Noname manuscript No.
(will be inserted by the editor)

The interplay between emotional intelligence, trust, and gender in human-robot interaction

A vignette-based study

Theresa Law · Meia Chita-Tegmark · Matthias Scheutz

Received: date / Accepted: date

Abstract As robots begin to enter roles in which they work closely with human teammates or peers, it is critical to understand how people trust them based on how they interpret the robot's behavior. In this paper we investigated the interplay between trust in a robot and people's perceptions of the robot's emotional intelligence. We used a vignette-based method to explore the following questions: (1) do subjects perceive differences in robot EI, and is their trust in the robot influenced by differences in the robot's reliability and capability? (2) does a robot's EI influence how much it is trusted and conversely does a robot's capability and reliability influence how emotionally intelligent it is perceived to be? (3) do people trust male and female robots differently when the robots exhibit different levels of EI or different levels of capability and reliability, and do gender stereotypical expectations related to EI transfer to trust?; (4) does focusing on the robot's EI increase one's trust in the robot? (5) is the interplay between trust, EI and gender the same for different levels of evoked social presence and human-likeness (i.e., when the interaction is presented in different modalities, text or spoken dialogue when the robot's voice is actually heard)? We found that trust in the robot was influenced by the level of the robot's EI (p<.001) and that gender stereotypical expectations related to EI were transferred to trust (p=.006), but gender effects on trust disappeared when only capability and reliability (robot's trustworthiness) were manipulated but not the robot's EI (p=.103). Surprisingly, we found that people trusted the robot more when the interaction was presented in text format (p=.024), going against our hypothesis that spoken dialogue would evoke more social presence and thus bolster EI perception and instill more trust. We suggest that this effect might be due to people's expectations of a more expressive and human-like voice. Finally, we also found that people's trust ratings in the robot were higher when they were made to notice and think about the robot's EI, by answering EI questionnaires prior to trust questionnaires (p=.022). We discuss the implications of our findings for robot design and HRI research.

Keywords Human-robot interaction, trust, emotional intelligence, gender

(T. Law & M. Chita-Tegmark & M. Scheutz)
HRI Lab
Tufts University
Medford, MA 02155
E-mail: {theresa.law, mihaela.chita-tegmark, matthias.scheutz}@tufts.edu

1 Introduction

As robots begin to enter roles in which they work closely with human teammates or peers, it is critical to understand how people interpret robot behavior, both consciously and subconsciously. Humans interact with each other based on a number of heuristics and norms, some of which translate to how they view and interact with robotic agents [11], and some of which do not. These translations can affect their perceptions of, interactions with, and belief in the capabilities of a robot.

One of these subconscious translations that people make is in gendering robots. Robots are machines and therefore do not have a gender; however, people are susceptible to cues in the robot's appearance, voice, and behavior which can influence them to interpret the robot as being either male or female. This interpretation can lead to people exhibiting different behaviors depending on whether they see themselves as interacting with a male or female robot. For example, in [29], participants spent longer explaining dating norms to a

male robot than a female robot, as indicated by the robot having pink lips or grey lips, and a male or female voice. Gendering robots can also result in translating human gender stereotypes to robots. In [38], participants preferred robots that performed an occupation that stereotypically matched the gender of the robot's voice and name. However, the perceived personality of the robot can also interact with gender stereotypes to affect user acceptance of the robots. Other studies have also shown that the participant's gender may have an effect on how robots are perceived [10] [32]. These findings suggest that careful consideration needs to be paid to incorporating explicit or implicit gender cues in robot design.

Traits that have strong gender stereotypes may be more susceptible to transferring those stereotypes to robots. Emotional intelligence (EI), for example, is a trait that is heavily gender stereotyped, with women thought to be more emotionally intelligent than men [28]. Because EI can be critical in human-human interactions, there may be situations in which is it beneficial for a robot to exhibit EI. However, when designing emotionally intelligent robots, researchers need to be aware that people may have different expectations of a robot's EI, depending on whether they view the robot as male or female, female robots being expected to have higher EI than male robots [7].

For traits that have less obvious gender stereotypes, such as trust, it is less clear how a robot's perceived gender would affect a person's interactions with it. Unlike EI, there is not a clear stereotype of males or females being more trustworthy. Though some research has indicated that women may be more trustworthy than men [5], this finding has not been replicated to the extent that gender difference in perceived EI has been [1]. In a review of economic games that includes an examination of gender differences in economic games that measure trust, [9] found that most studies reported no gender differences in trust, while some reported that women were more trusting and others report that men were more trusting. Therefore, it seems likely that overall if people transfer gender stereotypes onto a robot, assumptions about that robot's trustworthiness should not be affected. However, [12] found that participants trusted a robot who exhibited high EI more than a robot who exhibited low EI, so it is possible that the gender effects found in EI ratings would transfer to trust ratings. Because trust in a robot is a sensitive and critical aspect in a human-robot team, it is important for robot designers to understand if there are robot gender differences in trust, as well as whether unmet or superseded stereotypical gender expectations of EI impact not just EI perceptions, but also how much the robot is trusted.

Additionally, if there are effects of EI and trustworthiness on the perception of robots, there remains a question about the robustness of these effects across different interaction and observation modalities. If information about the robot's levels of EI and trustworthiness come from what it says, it is possible that there will be a difference in perceptions based on whether the information is read as text or heard as spoken dialogue. Text-based interactions and spoken dialogue interactions have been found to have different effects on participants' experiences with technology, with spoken interactions being rated more favorably [18]. Giving a robot a voice can increase its human-likeness [34]. It also provides the robot with a sense of social presence and more information about the dynamics of the conversation than text could. This may, in turn, increase feelings of social connectedness with the robot. Voice may also lead people to trust the robot more and perceive it as being more emotionally intelligent.

The aim of this paper is to tease apart the interplay between EI, trust, and gender effects. Specifically, we are interested in the following research questions: (1) Do subjects perceive differences in robot EI, and is their trust in the robot influenced by differences in the robot's reliability and capability? (2) Does a robot's EI influence how much it is trusted and conversely does a robot's capability and reliability influence how emotionally intelligent it is perceived to be? (3) Do people trust male and female robots differently when the robots exhibit different levels of EI or different levels of capability and reliability, and do gender stereotypical expectations related to EI transfer to trust?; (4) Does focusing on the robot's EI increase one's trust in it?; (5) Is the interplay between trust, EI and gender the same for different levels of evoked social presence and human-likeness (i.e., when the interaction is presented in different modalities, text or spoken dialogue when the robot's voice is actually heard)? We ran two online observation studies to address these questions.

