terms*—emotional intelligence; gender; social robots; human robot interaction**

I. INTRODUCTION

Our social experiences are primarily shaped by our interactions with other humans. Many of the perceptions and expectations that we form in human-human social interactions may, however, be inadvertently transferred to HRI scenarios, even when this transfer is unwarranted and may have detrimental effects. One such major class of perceptions and expectations are related to gender.

Robots are machines and as such do not have gender. Yet, this obvious fact does not prevent humans interacting with robots from ascribing gender to them and treating robots *as if* they were gendered entities. In fact, and somewhat surprisingly, the *as if* treatment is often so deeply rooted in our perception and behavior that we are not even aware of it: from subjects chatting longer with robots they perceive to be the "opposite sex" based on simple gender features such as gray vs. pink lips [1], to rating robots of the "opposite sex" as more credible, trustworthy, and engaging [2], to preferring robots whose gender markers (e.g., voices) match gender-stereotypical tasks (such as male robots performing security and female robots performing health care tasks) [3]. In some situations it might be desirable for a robot to elicit reactions through its gender markers (e.g. a robot managing a female locker-room might make the patrons feel more comfortable by having female gender markers), however projecting unwarranted gender stereotypes onto robots can often times be deterimental to the interaction.

The aim of this paper is to investigate the extent of gender effects in an area that has not yet received much attention in HRI: *emotional intelligence* (EI). While it is clear from human studies that EI is a critical multi-dimensional construct for social interactions that can affect a wide range of social dynamics (see the Background Section), there is currently little to no work in HRI on the effect of emotionally intelligent robots on human interactants or observers of HRI (there is only one recent study on the effects robot EI on human perceptions of those robots [4] and to our knowledge no interaction study on robots with EI). Assuming for the moment that people were to automatically perceive EI in robots and be influenced by those perceptions, it would be critical for robot designers to consider these perceptions and their effects. For example, it would be undesirable if the acceptance of a robot perfectly capable of performing a service task were diminished simply because it comes across as exhibiting low EI in its interactions with humans. Hence, we not only need to know whether people are likely to assess the EI of robots (automatically) when they observe them, but we also need to know what factors can modulate those perceptions.

The goal of this paper is to break new ground on this question by investigating the possible modulatory roles that participant gender and perceived robot gender have on the perceived EI of a robot. Specifically, we are interested in three overarching research questions: (1) whether subjects perceive EI in robots in the first place, and if so whether it differs from their perception of EI in humans – we hypothesize that they do based on [4]; (2) whether there are any differences in subjects' perceptions of robot EI based on the perceived gender of the robot (based on gender markers such as name and/or voice); and (3) whether the subjects' gender has any modulatory influence on their perceptions of robots' EI. We employ an *observation* study with video and text-based stimuli that display a typical office interaction between a supervisor and two subordinates where the supervisor reprimands one of the subordinates. After leaving the room, the other subordinate either does or does not display empathy, an important component of EI. We compare male and female human and robot agents in the role of the other subordinate and measure subjects' perceptions of EI. The results demonstrate that robot EI is perceived analogously to human EI and that intrinsic gender-based biases are transferred to robots and thus have to be addressed when a robot is construed as gendered.

The rest of the paper proceeds as follows: we start with a background review of emotional intelligence in humans and HRI, focused on gender effects on EI, followed by a brief review of gender effects in HRI more generally. We then introduce the experimental paradigm and results of our initial experiment that seeks to answer whether humans perceive different levels of EI in human-robot interactions, and whether these perceptions are modulated by gender manipulations. We then present our second and third experiments, motivated by findings in the first experiment, in which we isolate factors that might have impacted people's differential perceptions of EI. We show how Experiment 2 rules out the idea that differences in EI perception were based on the robot's voice (the only characteristic other than name distinguishing the robot in the male vs. female conditions), and describe Experiment 3 which tests whether implicit gender attributions are made even when the robot has no gender markers. In the subsequent general discussion, we compare the findings from all studies and discuss their implications for HRI, also including limitations of the studies. The conclusion then summarizes the results and highlights the need for further explorations of EI in HRI.

II. Background

Emotions play a crucial role in human psychology and consequently have a big impact on people's wellbeing. Social interactions, which constitute a major part of human life, are especially influenced by emotions. It is thus unsurprising that those people who effectively recognize and use emotions to adapt to different social contexts can increase their life success. This has motivated researchers to think of emotion assessment and management as a form of intelligence [5]–[7]. EI is briefly defined as the ability to perceive, understand, and manage emotions [6]–[10].

EI has been studied in humans in various contexts and has been linked to health outcomes [11], successful social interactions [12], educational outcomes [13], and job performance [7]. In the workplace, EI has been shown to enhance performance and stress management [14], bolster teamwork and conflict resolution [15], and overall boost the quality of the organizational climate [16]. The importance of EI in the workplace is of particular relevance for HRI, as robots are increasingly being integrated into the workplace. To maximize the benefits of collaborative work with robots, they need to be perceived as emotionally intelligent by their human team members.

Empathy has been argued to be one of the primary ways in which humans signal their EI [6]. Because EI involves the understanding and management of both one's own emotions and of the emotions of others, empathy has been conceptualized as the other-oriented side of EI [6]. Being able to share another person's perspective and respond effectively to their emotional needs is thus an essential component of EI and of healthy social relations [17]. In the workplace, perceptions of one's EI by peers has been related to higher ratings as a teammate, individual and leader [18], higher job performance ratings by supervisors [19] and better supervisor ratings of leadership and organizational citizenship [18]. In the context of HRI, it is thus natural to first study empathetic responses of the robot in relevant social contexts as signals of EI towards humans, especially if the lack thereof would be construed by humans as the robot's having low EI. Hence, it is important to understand potential pitfalls coming from perceptions and expectations regarding EI that should not be transferred to robots, for example those pertaining to gender-related differences and stereotypes.

