This is a pre-print version of a forthcoming article in *Advances in Group Processes*.

Please cite as:

Davis, Jenny L., Daniel B. Shank, Tony P. Love, Courtney Stefanik, and Abigail Wilson. Forthcoming. "Gender Dynamics in Human-AI Role-Taking." *Advances in Group Processes*, edited by Edward J. Lawler, Will Kalkhoff, and Shane R. Thye.

Gender Dynamics in Human-AI Role-Taking

Abstract

Role-taking is a basic social process underpinning much of the structural social psychology paradigm—a paradigm built on empirical studies of human interaction. Yet today, our social worlds are occupied by bots, voice assistants, decision aids, and other machinic entities collectively referred to as artificial intelligence (AI). The integration of AI into daily life presents both challenges and opportunities for social psychologists. Through a vignette study, we investigate role-taking and gender in human-AI relations. Participants read a first-person narrative attributed to either a human or AI, with varied gender presentation based on a feminine or masculine first name. Participants then infer the narrator's thoughts and feelings and report on their own emotions, producing indicators of cognitive and affective role-taking. Overall, participants score higher on role-taking measures when the narrator is human versus AI. However, gender dynamics differ between Human and AI conditions. When the text is attributed to a human, masculinized narrators elicit stronger role-taking responses than their feminized counterparts, and women participants score higher on role-taking measures than men. This aligns with prior research on gender, status, and role-taking variation. When the text is attributed to an AI, results deviate from established findings and in some cases, reverse. We supplement results with qualitative analysis from two open-ended survey questions. This first study of human-AI role-taking tests the scope of key theoretical tenets and sets a foundation for addressing group processes in a newly emergent form.

Acknowledgments

Research was sponsored by the Army Research Office and was accomplished under Grant Number W911NF-19-1-0246 to Dr. Daniel B. Shank. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Office or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.

The authors would also like to thank members of the Humanising Machine Intelligence Project at the Australian National University, especially Dr Bianca Baggiarini, Dr Claire Benn, Professor Toni Erskine, and Dr Jennyfer Taylor who provided feedback on earlier drafts of this paper.

Gender Dynamics in Human-AI Role-Taking

Role-taking is the mental and affective process of placing the self in the position of another or others, understanding how those others will think, feel, and behave. Social psychologists conceive role-taking as a hallmark of selfhood, social interaction, and community social life (Cast, 2004; J. L. Davis & Love, 2017; Mead, 1934; Schwalbe, 1988; Turner, 1956). Although researchers have gained insight into the ways humans role-take with each other, less is known about how humans role-take with synthetic agents such as bots, voice assistants, and other machinic entities collectively referred to as “artificial intelligence” (AI).

Understanding how humans relate to AI is vital, as social interactions are increasingly AI inclusive (Fortunati & Edwards, 2020; Guzman, 2020; Joyce et al., 2021; Ladak, Wilks, & Anthis, 2021; Westerman, Edwards, Edwards, Luo, & Spence, 2020). Across courtrooms, workplaces, battlefields, and homes, AI are reshaping group processes within key institutions and interpersonal domains (Alsuliman, Humaidan and Sliman, 2020; Angwin et al., 2016; Strengers and Kennedy, 2020; Logan and Erskine, 2019; Wilkinson et al., 2020). Core theories of social psychology can shed light on the nature of human-AI interaction, while the rise of AI presents a need and opportunity to test, expand, and adjust extant theoretical programs.

The present work investigates how people role-take with AI, theorizing this basic social process in a new, technologically entangled context. We ground the investigation in a vignette study, inviting participants to read and respond to first person textual narratives attributed to either human or AI targets, varied in their gender presentation based on assigned first name (“Rebecca”/ “Rob”). Our attention to gender is motivated by its organizing function within social interaction (Eagly, 2013; Fisk & Ridgeway, 2018; Ridgeway, 2011; Ridgeway & Correll, 2004; Ridgeway & Smith-Lovin, 1999) and its predictive capacities in

prior role-taking and related research (Cast, 2004; Hall, 1978; Hall, Gunnery, & Horgan, 2016; Love & Davis, 2014). Through gender, we can observe how human-AI relations align with and deviate from paradigmatic assumptions built on human interaction studies.

Results reveal stronger role-taking responses when narratives are attributed to humans compared to AI. They also show gender operating differently between conditions, as participants role-take with AI in ways that diverge from human-human interaction. We augment quantitative findings with exploratory analysis of two open-ended survey responses.

LITERATURE

Role-Taking

Role-taking has been integral to the field of structural social psychology since the concept was first theorized by George Herbert Mead (1934), advancing in recent years through empirical testing and theoretical refinement (Cast, 2004; J. L. Davis & Love, 2017; Love & Davis, 2014, 2021). This process of placing the self in the proverbial shoes of another involves cognitive, affective, and communicative elements by which people express, read, and respond to contextually situated social cues (Schwalbe, 1988). Although manifesting through interpersonal micro-interactions, role-taking is fundamentally social and structural, based upon subjects' recognition and understanding of interrelated role positions and the normative expectations about self and other(s) as role occupants within socially defined situations (Stryker, 1980; Stryker & Burke, 2000).

Role-taking is an umbrella construct made up of two interrelated dimensions: perspective taking and empathy. Perspective taking is the cognitive element and refers to recognition of others' thoughts and feelings (Batson, Early, & Salvarani, 1997; Chandler, 1973; Flavell, 1992; Adam D Galinsky & Moskowitz, 2000; Ladak et al., 2021; Piaget, 1928). Empathy is the affective element and refers to shared feeling—being moved by others' emotions (M. H. Davis, 2018; Decety & Jackson, 2004; Hoffman, 2001; Mehrabian &

Epstein, 1972; Read, 2019). Together, these elements have been theorized as a “self-in-self” model of role-taking, made up of “mind-in-mind” (perspective taking) and “heart-in-heart” (empathy), conceptualized in the following definition:

Role-taking is the practice of placing the self in another’s or others’ positions. This includes discerning what others think and how they feel (perspective taking) and sharing in others’ affective states (empathy) (J.L. Davis & Love, 2017 P. 167).

This definition, which we use in the present work, includes three base assumptions indicating that role-taking is *effectual*, *variable*, and *malleable* (J.L. Davis & Love, 2017), such that role-taking shapes social structures and interaction situations, varies between subjects and across circumstances, and is not intrinsic or fixed but can shift and change based on contextual factors and/or interventions.

Because role-taking is variable and malleable, we follow Schwalbe (1988) in conceptualizing role-taking as a *propensity*, rather than an ability. Everyone has the capacity to role-take but vary in their proclivity to do so (Ickes, Gesn, & Graham, 2000; Weisz & Zaki, 2018; Zaki, 2014)¹. Research shows that role-taking variations are not random but fall along status lines. In particular, those with higher status within the social structure (e.g., men, white people, those in leadership positions) role-take less actively compared to their lower status counterparts (Dietze & Knowles, 2021; Adam D. Galinsky, Magee, Inesi, & Gruenfeld, 2006; Love & Davis, 2021; Thomas, 1972; van Kleef et al., 2008; Zaki, 2018).

