Linda Onnasch, Eileen Roesler

Anthropomorphizing Robots: The Effect of Framing in Human-Robot Collaboration

Journal article | **Accepted manuscript (Postprint)**
This version is available at https://doi.org/10.14279/depositonce-9020

Accepted to *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*
(ISSN:1071-1813, eISSN:2169-5067)

Anthropomorphizing Robots: The Effect of Framing in Human-Robot Collaboration

Linda Onnasch[1] & Eileen Roesler[2],

[1] Engineering Psychology, Department of Psychology, Humboldt-Universität zu Berlin, Rudower Chaussee 18, 12489 Berlin, Germany

[2] Chair of Work, Engineering & Organizational Psychology, Department of Psychology and Ergonomics, Technische Universität Berlin, Marchstraße 12, 10587 Berlin, Germany

Abstract

Anthropomorphic framing of social robots is widely believed to facilitate human-robot interaction. In two subsequent studies, the impact of anthropomorphic framing was examined regarding the subjective perception of a robot and the willingness to donate money for this robot. In both experiments, participants received either an anthropomorphic or a functional description of a humanoid NAO robot prior to a cooperative task. Afterwards the perceived robot's humanlike perception and the willingness to "save" the robot from malfunctioning were assessed (donation behavior). Surprisingly, the first study revealed a negative effect of anthropomorphic framing on the willingness to donate. This negative effect disappeared if the robot's functional value for the task fulfillment was additionally made explicit (Study 2). In both studies, no effect of anthropomorphic framing on the humanlike perception of the robot was found. However, the behavioral results support the relevance of a functional awareness in social human-robot interaction.

INTRODUCTION

Anthropomorphism is a phenomenon with broad influence on the perception of nonhuman agents like animals, gadgets, or gods. The term itself is often used loosely for the tendency to attribute human characteristics to inanimate objects or animals in order to rationalize their actions. Furthermore, anthropomorphism goes beyond simply attributing life to nonliving objects, as it essentially includes "attributing capacities that people tend to think of as distinctly human to nonhuman agents, in particular humanlike mental capacities (e.g., intentionality, emotion, cognition)." (Waytz, Cacioppo, & Epley, 2010, p. 221). This mechanism gains increasing relevancy in nowadays technologized society, and especially for social robots as physically and socially embodied entities in a natural human environment (Kiesler & Hinds, 2004, Wainer Feil-Seifer, Shell, & Mataric, 2007). Thereby, applying anthropomorphic design aspects to robots aims to support meaningful social interaction and human acceptance of the robot (Duffy, 2003).

Current research focuses on human reactions to anthropomorphism of various social robots in diverse interaction scenarios. Bartneck, Verbunt, Mubin and Al Mahmud (2007) used an insect robot, equipped with light sensors, to investigate if "smarter" robots are perceived more alive. The participants interacted with the robot by using a flashlight to maneuver it through a dark room. Thereby, the robot´s intelligence was manipulated by changing its sensitivity to the light input. To test the assumption that robot´s more intelligent behavior leads to a more life-like perception of it, the willingness to "kill" the robot was assessed. After interacting with the robot, participants were asked to destroy the Crawling Microbug with a hammer. The results supported the assumed relationship between life-likeness and users' (non-) destructive behavior. However, the transferability of the results is not clearly given as the experiment included an atypical HRI scenario (destroying instead of cooperating) and the used robot was morphologically and interactively not comparable with actual social robots and their behavioral spectrum.

Nonetheless, the positive effect of anthropomorphism on human empathy has been supported by Riek, Rabinowitch, Chakrabarti, and Robinson (2009) as they showed that people empathize more with robots, which have a more humanlike than a mechanical appearance. In their web-based survey, participants watched movie clips featuring four robot protagonists, varying in their resemblance to humans. The clips contained either neutral interactions with humans or emotionally evocative ones, in which humans shouted at the robot or pushed it. The results suggested that people show higher empathy scores and willingness to rescue the robot in case of an earthquake, if the robot had a more humanlike design. Even though the robots used in the scenarios were more humanoid than the insect robot used by Bartneck et al. (2007) the stimuli depicting the robots were only two-dimensional and no actual interaction was performed. Furthermore, the willingness to rescue the robot in case of an earthquake recorded only a hypothetical (and very artificial) behavior.