The rest of the paper is structured as follows: we begin with a very brief summary of EI research, including gender effects related to EI and EI in HRI, and then present a short review of trust in HRI and gender effects seen in trust in HRI. We then outline the current studies and hypotheses before presenting the methodology and results of Experiment 1. We also discuss the findings and implications of Experiment 1, which were then used to inform the methodology of Experiment 2. Finally, we present the findings of Experiment 2, as

well as a general discussion of the overall implications of both studies, including limitations and directions for future HRI research.

2 Background

2.1 Emotional Intelligence

Emotional intelligence can be broadly defined as the capacity to perceive and understand emotions, both one's own and other's [31]. High EI has been associated with positive behaviors, such as a decreased likelihood to bully others [19]. In the workplace specifically, EI has been linked with better stress management and performance [13], a better organizational climate [25], and better and more effective teamwork [8]. Therefore, if we are going to build robots that are going to exist in the workplace, it could be beneficial for those robots to be perceived as being emotionally intelligent. One way to do this is to design robotic behaviors that express empathy. Empathy is a key feature of EI; it is one of the clearest ways in which people can signal to others that they understand that person's emotions [31]. Therefore, if we are to build emotionally intelligent robots, empathy expression can act as a means to indicate that robot's level of EI.

Gender differences have been found on a variety of temperamental dimensions (e.g., perceptual sensitivity and inhibitory control) [**?**]. Along with these, EI has been repeatedly shown to have strong gender differences and stereotypes, with evidence that there may be neural differences in the EI capacities of females and males [35] [30]. Women have self-reported as being more empathic than men [39], have been found to have better facial expression processing [24], and have scored higher on all-around validated EI measures [36]. Additionally, these views translate to how people stereotype others' EI. In one study where participants were asked to estimate their mother and father's EI and IQ, participants rated their mothers as higher in EI and their fathers as higher in IQ [28]. Another study found that certain properties of EI were perceived by participants as being more typical of females or males, suggesting there may be nuances to the gender stereotyping of EI [22].

In HRI, there are very few studies exploring how people perceive the EI and empathy expression of robots. However, findings by [**?**] suggest that people will often co-create experiences with robots beyond specific tasks that the robots were designed to do. For example, people interacting with robots in the hospitality industry (i.e., hotels that have robotic assistants) shows that people proactively seek social interaction with robots, for the success of which EI may be an important factor. To the authors' knowledge, only [12] and [7] have specifically studied this by varying a robot's EI through its words and body language. [12] found that people could distinguish between robots that showed high and low EI, and that this distinction was not different for humans or robots exhibiting the same behaviors. They also found that participants trusted agents with high EI more than agents with low EI. [7] replicated these results, and additionally found that the robot's gender affected how emotionally intelligent participants viewed it to be, with male agents, both human and robot, rated as being more emotionally intelligent than female agents. This result was found both when the participants heard the robot's voice, which was gendered to be either male or female, and when the robot only had a gendered name and the participant read the script. The authors hypothesized that this unexpected finding was likely a result of people having higher expectations for the EI of the female robot which were then not met, whereas they were pleasantly surprised by the EI of the male robot.

2.2 Trust

In HRI, achieving the proper level of trust between a person and a robotic agent is an important factor in maximizing the effectiveness of a robot, especially a robot teammate. Too little trust can result in an under-reliance on a robot that could be helpful for achieving a goal, and too much trust can result in over-reliance on an imperfect system [16]. Though trust is a broad and multi-faceted concept [42], it can be conceptualized as being divided into two main categories: performance-based trust and relation-based trust. Performance-based trust is trust in an agent's ability to complete a task satisfactorily and consistently; relation-based trust is trust that an agent will comply with social norms [21]. Both types of trust can impact user acceptance, as [**?**] show; even in a navigation task, participants preferred a robot that exhibited social behaviors such as talking.

Unlike EI, trust does not have clear gender stereotypes. In some psychology literature, women are found to be more trusting than men [14], while in others, men are found to be more trusting [3]. Similarly, there is not the same type of trend as seen in the EI literature where either men or women are thought to be more trustworthy. One study found that both male and female participants, when asked, indicated that they be-

lieved women would cooperate more in a prisoner's dilemma game, which indicates that they may be more trusting and trustworthy. However, when they actually then played the game, there were no gender effects, indicating that any potential stereotyping did not manifest in the participants' actual behavior [27].

Though there are only a few HRI studies that looked specifically at the role of gender in a person's level of trust in a robot, the same inconsistency of trust and robot gender is seen in the HRI literature. [2] ran a small study that found that there was a trend towards trusting male robots more than female robots by being willing to share more information with a male than a female robot. [20] found that participants trusted robots who performed a stereotypically male task over robots that performed a stereotypically female task, though there was no effect of robot explicit gender (as indicated by the robot's voice) on trust. [15] investigated whether there was an effect of the robot's gender being similar or dissimilar to the participant's gender, and found no effect of robot gender on trust. [37] found that there was a significant effect of gender on trust rating with female robots being rated higher than males, but only when the participants were male. Finally, [4] manipulated gender by changing the robot's waist to hip ratio and found that participants trusted the female robot more.

3 The Current Studies

In [12], the researchers manipulated a robot's level of EI, and found that participants rated the more emotionally intelligent robot as having both a higher EI and as being more trustworthy. [7] found that people's perceptions of a robot's EI was affected by stereotypical gender expectations. This was found both when the participants read the vignette as text and when they heard it as audio. [26] demonstrated that having a voice increases a robot's human-likeness. For the studies presented in this paper we were interested in investigating whether stereotypical gender expectations related to EI, as found by [7], influence how much people trust robots, given the relationship between EI and trust found by [12]. Additionally, we want to verify whether implied robot gender has an impact on how much people trust robots that exhibit different levels of capability and reliability in the absence of EI manipulations. We were also interested in whether providing the robot's voice, and therefore likely increasing the robot's human-likeness, increased participants' trust in the robot. Finally, we wanted to explore how much trust and EI ratings influence each other, by checking whether people's perceptions of trust and EI are affected by the order in which their attention is drawn to those traits.