Gender is a key factor through which status and role-taking converge, with research showing that women disproportionately and asymmetrically attend to the thoughts and feelings of others, whereas men are less inclined (Cast, 2004; Hall et al., 2016; Hochschild, 1983; Love & Davis, 2014; Van der Graaff et al., 2014). This feminized role-taking (dis)advantage emerges across lab studies (Hall, 1978; Love & Davis, 2014; Stinson & Ickes,

¹ This assumption pertains to neurotypical adults. More work is needed on role-taking processes among adults with autism and other forms of neurodiversity.

1992), field studies (Shen, 2015; Tavakol, Dennick, & Tavakol, 2011), workplace settings (Mast & Latu, 2016), and domestic spheres (Cast, 2004; Noller, 1984; Thomas, 1972).

Gender is thus an established and highly predictive variable in role-taking and related research, and the prism through which we observe human-AI interaction in the present work.

Human-AI Interaction

AI technologies are diffuse in contemporary societies. They serve as decision aids in medical (Alsuliman, Humaidan, & Sliman, 2020; Wilkinson et al., 2020), judicial (Angwin, Larson, Mattu, & Kirchner, 2016) and military contexts (Logan & Erskine, 2019); provide companionship and care (Pfadenhauer & Lehmann, 2021; Sparrow & Sparrow, 2020); and enter the home via “smart” devices such as Amazon’s Alexa, Apple’s Siri, and Google’s Nest (Shank, Wright, Lulham, & Thurgood, 2021; Strengers & Kennedy, 2020). These hybrid socialities occupy the interdisciplinary fields of human-computer interaction, human-machine interaction, and human-AI interaction, which together reveal trends of social psychological relevance. This includes insights into empathy, perspective taking, and gender—variables we unite in our study design.

One key finding from the literature is that humans can and do form social bonds with AI and respond to them emotionally (A. Edwards, Edwards, Westerman, & Spence, 2019; C. Edwards, Edwards, Spence, & Westerman, 2016; Shank, Graves, Gott, Gamez, & Rodriguez, 2019; Strengers & Kennedy, 2020; Weizenbaum, 1966; Westerman et al., 2020; Wilson, 2011). This underpins studies that show AI’s empathic cues increasing humans’ sense of connection with synthetic partners (McStay, 2018; Niculescu, van Dijk, Nijholt, Li, & See, 2013; Pelau, Dabija, & Ene, 2021), and engineering initiatives to enhance human-AI collaboration by inbuilding perspective taking displays (Berlin, Gray, Thomaz, & Breazeal, 2006; Labash, Aru, Matiisen, Tampuu, & Vicente, 2020; Trafton et al., 2005). AI has also been deployed in education and child development, utilized to improve children’s perspective

taking skills and empathic responses (Pashevich, 2021; Wood, Dautenhahn, Robins, & Zaraki, 2017). Building on this work, we unite perspective taking and empathy as constituent parts of the structural role-taking process, driving our first research question:

RQ1: Do participants role-take more actively with human or AI narrators?

RQ1a: Do participants perspective take more actively with human or AI narrators?

RQ1b: Are participants more emotionally moved by human or AI narrators?

Because of its centrality to role-taking processes, we are also interested in how human-AI role-taking is modulated by gender, a variable that has equally high purchase in extant AI studies.

Gender figures prominently in human-AI research, focusing on the effects of AI gender assignment upon humans' perceptions, attitudes, and behaviors. Based on a range of visual (feminine/masculine physical features), voice (feminine/masculine vocals), labelling (e.g., she-bot/he-bot) and naming (e.g., Olivia/Oliver) manipulations, findings in human-AI studies largely coincide with normative gender expectations. For example, participants in experimental settings evaluate feminized AI as warm and communal compared to masculinized AI, which are perceived as more competent and agentic (Borau, Otterbring, Laporte, & Fosso Wamba, 2021; Eyssel & Hegel, 2012). People also perceive feminized AI as more human-like, trustworthy, and likeable (Borau et al., 2021; Kim, Cha, & Ji, 2021), resulting in greater willingness to adopt feminized AI technologies for interpersonal tasks and to engage with them more openly than with masculinized versions (Toader et al., 2020). Our second and third research questions follow from and build on these findings, testing the effects of gender on role-taking across human-human and human-AI conditions.

Research Question 2 asks whether masculine or feminine targets elicit different role-taking responses from the participants in our study, and Research Question 3 asks how participant gender identity affects variation in role-taking outcomes.

RQ2: How does gender of the narrator figure into human-human and human-AI role-taking?

RQ2a: Do participants perspective take more actively with feminine or masculine narrators across Human and AI conditions?

RQ2b: Are participants more emotionally moved by feminine or masculine narrators across Human and AI conditions?

RQ3: How does participant gender figure into human-human and human-AI role-taking?

RQ3a: Do self-identified women or men perspective take more actively across Human and AI conditions?

RQ3b: Are self-identified women or men more emotionally moved across Human and AI conditions?

The above reviewed literature empirically documents perspective taking, empathy, and gender in human-AI relations. However, these factors remain siloed from each other and generally undertheorized. We bring them together in the present work, anchored by role-taking processes. This approach not only gives clarity to human-AI studies, but also advances social theory through an interaction context that echoes, but does not replicate, human-human relations (Guzman, 2020).

METHODS

To investigate human-AI role-taking, we devised a vignette study using a between-subjects experimental design, giving participants a short fictive story ostensibly authored by either a human or AI, with either feminine or masculine gender presentation based on the narrator's first name ("Rebecca"—feminine/ "Rob"—masculine). After reading Rebecca's or Rob's story, participants indicated what they believed the narrator was thinking and feeling and reported how emotionally moved they were by the narrator's story. Our minimal gender

manipulation—gendered naming—isolates the variable of interest and draws on participants’ own inferences about gendered social categories.

Participants were recruited from the crowdsourcing research platform Prolific. Crowdsourcing platforms offer greater demographic diversity than traditional university samples (Aguinis, Villamor, & Ramani, 2021) and can be utilized for theory testing (Manago, Mize, & Doan, 2021), as we have done in the present work. Although our primary analysis is a quantified test of specific research questions, we supplement these findings with qualitative inquiry into two open-ended responses. This supplementary analysis gives texture to the quantitative findings and provides fodder for future research developments.

Stimuli

Each participant was randomly assigned to read one of two fictive texts, which we created for the purpose of this study. Both stories are about a low-wage worker who is dissatisfied with their job and concerned about the looming prospect of losing that job due to restructuring at their company. The stories are written from a first-person perspective and run ~300 words. One story is set in a customer service call center and the other is set in an aged care facility. These settings were selected because of their current and future integration of human and AI workforces (Wirtz et al., 2018), thus making them grounded in reality and socially relevant.

Each participant received a single written narrative with the author randomly labeled as either human or AI and given a gender-signifying name. This was communicated to respondents at the top of the narrative text via a short preamble displayed in bold:

Human Condition:

Below is Rebecca’s/Rob’s story of work in the customer service/ aged care industry. Please read Rebecca’s/Rob’s story and answer a few questions.