Besides the anthropomorphic design of a robot´s appearance and behavior, the phenomenon of anthropomorphism can also be induced by framing. There are different ways to frame social robots anthropomorphic such as giving them human names (Keay & Graduand, 2011) or personified stories (Darling, 2017). Darling, Nandy and Breazeal (2015) used both for an experiment with an insect robot called Hexbug Nano to investigate the link between anthropomorphic framing and the willingness to destroy a robot. In the anthropomorphic condition, the Hexbug Nano was called "Frank" and his personality was described as friendly and easily distractible. Afterwards participants hesitated longer in destroying the small robot compared to a control condition without a story or name. As mentioned beforehand, the destruction of an insect-like robot is hardly transferrable to realistic social human-robot interaction.

A more recent study by Nijssen, Müller, van Baaren and Paulus (2019) investigated how anthropomorphic framing and the design of robots affect decision making in moral dilemmas. They assumed that anthropomorphic framing would increase the affinity humans feel with the robot as well as moral concerns toward the non-living object. In the online experiment with pictures of different agents (a human, a human-like robot and a machine-like robot), it was found, that regardless of the appearance, participants showed more moral concern and less willingness to sacrifice the agent when framed anthropomorphically. However, a drawback of this study is the data collection via an online survey, which again does not involve any real-life interactions.

Overall, the presented results suggest that anthropomorphic design and anthropomorphic framing change people´s way of interacting with robots. However, the transferability to a realistic and actual human-robot interaction is questionable. Although, Darling et al. (2015) elicit the impact of framing in human-robot interaction, the type of interaction studied in the experiment is atypical for daily human-robot interactions. Instead of being destroyed, robots are expected to socially cooperate with humans (Duffy, 2003), e.g. by providing support for human task fulfilment. Destroying a robot with a mallet is not accompanied with any advantage and rather unrealistic in everyday life interaction. Consequently, the question arises, whether a similar framing effect would occur in a cooperative task and more realistic environment.

Furthermore, the willingness to rescue a robot in case of an earthquake (Riek et al., 2009) or the willingness to sacrifice a robot in case of a social dilemma (Nijssen et al., 2019) raises additional doubts over the transferability to reality, as the scenarios used in these studies focus on fictitious decisions and not on actual behavior of the participants. Moreover, the scenarios are rather extreme (saving or sacrificing) and do not represent typical human-robot workplace scenarios (Bartneck et al., 2007; Darling et al., 2015, Riek et al., 2009). A more suitable situation would be the opportunity to “save” the robot from malfunctioning, e.g. a broken robot, that needs to be fixed. Considering such a situation, one way of operationalizing the act of saving is the assessment of donation behavior towards a repair of the robot.

Therefore, the two reported studies aimed to examine the effect of anthropomorphic framing by taking major drawbacks of the described experiments into consideration. First, a more realistic scenario was used to investigate if people tend to perceive anthropomorphically framed robots as more humanlike than mechanically framed robots. Second, the willingness to donate was assessed to examine whether anthropomorphic framing increases the willingness to support the robot.

STUDY 1

In the first experiment, the impact of framing the robot was assessed in the context of realistic human-robot interaction. A humanoid robot supported the participant to solve a mathematical puzzle. Thereby, the robot served as an assistance that indicated the most effective way to do the puzzle as well as a social companion cooperatively working together with the participant. In the *anthropomorphic framing* condition, the robot was described with a name and a personal story, including individual preferences like favorite color and hobbies. Instead of using a no-framing control condition, a *functional framing* condition was used to explicitly frame the already humanoid robot in a non-anthropomorphic tool-like manner. Therefore, the robot was characterized via technical specification like height and weight. The framing was applied prior to the actual interaction.

It was hypothesized that framing a robot anthropomorphic leads to a higher perceived humanness of the robot. It also was expected that anthropomorphic framing increases robot acceptance (Duffy, 2003) and therefore participants’ willingness to resign from their own maximum benefit in favor of the robot.

Method

Participants. For the first experiment, 40 participants were recruited from the Humboldt-Universität zu Berlin, resulting in a predominantly student sample with an age ranging from 18 to 55 years (*M*=26.5; *SD*=7.58). The majority were female (65%) and had no prior experience with robots (85%). Participants signed consent forms at the beginning of the experiment and received five Euros as compensation at the end of the experiment.

Task and materials. The aim of the human-robot collaboration was to solve a six-disk version of the Tower of Hanoi (Figure 1). In this mathematical game, a stack of disks has to be moved from the leftmost to the rightmost peg by moving only one disk at a time and never placing a larger disk on a smaller one in the fewest possible moves. The custom-made tower was situated approximately 30 cm away from the robot vis-à-vis the participant. The participants' task was to move the disks by following exactly the robot´s directives to solve the Tower of Hanoi in an optimal sequence. The robot served as an assistant in this task, simulating an interaction typically for human-robot social interaction, like teaching, entertaining, comforting or assisting people (Sheridan, 2016).