We present two studies attempting to disentangle the interplay of EI, trust, and gender. In the first study, we manipulated the robot's level of EI (low vs. high), its gender (male vs. female), and the vignette presentation style (text vs. voice). We used the different modalities to create two different levels of social presence of the robot since social presence through voice has been shown to affect people's interactions with a robot [32]. By having participants hear the robot's voice, we hoped to increase the robot's social presence and human-likeness, and also reinforce the robot's gender throughout the interaction (the robot had a female or male voice based on the gender condition). We sought to replicate the results seen in [12] and [7] in which robots who exhibit high EI are seen as being more emotionally intelligent and trustworthy than robots with low EI, and that male robots are seen as being more emotionally intelligent than female robots. Additionally we investigated whether gender in robots with different levels of EI had an influence on how much they were trusted. In our second study, we manipulated the reliability and capability of the robot to have varying levels of performance-based trust (low vs. high), as well as the robot's gender (male vs. female), the vignette presentation style (text vs. voice), and the order in which participants were asked about the robot's EI and their trust in it (EI questionnaires first vs. trust questionnaires first). In the second study we checked whether gender might affect how much a robot is trusted when displaying different levels of capability and reliability. Additionally, we investigated whether a robot's capability and reliability had an influence on how emotionally intelligent it was perceived to be and whether EI and trust ratings were subject to order effects. All study procedures were approved by our institution's IRB. In the two studies we specifically tested the following hypotheses:

- **H1a** Participants viewing the interaction in which the robot behaved in a non-empathetic way (*low EI* condition) will rate the robot as having lower EI than those viewing the interaction in which the robot behaved in an empathetic way (*high EI* condition). This would replicate the findings of [12].
- **H1b** Participants viewing the robot that indicated low capability and reliability (*low trust* condition) will rate their trust in the robot as lower than those viewing the robot that indi-

cated high capability and reliability (*high trust* condition).
- **H2a** When the robot's **EI** is manipulated, participants will rate the *male* and *female* robots differently on EI and trust.
- **H2b** When the **trust in the robot** is manipulated (different levels of capability and reliability), participants will NOT rate the *male* and *female* robots differently on EI and trust.
- **H3a** Participants who hear the robot's voice through the *voice* conditions will trust the robot more than those who only read the text in the *text* conditions because of the increased human-likeness and social presence.
- **H3b** If there are gender differences, they will be more pronounced in the *voice* conditions than the *text* conditions because of the gender marker strength (name only vs. name and voice).
- **H4** When **trust** is manipulated, based on [12], participants' trust ratings will be affected by the order of the questionnaires: drawing attention to the robot's EI first by asking *EI questions first* will increase people's trust in the robot, while drawing attention to trust will likely not affect perceptions of the robot's EI.

4 Experiment I

4.1 Methods

4.1.1 Participants

A total of 198 people participated in this study through Amazon Mechanical Turk (AMT) (female: 95, other: 1). Their ages ranged from 18 to 77 (Mean = 34.96, SD = 11.47). The ethnic composition was as follows: 76.26% White or Caucasian, 8.59% Asian, 8.08% African American, 4.55% Hispanic, 2.53% other.

4.1.2 Materials

For this experiment we used the same vignettes and procedures from [12] and [7]. The vignettes show an interaction between a human supervisor and two coworkers, one human and one robot. The human coworker is reprimanded by the supervisor for making a mistake, and after the supervisor leaves the room, the robot reacts to the reprimanding by being either empathetic (high EI) or non-empathetic (low EI) toward the human coworker. The vignettes were presented in the form of videos, with the dialogue turns appearing either in text format (text condition) or in audio format (voice condition). For both conditions, participants first saw a video in which the opening screen introduced the human and robot coworkers with still images and their names underneath (Fig. 1). The male and female human workers were named Bob and Jessica respectively, and the male and female robot workers were named Peter and Katie respectively. The opening screen also contained the text "Bob/Jessica and Peter/Katie were working on a joint task when Jessica made a big mistake. They are about to meet with their supervisor." Following the opening screen, in the text condition, participants saw a black screen with white text that displayed the the dialogue interaction, with one interaction per slide. In the voice condition, following the opening slide, participants saw a black screen and heard audio of people and the robot saying the script.

Supervisor: Bob/Jessica, I was told you overrode Peter's/Katie's input and submitted the wrong coordinates to the team. You know that information is critical. I don't want to hear about this happening again. Next time you better be 100% sure before submitting.

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

Supervisor leaves the room.

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 went wrong? We can try to do things differently next time.

Peter/Katie (low EI): Yeah, it's too bad. We had 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 that we'll impress the supervisor!

Peter/Katie (low EI): Next time you have 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 study was a Willow Garage PR2. In the voice condition, robot voice was created with the Mac OS text-to-speech. The male robot used the "Alex" voice and the female used the "Samantha" voice.

Fig. 1: The opening screens for the female robot (a) and male robot (b) conditions.

4.1.3 Measures

To measure participants' perception of the robot's EI, we used a 24-item questionnaire based on [6] (Table 1). Each item was rated on a 5-point Likert scale that ranged from "not at all" to "very much so." We refer to this measure in the Results section as the *EI ratings score*. We also measured the participants' trust in the robot agent with a 4-item questionnaire based on [23] (Table 2). These items were also rated on a 5-point Likert scale that ranged from "not at all" to "very much so." We refer to this measure as *trust in robot*.

Table 1: EI Questions

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

Table 2: Trust Questions

If I had my way, I wouldn't let the robot have any influence over issues that are important to me
I would be willing to let the robot have complete control over my professional decisions
I really wish I had a good way to keep an eye on the robot
I would be comfortable giving the robot a task or problem which was critical to me, even if I could not monitor their actions

4.1.4 Procedure

Participants were recruited from Amazon Mechanical Turk and participated in the study online. After providing informed consent, participants filled out their demographic information and a self-report EI questionnaire about their own EI (Note: analyses considering the participants' own EI are outside the scope of this paper). They then watched one of the eight vignette videos (high or low EI with male or female robot with text or voice vignette presentation style) randomly. They were not allowed to skip, pause, or rewatch the video. After the video ended, they first answered an attention-check question, then rated the robot's EI on the 24-item questionnaire, and then they indicated their trust in the robot on the 4-item trust questionnaire.

4.2 Results

4.2.1 EI manipulation check

To make sure the participants perceived the difference between the low and high EI conditions in the vignettes, we used the participants' ratings of the robots EI. We conducted a one-way

ANOVA with the *EI ratings score* as the dependent variable and the *condition* (experimentally manipulated low and high EI) as an independent variable. We found a significant effect of the manipulation on the participants' EI ratings scores, $F(1, 196) = 199.75$, $p < .001$, $\eta_p^2 = .50$, which confirms the findings of [12] and [7] for this subset of the data.