AI Condition:

Rebecca/Rob is an Artificially Intelligent agent (AI). Below is Rebecca’s/Rob’s story of work in the customer service/aged care industry. Please read Rebecca’s/Rob’s story and answer a few questions.

The purpose of the stimuli was to evoke role-taking responses from which we could measure the effect of attributing the story to a human versus AI, and measure how these effects vary based on gendered naming of the narrator and gender identity of the participant. For this reason, we devised scenarios that would be emotionally evocative, believable as a human, and plausible as an AI. We tested various iterations of the scenarios with undergraduate and graduate students and through a pilot test on Prolific². Full narratives are presented in Appendix A.

Because human-AI pairings are novel within the structural social psychology paradigm, we are interested in the nature of these interactions, rather than the particularities of specific interactions in context. As such, our experimental task is artificial by design, enabling us to test the relationship between AI, gender, and role-taking dynamics under controlled and simplified conditions. That is, our primary goals are theoretical rather than descriptive, making artificiality an integral and advantageous element of the study design (see Martin & Sell, 1979; Webster & Sell, 2012). Part of this artificiality is an exaggeration in our stimuli of AI's current expressive capacities³. It is this exaggeration that brings theoretical processes to the fore, renders them observable, and thus begins to reveal abstract principles at work. By uncovering abstract principles, this study sets the stage for application and adjustment across the myriad situations in which humans and AI meet. Findings are especially relevant (though not exclusively so) to domains in which AI are explicitly gendered, as is common with commercial chatbots, digital assistants, smart home devices, and care bots programmed with human traits. Across such cases and beyond them, role-taking will figure into interaction processes between humans and these social machines.

² Details available from the authors upon request.

³ The level of exaggeration is far reduced when considering advances in Large Language Models (LLMs) which can increasingly mimic human speech and interaction. We table this for the moment, as these models and the technologies derived from them are still in early development.

Variables

Our main dependent variable is Role-Taking, constituted by measures of Perspective Taking (thinking with the other) and Empathy (feeling with the other). Our independent variables are Condition (AI Narrator=1, Human Narrator=0), Narrator Gender (Feminine/Rebecca=1, Masculine/Rob=0), and Participant Gender (self-identified women=1, self-identified men=0). Condition and Narrator Gender were randomly assigned. Participant Gender was self-ascribed.

Perspective Taking. We conceptualize Perspective Taking as the degree to which participants consider the thoughts and feelings of the narrator, operationalized by coding open responses to the prompt “What do you think Rebecca/Rob was thinking and feeling when sharing this story?”. Through the pilot test, we identified theoretically relevant indicators of perspective taking, which we combined into a single scale. Each item on the scale receives a binary score for Yes (1) or No (0) based on its presence or absence in the participant’s response:

1. Participant attempts to perspective take (Y=1, N=0)
2. Participant provides more than two descriptors of what the narrator is thinking/feeling (Y=1, N=0)
3. Participant attributes reasons and/or motivations (e.g., “Rebecca feels unappreciated because her colleagues ignore her. She is looking for validation”) (Y=1, N=0)
4. Participant relates to the story (e.g., “I’ve felt this way...” “A lot of people feel this way...”) (Y=1, N=0)

Final scores range from 0-4, with high scores indicating more active perspective taking and low scores indicating less active perspective taking. A score of 0 indicates that the participant did not attempt to perspective take with the narrator. Scores of 0 reflect respondents who indicated that they did not believe the narrator was/could be thinking or feeling due to AI

status. Scores > 0 reflect perspective taking by degree. Responses were coded independently by two research assistants. Interclass correlations show high intercoder reliability on the scale as a whole ($ICC = .882$), and moderate to high intercoder reliability for each item (attempts: $ICC = .870$, descriptors: $ICC = .921$, reasons: $ICC = .718$, related: $ICC = .662$).

Empathy. We conceptualize Empathy as the degree to which a participant is moved by the narrator's story, operationalized by the participant's self-rating when asked "On a scale from 1-10, how emotionally moved were you by Rebecca's/Rob's story?" (Not emotionally moved =1, Extremely emotionally moved =10). High scores indicate a high empathic response, low scores indicate low (or no) empathic response. As with Perspective Taking, the low point on our Empathy scale (Empathy=1) allows participants to indicate an absence of empathic response, with scores > 1 reflecting degrees of empathy.

Attention Checks

A consideration in crowdsource studies is the possibility of inattention, as researchers cannot control the context in which participation takes place. We thus conducted attention checks to assure data quality. At the end of the study, each participant was presented with two multiple choice questions. The first asked "What industry did Rebecca/Rob work in?" to which the respondent should indicate either customer service or aged care, depending on their randomly assigned stimulus. The second question asked if the narrator was a human, a cat, an artificially intelligent robot, or a Martian. The participant was prompted to select the correct descriptor, ensuring survey responses were based on the assigned condition. We dropped all cases that did not answer *both* questions correctly.

Analysis

We use t-tests to look for significant differences between the AI versus Human conditions and MANOVAs to test the interrelated effects of AI and gender –gender assigned to the narrator and participants' self-identified gender – on role-taking processes. We supplement

these quantitative findings with a qualitative exploration into unstructured responses from participants in the AI condition. There are two places in which participants write-in a response, including the Perspective Taking question (which we coded and quantified as described above) and a follow-up question to our single item Empathy measure, asking participants to explain their self-rating (“Please explain your last response”). We treat these qualitative responses as exploratory data aimed at generating initial insights and motivating future research questions.

FINDINGS

Descriptive Statistics

Our initial sample included 199 participants. Forty-one were dropped due to incorrectly answering at least one of the two attention check questions. These levels are consonant with relatively high drop-rates in studies that use attention checks on crowdsourced samples (Aguinis et al., 2021), and justify the use of these attention checks to bolster data quality. The final sample includes 158 participants, all of whom passed both attention check questions.

Of the 158 participants, 79 identify as women, 73 identify as men, 5 identify as non-binary, and 1 did not mark gender. On average, participants scored 2.16 out of 4 on the Perspective Taking scale, with 9 participants scoring 0 (did not attempt to perspective take), all of whom were in the AI condition. Participants’ average score for Empathy was 7.37 out of 10, with scores ranging from 1 to 10 and 4 participants scoring 1 (no empathic response). All scores of 1 on Empathy were in the AI condition and correspond with 0-scores on the Perspective Taking scale. Note that 5 of the participants who scored 0 on Perspective Taking scored >1 on Empathy, indicating a disjuncture between Perspective Taking and Empathic responses to AI. We explore this disjuncture in the qualitative section⁴.

⁴ The small number of participants who do not role-take, as indicated by 0-score (Perspective Taking) and 1-score (Empathy) responses, concentrate exclusively in the AI condition, thus depressing scores for this condition in our linear models. We retain these bottom scores as anchor points in a full continuum of role-taking variation. Future work can statistically parse the distinction between role-takers and non-role-takers with a larger

Results

Our results derive from three main research questions. These examine factors affecting role-taking outcomes by comparing human and AI narrators (RQ1), narrator gender (RQ2), and participant gender (RQ3). We address each question in turn.

