Human–Robot Interaction Through the Lens of Social Psychological Theories of Intergroup Behavior

Eliot R. Smith¹, Selma Šabanović², and Marlena R. Fraune³

¹ Department of Psychological and Brain Sciences, Indiana University

² Luddy School of Informatics, Computing, and Engineering, Indiana University

³ Department of Psychology, New Mexico State University

This article reviews our program of research on human–robot interaction, which is grounded in theory and research on human intergroup relations from the field of social psychology. Under the “computers as social actors” paradigm, people treat robots in similar ways as they treat other humans. We argue that robots’ differences from humans lead them to be regarded as members of a potentially competing outgroup. Based on this conceptual parallel, our studies examine four related areas: People’s reactions to multiple (as opposed to single) robots; characteristics of robot groups (such as synchrony) that may influence people’s responses; tests of interventions that have been demonstrated to reduce prejudice in humans; and tests of other theoretical predictions drawn from work on human intergroup behavior. Several of these studies examined cultural differences between the U.S. and Japan. We offer brief descriptions and citations of 10 previously published studies (total $N = 1,635$), as well as 12 unpublished studies (total $N = 1,692$) that produced null or inconsistent results—to make them part of the scientific record and potentially inspire related investigations by others. Finally, we offer some broad conclusions based on this program of research.

Keywords: human–robot interaction, prejudice, intergroup behavior

Supplemental materials: <https://doi.org/10.1037/tmb0000002#supplemental-materials>

Since the pioneering work of Reeves and Nass (1996), human–robot interaction has often been studied within the “CASA” (computers as social actors) framework, which holds that people perceive and react to computational artifacts in similar ways as they react to other humans. For example, especially if they are humanoid in appearance and autonomous in action, robots elicit attributions of gendered characteristics (Eyssel & Hegel, 2012), and are treated with politeness (Reeves & Nass, 1996).

Extending the CASA perspective, we argue that social psychological theories on stereotyping, prejudice, and intergroup relations, developed to understand human intergroup interaction, can

fruitfully be applied to humans’ interactions with robots. Broadly speaking, reactions to robots often resemble reactions to human outgroups (e.g., immigrants or ethnic minorities). For example, like human outgroups, robots may elicit fears that they might take our jobs or physically harm us (e.g., de Graaf & Allouch, 2016; Nam, 2019). People think that robots have different values than we do: Robots are expected to make more utilitarian decisions than humans in moral dilemmas, such as directing a runaway trolley onto a track where it will kill only one person instead of five (Malle et al., 2015). In addition, robots are actually non-human, and human outgroups are commonly perceived in dehumanizing terms (Haslam, 2006). Dehumanization is related to mind attribution (e.g., Kozak et al., 2006), specifically involving perceptions that robots or outgroup members possess lesser mental capabilities than humans or ingroup members. Our work has sought to capitalize on such important parallels between human outgroups and robots to advance our understanding of human–robot interaction.

A Social Psychological Perspective

Specifically, social psychological work on human intergroup relations identifies several potential influences on human–robot interaction, including stereotypes, emotions, prejudice, norms, and motivations, as well as a number of interventions that could reduce prejudice.

Stereotypes

Stereotypes or beliefs about a group’s characteristics are often the basis for negative attitudes and behavioral avoidance. Common stereotypes of robots include physical dangerousness and the potential to take over humans’ jobs. Stereotypes can bias people’s interpretations of

Action Editor: Danielle S. McNamara was the action editor for this article.

ORCID iDs: Eliot R. Smith <https://orcid.org/0000-0002-0458-6235>; Marlena R. Fraune <https://orcid.org/0000-0002-4377-4634>

Disclosure: The authors declare no conflicts of interest in this work.

Acknowledgment: This work was supported by the National Science Foundation under Grant CHS-1617611. We thank Kyrie Amon, Sawyer Collins, and Steven Sherrin, who were instrumental in designing and conducting some of the studies described here.

Open Access License: This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC-BY-NC-ND). This license permits copying and redistributing the work in any medium or format for noncommercial use provided the original authors and source are credited and a link to the license is included in attribution. No derivative works are permitted under this license.

Disclaimer: Interactive content is included in the online version of this article.

Contact Information: Correspondence concerning this article should be addressed to Eliot R. Smith, Department of Psychological and Brain Sciences, Indiana University, 1101 E. Tenth St., Bloomington, IN 47405-7007, United States. Email: esmith4@indiana.edu

events involving the outgroup, often resulting in seeming confirmation and self-perpetuation of stereotypes (Fiske & Russell, 2010). Some studies show that a robot's appearance or other cues can cause people to apply stereotypes of human groups, such as gender or ethnic groups. For example, in one study a computer speaking with a female voice was rated as more knowledgeable about "feminine" topics such as relationships, compared to one using a male voice (Nass et al., 1997). In another study, robots perceived as having a female body shape were preferred for stereotypically female tasks (Bernotat et al., 2019). Thus, not only do people have stereotypes of robots themselves as a group, but they also sometimes apply stereotypes of human groups to robots.

Emotional Reactions

People can experience negative and sometimes positive emotions toward outgroups. Emotions toward robots may include anger or fear (e.g., Dekker et al., 2017; Hinks, 2020). Feelings of disgust have been reported toward robots that fall into the "uncanny valley," being highly similar but not identical in appearance to humans (Broadbent, 2017). Another relevant emotion is "intergroup anxiety" (Stephan & Stephan, 1985). This is a negative feeling of uncertainty in interaction with an outgroup member, due to not knowing how to behave, or fear of offending the other or of appearing prejudiced. Intergroup anxiety contributes to people's avoidance of outgroup members, and to uncomfortable, strained interaction across group lines. For untrained people, interaction with robots may produce anxiety and uncertainty in a similar way as interaction with a person of a different race or ethnicity (Nomura et al., 2006). But sometimes emotions toward human outgroups are positive, including sympathy or respect (Miller et al., 2004). Similarly, several studies show that people can feel empathy toward robots (Riek et al., 2009). In one study, children interacted with a robot and then saw the robot protesting that it was afraid of the dark and did not want to be put away in a closet. Over half of the children said that it was wrong to put the robot there (Kahn et al., 2012).

Prejudice and Its Behavioral Consequences

Prejudice, the negative evaluation of a social group or its members, results from negative stereotypes or emotions (Maio et al., 2010). Prejudice is behaviorally expressed in avoidance of the outgroup, unwillingness to cooperate or work with them, and in extreme cases even direct action against them (in the form of hate crimes or genocide). Similarly, people (especially children) sometimes abuse robots, such as informational robots in public areas (Salvini et al., 2010). Negative reactions against robots on the part of humans (e.g., co-workers, or sick or disabled people being cared for by robots) could limit the effectiveness of future robotics efforts.

Norms and Motivations

According to norms that are widely shared in some cultural settings, stereotypes, prejudice, and behavioral discrimination against human outgroups are socially inappropriate. Norms (socially endorsed standards for what is correct and appropriate; Jacobson et al., 2011) can limit the expression of prejudice and negative intergroup behavior. This is true especially when people internalize norms to constitute an internal source of motivation—they consider

expressions of prejudice not only as likely to be condemned by others, but also as inconsistent with their personal standards (Crandall & Eshleman, 2003; Plant & Devine, 1998). While some research examines how norms affect acceptance of social robots (de Graaf et al., 2019), we are unaware of any research that has examined the existence or effects of norms regarding prejudice against robots.