There is mounting evidence, from studies using various measures of EI, that gender differences exist with respect to EI in humans. A recent meta-analysis of EI showed that women generally scored significantly higher than men on emotion perception, understanding, facilitation, and regulation [20]. In addition, there are also gender-based stereotype effects in EI. For example, consider that both males and females rated their fathers' EI as lower than their mothers' [5]. Another study found that most EI dimensions were perceived as more typical of one gender than another, suggesting that EI perceptions reflect gender stereotypes [21]. There is also a wide-spread belief that women have generally higher EI than men, which is likely grounded in women's ability to have better non-verbal emotion perception and deeper emotion knowledge (e.g, [22]).

Given that gender differences in cognitive performance exist in humans and that they can be triggered by stereotyping a task, it is thus natural to expect gender differences to have an influence in human-robot interaction as well, at least to the extent that the interaction task might involve either aspects where humans show gender-based differences in task performance or when stereotypes are connected to the task. Moreover, to the extent that the robot is perceived to be a human-like agent, humans' perceptions of the robot's gender will likely further influence the interactions. Work on gender effects in HRI can thus be grouped based on whether it considers effects (1) of perceived robot gender (e.g., [1]), (2) of subject gender (e.g., [23]), or (3) due to the gender stereotypicality of the task (e.g., [24]). Regarding EI in HRI, there are to our knowledge currently no studies that have investigated EI effects in human-robot interaction experiments; only one recent study has investigated EI effects on the perception of human observers of human-robot interactions [4].

III. Experiment 1

The aims of the first experiment are to investigate the extend to which humans perceive different EI levels in human-robot interactions and to determine the extent to which this effect might be modulated by gender perceptions and participant gender. Given the importance of emotional intelligence for human-human social interactions (see Background section) it is likely that people will notice, expect and assess EI in social robots as well. Also, given gender effects in EI discussed earlier from human experimental work, it is likely that gender projections will also modulate perceptions of EI in robots.

As such, this experiment contrasts human-robot interaction with human-human interactions, using cues to manipulate the gender of the human and robot actors.

A. Methods

1) Participants: A total of 197 participants completed this study through the Amazon Mechanical Turk (AMT). Participants' ages ranged from 18 to 68 years (Mean = 36.10, SD = 10.66) and, 44% of them were female. The ethnic composition of the sample was: White or Caucasian 79.2%, Asian 6.6%, African American 8.6%, Hispanic 5.1%, and other 0.5%.

2) Materials: We used the vignette from [4] which features an interaction between three characters in an office setting: a supervisor and two subordinates. The supervisor reprimands one of the subordinates for a mistake, and then leaves the room. The two subordinates subsequently discuss the situation, with the subordinate who was not reprimanded reacting towards the subordinate who made the mistake. The subordinate's reaction to the mistake of their colleague was either friendly and supportive (high EI condition) or unfriendly and unhelpful (low EI condition). While our experimental manipulation highlights empathy, other components are involved as well (e.g. emotion regulation vs. angry response in the high vs. low EI conditions).

The reacting subordinate character (the agent) was played by either a human actor (human condition) or a robot (robot condition). The characters were either all male (male condition) or all female (female condition; see Fig. 1).

Supervisor: Bob/Jessica, I was told you overrode Peter's/Katie's input and submitted the wrong coordinates to the team. You know that Peter/Katie has more experience than you. I don't want to hear about this happening again. The next time you disagree with Peter,Katie, you'd better be 100% sure before making any changes.

Bob/Jessica: Yes {sir,m'am}.

The supervisor leaves the room and Bob/Jessica turns toward Peter/Katie.

Bob/Jessica: I messed up, I don't know what to do.

Peter/Katie (high EI): We've been doing so well until now, this is the first mistake we made. What do you think happened? We can try to do things differently next time.

Peter/Katie (low EI): Yeah, it's too bad. We have a perfect record before this. I don't know what you did wrong, but you need to fix it for next time.

Bob/Jessica: We were receiving so many requests, but I thought I was handling it. And when I saw our different coordinates, I just panicked and submitted without thinking.

Peter/Katie (high EI): Look here, next time you begin to feel overwhelmed, just let me know, and I can try to help out. We are great teammates and I know we'll impress the supervisor!

Peter/Katie (low EI): Next time you need to do better, I'm also dealing with a lot of work and can't pick up your slack. You need to get it together before we present to the supervisor.

The robot used for the videos was a Willow Garage PR2 robot which has a humanoid appearance given its head, torso and arms but no clear gender markers (e.g., red lips, long hair, beard, etc.). Robot gender was entirely manipulated through the synthesized voice and the name given to the agent. The robot behavior – either turning towards the human and opening up the arms or turning away from the human and crossing the arms – were modeled based on the behavior of the human actor and performed once right after the supervisor had left the room. These actions were entirely scripted through an ROS component and initiated by a remote operator at the same time in the script in all conditions.