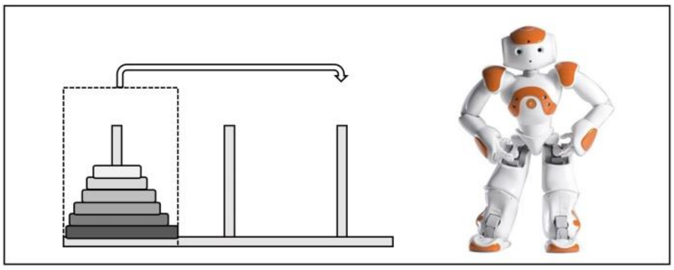

Figure 1. Depiction of the Tower of Hanoi with six disks as a joint task (left) and the NAO robot from Aldebaran Robotics (right).

The robot used in this study was a NAO Robot (Edition 4) from Aldebaran Robotics (Figure 1). The humanoid robot is 58 cm tall, weighs 5 kg and is equipped with a 1.6 GHz Intel ATOM Z530 processor and two high definition cameras. In order to program the required sequences of moves the software Choregraphe 2.1.4.13 was used. The programmed behavior of the robot included head and arm movements in the following chronology. First, the robot pointed towards one peg as a sign to remove the top disk from this peg (with head and arm movements). Subsequently the robot's movements pointed towards another peg as a prompt to place the previously picked disk there. Thereby, the robot's movement sequences always indicated the most efficient way to solve the Tower of Hanoi within 63 steps. Besides the consistent task relevant movements, no other interaction was possible. The robot’s functions and movement patterns were the same in both framing conditions.

Design. One group was framed anthropomorphic and the other one was framed functional via written instructions. In the anthropomorphic framing condition, participants received a description of NAO as a team member with humanlike characteristics like enjoying to play football, having a favorite color and being friendly. In the functional framing condition, the robot was described via technical information and specifications like height, weight and sensors. The task instruction matched the framing condition by using “NAO” and “him” in the anthropomorphic one instead of “the robot” and “it”. The framing manipulations resulted in a one factorial between-subjects design.

Dependent measures. To investigate how framing of the robot affected participant´s perception of the robot, a German version of the revised Godspeed questionnaire (Ho & MacDorman, 2010) with the dimensions *perceived humanness, eeriness,* and *attractiveness* was used.

Apart from this subjective measurement, the willingness to “save” the robot was assessed via pretended voluntary donations for a necessary repair of the robot. Thereby, participants were informed through a written information on the top of the donation sheet that one component of the robot was broken and money was collected to repair the malfunction. Participants were asked whether they would like to donate a part of their received money (max. five Euros) for this purpose. If participants were willing to donate, they had to fill in the donation list and state the exact amount they wanted to donate. However, no donations were actually collected and participants were debriefed before leaving.

Procedure. All subjects were assigned randomly in either the anthropomorphic or the functional framing condition and received corresponding written instructions (framing). Afterwards task instructions for the Tower of Hanoi were presented. The collaborative game lasted for approximately six minutes. Subsequently, subjects were asked to fill in the translated version of the revised Godspeed questionnaire (Ho & MacDorman, 2010). The monetary compensation was paid to pretend the end of the experiment after which they were told that the institute is collecting money to fix a part of the robot. In case of willingness to donate, participants completed a donation list, stating how much money they wanted to spend. After this, all participants were debriefed and obtained the complete amount of money. The entire experiment lasted 25 to 30 minutes.

Results

Perception of the robot. The overall score of the revised Godspeed indices revealed comparably high ratings in the functional (M=3.07, SD=0.29) and anthropomorphic framing (M=3.18, SD=0.57) condition. The highest ratings were given on the humanness scale with almost identical scores in the functional (M=3.94, SD=0.57) and anthropomorphic (M=3.89, SD=0.62) framing conditions. Ratings for the other two dimensions, eeriness and attractiveness, were lower but did not differ between the framing conditions either. Independent-samples t-test revealed neither significant differences between the framing conditions on the overall score (t(38)=0.75, p=.46) nor on the dimensions humanness (t(38)=-0.27, p=.79), eeriness (t(38)=0.37, p=.72) or attractiveness (t(38)=1.3, p=.19).