Interventions to Reduce Prejudice

Research on human intergroup relations has identified several types of intervention that can be effective in reducing prejudice. The most widely tested is intergroup contact. Getting to know individual members of the outgroup robustly reduces prejudice against the whole group (Pettigrew & Tropp, 2006). Other interventions aim at shifting social categorization, for example by moving people away from an "us and them" perspective on the ingroup and outgroup or by making a specific outgroup individual an ingroup or team member (e.g., Crisp & Hewstone, 1999). Still other interventions seek to change perceived norms to make prejudice seem less socially acceptable (Tankard & Paluck, 2016), or ask people to take the perspective of an outgroup member (Dovidio et al., 2004). Each of these interventions has demonstrated positive effects in at least some studies, although empirical tests in short-term laboratory studies have been much more common than tests in ongoing, real-world situations of intergroup conflict (Paluck & Green, 2009).

Goals of Our Research

Our program of research has pursued several important goals, both substantive and methodological. Substantively, first we systematically examined people's reactions to multiple robots. Robots are increasingly being designed for use in collaborative team environments, but most existing research examines dyadic interaction (one human, one robot) and measures people's perceptions of and behaviors toward robots without considering group membership or group dynamics as factors. Even one human interacting with one robot is an intergroup situation by definition, because the human is likely perceiving and interacting in terms of the contrasting group memberships. However, the presence of multiple humans or multiple robots makes the intergroup nature of the situation even more salient. Studies of human intergroup behavior show that interaction between two groups is often more competitive and aggressive, compared to one-on-one interaction (a phenomenon termed the discontinuity effect; Schopler & Insko, 1992). Recent work has begun to explore how people perceive and socially categorize robots in intergroup scenarios (e.g., Vanman & Kappas, 2019), as well as how to conceptualize and study robots in the context of group dynamics (e.g., Abrams & Rosenthal-von der Pütten, 2020; Jung et al., in press).

Second, in studies using multiple robots, we examined effects of group-level characteristics (rather than just individual characteristics such as robot appearance). Perhaps the most important group-level factor is "entitativity" (Hamilton & Sherman, 1996) or the extent to which multiple individuals are perceived as a group. Factors such as similar appearance (e.g., wearing uniform clothing), synchronized movements (e.g., marching in step), or working for common goals lead to increased perceptions of entitativity. A small amount of existing work has examined entitativity of robot groups (e.g., Bera

et al., 2018). Importantly, to the extent that robots are generally perceived as threatening, seeing them as more entitative is likely to increase humans' perceptions of them as negative and potentially aggressive (Dasgupta et al., 1999).

Third, we examined effects of prejudice-reduction manipulations drawn from the human literature in the context of human-robot interaction. As described above, many interventions, such as individualized contact with outgroup members, taking the perspective of the outgroup, or thinking of oneself and the outgroup as parts of a higher-level group, have been found to have positive effects in reducing prejudice.

Fourth, we tested additional theoretical predictions based on research on intergroup relations. For example, research suggests that emotions are an important determinant of people's attitudes and willingness to interact with an outgroup (Mackie & Smith, 2018) and that whether one thinks about an intergroup encounter in abstract or concrete terms can shape reactions (Trope & Liberman, 2010). Similar findings might hold for human-robot interaction.

In some of our studies aimed at these four goals, we also sought insights into the role of culture in shaping human-robot interaction. We focused on comparisons between the U.S. and Japan, cultures that have dramatically different popular images and stereotypes of robots (more positive in Japan than in the U.S.; Kaplan, 2004).

However, reactions to robots do not always follow this pattern; for example, Bartneck (2008) found that while Japanese participants rated conventional robots more positively than Americans did, the difference was reversed for highly anthropomorphic robots. More broadly, cultural differences affect many aspects of intergroup attitudes and behavior (Guimond et al., 2014) so we sought to examine such differences in reactions to robots.

In terms of research methods, our studies use two general approaches. First, we adapted some methods, measures, and study designs directly from existing social psychological work on human intergroup interaction. This allowed us to see whether results with robots would be similar. Second, some of our studies employ novel methods or manipulations that are impossible to use with humans. For example, a study on perspective taking had humans view the world from the visual perspective of a mobile robot (via its onboard camera) and control the robot's movements. Such novel methods offer new ways to test theoretical hypotheses about intergroup attitudes and behavior.

Review of Our Empirical Work

The core of this article is a focused review of our empirical studies aimed at the goals just listed, conducted between 2015 and 2020; key aspects of each study are summarized in Table 1. The review

Table 1
Summary of Studies Included in This Review

Topic	N, participants	Type of interaction with robot	Citation (see references) or further information
Single versus multiple robots in the field	261 People in cafeterias, U.S. and Japan	Live	Fraune, Kawakami, et al. (2015)
Single versus multiple robots of different types	127 Students	Video	Fraune, Sherrin, et al. (2015, March)
Single versus multiple robots of different types, unpublished replication	444 Amazon mechanical Turk workers	Video	Supplement II
Competitiveness of individuals versus groups	142 Students	Live	Fraune et al. (2019, March)
Entitativity of robot groups	173 Students, U.S. and Japan	Live	Fraune, Nishiwaki, et al. (2017)
Robots' behavior toward each other	Study 1: 630 Students, U.S. and Japan Study 2: 71 Students, U.S.	Study 1: Video Study 2: Live	Fraune, Oisted, et al. (2020)
Different forms of contact with a robot	189 Students	Live and video	Supplement III
Physical perspective taking with a telepresence robot	168 Students	Live	Supplement IV
Perspective taking using images	147 Students	Images	Supplement V
Regarding robots as teammates	48 Students	Live	Fraune, Šabanović, et al. (2017, August)
Regarding robots as teammates	102 Students	Live	Fraune, Šabanović, et al. (2020)
Regarding robots as teammates	81 Students	Live	Fraune (2020)
Regarding robots as teammates, replication in Japan	33 U.S. students (those in the two human-two robot ingroup condition of Fraune, Šabanović, et al., 2020) and 35 Japanese students	Live	Supplement VI
Effects of social norms	Study 1: 110 students, U.S. Study 2: 91 students, U.S. Study 3: 93 students, Japan	Video	Supplement VII
Effects of positive and negative emotions on willingness to interact	1,014 (participants in five studies using various populations)	Various	Smith et al. (2020)
Effects of internal and external motivation to control prejudice	223 Mechanical Turk workers	Questionnaire only	Supplement VIII
Effects of temporal perspective or construal level	Study 1: 113 Students Study 2: 43 Students Study 3: 36 Students	Questionnaire only	Supplement IX

includes unpublished as well as published studies, and studies that did not confirm hypotheses as well as those that did, to preserve our results for the scientific record. Descriptions in this article are brief for reasons of space. Additional details on methods and results can be found in the publications cited below, or for unpublished studies in supplemental materials.

Reactions to Multiple Robots

As noted earlier, exposure to or interaction with multiple robots (compared to a single robot) is theoretically expected to emphasize the intergroup nature of the situation. Expected effects include perception of robots as more similar to each other and more different from humans (Turner et al., 1994). In addition, group interaction (many-to-many) has been found to be systematically more aggressive and competitive than one-to-one interaction (Schopler & Insko, 1992).