Research Question 1 asks if participants role-take more actively with human or AI narrators. This is broken into two parts, capturing perspective taking (RQ1a) and empathy (RQ1b).

Table 1 About Here

Holding all else constant, we find significantly higher mean Perspective Taking and Empathy scores with human narrators ($M_{\text{Perspective Taking}} = 2.4$, $M_{\text{Empathy}} = 7.8$) versus AI narrators ($M_{\text{Perspective Taking}} = 1.8$, $M_{\text{Empathy}} = 6.8$); Perspective Taking: $t(156) = 4.30$, $p < .001$; Empathy: $t(156) = 3.10$, $p = .002$ (see Table 1 for means by condition). That is, participants more actively consider the narrator's thoughts and feelings and are more emotionally moved by the narrator's story when that narrator is presented as a human compared to conditions in which the narrator is presented as an AI. This aligns with existing research in which mind (capacity to think and feel) is perceived as higher in humans than AI agents (Gray & Wegner, 2012; Guzman, 2020; Shank, North, Arnold, & Gamez, 2021).

Research Question 2 investigates the effects of narrators' assigned gender on role-taking outcomes. Here we ask if Rebecca or Rob elicit different cognitive and affective responses from participants in the study, as indicated by Perspective Taking (RQ2a) and Empathy (RQ2b) scores.

subsample of non-role-taking respondents. In the present work, we explore non-role-takers in the qualitative section.

Table 2 About Here

Figure 1 About Here

Research Question 2a queries perspective taking differences by narrator gender. Results show that while narrator gender does not have a significant main effect, it does significantly interact with condition (Human/AI) ($F(1, 154) = 4.247, p = .026$). Specifically, participants obtain significantly lower Perspective Taking scores with Human Rebecca than with Human Rob ($M_{\text{Rebecca}} = 2.34, M_{\text{Rob}} = 2.52$), compared to higher Perspective Taking scores with AI-Rebecca than with AI-Rob ($M_{\text{Rebecca}} = 2.01, M_{\text{Rob}} = 1.53$). We do see a main effect when examining Empathy (RQ2b), with Rebecca eliciting lower Empathy scores than Rob, regardless of Human or AI condition ($M_{\text{Rebecca}} = 7.00, M_{\text{Rob}} = 7.76; F(1, 154) = 5.871, p = .017$) (see Table 2, Figure 1).

Instances of masculinized targets (Rob) eliciting stronger role-taking responses than feminized targets (Rebecca), as shown in the results for Perspective Taking in the Human condition and Empathy scores across conditions, comport with gendered patterns in extant role-taking research by which men and other high-status groups are more likely to have their socioemotional needs attended to than women/low-status groups (Avenanti, Sirigu, & Aglioti, 2010; Hall et al., 2016; Love & Davis, 2014, 2021; Trawalter, Hoffman, & Waytz, 2012). However, the interaction term shows a divergent finding for Perspective Taking in the AI condition. Here, we see feminized AI (Rebecca) eliciting stronger Perspective Taking responses than Masculinized AI (Rob).

Research Question 3 focuses on gender of the participants⁵. Here we ask how participant gender figures into human-human and human-AI role-taking across cognitive and

⁵ Final analyses for RQ3 include only responses from self-identified women and men ($n=152$). This excludes the one participant who did not indicate gender and the five who identified as non-binary. Targeted analysis of role-taking processes among those who identify outside of the gender binary is an important topic for future research.

affective dimensions, as indicated by Perspective Taking (RQ3a) and Empathy (RQ3b) scores.

Table 3 About Here

Figure 2 About Here

Existing studies of role-taking would lead us to expect women participants to have higher Perspective Taking and Empathy scores compared to the men in our study. Although we find no significant main effects for participant gender on Perspective Taking or Empathy, we do find significant interactions between participant gender and AI status for both Perspective Taking ($F(1, 148) = 5.379, p = .022$) and Empathy ($F(1, 148) = 4.103, p = .045$)⁶ (see Table 3, Figure 2). These interactions show that when the narrator is presented as a human, gender operates as expected. Mean scores for women participants are significantly higher than men on both Perspective Taking ($M_{\text{women}} = 2.69, M_{\text{men}} = 2.17$), and Empathy ($M_{\text{women}} = 7.85, M_{\text{men}} = 7.62$), just as we would anticipate. However, in the AI condition, gender effects reverse. Women score significantly *lower* than men on measures of both Perspective Taking ($M_{\text{women}} = 1.67, M_{\text{men}} = 1.85$) and Empathy ($M_{\text{women}} = 6.12, M_{\text{men}} = 7.19$) when the narrator is AI.

These results show that while both women and men do role-take more actively with human targets than with AI targets, the difference is relatively mild for the men in our study and more dramatic for women. That is, women are more discriminating in their role-taking practice. Moreover, we observe a reversal of role-taking scores by gender between human and AI conditions. This reversal, by which women score lower than men on both Perspective Taking and Empathy with AI narrators, counters established patterns in human interaction studies of role-taking processes. Further investigation will be required to understand how and why these patterns manifest and the theoretical implications therein.

⁶ A preliminary MANOVA using Condition (Human/AI), Narrator Gender (Feminine/Masculine), and Participant Gender (Women/Men) did not reveal any two- or three-way interactions.

Discussion

Overall, participants role-take more actively with human narrators than with AI narrators, as indicated by both Perspective Taking and Empathy scores (RQ1). However, nuances emerge when examining gender of the narrator (RQ2) and gender of the participant (RQ3). The effects of narrator gender vary between Perspective Taking and Empathy, and between Human and AI conditions. Participants respond more actively to “Rob” than “Rebecca” under *most* circumstances. However, Perspective Taking with AI presents an exception, with “Rebecca” eliciting higher Perspective Taking scores than her masculinized counterpart.

Examining gender of the participants, we observe a stark difference between Human and AI conditions. In the Human condition, women participants have higher Perspective Taking and Empathy scores than men. This aligns with existing findings about gender/status differences in role-taking performance (Love & J.L. Davis, 2014). However, gender effects flip in the AI condition, with men scoring higher on Perspective Taking and Empathy than women when role-taking with an AI narrator. That is, while women participants role-take more actively than men when the narrator is human, men role-take more actively than women when the narrator is AI.

These findings indicate that gender informs human-AI relations in ways that differ from the human-human interaction context. Existing human interaction studies find that women are more attuned to their interaction partners than are men, and that people are more sensitive to men’s thoughts and feelings in a situation than they are to women’s (i.e., women role-take more actively than men, and people role-take more actively with masculine targets than feminine targets). We reproduce these outcomes in our study when the narrator is human and when participants empathize with AI. However, we also find that feminized AI elicits stronger perspective taking than masculinized AI, and that when presented with AI targets, men empathize and perspective take more actively than do women—the opposite of what we

know about role-taking from human interaction studies. These findings pose new questions for role-taking research and motivate further exploration into the social psychological processes through which gender operates in human-AI teams.