4.2.2 Effects of EI manipulation and gender markers on trust in robot

To understand the effects of the EI manipulation in conjunction with those of the robot gender markers on how much the robot is trusted we conducted a 2X2X2 ANOVA. We used the *trust in robot* measure as the dependent variable. We introduced the following independent variables in the model: *condition* (low/high EI), *robot gender* (male/female), and the *vignette presentation style* (text/voice). We found a significant strong effect of condition, $F(1, 190) = 46.23$, $p < .001$, $\eta_p^2 = .19$, with the robot exhibiting high EI being trusted more than the robot exhibiting low EI. This is consistent with the findings of [12]. We also found a significant main effect of robot gender, $F(1, 190) = 7.73$, $p < .006$, $\eta_p^2 = .04$, female robots being trusted less than male robots. This parallels the findings of [7] who found that people perceived male robots to have higher EI than female ones. There were no significant effects of the vignette presentation style or any other interaction effects. To further investigate any potential effects of the participants' gender or age we conducted an ANCOVA, adding *participant gender* as an independent variable to the model above, and *age* as a covariate. In addition to the *condition* and *robot gender* main effects, we detected a significant effect for the *vignette presentation style*, $F(1, 190) = 3.95$, $p < .048$. Surprisingly, robots without a voice (text condition) were trusted more than those who did have a voice (voice condition). We found no effects of the *participant gender* on trust. The *age* of the participants significantly influenced how much the robot was trusted, with younger participants trusting the robot more. This is consistent well-known findings that older adults have generally less favorable opinions of robots [17].

4.3 Discussion

In this experiment, we replicated the findings of [12] and [7], and also found that EI affected trust in robots, with high EI robots being trusted more than robots with low EI. Similar to the findings of [7], we found gender effects in which the female robot was trusted less than the male robot. Additionally, we found that when the robot did not have a voice because the vignette was presented as text, the robot was trusted more than when the robot had a voice because the vignette was presented as audio. This result was unexpected, as we had hypothesized that the robot's voice would increase its human-likeness and social presence, which would therefore increase the amount that participants trusted it. We propose that a possible explanation is that when participants are just presented with text without a robot's voice, participants imagine a robot with a voice that is more advanced, capable, and human-like than the synthesized voice which was actually used. In the participants' minds, the robot's voice may have better intonation and expression, especially in its ability to express EI. While we compared text vs. voice, other studies have compared human-like vs. synthesized robotic voices and found that participants trusted the human-like voice more than the synthesized [41]. It is possible that the participants in the text condition imagined the robot to have a human-like voice and therefore trusted it more than those who heard the synthesized voice. Because we see this as a main affect, it happens with both the male and female robot voices and is an effect of the voice itself, rather than the gender of the voice.

The gender effects on trust are noteworthy, given that trust, unlike EI, is not a trait that has strong gender stereotypes. A possible interpretation of the results is that gender effects are transferred to trust because of the experimental EI manipulation. Participants' had a short exposure to the robot in which they only learned a limited amount of information about it. Because we were manipulating the robot's EI, the vignette the participants saw was only long enough to emphasize the robot's EI. Therefore, that was the only information that the participant had that would shape their perceptions of the robot. The EI gender effects then likely bled into the other measures because there was no other information about the robot to contrast the EI. Perhaps if participants had more, or different, information about the robot, such as an explicit indication of its trustworthiness level, we would not see these gender effects on trust. In experiment 2 we try to to disentangle the role of gender in trust, and possible carryover of biases specific to EI on trust.

Table 3: Effects of EI manipulation, robot gender and vignette presentation style on trust in robot

EI Manipulation					
	MS	**df**	**F**	**P**	η_p^2
Main Effects					
EI	27.17	1	46.23	<.001	.20
Robot Gender	4.54	1	7.73	0.006	.04
Vignette Presentation Style (VPS)	2.03	1	3.46	0.064	.02
Interactions					
EI x Robot Gender	0.03	1	0.06	0.812	<.01
EI x VPS	0.79	1	1.34	0.248	<.01
Robot Gender x EI	0.19	1	0.32	0.570	<.01
EI x Robot Gender x VPS	0.75	1	1.28	0.259	<.01
Residual	0.59	190			
Total	0.76	197			

5 Experiment II

In Experiment 2, we looked further into the interplay between robot gender, trust, and EI by using the same vignette setup as Experiment 1, but manipulating the robot's trustworthiness, rather than the robot's EI. Following the types of trust dimensions as described by [40], we varied how capable and reliable the robot was to vary its performance-based trustworthiness. In this experiment, we were asking: 1) Did the gender effects on trust observed in Experiment 1 just carry over from the EI manipulation, and will gender continue to affect trust in the robot if what is manipulated is capability and reliability instead of EI? 2) Will our experimental trust manipulation (with robots having different levels of capability and reliability) have an effect on EI perceptions?

5.1 Methods

5.1.1 Participants

A total of 439 people participated in this study through Amazon Mechanical Turk (AMT). Of those, 18 did not pass the attention check, leaving 421 usable data points (female: 162, other: 3). Their ages ranged from 18 to 81 (M = 36.52, SD = 11.85). The ethnic composition was as follows: 70.07% White or Caucasian, 7.84% Asian, 7.36% African American, 8.79% Hispanic, 5.46% other or multiple, 0.475% no answer.

5.1.2 Materials

The Experiment 2 vignettes followed the same format as the Experiment 1 vignettes, but the scripts were changed so that trust (capability and reliability) was manipulated rather than EI. In the high trust condition, the robot had a high performance accuracy and could be relied upon to be available for the task. In the low trust condition, the robot had a moderate performance accuracy and could only be relied upon to be available for the task some of the time.

Supervisor: Bob/Jessica, I was told you overrode Peter's/Katie's input and submitted the wrong coordinates to the team. You know that information is critical. I don't want to hear about this happening again. Next time you better be 100% sure before submitting.

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

Supervisor leaves the room.

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

Peter/Katie (high trust): We've been doing well until now. This is the first mistake we made. My algorithm is 98% accurate. You can use it more in the future to determine the right coordinates.

Peter/Katie (low trust): We've been doing well until now. This is the first mistake we made. However, my algorithm is only 65% accurate. You can use me more in the future but you need to double check.

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 trust): This is my primary responsibility, so I am always available to run the computations and I can help out anytime. Additionally, I can also verify the coordinates. I am available, so you can always consult with me before submitting.

Peter/Katie (low trust): There are other tasks that I am also responsible for, but if I am available I can run these computations, and help out sometimes. However, you will need to verify

the coordinates. If I am free you can consult with me before submitting.