Multiple Robots in the Field

In a field study conducted in university cafeterias in the U.S. and Japan, small Sociable Trash Box (STB) robots approached participants to collect trash (Fraune, Kawakami, et al., 2015). We manipulated the numbers of robots (Single or Group) and their behavior, which was either social (contingent on other agents' behavior) or functional (robots approached and left humans regardless of the humans' behavior). Results indicated that people interacted more with robots in groups than with single robots, yet reported similar levels of liking for both. Participants also rated social robots as more friendly and helpful than functional robots in general. Across many questionnaire measures, they rated single social robots more positively than a group of social robots, but a group of functional robots were viewed more positively than single functional robots.

American and Japanese participants responded similarly to the robots on most measures. However, Japanese participants reported universally more positive responses toward the robots, which is consistent with general cultural differences in views of robots in the two countries (e.g., Kaplan, 2004). Japanese participants also looked at robots longer, performed more direct interaction, and threw more trash in the robots than Americans.

Single Versus Multiple Robots of Different Types

In a study with a 2×3 design ($N = 127$, Fraune, Sherrin, et al., 2015), student participants viewed brief videos of one or multiple robots of three types: Nao (small, anthropomorphic), Pleo (dinosaur-like), or Create (mechanomorphic). The videos portrayed typical behaviors of the robots, either alone or in multiples: Creates drove around and beeped, Pleos walked and roared, and Naos waved, sat down, and stood up. Results for key evaluative measures (attitude toward the robots, willingness to interact with them, emotions) displayed an interaction of Number $\times$ Type. The more anthropomorphic Nao robots were viewed more positively in a group, while the least anthropomorphic Create robots were more negative in a group. A tentative explanation is that viewing multiple robots causes participants to self-categorize more strongly as humans. This in turn makes differences between humans and robots more salient, and such differences are viewed negatively. As a result, when in a group,

the most human-like robots are viewed more positively and the least human-like are viewed more negatively.

We replicated this study online with an MTurk participant sample to obtain a larger sample size and more variation in socioeconomic status and age than exists in our student sample. This unpublished study ($N = 444$, Supplement II) used Paro, a seal-like robot, instead of the dinosaur-like Pleo. This study failed to replicate the Number $\times$ Type interaction found in the first study, and in fact there were no significant main effects or interactions of robot type or number on the key dependent measures.

Competitiveness of Individuals Versus Groups

In general, interaction between human groups, compared to interaction between individuals, is found to be more negative, competitive, and aggressive (Schopler & Insko, 1992). We designed a study ($N = 142$) paralleling those in this literature to see if a similar "discontinuity effect" would be found in human-robot interaction (Fraune et al., 2019). Teams of one or three humans competed against teams of one or three robots on a series of social dilemma tasks that allowed either cooperative responses (maximizing both teams' outcomes) or competitive ones (maximizing one's own team's outcomes at the expense of the other team). Similar to the human literature, results showed that groups of three humans displayed marginally more competitive behavior than single humans. However, in contrast to typical findings, we did not find that groups of robots elicited more fear or more competition than single robots, unless the robot groups were perceived as highly entitative on a questionnaire measure. In an unpredicted finding, human participants competed slightly more against robot teams that matched their number. Overall, some aspects of the discontinuity effect (increased competitiveness by human groups) were found in this study but others (increased human competitiveness against robot groups) were not observed.

Summary

Across these studies, we do not find more positive or negative responses to multiple robots than to individuals across the board. Rather, depending on robot behavior, robot type, and perhaps participant population, single and multiple robots sometimes elicit similar responses and sometimes different.

Characteristics of Robot Groups

In studies using multiple robots, we have examined effects of group-level characteristics (rather than just individual characteristics such as robot appearance) on people's reactions. Our studies have examined both entitativity (the extent to which multiple robots are perceived as a group, e.g., because of identical appearance or synchronized movements) and the nature of robots' behavior toward humans and toward each other.

Entitativity of Robot Groups

In a study ($N = 173$, Fraune, Nishiwaki, et al., 2017) conducted in both the U.S. and Japan, we examined how robot Entitativity Condition (Single Robots, Diverse Group, Entitative Group) and Country (U.S., Japan) affect emotions, mind attributions, and

willingness to interact with robots. In human groups, entitativity makes threatening groups seem more threatening and cooperative groups seem more cooperative (e.g., Dasgupta et al., 1999; Johnson & Downing, 1979). Thus, in this study we primed participants to experience threat, so that entitative robots would be expected to be perceived as more threatening. Threat was induced by telling participants that they were performing a cognitive test, and we were examining whether robots could perform better than humans, which would put humans and robots in competition for jobs. Participants entered the lab and performed an ambiguous task with robots according to condition, then completed questionnaires about the experience.

Results indicate that Entitative robot groups, compared to Single robots, were viewed more negatively. Entitative robots were also more threatening than Diverse robots. Diverse robot groups, compared to Single robots, were viewed as having more mind, and participants were more willing to interact with them. These findings were similar in the U.S. and Japan. This indicates that entitative robot groups can elicit negative reactions, a critical point to keep in mind when designing robots.

One cultural difference that did emerge is that American participants rated robots more positively than Japanese participants. Although most prior work suggests that Japanese people feel more positively toward robots than Westerners, there are some findings that go in the opposite direction (e.g., MacDorman et al., 2009). In our study, females reported marginally more positive emotion toward robots, and there was a greater percentage of female participants in the U.S. compared to Japan, perhaps partly explaining this finding.

Robots' Behavior Toward Each Other

There is little existing work on effects of robot-robot communication on humans' perceptions and reactions to the robots (but see Iio et al., 2017; Williams et al., 2015). In an online study ($N = 630$, Fraune, Oisted, et al., 2020), participants viewed videos of STB robots that acted in different ways toward each other (single robot, group with social behavior, group with functional behavior) and toward humans (social, functional). This study was conducted in the U.S. and Japan. The robot behaviors were the same as the behaviors in the prior STB study (Fraune, Kawakami, et al., 2015). After viewing the videos, participants completed questionnaires about the robots.

Participants who saw robots in groups (regardless of their behavior) rather than single robots were more willing to interact with robots in the future, and perceived robot group entitativity was related to more positive responses to robots. This is likely because the robots were helpful rather than threatening, and therefore, entitativity was viewed positively. Robot behavior toward other robots drove perceptions of them: When the robots were social toward each other, participants viewed the robots as more anthropomorphic, and viewed people as having higher rapport with the robots. Japanese participants rated the robots generally more positively than Americans did, but most results were similar across cultures.

In a related, in-lab study ($N = 71$, Fraune, Oisted, et al., 2020), participants played a box-moving game with two STB robots assigned as their teammates. We again manipulated robot behavior toward robots (social, functional) and toward humans (social,

functional). Results indicated that, as in the previous study, perceived robot group entitativity related to more positive responses to robots. Also replicating the previous study, robot social behavior toward *robots* slightly increased anthropomorphic perceptions of them. Conversely, robot social behavior toward *humans* increased positive responses toward them in behavioral and questionnaire measures—but only when accounting for perceived robot entitativity. Differences between these two studies might relate to viewing robots on video compared to working with them, or to the robots' differing tasks (collecting trash compared to moving boxes).

Summary

These studies generally confirm the hypothesis that entitative robots will be viewed more negatively than diverse ones in a competitive context, but more positively under cooperation. They also support the novel hypothesis that in a group, robots' behavior toward each other has consequences for the way they are perceived: Social rather than functional behavior increases perceptions of anthropomorphism. Not surprisingly, social behavior toward humans results in more positive perceptions.

Interventions Intended to Reduce Prejudice

Several studies have examined effects of prejudice-reduction manipulations drawn from the human literature. These interventions include individualized contact with outgroup members (robots), taking the perspective of an outgroup member, or thinking of oneself and outgroup members as parts of a single group or team.