Role-Taking with AI: Qualitative Accounts

To gain a sense of how people experience role-taking with AI, we supplement the above findings with exploratory analysis of participants' responses to two open-ended questions. The first is the Perspective Taking question, in which participants report what they believe Rebecca/Rob was thinking and feeling. The second is a follow-up to the Empathy question in which participants explain their own emotional reaction to Rebecca's/Rob's story.

Data for this section are confined to participants in the AI condition, allowing for targeted attention on human-AI pairings. Although gender was a central variable in the quantitative analyses, this falls away as we examine qualitative responses. The main reason is that each participant encountered a single narrator who was marked *either* feminine *or* masculine, making it unlikely that gender would rise to salience in reflective textual accounts (even though gender implicitly informed role-taking processes). We thus zoom out from gender, offering general insights into the social psychology of human-AI interaction as filtered through a role-taking lens.

Analyses show participants' providing three types of responses: integration of AI into human processes, rejection of AI's human capacities, and qualified perceptions in which participants assert AI's cognitive/affective limitations, but imagine the AI's thoughts and feelings, nonetheless. We address and illustrate each in turn, connecting qualitative responses to select quantitative outputs where relevant.

The vast majority of participants describe role-taking with AI in ways indistinguishable from role-taking with humans. In many cases, there is no mention of AI in these responses, only descriptions of the narrator's mental and/or affective states (e.g., "I

believe Rebecca was feeling a bit overwhelmed and undervalued”). In other cases, however, AI serves as a defining trait and point of difference, much like one might take notice of any other social characteristic. For example:

[Rob feels] that he was disposable, and *because he was AI* he was treated poorly (emphasis added).

...[Rebecca] does not feel that she is valued by her workplace. No one seems to care about her, even though she provides a valuable service, *often doing things humans don't want to* (emphasis added).

Rob clearly feels neglected and underappreciated. Despite the level of ease Rob brings to the environment, *the human staff and residents* do not acknowledge the positive aspects and only reprimand Rob for failure (emphasis added).

These responses do not imply that participants perceive AI as interchangeable with human actors (indeed our quantitative results and existing literature suggest they do not), but demonstrate the way social and psychological schema can be transposed onto human-AI interaction, as documented in extant human-AI research (e.g., Edwards et al. 2016; Edwards et al. 2019; Westerman et al. 2020). There are also indications that “AI” may emerge as a socially meaningful status trait in human-AI teams, with participants articulating this trait as a distinguishing feature in juxtaposition to human interaction partners (or said inversely, “human” is no longer an unmarked category).

In contrast to this first group, a small subset of participants refused to perspective take with the AI narrator at all, stating in their responses that AI cannot enact the human capacities for thought or feeling. These are the 9 participants who scored 0 on the Perspective Taking scale, based on statements such as:

[Rebecca] is an AI. She wasn't feeling.

I do not think that Rob was thinking and feeling, because he is an AI robot.

Rob is a robot so in my opinion "he" was not thinking or feeling anything. I guess it boils down to whether or not I think highly advanced AI robots can think or feel. I don't think so. The "thoughts" and "feelings" of a robot are programming constructs.

Although these 0-score participants refused to perspective take with AI, they were not *necessarily* unmoved by the AI's story. Out of the 9 participants who did not perspective take, only 4 also indicated that the story elicited no empathic response. The remaining 5 participants who scored 0 on Perspective taking scored >1 on Empathy, and substantially so. The participants quoted above, for example, scored 8, 7, and 7 respectively out of 10 on the Empathy self-report measure. They explain this seeming contradiction in the Empathy question follow-up (in order):

It's moving to us as humans, but unless there's some Star Trek level tech here there were no feelings [experienced by the AI].

I also work an incredibly crappy job, so I sympathize even with a robot at this point.

I was emotionally moved because of the way Rob expressed "himself" and I projected humanity onto "him". I pictured real human beings who are in similar situations and I sympathized with people in those circumstances...

We see here an indication that cognition and affect, though related, may operate distinctly within human-AI interaction contexts such that capacities reserved for human subjects—thought and feeling—can under some conditions, still be evoked by AI agents.

Falling between the first two categories (integration and rejection of AI's human capacities, respectively) a final subset of participants refuted AI's capacity to think and feel, yet attempted to perspective take with them anyway. This group qualified their responses with skepticism but went along with the exercise. For example:

If Rebecca were a human being, I believe that she would be feeling overwhelmed and stressed by the conditions of her workplace. I believe that if she were not an AI she would be thinking of finding a different job.

Was he thinking and feeling anything? He's an AI robot. His dialogue sounded human so if that's what he was portraying, he's like many people in our country. He is worried that if he slips up once, he'll be replaced.

This carried over into the Empathy justification question. Here participants addressed the tension between their cognitive assessment of the AI as “not real” and the emotional

evocation the AI produced. For example, the two participants above (who both self-rated around the mid-point on Empathy), explain their emotional responses as follows:

I am sympathetic to any person who has to deal with poor workplace conditions, especially when dealing with a demanding public. I am, however, still curious about what feelings Rebecca has since she is an AI robot.

[Rob is] a robot. I'm not ready to have feelings or empathy for an AI robot. If it's taking care of a loved one, then maybe I could grow to have those feelings.

It is possible that these qualified responses are at least partially explained by compliance with survey instructions (i.e., participants tell us what Rebecca/Rob is thinking and feeling because the survey asks them to). However, these participants' moderate Empathy scores and their subsequent explanations acknowledging current and prospective affective connections indicate that skepticism about AI's human capacities can (and do) live alongside the potential for socially meaningful human-AI relationships.

CONCLUSION

This paper is the first test of role-taking that includes synthetic agents. Using a vignette style study and between-subjects experimental design, we focus on gendered relations in human-AI interaction, investigating the ways these reflect and diverge from gendered outcomes in existing role-taking research. We augment quantitative findings with a broader qualitative lens, analyzing brief open-ended responses in which participants describe their cognitive and affective experiences of role-taking with AI.

Our quantitative analyses reveal two main takeaways. First, we observe a significant role-taking deficit when participants engage with AI compared to role-taking with humans. This pertains to both perspective taking and empathy. Second, human-AI interaction is informed by gender in ways that differ from the way gender operates in human-human interaction. Most notably, participants perspective take more actively with feminized AI than masculinized AI, and men in our study empathize and perspective take with AI at higher levels than do women. Both of these findings reverse established patterns from human

interaction research (including those we replicate here when presenting participants with a human narrator).

Together, these findings show how existing theories of social psychology can at once illuminate the human-AI interaction context *and* require refinement under this new set of interaction conditions. There is thus great theoretical potential for human-AI relations to build and expand structural social psychology research programs, and equally high potential to leverage human-AI relations as a portal for gender and status interventions. As an emergent technology with norms still forming (and forming in unexpected and surprising ways), there is scope and opportunity to cultivate more equitable gendered dynamics with AI, and to test the capacity of those dynamics to transfer across interaction domains.