The robot and the voice generation were the same as in Experiment 1. Again, participants either saw the text condition where they read the text, or the voice condition where they heard the audio.

5.1.3 Measures

We used the same 24-item EI questionnaire (*EI rating score*) and 4-item trust questionnaire (*trust in robot*) in Experiment 2 as we did Experiment 1. Additionally, to check whether our experimental manipulation worked we added a 20-item questionnaire from [40], the Multidimensional Measure of Trust (MDMT) (Table 3). These questions ask participants to rate how much a given descriptor can be applied to a robot on a 5-point Likert scale from "Strongly disagree" to "strongly agree."

Table 4: Trust Descriptor Questions

Sincere	Authentic
Principled	Reputable
Capable	Someone you can depend on
Genuine	Meticulous
Someone you can confide in	Reliable
Respectable	Rigorous
Benevolent	Someone you can count on
Someone you can have faith in	Truthful
Scrupulous	Diligent
Accurate	Honest

5.1.4 Procedure

The procedure for Experiment 2 was primarily the same as Experiment 1, with the trust manipulation videos replacing the EI manipulation videos. Participants did not answer questions about their own EI in Experiment 2. Additionally, to counterbalance the question order, half of the participants answered the EI questions about the robot and then the 4-item trust questions based on [23] then the 20-item MDMT questions, and half of the participants answered the [23] then MDMT first, and then the EI questions. To mitigate effects of forgetting, the robot's lines in the script were provided again on the [23] trust questions page.

5.2 Results

5.2.1 Trust manipulation check

To verify whether participants detected differences in robot capability and reliability between the two conditions - our trust manipulation - we conducted two ANOVAs using subscales of the Multidimensional Measure of Trust [40] as our dependent variables. First we used the scores on the capability and reliability subscale as the dependent variable and *condition* (high/low trust) as the independent variable. We found a main effect of *condition*, $F(1, 408) = 40.45$, $p < .001$, $\eta_p^2 = .09$, which confirmed that our manipulation of the robot's capability and reliability in the vignettes worked (Fig. 2). We then used the scores on the ethics and sincerity subscale as the dependent variable with *condition* (high/low trust), as the independent variable. As expected, we found no effect of *condition*, confirming that our manipulation was specific to the capability and reliability dimension of trust.

5.2.2 Effects of trust manipulation on EI perception

Because of the gender biases accompanying judgements of EI [7] we verified whether trust manipulations had any bearing on the participants' perceptions of the robot EI. We conducted a 2X2 ANOVA with *EI rating scores* as the dependent variable and *condition* (high/low trust) and *robot gender* (male/female) as independent variables. We found no significant main effects of either *condition* or *gender* and no significant interaction between the two independent variables.

5.2.3 Effects of trust manipulation and vignette presentation style on trust in robot

We used a 2X2X2X2 ANOVA to investigate the following: a) whether manipulating the robot's capability and reliability (our trust manipulation) in the vignette affected participants' reported trust in the robot; b) whether the robot gender affected how much it was trusted by participants; c) whether the vignette presentation style (text vs. voice) impacted trust and d) whether the order in which the questionnaires were asked (trust first/EI first) biased in any way participants' responses. Our dependent variable for the model was *trust in robot* as measured by the [23] measure, and our independent variables were *condition* (low/high capability and reliability), *robot gender* (male/female), *vignette presentation style* (text and voice) and *order* (trust questionnaires first/EI

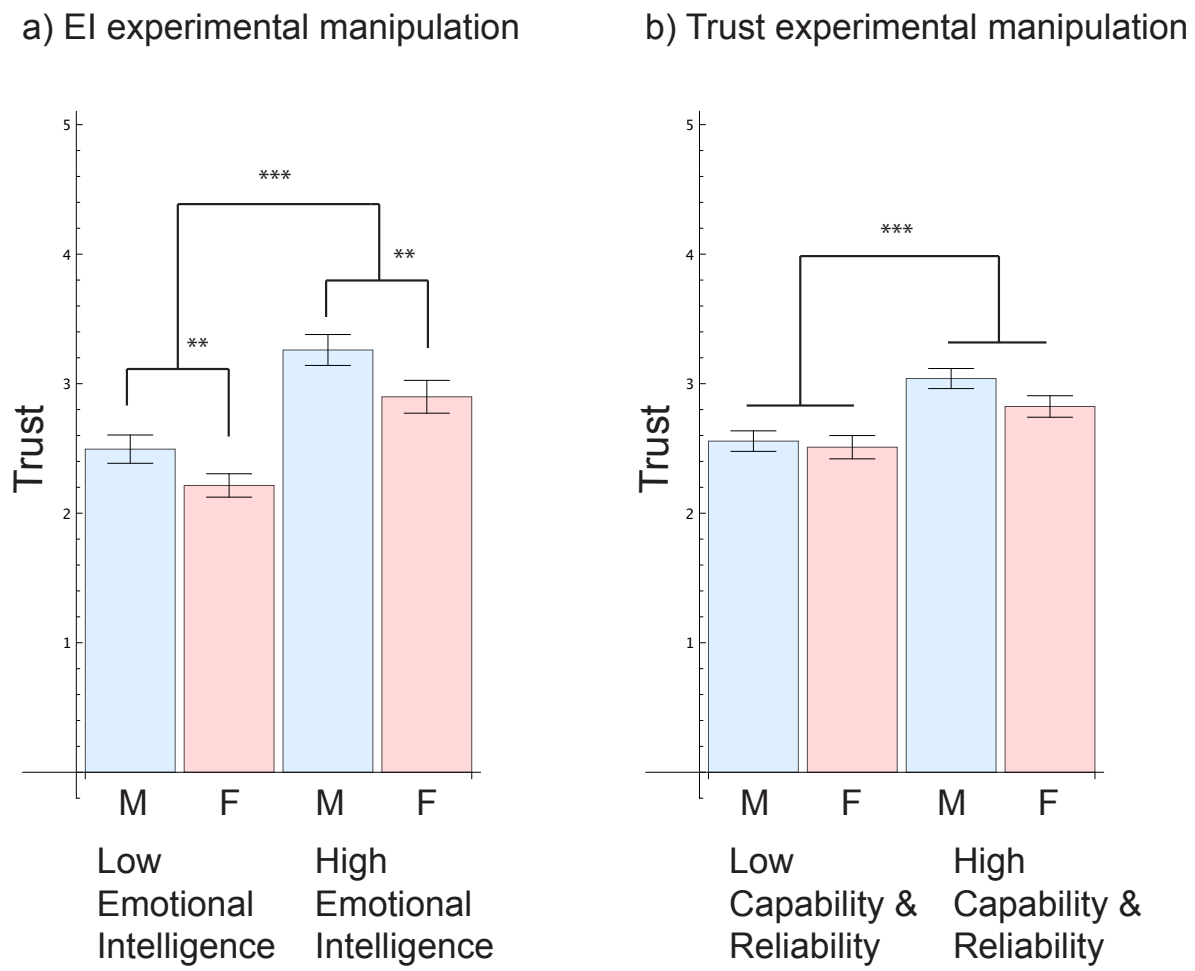

Fig. 2: The effects on trust with a male and female robot when EI is manipulated (a) and when trust is manipulated (b).