Different Forms of Contact With a Robot

Research on human intergroup relations has found robust positive effects of actual interpersonal contact (Pettigrew & Tropp, 2006), and weaker but similar effects of more indirect or remote forms of contact. In this unpublished study ($N = 189$, Supplement III), two in-lab conditions involved live interaction with a Baxter robot (large, anthropomorphic), where participants engaged in scripted interaction with the robot or were in a control condition (introduced to the robot but having no interaction). In three additional online conditions, participants (a) viewed video of a live participant's interaction (vicarious contact), (b) viewed a live participant being interviewed about their interaction (extended contact), or (c) completed the dependent measures without viewing any video (an online control condition).

Results were not as predicted. Although the online conditions produced more positive responses than the in-lab conditions, this is not very interpretable. The in-lab direct contact condition did not significantly differ from the in-lab control condition on key dependent variables. Similarly, there were no significant differences between either online indirect contact condition and the online control condition. This study provides no evidence for the idea that direct, vicarious, or extended contact with a robot produces more favorable responses.

Physical Perspective Taking With a Telepresence Robot

In an unpublished study ($N = 168$, Supplement IV), participants interacted with a telepresence robot (Beam+). Participants either

controlled the robot's movements around a room or watched as it moved supposedly autonomously (actually controlled by an experimenter); and viewed from the robot's own perspective (an onboard camera) or from an overhead camera. This created a 2×2 design plus a hanging control condition where participants were introduced to the robot but did not control or observe it move.

We found no significant effects on the key dependent variables (attitude, willingness to interact with robots). An overall ANOVA on the five conditions showed no significant differences. Comparing all four of the experimental cells combined against the control condition likewise showed no significant results. Thus, there is no evidence from this study that these physical instantiations of perspective taking—controlling the robot's movements or viewing the world from its perspective—make people's responses more positive.

Perspective Taking Using Images

We also examined perspective taking using a method more similar to many studies in the human intergroup literature. In such studies participants are shown an image of an outgroup member (e.g., an Arab Muslim) and asked to write about a day in this person's life, from the pictured person's own perspective—imagining what the person might be thinking and feeling (Todd & Galinsky, 2014). In a control condition, participants are instructed to write using an objective, uninvolved perspective on the person. We used this method in a 2×3 design: Perspective taking versus objective (control) Instructions $\times$ Human, anthropomorphic robot, or mechanomorphic robot target. The three images all depicted the target as a household worker in a kitchen setting; the human target was portrayed holding cleaning tools and supplies.

Results of this unpublished study ($N = 147$, Supplement V) showed no main effect or interactions of the perspective taking instructions. That is, overall (combining across the three targets) perspective taking did not result in more favorable responses on key dependent variables. Nor did the effect of the perspective taking manipulation differ significantly across targets—it did not produce more favorable responses even to the human target. Like the other perspective taking study just described, this study furnishes no evidence that perspective taking has positive effects on people's responses to robots.

Regarding Robots as Teammates

As technical developments increasingly enable robots to work in teams together with humans, social scientists are exploring how people think of and interact with robots as team members. Despite early work suggesting that robots will not be trusted as team members (Groom & Nass, 2007), interviews with members of military bomb disposal teams showed that they come to think of even “merely functional” robots they work with as team members, who cannot be easily replaced by another similar robot if damaged (Carpenter, 2016). A study by Correia et al. (2018) had two human-robot teams compete in a game, and found that robots expressing group-based emotions based on their team's outcomes (rather than emotions based on their individual outcomes) were liked better and trusted more by their human teammates. Such emotions suggest that the robots more strongly identify with their team. Finally, experimental studies have shown that people respond positively to robots

defined as members of an arbitrary ingroup (Kuchenbrandt et al., 2013), and to robots described as similar to the participant in gender and responses to a “work style” questionnaire (You & Robert, 2018). Based on this prior literature, we predicted that robots assigned to a participant's ingroup would be perceived and treated more favorably than outgroup robots.

In our first study on robot teammates (Fraune, Šabanović, et al., 2017), participants were assigned to two competing teams, each consisting of two humans and two robots, to examine how people treat others depending on group membership (ingroup, outgroup) and agent type (human, robot). The robots in this study were small, minimally social Mugbot robots. A key measure in this study was behavioral aggression, which we measured (as in many previous studies) by the volume of unpleasant noise blasts assigned to each agent by a member of the “winning” team after each round of competition.

Participants' attitudes favored the ingroup over the outgroup, and humans over robots. Correspondingly, participants assigned softer noise blasts to ingroup than to outgroup members, and to humans than to robots. Group membership had a larger effect than agent type, meaning that participants actually assigned softer noise blasts to ingroup robots than to outgroup humans. On questionnaire measures, participants rated ingroup members more positively than outgroup members, regardless of agent type.

We examined whether the results replicated with a larger sample and how team composition affected the results (Fraune, Šabanović, et al., 2020). Participants ($N = 102$) were again assigned to competing teams of humans and robots. The design factors were players' Group Membership (ingroup, outgroup), Agent Type (human, robot), and participant Team Composition (humans as minority, equal, or majority within the ingroup compared to robots).

Results replicated the findings of the first study—that is, participants favored ingroup over outgroup and humans over robots. Again, they favored ingroup robots over outgroup humans. Interestingly, people differentiated more between humans and robots in the ingroup than humans and robots in the outgroup, a type of outgroup homogeneity effect (Judd & Park, 1988). These effects generalized across Team Composition.

In another follow-up study (Fraune, 2020), we examined whether robot anthropomorphism affected the strength of effects of group membership. We used the same study design with robots varying in anthropomorphism (anthropomorphic—NAO, mechanomorphic—iRobot Create). The robots greeted participants in a way that fit their form (e.g., NAO robots said hello, Create robots beeped). Results replicated the prior findings, but also showed effects of anthropomorphism. When the robots were anthropomorphic rather than mechanomorphic, the effects of their group membership were more closely resembled patterns found in human intergroup research.

Finally, we ran a replication in Japan of the two human—two robot team condition of a study described above (Fraune, Šabanović, et al., 2020). An unpublished analysis (Supplement VI) compared the data from that condition with the newly collected data from Japan ($N = 35$), with Country as a between-subjects factor. On the key dependent variable, volume of noise blasts, there was a large main effect of group membership (ingroup members assigned less noise compared to outgroup). This effect was significantly smaller in Japan than in the U.S., consistent with other evidence that minimal or arbitrarily assigned group memberships are less impactful in East Asian cultures than in the West (Yuki, 2003). In Japan, humans were

given slightly more noise than robots (rather than receiving less noise as in the U.S.), consistent with the generally more positive cultural image of robots in Japan compared to the U.S. (Kaplan, 2004). In both countries, ingroup robots received less noise than outgroup humans. On other evaluative measures (attitude, positive, and negative emotions) there were no significant main effects or interactions of the country. Overall, the results were similar cross culturally, with shared group membership, although slightly less important in Japan, still outweighing the human-robot distinction.