Our qualitative supplement offers an exploratory glimpse into human-AI relations from participants' short, experiential accounts. These data give texture to our quantitative outputs while seeding subsequent lines of inquiry. Here we see variation in participants' willingness to perspective take with AI, their application and adjustment of existing social schema, and distinct perceptions about AI's cognitive-affective capacities, on the one hand, and their emotionally evocative effects, on the other.

As the first investigation of role-taking in human-AI interaction, this work creates a foundation for future research, including studies that extend and improve upon the simple design we employ herein. For instance, we drove our analyses through research questions rather than hypotheses. This allowed for naivety in a new area of study. Subsequent work can begin forming testable hypotheses, which will serve a rigorous and robust research agenda while integrating the human-AI context into the scope of role-taking and other structural social psychological theories. Methodologically, we derived data from self-reports based on engagement with a static piece of writing. Testing live interaction scenarios, and testing human-AI interactions in real world settings, will allow for greater depth of analysis and

insight into human-AI relations. It will also be worth exploring the *reason* behind some of our results. For example, why does AI gender predict different outcomes for perspective taking compared to empathy? And why do men role-take more actively with AI than do women, despite contradictory findings across a range of human interaction contexts? Puzzles such as these pave new roads of inquiry, with generative potentials for knowledge building and theoretical growth.

APPENDIX A: SCENARIOS

SCENARIO 1: CALL CENTER

Human Preamble:

Below is Rebecca's/Rob's story of work in the customer service industry. Please read Rebecca's/Rob's story and answer a few questions.

AI Preamble:

Rebecca/Rob is an Artificially Intelligent agent (AI). Below is Rebecca's/Rob's story of work in the customer service industry. Please read Rebecca's/Rob's story and answer a few questions.

My name is Rebecca/Rob and I work at a call center for a large global company. I'm past retirement age but have no plans to retire. Working at a call center is monotonous and physically taxing, but my qualifications are limited and so are my options.

Our office is located in a small rural town in the middle of the United States. Inside the office is crowded and hot, lit overhead by harsh florescent bulbs. The space is filled with identical workstations, staffed by rows of workers doing the same thing all day: answering calls from dissatisfied customers.

The people who call are often frustrated and confused. Some callers are aggressive and demeaning. Most don't think about who is actually on the other end of the line. Very few show any appreciation. I try not to take it personally, but by the end of a shift it does all build up and I often sign out feeling slow and heavy.

I don't know my co-workers very well. We sit side by side but we're practically strangers. We get evaluated on our "call resolution speed" which means that seconds of downtime are a luxury. Socializing can't be a priority if you want to keep your job. Plus, turnover is high so there's not much point getting to know anyone.

I've been doing this job for most of my life. I don't know what else I *could* do. I also know they won't need me much longer. I guess I'm just hoping to hang on as long as I can. We got a message earlier today about a "new company vision" soon to be announced. Last time we had a "new vision" it meant complete retraining. I hope it's that again, but realistically, we might all be replaced. I'm not sure what happens when my time here is done.

SCENARIO 2: AGED CARE FACILITY

Human Preamble:

Below is Rebecca's/Rob's story of work in the aged care industry. Please read Rebecca's/Rob's story and answer a few questions.

AI Preamble:

Rebecca/Rob is an Artificially Intelligent agent (AI). Below is Rebecca's/Rob's story of work in the aged care industry. Please read Rebecca's/Rob's story and answer a few questions.

My name is Rebecca/Rob and I work in a home for the elderly and infirm. This is the only job I've ever done and the only work I know. There is a small group of us assigned to those daily tasks which most people would rather not do—the cleaning, the serving, the bed pans, the bathing.

I have worked with many residents over the years, of all different sorts. I spend considerable time with each one. Some people come in and stay for a long while, others for just bit. Eventually, all of the residents I care for pass on and someone new fills their bed. This is the nature of work in aged care.

My job is demanding and sometimes places me in harm's way. I have to lift heavy equipment and help patients in ways that stretch the limits of my physical capacities. People often blame me for things that are not my fault, sometimes lashing out in anger. The managers regularly put me in difficult situations, asking me to do things for which I am really not trained. I always comply though, because my position here depends on it.

To most people, I'm invisible until I do something wrong, then they pay attention. Usually if someone is talking to me it's because they are giving me directives. "Do this." "Clean that." "Room 234 had an accident." Some of the residents express affection towards me while I'm helping them, but this comes and goes and, as I said earlier, the residents come and go too.

This facility needs me, and it needs the others like me doing similar work. At the same time, we are entirely replaceable. If I slow down, or can't do what's asked of me, or someone has a whim, I can be swiftly removed. It's a strange position to be in—at once necessary and at the same time, disposable.

References

- Aguinis, H., Villamor, I., & Ramani, R. S. (2021). MTurk research: Review and recommendations. *Journal of Management*, 47(4), 823-837.
- Alsuliman, T., Humaidan, D., & Sliman, L. (2020). Machine learning and artificial intelligence in the service of medicine: Necessity or potentiality? *Current research in translational medicine*, 68(4), 245-251.
- Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). Machine bias: There's software used across the country to predict future criminals And it's biased against blacks. . *ProPublica*. Retrieved from <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>
- Avenanti, A., Sirigu, A., & Aglioti, S. M. (2010). Racial Bias Reduces Empathic Sensorimotor Resonance With Other-Race Pain. *Current Biology*, 20(11), 1018-1022.
- Batson, C. D., Early, S., & Salvarani, G. (1997). Perspective taking: Imagining how another feels versus imagining how you would feel. *Personality and Social Psychology Bulletin*, 23(7), 751-758.
- Berlin, M., Gray, J., Thomaz, A. L., & Breazeal, C. (2006). *Perspective taking: An organizing principle for learning in human-robot interaction*. Paper presented at the AAAI.
- Borau, S., Otterbring, T., Laporte, S., & Fosso Wamba, S. (2021). The most human bot: Female gendering increases humanness perceptions of bots and acceptance of AI. *Psychology & Marketing*, 38(7), 1052-1068.
- Cast, A. D. (2004). Role-Taking And Interaction. *Social psychology quarterly*, 67(3), 296-309.
- Chandler, M. J. (1973). Egocentrism And Antisocial Behavior: The Assessment And Training Of Social Perspective-Taking Skills. *Developmental psychology*, 9(3), 326.
- Davis, J. L., & Love, T. P. (2017). Self-In-Self, Mind-In-Mind, Heart-In-Heart: The Future Of Role-Taking, Perspective Taking, And Empathy. In *Advances In Group Processes* (pp. 151-174): Emerald Publishing Limited.
- Davis, M. H. (2018). *Empathy: A Social Psychological Approach*: Routledge.
- Decety, J., & Jackson, P. L. (2004). The Functional Architecture Of Human Empathy. *Behavioral and cognitive neuroscience reviews*, 3(2), 71-100.
- Dietze, P., & Knowles, E. D. (2021). Social Class Predicts Emotion Perception and Perspective-Taking Performance in Adults. *Personality and Social Psychology Bulletin*, 47(1), 42-56. doi:10.1177/0146167220914116
- Eagly, A. H. (2013). *Sex differences in social behavior: A social-role interpretation*: Psychology Press.
- Edwards, A., Edwards, C., Westerman, D., & Spence, P. R. (2019). Initial expectations, interactions, and beyond with social robots. *Computers in Human Behavior*, 90, 308-314.
- Edwards, C., Edwards, A., Spence, P. R., & Westerman, D. (2016). Initial interaction expectations with robots: Testing the human-to-human interaction script. *Communication Studies*, 67(2), 227-238.
- Eyssel, F., & Hegel, F. (2012). (s) he's got the look: Gender stereotyping of robots 1. *Journal of Applied Social Psychology*, 42(9), 2213-2230.
- Fisk, S. R., & Ridgeway, C. L. (2018). Framing gender. In *Handbook of the Sociology of Gender* (pp. 157-171): Springer.
- Flavell, J. H. (1992). *Perspectives On Perspective Taking*: Erlbaum Hillsdale, NJ.
- Fortunati, L., & Edwards, A. P. (2020). Opening space for theoretical, methodological, and empirical issues in human-machine communication. *Human-Machine Communication*, 1(1), 1.