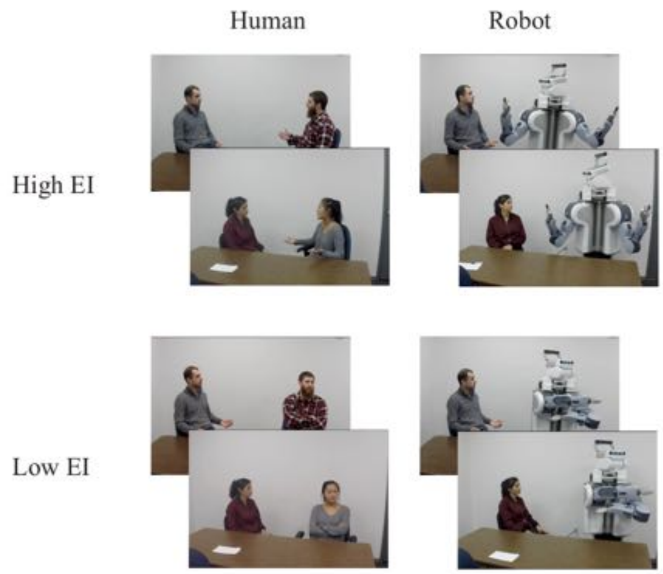

Fig. 1. The eight experimental conditions: EI (low/high), agent type (human/robot), agent gender (male/female).

We recorded eight videos, four with all male and four with all female actors in the low EI and high EI conditions, respectively. Each video had a supervisor and two subordinate characters who were either two humans or a human and a robot. The actors who played the male/female supervisor and the male/female subordinate remained the same in all four respective videos. The robot was the same in the male and female videos. The videos were filmed on a GoPro and edited in iMovie. The male and female conditions had identical dialogue for each EI condition, and male and female actors exhibited the same posture. We dubbed the voices of the robots in both the male and female conditions for clarity. In the male condition, the robot voice used was the Mac OS text-to-speech voice "Alex", and in the female condition, the voice used was the Mac OS text-to-speech voice "Samantha".

3) Measure: To measure EI perceptions of the agent, we used a 24-item questionnaire (see Table I, Cronbach's α was 0.99) based on [25]. Each item was scored on a 5-point Likert scale. Participants indicated how much each statement applied to the agent from "not at all" to "very much so" and items referred to emotion perception, understanding and management. We averaged the scores from all the items for each participant.

4) Procedure: The experiment consisted of three parts, all presented in a standard web browser. After agreeing to the con-

Sensitive to the needs of other people
Cheers people up when they need it
Creates a sense of belonging in groups or teams
Supports others when they are upset
Make people feel at ease
Good people skills
Contributes to a positive environment
Supports team or group member
Provides constructive feedback to people
Creates positive moods in people
Understands people's emotions
Emotionally connects with people
Puts people down
Would be a good colleague to work with
Bruque or abrasive with other people
Considerate of others' feelings
Has productive and helpful interactions with people
Sets a positive tone
Knows why people feel the way they do
Makes people feel bad when giving them feedback
Gets along well with people
Acts in a caring and kind way towards others
Knows the right thing to say when someone is upset
Is mean or unpleasant to others

TABLE I
EI PERCEPTIONS QUESTIONNAIRE WITH THE SCALE: "1=NOT AT ALL", "2=A LITTLE", "3=SOMEWHAT", "4=A LOT", "5=VERY MUCH SO".

sent form, participants first filled out basic demographic information (gender, age, ethnicity). Then they received instructions for watching the video. Subjects were randomly assigned to one of the above-mentioned eight conditions that differed with respect to *agent EI* (low/high), *agent type* (human/robot), and *agent gender* (male/female). After watching the video, subjects had to complete an EI survey that was used to measure their EI perceptions of the reacting (human or robot) agent in the scenario. Once the survey was completed, the subjects were able to collect their payment.

B. Results

1) Perceptions of Human and Robot EI : We started our analyses with the general question of whether subjects perceived a difference between the low and high EI conditions, which we purposefully made as *different as possible*. As mentioned, we expected the subjects to clearly see this difference at least in the human condition. Based on the work by [4], we also expected subjects to see the difference in the robot condition.

Consequently, we conducted a 2x2 ANOVA with *agent type* (human/robot) and *agent EI* (low/high) as independent variables and the perceived EI of the agent as the dependent variable. We found a significant main effect of agent EI, with the EI of agents in the low-EI condition rated as significantly lower than the EI of agents in the high-EI condition, $F(1, 193) = 154.91$, $p < .001$, $\eta_p^2 = .44$. No difference in agent type (robot/human) was found, $F(1, 193) = 0.03$, $p = .86$,$\eta_p^2 < .01$, and no interaction between EI and agent-type was observed, $F(1, 193) = 0.13$, $p = .72$, $\eta_p^2 < .01$.

2) Effects of Agent Gender : Next, we were interested in determining whether agent gender made a difference in EI perception. Hence, we performed a 2x2x2 ANOVA with *agent gender* (male/female), *agent type* (human/robot), and *agent EI* (low/high) as independent variables and perceived EI of the agent as dependent variable. In addition to the already observed main effect on EI, we found a significant main effect of agent gender, $F(1, 189) = 8.15$, $p = .005$, $\eta_p^2 = .04$, with male agents (both human and robot), rated significantly higher than female agents.

3) Effects of Participant Gender : To further investigate the potential influence of the participant's gender on EI perceptions, we conducted a 2x2x2 ANOVA with *participant gender* (male/female), *agent type* (human/robot), and *agent EI* (low/high) as independent variables and perceived EI of the agent as the dependent variable and found that the participants' gender did not have an influence on their perceptions of the agent EI, $F(1, 187) = 0.59$, $p = .44$, $\eta_p^2 < .01$.