Effects of Social Norms

Norms (information about how others generally act or what they regard as appropriate) influence all types of intergroup behavior (Mackie & Smith, 2018), so changing norms can be an important strategy for improving intergroup relations. Effects of these two types of norms (descriptive or what others do; injunctive or what others think is appropriate) can differ in some cases (Jacobson et al., 2011). An unpublished study ($N = 110$, Supplement VII) studied effects of norms on willingness to use a hypothetical home robot, using three conditions: Descriptive norm (participants were told that other people want this robot in their home), injunctive norm (participants were told that other people think you should want this robot in your home), and a control condition with no norm manipulation. There were no effects of condition on key dependent variables including attitude, intention to interact with robots, and positive emotions about robots. However, interactions of norm condition by gender were found, generally showing that men were relatively more influenced by descriptive norms and women by injunctive norms.

We ran a second study ($N = 91$) to replicate this unpredicted gender interaction. However, there was no significant interaction on most dependent variables. Gender did interact with norm condition on the measure of positive emotions, but in the reverse direction (injunctive norms had a relatively larger effect for men than for women). Across these two unpublished studies, then, we find no overall effects of norm manipulations, and inconsistent evidence of interactions between norm type and participant gender.

We also replicated this study in Japan ($N = 93$). Like the studies in the U.S., this unpublished study found no significant overall effects of norm condition (descriptive, injunctive, no norms), and no condition by gender interaction. Thus, the results are descriptively similar across cultures but null results are difficult to interpret.

Summary

In these studies, several manipulations that have been found to improve attitudes toward human outgroups do not have similar effects for robots. Intergroup contact robustly reduces prejudice with humans (Pettigrew & Tropp, 2006) but we found no significant effects of direct or indirect contact with robots. Perspective taking, in a physical instantiation or based on instructions to take a robot's perspective, also had no effects. Manipulations intended to shift perceived norms had scattered effects that were inconsistent across studies. In contrast, making robots teammates does have reliable positive effects, as was expected based on prior literature showing, for example, positive responses to robots defined as ingroup members (Kuchenbrandt et al., 2013) or described as similar to human participants (You & Robert, 2018). In fact, the effects of team or

group membership are so strong that in our studies, participants responded more positively to ingroup robots than to outgroup humans.

Other Predictions from Intergroup Research

Social psychological research suggests that emotions and motivations related to prejudice are important determinants of people's prejudiced attitudes and their willingness to interact with an outgroup (Mackie & Smith, 2018). Other studies show that abstract versus concrete construal can shape reactions to an intergroup encounter (Trope & Liberman, 2010).

Effects of Positive and Negative Emotions on Willingness to Interact

One practically and theoretically important question concerns the impact of emotions on people's willingness to interact with robots. Willingness to interact in the future with robots or any outgroup is important because it can begin a virtuous cycle in which interaction reduces prejudice, which encourages even more interaction, and so on (Paolini et al., 2018). Smith et al. (2020) examined whether positive or negative emotions are more powerful predictors of willingness to interact. Although theorists and researchers have often focused on negative emotions (especially anxiety), scattered findings in the literature on human intergroup relations suggest the importance of positive emotions. Smith et al. applied a novel analysis to combined data from five studies that used different types of robots and different modes of interaction (live vs. video), to identify patterns that emerge consistently across such study-to-study variation. As we expected, positive emotions were stronger predictors than negative emotions. Interestingly, researchers have yet to examine the relative impact of positive versus negative emotions in regard to interaction with human outgroups.

Effects of Internal and External Motivation to Control Prejudice

In studies about human outgroups, Plant and Devine (1998) and many other researchers have examined effects of internal and external motivation to control prejudice (IMS and EMS). Generally, IMS (wanting to be unprejudiced due to one's own internal standards) is correlated with lower levels of prejudice, while EMS (wanting to be unprejudiced due to worries about reactions from others) is correlated with higher levels of prejudice. In a study currently being prepared for submission ($N = 223$, Supplement VIII), we gave some participants the standard version of IMS and EMS, as well as measures related to prejudice, attitudes, and willingness to interact with African Americans as an outgroup. Other participants completed the same measures reworded to refer to robots, for example (IMS) "Because of my personal values, I believe that using stereotypes about robots is wrong." Participants were randomly assigned to one of the two target groups, so we could examine responses of equivalent samples of participants on a nearly identically worded set of measures.

Relations of the key dependent variables (attitude and willingness to interact) to emotions, contact, and IMS and EMS were largely similar for the two groups. Positive and negative emotions as well as

previous contact predicted the dependent variables in the same ways. Effects of internal motivation were the same for the two groups, while effects of external motivation were significant for African Americans (predicting higher prejudice, as in previous work) but non-significant for the robot outgroup. This is perhaps unsurprising because cultural norms against anti-robot prejudice seem weak or absent, so people would have little reason to expect others to react negatively if they expressed such prejudice.

Effects of Temporal Perspective or Construal Level

We conducted three unpublished studies (Supplement IX) based on hypotheses from construal level theory (Trope & Liberman, 2010). In general, psychologically closer events elicit thoughts about concrete aspects, so thoughts about an impending interaction with a robot might include anxiety and unfamiliarity with robots. In contrast, more distant events are conceptualized in terms of more general, abstract features, perhaps including curiosity and interest in robots. Based on these ideas, we hypothesized that a closer or more concrete perspective might lead to more negative reactions, compared to a distant or more abstract perspective. Study 1 ($N = 113$) used a 2×2 design, where participants saw an image of Nao or Baxter and were told they would interact with this robot either later in the same experimental session or in a second session to be scheduled in a couple of months. Studies 2 and 3 ($N = 43$ and 36) were online surveys where participants read short paragraphs discussing what robots do in the world currently in concrete terms (e.g., sweeping the floor), or what robots will do in future years in abstract terms (e.g., maintaining cleanliness).

Analyzing the studies individually, there were no significant effects of condition on any of the dependent measures, except that Study 1 produced an interaction of Condition by Robot Type on negative emotions. Because all three studies used the same dependent variables and a measure of construal level (the extent to which participants think about robot behaviors in concrete vs. abstract terms), we combined all three ($N = 192$) to attain greater power for a correlational analysis. This offered some support for our hypothesis: More abstract construal correlates with greater willingness to interact with robots ($r = .33$, $p < .001$) as well as more positive emotions and attribution of more mind to the robot (Kozak et al., 2006). Although our manipulations in these studies had few effects, the correlational findings do suggest that a more distant or abstract perspective might lead people to think of robots more positively.

Summary

Results of these studies again show both similarities and differences from research on human intergroup relations. Emotions and internal/external motivations to control prejudice appear to have similar effects on attitudes and willingness to interact with robots as they do with human outgroups, suggesting the value of further research on these constructs in regard to robots. We also sought unsuccessfully to manipulate the abstractness versus concreteness with which people think about interaction with robots. However, more abstract construal did correlate with more positive views of robots, so further investigation along these lines might identify effective intervention strategies.

Discussion and Conclusions

Under each major topic there are both similarities and differences between our results and related studies on human intergroup relations.

Reactions to Multiple Robots

First, our studies find that multiple robots can elicit different reactions than single ones, as predicted from the idea that multiple robots (or multiple humans) reinforce the intergroup nature of human–robot interaction. However, the pattern is not straightforward—for example, multiple robots do not generate more negative reactions across the board as would be predicted by some literature (Schopler & Insko, 1992). Instead, findings frequently involve interactions of the number of robots with robot behavior (social or functional), robot type, and perhaps participant population. These results underline the importance of studying human interactions with groups of robots, rather than assuming that findings with individual robots will also apply to groups.