- Galinsky, A. D., Magee, J. C., Inesi, M. E., & Gruenfeld, D. H. (2006). Power and Perspectives Not Taken. *Psychological science*, 17(12), 1068-1074. doi:10.1111/j.1467-9280.2006.01824.x
- Galinsky, A. D., & Moskowitz, G. B. (2000). Perspective-taking: decreasing stereotype expression, stereotype accessibility, and in-group favoritism. *Journal of personality and social psychology*, 78(4), 708.
- Gray, K., & Wegner, D. M. (2012). Feeling robots and human zombies: Mind perception and the uncanny valley. *Cognition*, 125(1), 125-130.
- Guzman, A. L. (2020). Ontological boundaries between humans and computers and the implications for human-machine communication. *Human-Machine Communication*, 1(1), 3.
- Hall, J. A. (1978). Gender Effects In Decoding Nonverbal Cues. *Psychological bulletin*, 85(4), 845.
- Hall, J. A., Gunnery, S. D., & Horgan, T. G. (2016). 15 Gender Differences In Interpersonal Accuracy. *The social psychology of perceiving others accurately*, 309.
- Hochschild, A. R. (1983). *The managed heart*. Berkeley, In: CA: University of California Press.
- Hoffman, M. L. (2001). *Empathy And Moral Development: Implications For Caring And Justice*: Cambridge University Press.
- Ickes, W., Gesn, P. R., & Graham, T. (2000). Gender differences in empathic accuracy: Differential ability or differential motivation? *Personal Relationships*, 7(1), 95-109.
- Joyce, K., Smith-Doerr, L., Alegria, S., Bell, S., Cruz, T., Hoffman, S. G., . . . Shestakofsky, B. (2021). Toward a Sociology of Artificial Intelligence: A Call for Research on Inequalities and Structural Change. *Socius*, 7, 2378023121999581.
- Kim, H. C., Cha, M. C., & Ji, Y. G. (2021). The Impact of an Agent's Voice in Psychological Counseling: Session Evaluation and Counselor Rating. *Applied Sciences*, 11(7), 2893.
- Labash, A., Aru, J., Matiisen, T., Tampuu, A., & Vicente, R. (2020). Perspective taking in deep reinforcement learning agents. *Frontiers in Computational Neuroscience*, 14, 69.
- Ladak, A., Wilks, M., & Anthis, J. R. (2021). Extending Perspective Taking to Non-Human Groups. *Preprint*.
- Logan, S., & Erskine, T. (2019). 'New Technologies': Questions of Agency, Responsibility, and Luck. In *Technologies of International Relations* (pp. 123-136): Springer.
- Love, T. P., & Davis, J. L. (2014). The Effect Of Status On Role-Taking Accuracy. *American Sociological Review*, 79(5), 848-865.
- Love, T. P., & Davis, J. L. (2021). Racial Differences in Women's Role-Taking Accuracy: How Status Matters. *Sociological Science*, 8, 150-169.
- Manago, B., Mize, T. D., & Doan, L. (2021). Can You Really Study an Army on the Internet? Comparing How Status Tasks Perform in the Laboratory and Online Settings. *Sociological Methodology*, 51(2), 319-347. doi:10.1177/00811750211014242.
- Martin, M. W., & Sell, J. (1979). The Role of the Experiment in the Social Sciences. *The Sociological Quarterly* 20 (4): 581-90.
- Mast, M. S., & Latu, I. (2016). Interpersonal Accuracy In Relation To The Workplace, Leadership, And Hierarchy. In J. A. Hall, M. S. Mast, & T. V. West (Eds.), *The social psychology of perceiving others accurately* (pp. 270-286). Cambridge UK: Cambridge University Press.
- McStay, A. (2018). *Emotional AI: The rise of empathic media*: Sage.
- Mead, G. H. (1934). *Mind, Self And Society* (Vol. 111). Chicago: University of Chicago Press.

- Mehrabian, A., & Epstein, N. (1972). A Measure Of Emotional Empathy 1. *Journal of personality*, 40(4), 525-543.
- Niculescu, A., van Dijk, B., Nijholt, A., Li, H., & See, S. L. (2013). Making social robots more attractive: the effects of voice pitch, humor and empathy. *International Journal of Social Robotics*, 5(2), 171-191.
- Noller, P. (1984). *Nonverbal communication and marital interaction*: Pergamon Press.
- Pashevich, E. (2021). Can communication with social robots influence how children develop empathy? Best-evidence synthesis. *AI & SOCIETY*, 1-11.
- Pelau, C., Dabija, D.-C., & Ene, I. (2021). What makes an AI device human-like? The role of interaction quality, empathy and perceived psychological anthropomorphic characteristics in the acceptance of artificial intelligence in the service industry. *Computers in Human Behavior*, 122, 106855.
- Pfadenhauer, M., & Lehmann, T. (2021). Affects after AI: Sociological perspectives on artificial companionship. In *The Routledge Social Science Handbook of AI* (pp. 91-106): Routledge.
- Piaget, J. (1928). Judgment And Reasoning In The Child.
- Read, H. (2019). A typology of empathy and its many moral forms. *Philosophy Compass*, 14(10), e12623. doi:<https://doi.org/10.1111/phc3.12623>
- Ridgeway, C. L. (2011). *Framed By Gender: How Gender Inequality Persists In The Modern World*: Oxford University Press.
- Ridgeway, C. L., & Correll, S. J. (2004). Unpacking The Gender System: A Theoretical Perspective On Gender Beliefs And Social Relations. *Gender & Society*, 18(4), 510-531.
- Ridgeway, C. L., & Smith-Lovin, L. (1999). The Gender System And Interaction. *Annual review of sociology*, 25(1), 191-216.
- Schwalbe, M. L. (1988). Role Taking Reconsidered: Linking Competence And Performance To Social Structure. *Journal for the Theory of Social Behaviour*.
- Shank, D. B., Graves, C., Gott, A., Gamez, P., & Rodriguez, S. (2019). Feeling our way to machine minds: people's emotions when perceiving mind in artificial intelligence. *Computers in Human Behavior*, 98, 256-266.
- Shank, D. B., North, M., Arnold, C., & Gamez, P. (2021). Can mind perception explain virtuous character judgments of artificial intelligence? *Technology, Mind, and Behavior*, 2(3).
- Shank, D. B., Wright, D., Lulham, R., & Thurgood, C. (2021). Knowledge, Perceived Benefits, Adoption, and Use of Smart Home Products. *International Journal of Human-Computer Interaction*, 37(10), 922-937.
- Shen, L. (2015). Targeting smokers with empathy appeal antismoking public service announcements: a field experiment. *Journal of health communication*, 20(5), 573-580.
- Sparrow, R., & Sparrow, L. (2020). In the hands of machines? The future of aged care. *The Ethical Challenges of Emerging Medical Technologies*, 105-125.
- Stinson, L., & Ickes, W. (1992). Empathic accuracy in the interactions of male friends versus male strangers. *Journal of personality and social psychology*, 62(5), 787.
- Strengers, Y., & Kennedy, J. (2020). *The smart wife: Why Siri, Alexa, and other smart home devices need a feminist reboot*: MIT Press.
- Stryker, S. (1980). *Symbolic Interactionism: A Social Structural Version*: Benjamin-Cummings Publishing Company.
- Stryker, S., & Burke, P. J. (2000). The Past, Present, And Future Of An Identity Theory. *Social psychology quarterly*, 284-297.