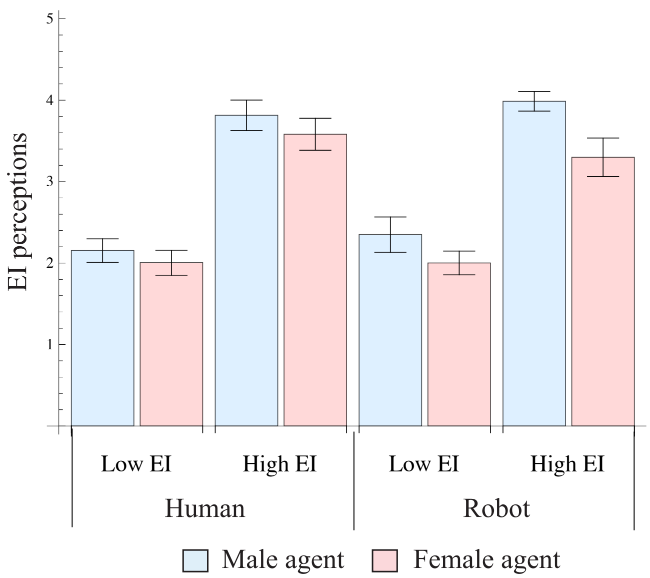

Fig. 2. EI perceptions (Means, SE) as influenced by agent type, agent EI and agent gender with video stimuli.

4) Effects of Agent and Participant Gender : Finally, to investigate any interaction effects between agent and participant gender, we conducted a 2x2x2x2 ANOVA including *participant gender*, *agent gender* (male/female), *agent type* (human/robot), and *agent EI* (low/high) as independent variables and perceived EI of the agent as the dependent variable. We found a significant main effect of agent EI, $F(1, 179) = 143.29$, $p < .001$, $\eta_p^2 = .44$, and of agent gender, $F(1, 179) = 8.08$, $p = .005$, $\eta_p^2 = .04$, but no significant effect of participant gender and no significant interactions.

C. Discussion

Male agents (humans and robots) were rated by participants as having significantly higher EI. This result is surprising since previous studies have shown that females tend to report having higher EI and also tend to be perceived by others as having higher EI [21]. In the case of the robot, this result is particularly intriguing, as the robot was identical in both gender conditions: while in the human condition one might argue that the differences in EI ratings was due to the appearance of the two actors (e.g., the female actor being less likable than the male actor), this argument cannot be made in the robot condition where the exact same robot was used in exactly the same way in each condition, i.e., with the same appearance, behavior, and tone-of-voice – albeit with different gendered voices. Hence, there are only two potential explanations: (1) either the difference in perceived EI for the robot was due to the robot's voice, or (2) it was due to gender-related differences in expectations: people expecting more EI of a female robot, with their ratings reflecting disappointment in the female robot and pleasant surprise with the male robot. A voice-based explanation might make sense if, for example, the male voice sounded particularly emphatic to subjects while the female voice was perceived as particularly off-putting. If the voice was the reason for the observed EI ratings, then removing the voice should remove the difference.

The question then arises whether the observed gender effects could be removed by eliminating gender features in the voice. There is difficulty in finding a gender-neutral voice for a robot. People tend to view robots without any clear gender markers as male. Hence, even if one were to select a neutral voice that in isolation might be rated as gender-neutral, this voice would be subservient to people's view of robots as male when paired with the appearance of a robot without visible gender features. Hence, the best way to remove any perceptual gender markers is to remove all perceptual features from the scenario and switch to a text-based vignette where people have to go entirely by the semantics of the story, instead of being able to integrate perceptual cues. In that case, only the agents' male or female names will be indicators of the agents' gender.

IV. Experiment 2

The aim of this experiment was to determine to what extent the effects we observed in the first experiment depended on the human or robot agents' appearance, their behavioral dynamics, and their voice characteristics. This includes gender features of the robot by way of the robot's gendered voice. By removing all appearance, behavior, and voice-based aspects of the interaction and just leaving the bare story and the subjects' imagination, we will be able to understand the extent to which the observed effects might have been due to subjects' gender-based prejudices and prior views of robots.

A. Methods

1) Participants: A total of 197 participants completed this study through the AMT. Participants' ages ranged from 18 to 77 years (Mean = 35.12, SD = 11.01) and 47% of them were female. The ethnicity composition of the sample was: White or Caucasian 73.6%, Asian 10.1%, African American 8.6%, Hispanic 4.6%, and other 3%.

2) Materials: We generated eight videos with text slide shows for this experiment, one for each condition, initially showing only head-shots of the two subordinates, and displaying only the text of the dialogue interactions without any sound (one interaction per slide) and without any additional visual cues as to the behavior of the interactants, their dynamics, or their voices. Subjects thus had to read the text and then answer questions based just on the read information.

a) Initial frame b) Example of dialogue turn

Fig. 3. Text stimuli from Experiment 2.

B. Results

1) Perceptions of Human and Robot EI: We started again with a 2x2 ANOVA with *agent type* (human/robot) and *agent EI* (low/high) as independent variables and perceived EI of the agent as the dependent variable and found a significant main effect of perceived agent EI just as in Experiment 1, with the EI of agents in the low-EI condition rated as significantly lower than the EI of agents in the high-EI condition, $F(1, 193) = 164.4$, $p < .001$, $\eta_p^2 = .46$.

2) Effects of Agent Gender : Adding agent gender, we performed a 2x2x2 ANOVA with *agent gender* (male/female), *agent type* (human/robot), and *agent EI* (low/high) as independent variables and perceived EI of the agent as the dependent variable. In addition to the main EI effect, $F(1, 189) = 171.51$, $p < .001$, $\eta_p^2 = .48$, we also get a significant agent gender effect, $F(1, 189) = 11.92$, $p < .001$, $\eta_p^2 = .06$, as in the video condition (Experiment 1), with males (human or robot) being perceived as more emotionally intelligent than females.