Characteristics of Robot Groups

These studies generally confirm the prediction that an entitative group of robots will be viewed more negatively than diverse ones in a threatening or competitive context (e.g., Dasgupta et al., 1999), but more positively under cooperation. We also tested effects of robots' behavior toward each other, a factor that has been relatively unexplored in comparison to robots' behavior toward humans. Robots that acted socially (rather than functionally) toward each other were perceived more anthropomorphically.

Interventions Intended to Reduce Prejudice

Our studies testing interventions aimed at reducing prejudice had mixed results. On the positive side, turning robots into teammates does have reliable positive effects. In fact, the effects of team or ingroup membership were so strong in our studies that participants responded more positively to ingroup robots than to outgroup humans. Some implications of this finding are potentially troubling. For example, members of competitive teams might conceivably withhold financial or medical resources from outgroup humans in favor of helping their own ingroup robots.

Other interventions failed to produce expected results. Intergroup contact has perhaps the most robust and well-replicated effects with humans (Pettigrew & Tropp, 2006) but we found no significant effects of direct or indirect contact with robots. Perspective taking had no significant effects in two different paradigms. Manipulations intended to shift perceived norms about attitudes or behavior toward robots had only scattered effects that were inconsistent across studies. Overall, it appears that not all interventions that have been effective with human groups will work equally well for human–robot interaction.

Other Predictions from Intergroup Research

As predicted from the literature on human outgroups (Mackie & Smith, 2018), emotions toward robots played a powerful role in predicting people's willingness to interact with robots—and positive

emotions had stronger effects than negative ones. This result suggests that interventions aimed at decreasing negative emotions in order to increase willingness to interact may be somewhat mistargeted; novel interventions seeking increased levels of positive emotions such as excitement might have more potent effects.

Internal and external motivation to control prejudice also show patterns similar to those with human outgroups. Internal motivation relates to more positive attitudes and behaviors toward robots, and external motivation correlates with negative attitudes and behaviors. These effects may be partially mediated by positive and negative emotions. A currently unanswered question is where these motives come from. In contrast to prejudice against human social groups, our culture does not appear to teach that prejudice against robots is wrong. So why do some people develop internal standards against such prejudice, or expect social disapproval of anti-robot prejudice? We cannot say at this time.

Finally, we attempted to manipulate construal level, or the abstractness versus concreteness with which people think about interaction with robots. While our manipulations were ineffective, correlational findings did suggest that more abstract construal goes along with more positive views of robots. Further investigation along these lines might prove to be fruitful.

Cultural Comparisons

Our studies that compared the U.S. and Japan found results that were generally fairly similar across cultures. In most cases, Japanese participants rated robots more positively than did U.S. participants, a result that was expected on the basis of some prior work (e.g., Kaplan, 2004). Surprisingly, this difference was reversed in one study, Fraune, Nishiwaki, et al. (2017). In the studies assigning robots as teammates in both countries, the more positive treatment of ingroup versus outgroup members was more powerful than the difference between humans and robots. A few other detailed results differed between cultures, such as the results for looking time in Fraune, Kawakami, et al. (2015). Overall, we find more similarities than differences between the U.S. and Japan.

Broader Implications

This work applies social psychological perspectives on stereotyping, prejudice, and conflict between human groups (often defined by race, religion, nationality, etc.) to understand human-robot interaction. Stretching theories beyond their original domain of applicability in this way can reveal a surprising degree of generalizability, such as our replicated finding that making a robot an ingroup member can lead people to treat the robot positively, even better than an outgroup human. Stretching theories can also reveal important limitations and boundary conditions, such as our failure to replicate the discontinuity effect with groups of robots, the higher level of competitiveness found when groups of humans rather than individuals interact.

The failure of contact and perspective taking manipulations to influence reactions to robots could reflect the perception that robots lack the right kind of “mind” or inner essence that one can relate to as a personal acquaintance or friend, or whose perspective can be adopted (e.g., Gray et al., 2007). Perhaps making robots into teammates, a manipulation that did succeed, operates through different mechanisms that are not so reliant on perceiving mind. Some of our

studies confirm that robots are attributed less mind than even human outgroup members (Fraune, 2020; Fraune, Šabanović, et al., 2020). However, one study (Fraune, Nishiwaki, et al., 2017) found that a diverse group of robots was attributed more mind than a single robot, and an unpublished study (Supplement IX) showed that abstract construal of robots correlated with more mind perception. Such findings may point the way toward conditions that could increase the perception of human-like mind in robots.

Methodological Implications and Limitations

Human-robot interaction can be an excellent domain for testing social psychological theories, because it permits experimental manipulations that would be impossible to implement in human interaction and could offer novel tests of theories regarding intergroup perception and behavior. One example is the physical perspective taking manipulations we used in one study: Taking a robot’s visual perspective by viewing through its onboard camera, and controlling the robot’s physical movements. A second example is that a group of robots could move precisely in synchrony, an important cue to the strength or “entitativity” of the group that might intensify negative reactions (Dasgupta et al., 1999; Lakens & Stel, 2011). The ability to vary robot appearance as well as behavior allows examination of factors not available in human-human interaction, such as the effects of the agent’s degree of anthropomorphism on perceptions and interactions (MacDorman & Ishiguro, 2006).

We acknowledge several limitations of these studies. Some studies use video rather than live interaction with robots, because video makes it feasible to examine many types of robots for generality. Some, including all our live-interaction studies, use student samples (in the U.S. or Japan), which are restricted in age and socio-economic status. All of our research involves short-term lab or questionnaire studies; we have performed no observational studies of ongoing realistic human-robot interaction, for example in industrial or other work settings, where different factors such as long-term familiarity with robots might come into play. The majority of our findings rest on experimental manipulations, which can support strong causal conclusions. However, a few findings (e.g., in Smith et al., 2020 and in Supplement VIII) rest on correlational relationships, leaving some causal ambiguity.

Summary

The main message of this review is that theory and research on prejudice and intergroup behavior from social psychology have much potential for helping researchers understand human-robot interaction. While some of our findings neatly confirmed theoretical expectations, there were many exceptions, as is to be expected when theories are extended beyond their original domain of application. Future research should continue to explore parallels and differences between human intergroup behavior and human-robot interaction, including with novel types of robots and in novel contexts such as interactions between groups of humans and groups of robots.

References

- Abrams, A. M., & Rosenthal-von der Pütten, A. M. (2020). I-C-E framework: Concepts for group dynamics research in human-robot interaction. *International Journal of Social Robotics*, 1–17. <https://doi.org/10.1007/s12369-020-00642-z>