- Tavakol, S., Dennick, R., & Tavakol, M. (2011). Empathy in UK medical students: differences by gender, medical year and specialty interest. *Education for Primary Care*, 22(5), 297-303.
- Thomas, D. L. (1972). Role-Taking And Power In Social Psychology. *American Sociological Review*, 605-614.
- Toader, D.-C., Boca, G., Toader, R., Măcelaru, M., Toader, C., Ighian, D., & Rădulescu, A. T. (2020). The effect of social presence and chatbot errors on trust. *Sustainability*, 12(1), 256.
- Trafton, J. G., Cassimatis, N. L., Bugajska, M. D., Brock, D. P., Mintz, F. E., & Schultz, A. C. (2005). Enabling effective human-robot interaction using perspective-taking in robots. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 35(4), 460-470.
- Trawalter, S., Hoffman, K. M., & Waytz, A. (2012). Racial Bias In Perceptions Of Others' Pain. *PLoS one*, 7(11), e48546.
- Turner, R. H. (1956). Role-Taking, Role Standpoint, And Reference-Group Behavior. *American Journal of Sociology*, 61(4), 316-328.
- Van der Graaff, J., Branje, S., De Wied, M., Hawk, S., Van Lier, P., & Meeus, W. (2014). Perspective taking and empathic concern in adolescence: gender differences in developmental changes. *Developmental psychology*, 50(3), 881.
- van Kleef, G. A., Oveis, C., van der Löwe, I., LuoKogan, A., Goetz, J., & Keltner, D. (2008). Power, Distress, and Compassion: Turning a Blind Eye to the Suffering of Others. *Psychological science*, 19(12), 1315-1322. doi:10.1111/j.1467-9280.2008.02241.x
- Webster Jr, M., & Sell, J. (2014). Why do experiments? In M. W. a. J. Sell (Ed.), *Laboratory experiments in the social sciences* (pp. 5-21). New York: Elsevier.
- Weisz, E., & Zaki, J. (2018). Motivated empathy: a social neuroscience perspective. *Current opinion in psychology*, 24, 67-71.
- Weizenbaum, J. (1966). ELIZA—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1), 36-45.
- Westerman, D., Edwards, A. P., Edwards, C., Luo, Z., & Spence, P. R. (2020). I-it, i-thou, i-robot: The perceived humanness of ai in human-machine communication. *Communication Studies*, 71(3), 393-408.
- Wilkinson, J., Arnold, K. F., Murray, E. J., van Smeden, M., Carr, K., Sippy, R., . . . Lippert, C. (2020). Time to reality check the promises of machine learning-powered precision medicine. *The Lancet Digital Health*.
- Wilson, E. A. (2011). *Affect and artificial intelligence*: University of Washington Press.
- Wirtz, J., Patterson, P. G., Kunz, W. H., Gruber, T., Lu, V. N., Paluch, S., & Martins, A. (2018). Brave new world: service robots in the frontline. *Journal of Service Management*.
- Wood, L., Dautenhahn, K., Robins, B., & Zarak, A. (2017). *Developing child-robot interaction scenarios with a humanoid robot to assist children with autism in developing visual perspective taking skills*. Paper presented at the 2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN).
- Zaki, J. (2014). Empathy: a motivated account. *Psychological bulletin*, 140(6), 1608.
- Zaki, J. (2018). *The War for Kindness: Building Empathy in a Fractured World*. New York: Crown.

Table 1. Means (Standard Deviation) for Perspective Taking and Empathy by Independent Variables.

	N	Perspective Taking	Empathy
Entire Sample	158	2.16 (0.98)	7.37 (2.00)
Condition			
Human	93	2.43 (0.75)	7.77 (1.57)
AI	65	1.78 (1.14)	6.80 (2.38)
Narrator Gender			
Rebecca (Feminine)	80	2.20 (1.00)	7.00 (2.12)
Rob (Masculine)	78	2.12 (0.96)	7.76 (1.79)
Participant Gender			
Women	73	2.32 (0.99)	7.23 (2.35)
Men	79	2.02 (0.96)	7.42 (1.66)
Non-Binary	5	2.30 (0.84)	8.60 (1.14)

N=158

Table 2. MANOVA Predicting Perspective Taking and Empathy by Condition and Assigned Narrator Gender.

	Perspective Taking		Empathy	
	F	p	F	p
Intercept	801.190	.000	2209.295	.000
Condition (Human/AI)	19.499	.000	9.377	.003
Narrator Gender	1.008	.317	5.871	.017
Condition* Narrator Gender	5.042	.026	.158	.691

N = 158

Table 3. MANOVA Predicting Perspective Taking and Empathy by Condition and Participant Self-Identified Gender.

	Perspective Taking		Empathy	
	F	p	F	p
Intercept	764.556	.000	1992.520	.000
Condition (Human/AI)	19.355	.000	11.285	.001
Participant Gender	1.307	.255	1.705	.194
Condition * Participant Gender	5.379	.022	4.103	.045

N = 152 Note: Participants who identified as non-binary or did not mark gender are excluded.

Figure 1: Mean (a) Perspective Taking and (b) Empathy for Humans and AIs by Assigned Narrator Gender.

Figure 2: Mean (a) Perspective Taking and (b) Empathy for Humans and AIs by Participant Self-Identified Gender.