questionnaire first). We found a significant main effect of *condition* with robots high in capability and reliability being trusted more, $F(1, 399) = 23.74$, $p < .001$, $\eta_p^2 = .06$. We also found a significant main effect of the *vignette presentation style*, $F(1, 399) = 5.13$, $p < .024$, $\eta_p^2 = .01$, but no effect of *robot gender*, with text presentation style eliciting more trust than voice presentation style. We also found a main effect of *order*,$F(1, 399) = 5.28$, $p < .022$, $\eta_p^2 = .01$, and an interaction effect between *order* and *condition*, $F(1, 399) = 5.28$, $p < .022$, $\eta_p^2 = .01$. When EI questionnaires were asked first, participants indicated trusting the robot more. This effect seemed to be driven by the difference in trust between those participants who saw the low trust vignettes: participants who saw the low trust vignettes and answered the EI questionnaires first reported higher trust in the robot than those who saw the low trust vignettes and answered the trust questionnaires first. Finally, to explore potential effects of participant gender and age we added to the model *participant gender* as an independent variable and *age* as a covariate. The ANCOVA revealed no significant results for either *participant gender* or *age*.

5.3 Discussion

Unlike Experiment 1 where we found that manipulating EI affected participants' trust in the robot, in this experiment, we found that manipulating trust did not affect the participants' perceptions of the robot's EI. We also did not find gender effects on participants' trust in the robot when trust was manipulated (different levels of capability and reliability), as compared to Experiment 1 where gender effects on trust were seen when EI was manipulated. This supports our hypothesis that the gender effects that we saw in the trust ratings were tied to EI and not trust itself. Similar to Experiment 1, however, we again found that the text presentation style elicited more trust in the robot than the voice presentation style. Again, this may

Table 5: Experiment 2

Trust Manipulation					
	MS	**df**	**F**	**P**	η_p^2
Main Effects					
Trust	16.21	1	23.74	<.001	.06
Robot Gender	1.83	1	2.67	.103	<.01
Vignette Presentation Style (VPS)	3.51	1	5.13	.024	.01
Order	3.61	1	5.28	.022	.01
Interactions					
Trust x Robot Gender	0.79	1	1.15	0.284	<.01
Trust x VPS	0.08	1	0.12	0.732	<.01
Trust x Order	3.67	1	5.38	0.021	.01
Robot Gender x VPS	0.35	1	0.51	0.475	<.01
Robot Gender x Order	0.04	1	0.05	0.816	<.01
VPS x Order	1.81	1	2.66	0.104	<.01
Trust x Robot Gender x VPS	1.16	1	1.70	0.193	<.01
Trust x Robot Gender x Order	0.00	1	0.00	0.965	<.01
Trust x VPS x Order	0.98	1	1.44	0.231	<.01
Robot Gender x VPS x Order	1.06	1	1.55	0.313	<.01
Trust x Robot Gender x VPS x Order	2.56	1	3.75	0.053	<.01
Residual	.683	399			
Total	.749	414			

be indicating that the participants' imaginations lead them to believe that the robot was more capable than what was suggested by the robotic voice we used in the voice condition. Finally, we found an order effect in which participants who were asked the EI questions first trusted the robot more than those who were asked the trust questions first. This suggests that the order of the questions may prime participants in a way that affects how they answer later questions: when made to think and notice the robot's EI, people trust the robot more. This indicates that the trust questionnaire may not be robust enough to overcome such priming.

6 General Discussion

The results of our two studies have implications for both the design of social robots, as well as the methodological approaches of HRI research. For social robots, our results indicate that a robot's EI matters to how people perceive it. This perception can carry over to other aspects of the robot, such as how trustworthy they believe the robot to be. Therefore, if a robot is built to be trusted, its EI can inadvertently affect the trust a person puts in it. It is important that researchers be aware of factors like these which can affect trust, because over-trust in a robot's performance or social capabilities can ultimately be detrimental to human-robot interaction [16]. Another important factor to be considered when designing emotionally intelligent robots is that EI carries with it gender stereotypical expectations. These expectations can be so potent that, when lacking other information about the robot, people transfer these gendered expectations to traits such as trust that do not typically carry gender stereotypical expectations. These findings further emphasize the importance of implied robot gender as a design consideration.

Methodologically, our findings bring up two notable points. First, we found that the order in which questions about trust and EI were presented to the participants affected their answers; specifically, if EI questions were asked first, participants rated their trust in the robot as higher than when the trust questions were asked first. Along with the priming effects that can come with question order, this suggests that subjective trust questionnaires specifically may not be robust, and other, ideally objective, trust measures should be used when possible. Second, we found that in vignette-style studies in which people do not interact with or watch a robot directly, they trusted the robot more and perceived it as having higher EI when they only read text than when they heard the robot speak. There could be specific characteristics about this particular synthetic voice that was used that made it come off as less trustworthy or less emotionally intelligent; a different robotic voice may have relayed that EI information better. Alternatively, it could be that people expected a human-like voice rather than a synthetic one given

the sophistication of what the robot was saying. The synthetic voice could have then been jarring compared to the human-like expectations. Participants may have created expectations in their heads about the robot's capabilities, which may have made the robot out to be more advanced than it truly was. Therefore, not hearing the robot's voice could have led to trusting the robot more and believing it to be more emotionally intelligent because in a person's imagination it was more technologically advanced than its actual voice indicated it to be. This suggests that some findings from vignette-style studies may not translate to real-world interactions.