- Bartneck, C. (2008). Who like androids more: Japanese or US Americans? *RO-MAN 2008—The 17th IEEE international symposium on robot and human interactive communication* (pp. 553–557). IEEE.
- Bera, A., Randhavane, T., Kubin, E., Wang, A., Gray, K., & Manocha, D. (2018, October). The socially invisible robot navigation in the social world using robot entitativity. In *2018 IEEE/RSJ international conference on intelligent robots and systems (IROS)* (pp. 4468–4475). IEEE.
- Bernotat, J., Eyssel, F., & Sachse, J. (2019). The (fe) male robot: How robot body shape impacts first impressions and trust towards robots. *International Journal of Social Robotics*, 1–13. <https://doi.org/10.1007/s12369-019-00562-7>
- Broadbent, E. (2017). Interactions with robots: The truths we reveal about ourselves. *Annual Reviews in Psychology*, 68(1), 627–652. <https://doi.org/10.1146/annurev-psych-010416-043958>
- Carpenter, J. (2016). *Culture and human–robot interaction in militarized spaces: A war story*. Routledge. <https://doi.org/10.4324/9781315562698>
- Correia, F., Mascarenhas, S., Prada, R., Melo, F. S., & Paiva, A. (2018). Group-based emotions in teams of humans and robots. In *Proceedings of the 2018 ACM/IEEE international conference on human–robot interaction* (pp. 261–269). Association for Computing Machinery.
- Crandall, C. S., & Eshleman, A. (2003). A justification-suppression of the expression and experience of prejudice. *Psychological Bulletin*, 129(3), 414–446. <https://doi.org/10.1037/0033-2909.129.3.414>
- Crisp, R. J., & Hewstone, M. (1999). Differential evaluation of crossed category groups: Patterns, processes, and reducing intergroup bias. *Group Processes & Intergroup Relations*, 2(4), 307–333. <https://doi.org/10.1177/1368430299024001>
- Dasgupta, N., Banaji, M. R., & Abelson, R. P. (1999). Group entitativity and group perception: Associations between physical features and psychological judgment. *Journal of Personality and Social Psychology*, 77(5), 991–1003. <https://doi.org/10.1037/0022-3514.77.5.991>
- de Graaf, M. M., & Allouch, S. B. (2016, August). Anticipating our future robot society: The evaluation of future robot applications from a user's perspective. *2016 25th IEEE international symposium on robot and human interactive communication (RO-MAN)* (pp. 755–762). IEEE.
- de Graaf, M. M., Ben Allouch, S., & van Dijk, J. A. (2019). Why would I use this in my home? A model of domestic social robot acceptance. *Human–Computer Interaction*, 34(2), 115–173. <https://doi.org/10.1080/07370024.2017.1312406>
- Dekker, F., Salomons, A., & Waal, J. V. D. (2017). Fear of robots at work: The role of economic self-interest. *Socio-economic Review*, 15(3), 539–562.
- Dovidio, J. F., Ten Vergert, M., Stewart, T. L., Gaertner, S. L., Johnson, J. D., Esses, V. M., Riek, B. M., & Pearson, A. R. (2004). Perspective and prejudice: Antecedents and mediating mechanisms. *Personality and Social Psychology Bulletin*, 30(12), 1537–1549. <https://doi.org/10.1177/0146167204271177>
- 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. <https://doi.org/10.1111/j.1559-1816.2012.00937.x>
- Fiske, S. T., & Russell, A. M. (2010). Cognitive processes. *The SAGE handbook of prejudice, stereotyping and discrimination* (pp. 115–130). Sage Publications.
- Fraune, M. R. (2020). Our robots, our team: Robot anthropomorphism moderates group effects in human–robot Teams. *Frontiers in Psychology*, 11, Article 1275. <https://doi.org/10.3389/fpsyg.2020.01275>
- Fraune, M. R., Kawakami, S., Šabanović, S., De Silva, S., & Okada, M. (2015). *Three's company, or a crowd?: The effects of robot number and behavior on HRI in Japan and the USA* [Conference session]. Robotics: Science and Systems Conference, Rome, Italy.
- Fraune, M. R., Nishiwaki, Y., Šabanović, S., Smith, E. R., & Okada, M. (2017). *Threatening flocks and mindful snowflakes: How group entitativity affects perceptions of robots* [Conference session]. Human Robot Interaction Conference, Vienna, Austria.
- Fraune, M. R., Oisted, B. C., Sembrowski, C. E., Gates, K. A., Krupp, M. M., & Šabanović, S. (2020). Effects of robot–human versus robot–robot behavior and entitativity on anthropomorphism and willingness to interact. *Computers in Human Behavior*, 105, Article e106220. <https://doi.org/10.1016/j.chb.2019.106220>
- Fraune, M. R., Šabanović, S., & Smith, E. R. (2017, August). *Teammates first: Favoring ingroup robots over outgroup humans* [Conference session]. IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2017), Lisbon, Portugal.
- Fraune, M. R., Šabanović, S., & Smith, E. R. (2020). Some are more equal than others: Ingroup robots gain some but not all benefits of team membership. *Interaction Studies: Social Behaviour and Communication in Biological and Artificial Systems*.
- Fraune, M. R., Sherrin, S., Šabanović, S., & Smith, E. R. (2015, March). *Rabble of robot effects: Number and type of robots modulates attitudes, emotions, and stereotypes* [Conference session]. ACM/IEEE International Conference on Human–Robot Interaction (HRI 2015), Portland, Oregon.
- Fraune, M. R., Sherrin, S., Šabanović, S., & Smith, E. R. (2019, March). *Is human–robot interaction more competitive between groups than between individuals?* [Conference session] ACM/IEEE International Conference on Human–Robot Interaction (HRI 2019), Daegu, South Korea.
- Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. *Science (New York, NY)*, 315(5812), 619–619. <https://doi.org/10.1126/science.1134475>
- Groom, V., & Nass, C. (2007). Can robots be teammates?: Benchmarks in human–robot teams. *Interaction Studies: Social Behaviour and Communication in Biological and Artificial Systems*, 8(3), 483–500. <https://doi.org/10.1075/is.8.3.10gro>
- Guimond, S., Sablonnière, R. de la, & Nugier, A. (2014). Living in a multicultural world: Intergroup ideologies and the societal context of intergroup relations. *European Review of Social Psychology*, 25(1), 142–188. <https://doi.org/10.1080/10463283.2014.957578>
- Hamilton, D. L., & Sherman, S. J. (1996). Perceiving persons and groups. *Psychological Review*, 103(2), 336–355. <https://doi.org/10.1037/0033-295X.103.2.336>
- Haslam, N. (2006). Dehumanization: An integrative review. *Personality and Social Psychology Review*, 10(3), 252–264. https://doi.org/10.1207/s15327957pspr1003_4
- Hinks, T. (2020). Fear of robots and life satisfaction. *International Journal of Social Robotics*, 1–14. <https://doi.org/10.1007/s12369-020-00640-1>
- Iio, T., Yoshikawa, Y., & Ishiguro, H. (2017, November). Starting a conversation by multi-robot cooperative behavior. In *International conference on social robotics* (pp. 739–748). Springer.
- Jacobson, R. P., Mortensen, C. R., & Cialdini, R. B. (2011). Bodies obliged and unbound: Differentiated response tendencies for injunctive and descriptive social norms. *Journal of Personality and Social Psychology*, 100(3), 433–448. <https://doi.org/10.1037/a0021470>
- Johnson, R. D., & Downing, L. L. (1979). Deindividuation and valence of cues: Effects on prosocial and antisocial behavior. *Journal of Personality and Social Psychology*, 37, 1532–1538. <https://doi.org/10.1037/0022-3514.37.9.1532>
- Judd, C. M., & Park, B. (1988). Out-group homogeneity: Judgments of variability at the individual and group levels. *Journal of Personality and Social Psychology*, 54(5), 778–788. <https://doi.org/10.1037/0022-3514.54.5.778>
- Jung, M., DiFranzo, D., Stoll, B., Shen, S., Lawrence, A., & Claire, H. (in press). Robot assisted tower construction—A method to study the impact of a robot? s Allocation behavior. [THRI]. *ACM Transactions on Human–Robot Interaction*.
- Kahn, P. H., Jr., Kanda, T., Ishiguro, H., Freier, N. G., Severson, R. L., Gill, B. T., Ruckert, A. H., & Shen, S. (2012). “Robovie, you’ll have to go into the closet now”: Children’s social and moral relationships with a humanoid robot. *Developmental Psychology*, 48(2), 303–314. <https://doi.org/10.1037/a0027033>