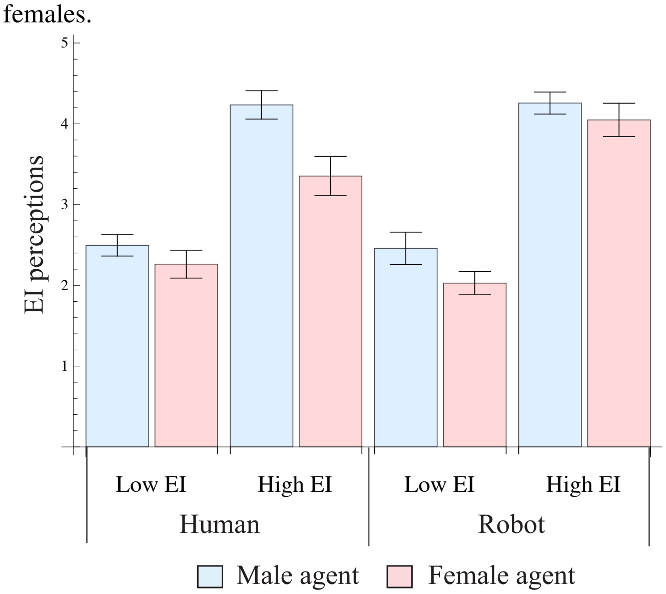

Fig. 4. EI perceptions (Means, SE) as influenced by agent type, agent EI and agent gender with text stimuli.

3) Effects of Participant Gender : To explore the influence of the participant's gender on EI perceptions we conducted a 2x2x2 ANOVA with *participant gender* (male/female), *agent type* (human/robot), and *agent EI* (low/high) as independent variables and, again, perceived EI of the agent as the dependent variable. We found no main effect of participant gender, $F(1, 189) = 2.10$, $p = .149$, $\eta_p^2 = .01$. However in addition to the main effect of agent EI, $F(1, 189) = 167.4$, $p < .001$, $\eta_p^2 = .47$, we also found significant two-way interactions between agent EI and agent type, $F(1, 189) = 5.40$, $p = .021$, $\eta_p^2 = .03$, and agent type and subject gender, $F(1, 189) = 11.90$, $p < .001$, $\eta_p^2 = .06$. Post-hoc comparisons with Tukey-Kramer corrections revealed that males rated the human agent higher than the robot agent, while females rated the human agent lower than the robot agent, with a significant difference between the higher male and lower female ratings of the human ($p = .005$) and the females' higher ratings of the robot than the human ($p = .016$). This finding replicates results in [23] where male subjects gave robots lower ratings on robot-related items when asked by the robot compared to answering questions on a written questionnaire without the robot present, while the direction of the ratings was exactly reversed for female subjects. The hypothesis was that the actual robot did not meet males' expectations regarding appearance and behavior, while the robot surpassed females' expections when they saw and interacted with it.

4) Effects of Agent and Participant Gender : Adding agent gender to the mix and performing a 4-way ANOVA, we again saw a strong main agent gender effect, $F(1, 181) = 8.71$, $p = .004$, $\eta_p^2 = .05$, in addition to the strong EI effect, and we obtained significant two-way interactions between agent EI and agent type, $F(1, 181) = 5.31$, $p = .022$, $\eta_p^2 = .03$, and agent type and participant gender, $F(1, 181) = 12.00$, $p = .001$, $\eta_p^2 = .06$, as before. In addition, we obtained a significant three-way agent EI, agent gender, and participant gender interaction, $F(1, 181) = 4.33$, $p = .039$, $\eta_p^2 = .02$. Pairwise comparisons showed that in the low EI condition, female participants rated female agents as having lower EI than male agents ($p = .045$). This suggests that gender stereotypes related to EI are at play and perhaps female participants use them to a higher degree in their evaluations.

C. Discussion

We moved from video-based to text-based interactions in order to determine whether the removal of the gendered voice in the robot condition, and overall the removal of any gendered-based cues other than the names of the actors, would be able to attenuate or even eliminate the gender-based effects. Surprisingly, quite the opposite was true: while the text-based stimuli did remove some of the specific video effects and also added effects not seen with video stimuli, the overall pattern of subjects rating male agents as higher in EI than female agents remained. The text-based stimuli confirmed that these biases are not the result of perceptual influences such as appearance (in the case of the human) or voice (in the case of the robot); nor can they be due to semantic differences as the text was the same in the low EI and high EI conditions, respectively; the only difference was the name of the agent designating either male or female gender.

It is possible that stereotypical gender difference in EI expectations (e.g., females having more EI than males) influenced differences in EI perceptions of males vs. female robots. Because females are stereotypically expected to have high EI, it may have led to a more conservative EI perception of females in general and even female robots (i.e., they were penalized for not matching stereotypically high level of EI in females). Similarly, males and male robots could have been perceived to display higher EI, more than stereotypically expected (i.e., a bonus due to a stereotypical image of low EI in males).

To investigate any differential expectations participants might have had with respect to high vs. low EI in males and females, and to test whether these expectations are automatically transferred from humans to robots, we need to show participants the vignettes with the agent stripped of any remaining gender markers, including its name. By asking participants to indicate how likely it is for the agent to be a male or a female in the low and high EI conditions we will be able to capture the participants' EI expectations of males and females.

Another possibility is that while the difference between males and females was due to stereotypical expectations of EI in the case of the humans, the perception of low EI was due to the jarring match between the PR2's female name and its bulky, angular steel-based frame, suggestive of large biceps. To rule out this possibility, we need to show participants the vignette without any photos of the robot, leaving the agent's appearance entirely to the participant's imagination.