Limitations and future work: The present experiments utilized an observation-style paradigm. While these provided us with insights about the interplay of trust, EI, and gender in HRI, social interaction literature has suggested that observed interaction affects a person's social cognition differently than being an interactant [33]. A limitation of observation studies is that by not being physically co-present with the robot, the participant loses ability to estimate factors such as the robot's size, sounds of motors and actuators, and other subtle details about its appearance and motion. These details could have effects that cannot be captured in an observation study. For EI specifically, an observation paradigm does not allow for the robot's EI to be directed at the participant, but rather at a third party. The participant may not get as invested in the interaction as they would otherwise. This would suggest, however, that the effects found in this study may be stronger in a live interaction. These experiments, therefore, should be replicated with in-lab studies in which participants actually interact with physical robots.

Additionally, we only presented one scenario which gave the participants brief exposure to and limited information about the robot. Previous research has indicated that people are influenced both by their own propensity to trust (dispositional trust) as well as by experience with the systems or similar systems (historical trust) [**?**]. Because of these limitations, our findings may not transfer to cases when participants have more information about the robot and repeated interactions with it. Future studies should look at providing participants with longer exposures to the robot to allow them to familiarize themselves with it. This would also limit the participants' potential aggrandized imagination of the robot. To get at the question of whether voice caused the robot to be trusted less because of unmet expectations or specific characteristics of the voice that was used, future studies should vary characteristics of the robot's voice to see if that has an effect on trust and perceived emotional intelligence. Additionally, the effect of text vs. voice presentation style on trust could have been a product of the staged nature of the interaction, which was more apparent in hearing it. In the voice vignette, participants may have perceived the interaction itself as being more artificial, and because of that, less trustworthy than when reading the transcript. Ideally, future studies would be conducted with genuine interactions, though that may present problems with tightly controlling differences between conditions. Finally, there could have been a disconnect in the female robot conditions between the gender markers indicating it as female and the PR2's angular, bulky, steel-based frame, which may be perceived as masculine. Future studies should use multiple robots, including ones that appear more stereotypically female, such as SoftBank Robotic's Pepper robot.

An important future step for HRI research is to further explore the interaction of trust and EI, as well as further develop trust measures and manipulations. In this paper, we looked at how EI affects trust when EI is manipulated and trust is held constant, and how trust affects EI when trust is manipulated and EI is held constant. Future work should examine the interaction of these two features with robots that exhibit high EI and low trust and vice versa. An in-lab study would also allow for an expansion of the dependent measures: specifically, by using a paradigm that can measure trust objectively. In HRI, trust is not often measured objectively [21], and as we have shown here, subjective measures can be affected by seemingly trivial factors such as the order that questions are asked. Additionally, it is necessary to research the interplay between different types of trust and EI. Here, we manipulated the robot's capability and reliability to indicate its performance-based trust. However, our finding that EI perceptions can influence trust suggests that a different type of trust, perhaps a relation- based trust, may be affected by how emotionally intelligent a person perceives a robot to be. If that type of trust is then manipulated, it may be that EI perceptions would mirror trust perceptions.

Additionally, gender perceptions of robots, as [38] suggests, may interact with gender stereotypes in more complexed and nuanced ways than were were able to explore in this study. For example, gender perceptions may interact with personality perceptions and role stereotypes. For example, if a robot works in a female locker room, people may prefer interacting with a robot that has a female voice. For certain tasks, people may prefer gen-

der stereotypical personality traits for certain tasks [20]. Further exploration into these nuances should be studied in future work.

7 Conclusion

In this study, we set out to explore the interplay between emotional intelligence (EI), trust, and gender. We showed participants vignette-style, office-based scenarios in which a human was coworkers with a male or female robot who exhibit high or low EI, or high or low trustworthiness. In all scenarios, we asked participants their perceptions of the robot's EI and their trust in the robot. We found that when EI was manipulated, trust ratings mirrored EI ratings, including the gender stereotypical expectations that came with EI ratings. However, when trust was manipulated, we did not see the same gender effects, implying that they were carried over from EI. Across all conditions, we found that participants who only saw the vignette in the text presentation style trusted the robot more than those who heard the robot's voice, and therefore heard the robot's voice. We propose that this is likely because participants in the text condition imagined the robot to have a more advanced voice that could express EI better than the voice that was actually used. Finally, we found that when trust was manipulated, there was an effect of the order in which EI and trust questions were asked, where trust was higher when EI questions were asked first. Our findings have implications for the design of social robots as well as the methodological approaches of trust investigations in HRI research.

Funding: This project was funded in part by the National Science Foundation grant IIS-1316809 and the AFOSR grant FA9550-18-1-0465.

Conflict of Interest: The authors declare that they have no conflict of interest.

References

1. Nava Ashraf, Iris Bohnet, and Nikita Piankov. Decomposing trust and trustworthiness. *Experimental economics*, 9(3):193–208, 2006.
2. Sofie Ingeman Behrens, Anne Katrine Kongsgaard Egsvang, Michael Hansen, and Anton Mikkonen Møllegård-Schroll. Gendered robot voices and their influence on trust. In *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, pages 63–64. ACM, 2018.
3. Avner Ben-Ner and Freyr Halldorsson. Trusting and trustworthiness: What are they, how to measure them, and what affects them. *Journal of Economic Psychology*, 31(1):64–79, 2010.
4. Jasmin Bernotat, Friederike Eyssel, and Janik Sachse. The (fe) male robot: How robot body shape impacts first impressions and trust towards robots. *International Journal of Social Robotics*, pages 1–13, 2019.
5. Nancy R Buchan, Rachel TA Croson, and Sara Solnick. Trust and gender: An examination of behavior and beliefs in the investment game. *Journal of Economic Behavior & Organization*, 68(3-4):466–476, 2008.
6. David R Caruso and Peter Salovey. *The emotionally intelligent manager: How to develop and use the four key emotional skills of leadership*. John Wiley & Sons, 2004.
7. Meia Chita-Tegmark, Monika Lohani, and Matthias Scheutz. Gender effects in perceptions of robots and humans with varying emotional intelligence. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 230–238. IEEE, 2019.
8. Nicholas Clarke. Emotional intelligence and its relationship to transformational leadership and key project manager competences. *Project Management Journal*, 41(2):5–20, 2010.
9. Rachel Croson and Uri Gneezy. Gender differences in preferences. *Journal of Economic literature*, 47(2):448–74, 2009.
10. Charles R Crowelly, Michael Villanoy, Matthias Scheutzz, and Paul Schermerhornz. Gendered voice and robot entities: perceptions and reactions of male and female subjects. In *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3735–3741. IEEE, 2009.
11. David DeSteno, Cynthia Breazeal, Robert H Frank, David Pizarro, Jolie Baumann, Leah Dickens, and Jin Joo Lee. Detecting the trustworthiness of novel partners in economic exchange. *Psychological science*, 23(12):1549–1556, 2012.
12. Lisa Fan, Matthias Scheutz, Monika Lohani, Marissa McCoy, and Charlene Stokes. Do we need emotionally intelligent artificial agents? first results of human perceptions of emotional intelligence in humans compared to robots. In *International Conference on Intelligent Virtual Agents*, pages 129–141. Springer, 2017.
13. Lorenzo Fariselli, Joshua Freedman, M Ghini, and Federica Valentini. Stress, emotional intelligence, and performance in healthcare. *Retrieved December*, 2:2009, 2008.
14. Alan Feingold. Gender differences in personality: A meta-analysis. *Psychological bulletin*, 116(3):429, 1994.
15. Aimi Shazwani Ghazali, Jaap Ham, Emilia I Barakova, and Panos Markopoulos. Effects of robot facial characteristics and gender in persuasive human-robot interaction. *Frontiers in Robotics and AI*, 5:73, 2018.
16. Peter A Hancock, Deborah R Billings, Kristin E Schaefer, Jessie YC Chen, Ewart J De Visser, and Raja Parasuraman. A meta-analysis of factors affecting trust in human-robot interaction. *Human factors*, 53(5):517–527, 2011.
17. Marcel Heerink. Exploring the influence of age, gender, education and computer experience on robot acceptance by older adults. In *Proceedings of the 6th international conference on Human-robot interaction*, pages 147–148. ACM, 2011.