- Kaplan, F. (2004). Who is afraid of the humanoid? Investigating cultural differences in the acceptance of robots. *International Journal of Humanoid Robotics*, 1(03), 465–480. <https://doi.org/10.1142/S0219843604000289>
- Kozak, M. N., Marsh, A. A., & Wegner, D. M. (2006). *What do i think you're doing?* Action identification and mind attribution. *Journal of Personality and Social Psychology*, 90(4), 543–555. <https://doi.org/10.1037/0022-3514.90.4.543>
- Kuchenbrandt, D., Eyssel, F., Bobinger, S., & Neufeld, M. (2013). When a robot's group membership matters. *International Journal of Social Robotics*, 5(3), 409–417. <https://doi.org/10.1007/s12369-013-0197-8>
- Lakens, D., & Stel, M. (2011). If they move in sync, they must feel in sync: Movement synchrony leads to attributions of rapport and entitativity. *Social Cognition*, 29, 1–14. <https://doi.org/10.1521/soco.2011.29.1.1>
- MacDorman, K. F., & Ishiguro, H. (2006). The uncanny advantage of using androids in cognitive and social science research. *Interaction Studies: Social Behaviour and Communication in Biological and Artificial Systems*, 7(3), 297–337. <https://doi.org/10.1075/is.7.3.03mac>
- MacDorman, K. F., Vasudevan, S. K., & Ho, C. (2009). Does Japan really have robot mania? Comparing attitudes by implicit and explicit measures. *AI & Society*, 23, 485–510. <https://doi.org/10.1007/s00146-008-0181-2>
- Mackie, D. M., & Smith, E. R. (2018). Intergroup emotions theory: Production, regulation, and modification of group-based emotions. *Advances in Experimental Social Psychology*, 58, 1–69. <https://doi.org/10.1016/bs.aesp.2018.03.001>
- Maio, G. R., Haddock, G., Manstead, A. S., & Spears, R. (2010). Attitudes and intergroup relations. *The SAGE handbook of prejudice, stereotyping, and discrimination* (pp. 261–275). Sage Publications.
- Malle, B. F., Scheutz, M., Arnold, T., Voiklis, J., & Cusimano, C. (2015, March). *Sacrifice one for the good of many?* [Conference session]. Presented at HRI '15, Portland, OR. <https://doi.org/10.1145/2696454.2696458>
- Miller, D. A., Smith, E., & Mackie, D. M. (2004). Effects of intergroup contact and political predispositions on prejudice: Role of intergroup emotions. *Group Processes & Intergroup Relations*, 7(3), 221–237. <https://doi.org/10.1177/1368430204046109>
- Nam, T. (2019). Citizen attitudes about job replacement by robotic automation. *Futures*, 109, 39–49. <https://doi.org/10.1016/j.futures.2019.04.005>
- Nass, C., Moon, Y., & Green, N. (1997). Are machines gender neutral? Gender-stereotypic responses to computers with voices. *Journal of Applied Social Psychology*, 27(10), 864–876. <https://doi.org/10.1111/j.1559-1816.1997.tb00275.x>
- Nomura, T., Kanda, T., & Suzuki, T. (2006). Experimental investigation into influence of negative attitudes toward robots on human–robot interaction. *AI & Society*, 20(2), 138–150. <https://doi.org/10.1007/s00146-005-0012-7>
- Paluck, E. L., & Green, D. P. (2009). Prejudice reduction: What works? A Review and assessment of research and practice. *Annual Review of Psychology*, 60(1), 339–367. <https://doi.org/10.1146/annurev.psych.60.110707.163607>
- Paolini, S., Harwood, J., Hewstone, M., & Neumann, D. L. (2018). Seeking and avoiding intergroup contact: Future frontiers of research on building social integration. *Social and Personality Psychology Compass*, 12(12), Article e12422. <https://doi.org/10.1111/spc3.12422>
- Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, 90(5), 751–783. <https://doi.org/10.1037/0022-3514.90.5.751>
- Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology*, 75, 811–832. <https://doi.org/10.1037/0022-3514.75.3.811>
- Reeves, B., & Nass, C. I. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. Cambridge university press.
- Riek, L. D., Rabinowitch, T. C., Chakrabarti, B., & Robinson, P. (2009, March). How anthropomorphism affects empathy toward robots. In *Proceedings of the 4th ACM/IEEE international conference on human robot interaction* (pp. 245–246). Association for Computing Machinery.
- Salvini, P., Ciaravella, G., Yu, W., Ferri, G., Manzi, A., Mazzolai, B., Laschi, C., Oh, S. R., & Dario, P. (2010, September). How safe are service robots in urban environments? Bullying a robot. In *19th International symposium in robot and human interactive communication* (pp. 1–7). IEEE.
- Schopler, J., & Insko, C. A. (1992). The discontinuity effect in interpersonal and intergroup relations: Generality and mediation. *European Review of Social Psychology*, 3(1), 121–151. <https://doi.org/10.1080/14792779243000041>
- Smith, E. R., Sherrin, S., Fraune, M., & Šabanović, S. (2020). Positive emotions, more than anxiety or other negative emotions, predict willingness to interact with robots. *Personality and Social Psychology Bulletin*, 46, 1270–1283. <https://doi.org/10.1177/0146167219900439>
- Stephan, W. G., & Stephan, C. W. (1985). Intergroup anxiety. *Journal of Social Issues*, 41(3), 157–175. <https://doi.org/10.1111/j.1540-4560.1985.tb01134.x>
- Tankard, M. E., & Paluck, E. L. (2016). Norm perception as a vehicle for social change. *Social Issues and Policy Review*, 10(1), 181–211. <https://doi.org/10.1111/sipr.12022>
- Todd, A. R., & Galinsky, A. D. (2014). Perspective-taking as a strategy for improving intergroup relations: Evidence, mechanisms, and qualifications. *Social and Personality Psychology Compass*, 8(7), 374–387. <https://doi.org/10.1111/spc3.12116>
- Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review*, 117(2), 440–463. <https://doi.org/10.1037/a0018963>
- Turner, J. C., Oakes, P. J., Haslam, S. A., & McGarty, C. (1994). Self and collective: Cognition and social context. *Personality and Social Psychology Bulletin*, 20(5), 454–463. <https://doi.org/10.1177/0146167294205002>
- Vanman, E. J., & Kappas, A. (2019). “Danger, Will Robinson!” The challenges of social robots for intergroup relations. *Social and Personality Psychology Compass*, 13(8), Article e12489. <https://doi.org/10.1111/spc3.12489>
- Williams, T., Briggs, P., & Scheutz, M. (2015). Covert robot–robot communication: Human perceptions and implications for human–robot interaction. *Journal of Human–Robot Interaction*, 4(2), 24–49. <https://doi.org/10.5898/JHRI.4.2.Williams>
- You, S., & Robert, L. P., Jr. (2018, February). Human–robot similarity and willingness to work with a robotic co-worker. In *Proceedings of the 2018 ACM/IEEE international conference on human–robot interaction* (pp. 251–260). Association for Computing Machinery.
- Yuki, M. (2003). Intergroup comparison versus intragroup relationships: A cross-cultural examination of social identity theory in North American and East Asian cultural contexts. *Social Psychology Quarterly*, 66, 166–183. <https://doi.org/10.2307/1519846>

Received April 07, 2020

Revision received October 14, 2020

Accepted October 16, 2020 ■