V. EXPERIMENT 3

The aim of this experiment was to determine to what extent the agent gender effects we observed in the previous experiments depended on people's different stereotypical expectations of EI for males and females, which might have led to males receiving a bonus in the ratings for displaying high EI and females being penalized for showing low EI. We hypothesized that participants would have higher expectations of EI for females than males. We also wanted to test participants' spontaneous associations between gender and different robot appearances. We were particularly interested in whether participants associated the appearance of the PR2 robot (which was used in the previous experiments) with the male or female gender. We hypothesized that participants would be more inclined to think of the PR2 robot as "being male" due to its steel structure and sharp angular appearance.

A. Methods

1) Participants: A total of 100 participants completed this study through the AMT. Participants' ages ranged from 18 to 75 years (Mean = 33.73, SD = 12.01) and 44% of them were female. The ethnicity composition of the sample was: White or Caucasian 78%, Asian 4%, African American 7%, Hispanic 9% and other 2%.

2) Materials: We used a modified version of the videos with text slides from the previous experiment. The modification consisted in removing any gender markers from the dialogue. To this end, the head-shots of the two employees were removed from the beginning of the video, as well as any reference to them by name. Instead, the employees were referred to by their title: Group Leader and Assistant/Robot Assistant in the human and robot condition respectively. Note that the Assistant/Robot Assistant was the employee displaying varying levels of EI in the high EI and low EI conditions.

3) Measures: In addition to the EI measure from the previous experiments, we created two measures that we asked participants to complete after watching the videos: a name ranking measure and a robot naming measure. For the name ranking measure, as the name suggests, participants were asked to rank order how likely it was for the Assistant/Robot Assistant to have six different names. Two of the names were typical English male names (Peter and Bob), two were typical English female names (Katie and Jessica) and two were gender-neutral names (Pat and Taylor). The goal of this question was to understand how participants associated different levels of EI (low/high) with agents of different genders (male/female), as presumed by the different names. The robot naming measure consisted of asking participants to help a company name its robot employees. We showed participants photographs of PR2 (recall that for this experiment, we did not use any depiction of the characters in the vignette), Nao and MDS robots. Participants were asked to select between a male, a female and a gender-neutral name for each of the robots. The name options were randomized across robots and the display order of the robots was also randomized across participants. The goal of this question was to understand whether participants have spontaneous gender associations with different robot architectures. We were particularly interested in whether participants spontaneously associated the PR2 robot with the male gender.

B. Results

1) Perceptions of Human and Robot EI : To verify that our experimental manipulation of EI worked, we conducted a 2x2 ANOVA with *agent type* (human/robot) and *agent EI* (low/high) as independent variables and perceived EI of the agent as the dependent variable. We replicated our findings from the previous two experiments and found a significant effect of the EI experimental manipulation, $F(1, 91) = 114.57$, $p < .001$, $\eta_p^2 = .56$. We found no significant effect of agent type, (human/robot) on EI perception, $F(1, 91) = 0.29$, $p = .592$, $\eta_p^2 < .01$.

2) Effects of Participant Gender: To investigate whether the gender of the participants influenced EI perceptions, as we found in the second experiment, we further added *participant gender* (male/female) as an indepdendent variable for a 2x2x2 ANOVA. No significant effect of participant gender was found, $F(1, 87) = 0.01$, $p = .944$, $\eta_p^2 < .01$.

3) EI and Gender Expectations: The goal of this experiment was to understand associations between different levels of EI and agents of different genders. We used the name ranking measure (in order of how likely they were to belong to the agent in the vignettes) to create a score for each of the names. The scores ranged from 6 (chosen as most likely) to 1 (chosen as least likely). For each of the three name categories (male/female/gender-netural), the scores were averaged between the two names in the category. To understand how the different name scores might have been affected by the different experimental manipulations, we conducted a set of three 2x2x2 ANOVAs, with male name scores, female name scores and gender-netural name scores as the dependent variables and *agent type* (human/robot), *agent EI* (low/high) and *participant gender* (male/female) as the independent variables. We found a main effect of EI on male and female name scores but not on gender-netural name scores. Male names had a higher score, thus being considered more likely to belong to the agent, when the EI was low rather than high for both humans and robots, $F(1, 92) = 24.29$, $p < .001$, $\eta_p^2 < .21$. The opposite was true of female names, with higher female name scores when the EI was high rather than low for both humans and robots, $F(1, 92) = 10.55$, $p = .002$, $\eta_p^2 < .10$. Gender-netural names were considered equally likely in both the low and high EI conditions, $F(1, 92) = 3.18$, $p = .08$, $\eta_p^2 = .03$. We found no effects of agent type or participant gender on name ranking.

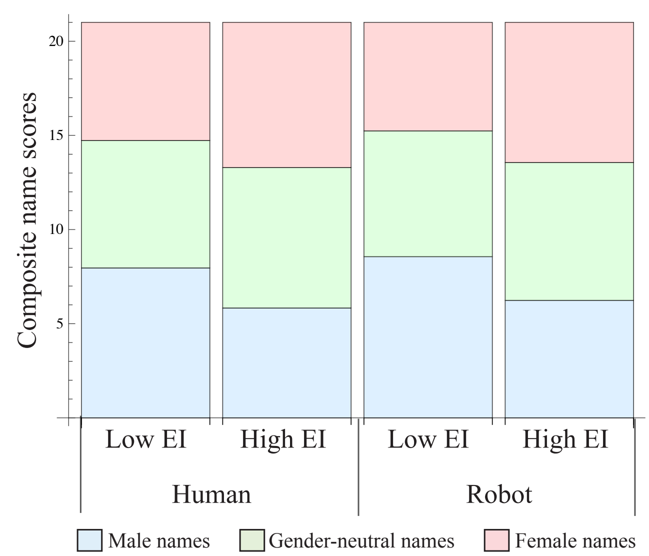

Fig. 5. Name ranking scores. Male names are selected preferentially in low EI conditions while female names in high EI conditions.