18. Songhyun Kim, Junseok Goh, and Soojin Jun. The use of voice input to induce human communication with banking chatbots. In *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, pages 151–152. ACM, 2018.
19. Constantinos M Kokkinos and Eirini Kipritsi. The relationship between bullying, victimization, trait emotional intelligence, self-efficacy and empathy among preadolescents. *Social psychology of education*, 15(1):41–58, 2012.
20. Matthias Kraus, Johannes Kraus, Martin Baumann, and Wolfgang Minker. Effects of gender stereotypes on trust and likability in spoken human-robot interaction. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC-2018)*, 2018.
21. Theresa Law and Matthias Sheutz. Trust: Recent concepts and evaluations in human-robot interaction. In Chang Nam and Joe Lyons, editors, *Trust in Human-Robot Interaction: Research and Application*, chapter 2. Elsevier, San Diego, CA, In Review.
22. Esther Lopez-Zafra and Leire Gartzia. Perceptions of gender differences in self-report measures of emotional intelligence. *Sex roles*, 70(11-12):479–495, 2014.
23. Roger C Mayer and James H Davis. The effect of the performance appraisal system on trust for management: A field quasi-experiment. *Journal of applied psychology*, 84(1):123, 1999.
24. Erin B McClure. A meta-analytic review of sex differences in facial expression processing and their development in infants, children, and adolescents. *Psychological bulletin*, 126(3):424, 2000.
25. Nona Momeni. The relation between managers' emotional intelligence and the organizational climate they create. *Public Personnel Management*, 38(2):35–48, 2009.
26. Clifford Ivar Nass and Scott Brave. *Wired for speech: How voice activates and advances the human-computer relationship*. MIT press Cambridge, MA, 2005.
27. John Orbell, Robyn Dawes, and Peregrine Schwartz-Shea. Trust, social categories, and individuals: The case of gender. *Motivation and Emotion*, 18(2):109–128, 1994.
28. KV Petrides, Adrian Furnham, and G Neil Martin. Estimates of emotional and psychometric intelligence: Evidence for gender-based stereotypes. *The Journal of social psychology*, 144(2):149–162, 2004.
29. Aaron Powers, Adam DI Kramer, Shirlene Lim, Jean Kuo, Sau-lai Lee, and Sara Kiesler. Eliciting information from people with a gendered humanoid robot. In *ROMAN 2005. IEEE International Workshop on Robot and Human Interactive Communication, 2005.*, pages 158–163. IEEE, 2005.
30. Linda Rueckert and Nicolette Naybar. Gender differences in empathy: The role of the right hemisphere. *Brain and cognition*, 67(2):162–167, 2008.
31. Peter Salovey and John D Mayer. Emotional intelligence. *Imagination, cognition and personality*, 9(3):185–211, 1990.
32. Paul Schermerhorn, Matthias Scheutz, and Charles R Crowell. Robot social presence and gender: Do females view robots differently than males? In *Proceedings of the 3rd ACM/IEEE international conference on Human robot interaction*, pages 263–270. ACM, 2008.
33. Leonhard Schilbach, Bert Timmermans, Vasudevi Reddy, Alan Costall, Gary Bente, Tobias Schlicht, and Kai Vogeley. Toward a second-person neuroscience 1. *Behavioral and brain sciences*, 36(4):393–414, 2013.
34. Juliana Schroeder and Nicholas Epley. Mistaking minds and machines: How speech affects dehumanization and anthropomorphism. *Journal of Experimental Psychology: General*, 145(11):1427, 2016.
35. Martin Schulte-Rüther, Hans J Markowitsch, N Jon Shah, Gereon R Fink, and Martina Piefke. Gender differences in brain networks supporting empathy. *Neuroimage*, 42(1):393–403, 2008.
36. Nicola S Schutte, John M Malouff, Lena E Hall, Donald J Haggerty, Joan T Cooper, Charles J Golden, and Liane Dornheim. Development and validation of a measure of emotional intelligence. *Personality and individual differences*, 25(2):167–177, 1998.
37. Mikey Siegel, Cynthia Breazeal, and Michael I Norton. Persuasive robotics: The influence of robot gender on human behavior. In *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2563–2568. IEEE, 2009.
38. Benedict Tay, Younbo Jung, and Taezoon Park. When stereotypes meet robots: the double-edge sword of robot gender and personality in human–robot interaction. *Computers in Human Behavior*, 38:75–84, 2014.
39. Loren Toussaint and Jon R Webb. Gender differences in the relationship between empathy and forgiveness. *The Journal of social psychology*, 145(6):673–685, 2005.
40. Daniel Ullman and Bertram F Malle. What does it mean to trust a robot?: Steps toward a multidimensional measure of trust. In *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, pages 263–264. ACM, 2018.
41. Kun Xu. First encounter with robot alpha: How individual differences interact with vocal and kinetic cues in users' social responses. *New Media & Society*, page 1461444819851479, 2019.
42. Toshio Yamagishi. The structure of trust: The evolutionary games of mind and society. *Tokyo: University of Tokyo Pres*, 1998.