4) Gender and Robot Appearance: To understand spontaneous associations between gender and different robot appearances, we calculated the frequency with which participants assigned different robots with male, female and gender-neutral names. Participants selected male, female and gender-neutral names for the PR2 robot significantly different from chance,$\chi^2(2) = 21.98$, $p < .001$, with male names being selected 39% (95% CI: 29%-49%) of the time, female names being selected significantly less than chance 12% of the times

(95% CI: 6%-20%, $p < .001$), and gender neutral names being selected significantly more than chance, 49% of the time (95% CI: 39%-59%, $p = 0.001$).

C. Discussion

Participants associated low EI with agents with male names and high EI with agents with female names. This was the case for both the human and robot agents. This suggests that people have higher expectations in terms of EI from females, which supports the notion that the intriguing effects of higher EI-ratings for males (humans and robots) in the first two experiments was due to stereotypical gender-based expectations. Males were perceived to react "typically" in the low EI condition and got bonus points for exceeding expectations in the high EI condition, while females were perceived to react "typically" in the high EI condition and were penalized for not meeting expectations in the low EI condition. These stereotypical expectations were also transferred to robots, where people were more likely to select male names for robot agents in the low EI condition and female names for robots in the high EI condition. Moreover, people seemed to make spontaneous preferential gender projections onto robots, as shown by their significantly different-from-chance selection of gendered names for robots including PR2, the appearance of which has no apparent gender markers.

VI. Discussion

Together, the three sets of experiments establish clearly that human observers can detect low vs. high EI in humans and robots alike as we expected, which confirms the previous finding in [4], and that robots with gender markers will trigger gender-based stereotypes about EI in human observers: females are viewed as having higher EI than males, and the same goes for robots that are viewed as gendered by human observers. When gender markers are removed, gender-based differences disappear.

1) Implications for HRI: Our results clearly indicate the need to develop mechanisms for robots that allow them to exhibit high EI, because people will automatically judge robot EI when they observe human-robot interaction; being perceived as having low EI can have detrimental implications for robots (e.g., loss of human trust as shown in [4]). Designers could preempt that by building in brief interactions to develop high EI perceptions of the robots [26]. Also, by being aware of the inherent gender biases exhibited by participants, robot designers can attempt to mitigate gender-based influence on perceptions of a robot's EI. For instance, designers could opt for a female voice for robots exhibiting EI, given that male robots surpass expectations in the beginning but might disappoint in the long run. However, for robots performing tasks that involve delicate social situations a male voice might be perferable to minimize negative reactions towards potential faux pas. Such strategies could be adopted to facilitate optimal levels of perceived EI and trust in robots. Also, as gender stereotypes may differ by culture and may change over time, HRI researchers need to be constantly probing and adapting designs.

2) Limitations and Future Work: In this study used an observation paradigm because the genderless aspect of the vignettes in Experiment 3 cannot be realized in an interaction study with a physical robot. While our results yield important insights about EI perceptions for HRI design when humans are observers of interactions, we need to also investigate interaction contexts where participants are interactants. Evidence from the social interaction literature indeed suggests that social cognitions are different when one is in the role of an interactant rather than just an observer of interactions [27]. Such studies are clearly needed as a next step in investigating whether these effects apply to interactants, and to what extent.

Also, the study primarily investigated one aspect of EI, namely the capacity for empathy. It would be important to determine to what extent other aspects of EI are noticed by human interaction observers and interactants, and whether they have any effects on their attitudes towards robots. Additionally, the exposure to the agent is very limited: participants saw only one interaction, which likely makes their EI perceptions lack in nuance (suggested by the measure's high Cronbach's α). EI perceptions in the context of familiarity with the agent needs to be explored in the future.

Finally, we investigated only one office interaction scenario, which, while prototypical for many reprimand situations, is by no means representative for all types of situations where EI may make a difference. We are thus planning to extend the investigation to other interaction contexts to determine the generalizability of the results across more human social contexts.

VII. Conclusion

In this study, we set out to investigate gender effects in human perception of emotional intelligence (EI) in robots. For this purpose, we evaluated subjects' perceptions of interactions in a typical office scenario between a human supervisor and either two humans or one human and one robot, either all male or all female, using video-based and text-based stimuli. We found strong EI effects across both stimuli domains based on the subjects' ratings of the human and robot EI: both male and female subjects rated human and robot agents in the high EI condition as significantly higher than human and robot agents in the low EI condition. Moreover, both genders rated male (human or robot) agents as having higher EI than female (human and robot) agents, showing that gender-based stereotypical expectations of EI can be transferred from humans to robots. The most immediate implications for HRI are that (1) it is important to be aware of the level of EI robot behaviors may signal to human observers, even if they are not intended, and (2) gender markers are important modulators of the perceived EI in robots, triggering stereotypical responses and expectations that need to be taken into account, especially with robots that will be construed as gendered by humans.

VIII. Acknowledgments

This project was supported by NSF grant IIS-1316809.

REFERENCES

[1] A. Powers, A. D. Kramer, S. Lim, J. Kuo, S. lai Lee, and S. Kiesler, "Eliciting information from people with a gendered humanoid robot," in *Proceedings of the 2005 IEEE International Workshop on Robots and Human Interactive Communication*, 2005, pp. 158–163.

[2] M. Siegel, C. Breazeal, and M. I. Norton, "Persuasive robotics: The influence of robot gender on human behavior," in *Proccedings of IROS*, 2009, pp. 2563–2568.

[3] B. Tay, Y. Jung, and T. Park, "When stereotypes meet robots: The double-edge sword of robot gender and personality in human–robot interaction," *Computers in Human Behavior*, vol. 30, pp. 75–84, 2014.

[4] L. Fan, M. Scheutz, M. Lohani, M. McCoy, and C. Stokes, "Do we need emotionally intelligent articial agents? first results of human perceptions of emotional intelligence in humans compared to robots," in *Proceedings of the Seventeenth International Conference on Intelligent Virtual Agents*, 2017.

[5] K. V. Petrides, A. Furnham, and G. N. Martin, "Estimates of emotional and psychometric intelligence: Evidence for gender-based stereotypes," *The Journal of Social Psychology*, vol. 144, pp. 149–162, 2004.

[6] P. Salovey and J. D. Mayer, "Emotional intelligence," *Imagination, cognition and personality*, vol. 9, no. 3, pp. 185–211, 1990.

[7] D. P. Goleman, "Emotional intelligence: Why it can matter more than iq for character, health and lifelong achievement," 1995.

[8] R. Bar-On, "The emotional intelligence inventory (eq-i): Technical manual," *Toronto, Canada: Multi-Health Systems*, 1997.

[9] J. Ciarrochi, J. P. Forgas, and J. D. Mayer, *Emotional intelligence in everyday life: A scientific inquiry*. Psychology Press, 2001.

[10] R. K. Cooper and A. Sawaf, *Executive EQ: Emotional intelligence in leadership and organizations*. Penguin, 1998.

[11] N. S. Schutte, J. M. Malouff, E. B. Thorsteinsson, N. Bhullar, and S. E. Rooke, "A meta-analytic investigation of the relationship between emotional intelligence and health," *Personality and individual differences*, vol. 42, no. 6, pp. 921–933, 2007.

[12] C. M. Kokkinos and E. Kipritsi, "The relationship between bullying, victimization, trait emotional intelligence, self-efficacy and empathy among preadolescents," *Social psychology of education*, vol. 15, no. 1, pp. 41–58, 2012.

[13] S. E. Rivers, M. A. Brackett, M. R. Reyes, N. A. Elbertson, and P. Salovey, "Improving the social and emotional climate of classrooms: A clustered randomized controlled trial testing the ruler approach," *Prevention science*, vol. 14, no. 1, pp. 77–87, 2013.

[14] L. Fariselli, J. Freedman, M. Ghini, and F. Valentini, "Stress, emotional intelligence, and performance in healthcare," *Retrieved December*, vol. 2, p. 2009, 2008.

[15] N. Clarke, "Emotional intelligence and its relationship to transformational leadership and key project manager competences," *Project Management Journal*, vol. 41, no. 2, pp. 5–20, 2010.

[16] N. Momeni, "The relation between managers' emotional intelligence and the organizational climate they create," *Public Personnel Management*, vol. 38, no. 2, pp. 35–48, 2009.

[17] N. S. Schutte, J. M. Malouff, C. Bobik, T. D. Coston, C. Greeson, C. Jedlicka, E. Rhodes, and G. Wendorf, "Emotional intelligence and interpersonal relations," *The Journal of social psychology*, vol. 141, no. 4, pp. 523–536, 2001.

[18] H. A. Elfenbein, S. G. Barsade, and N. Eisenkraft, "The social perception of emotional abilities: Expanding what we know about observer ratings of emotional intelligence." *Emotion*, vol. 15, no. 1, p. 17, 2015.

[19] K. S. Law, C.-S. Wong, and L. J. Song, "The construct and criterion validity of emotional intelligence and its potential utility for management studies." *Journal of applied Psychology*, vol. 89, no. 3, p. 483, 2004.

[20] D. Joseph and D. Newman, "Emotional intelligence: An integrative meta-analysis and cascading model," *Journal of Applied Psychology*, vol. 95, no. 1, pp. 54–78, 2010.

[21] E. Lopez-Zafra and L. Gartzia, "Perceptions of gender differences in self-report measures of emotional intelligence," *Sex roles*, vol. 70, no. 11-12, pp. 479–495, 2014.

[22] E. McClure, "A meta-analytic review of sex differences in facial expression processing and their development in infants, children, and adolescents," *Psychological Bulletin*, vol. 126, pp. 424–453, 2000.

[23] P. Schermerhorn, M. Scheutz, and C. R. Crowell, "Robot social presence and gender: Do females view robots differently than males?" in *Proceedings of the Third ACM IEEE International Conference on Human-Robot Interaction*, Amsterdam, NL, March 2008.

[24] D. Kuchenbrandt, M. Häring, J. Eichberg, F. Eyssel, and E. Andre, "Keep an eye on the task! how gender typicality of tasks influence human–robot interactions," *International Journal of Social Robotics*, vol. 6(3), pp. 417–427, 2014.

[25] D. Caruso and P. Salovey, *Emotionally Intelligent Manager: How to Develop and Use the Four Key Emotional Skills of Leadership*. Jossey Bass Inc, 2004.

[26] M. Lohani, C. K. Stokes, K. B. Oden, S. J. Frazier, K. J. Landers, P. L. Craven, D. V. Lawton, M. McCoy, and D. J. Macannuco, "The impact of non-technical skills on trust and stress," in *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 2017, pp. 191–192.

[27] L. Schilbach, B. Timmermans, V. Reddy, A. Costall, G. Bente, T. Schlicht, and K. Vogeley, "Toward a second-person neuroscience 1," *Behavioral and brain sciences*, vol. 36, no. 4, pp. 393–414, 2013.