Willingness to donate. 18 out of 20 participants were willing to donate in the functional framing condition. In the anthropomorphic framing condition, only 10 out of 20 were willing to resign from maximum benefit (Figure 2). The results of a chi-square test showed a significant association between framing and whether or not participants donated ($\chi^2(1,N=40)=7.619$, p=.006).

Donation amount. The donation amount was analyzed for participants who were willing to donate money (N=28). The individually donated amount differed between one and five Euro (which was the amount participants received for participation). Participants of the functional framing condition donated more (M=2.25, SD=1.63) than participants in the anthropomorphic framing condition (M=1.9, SD=1.2). This difference between framing conditions (Figure 2) was not statistically significant as an independent-samples t-test revealed (t(26)=0.59, p=.09).

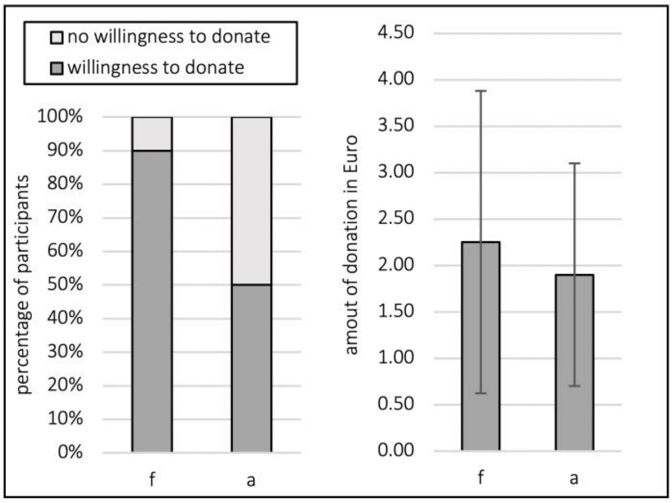

Figure 2. Results from study 1: Participant's willingness to donate (left); means and standard deviation of donation amount in Euro for functional (f) and anthropomorphic (a) framing condition (right).

Discussion

The aim of the first experiment was to examine the effect of anthropomorphic framing in a realistic social human-robot interaction scenario. Thereby, the humanlike perception of the robot as well as the actual human donation behavior in order to repair the either functionally or anthropomorphically framed robot was analyzed.

The expected differences in human perception induced by anthropomorphic framing were not supported by the results. Participants did not perceive the anthropomorphically framed robot as more humanlike, eerie or attractive compared to the functional framed one. A major reason for this result might be that the robot used in the study was an already humanoid robot. Findings of Riek et al. (2009) indicate that robots, classified as humanoid like NAO, make people feel more empathic towards the robot and lead to attributions of mental and emotional states to the robot. The overall high mean scores of humanness above three on a one to five point Likert scale in the current study support this presumption.

Surprisingly, and contrary to previous findings (Darling, 2017; Nijssen et al., 2019; Riek et al., 2009), more participants in the functional, not anthropomorphic, condition were willing to donate their money for the robot repair (18 out of 20).

Therefore, the results do not support the assumption that anthropomorphism is always beneficial in HRI. Participants might have associated more value as a technical sophisticated research tool to the functionally framed robot. Thus, the anthropomorphic framing might have masked the robot´s tool-like status and importance for task fulfillment. Therefore, in some contexts of HRI, it might be beneficial to have countermeasures to anthropomorphism in order to maintain a certain tool status of robots.

To test the post-hoc interpretation of results, a second experiment was conducted to investigate framing effects when explicitly stating the functional value of the robot.

STUDY 2

Following on from Study 1, the second experiment had the aim to investigate framing effects if the functional value of the robot is additionally explicated. Therefore, the framing was supplemented with a paragraph about the robot’s functional value and relevance in the context of task fulfillment. It was assumed that the not-intended effects of anthropomorphism (loss of tool-status and therefore not perceiving its functional value) could be reduced by an explicit description of value.

Method

The task and the robot as well as the design, dependent measurements and procedure were the same as in Study 1. The only difference was the attachment of additional information on the written sheet used for the induction of framing.

Participants. For the second experiment, again 40 participants, who had not taken part in the first study, were recruited from the Humboldt-Universität zu Berlin. The participants age ranged from 18 to 35 (*M*=25.83, *SD*=6.67). Again, the majority were female (65%) and had no prior experience with the robot used in this experiment (95%). All received five Euros as compensation at the end of the experiment.

Additional materials. The information about the robot was supplemented in both framing conditions by a paragraph that stated the *robot´s value*, i.e. the robot´s relevance for the successful task execution and the research activities of the lab in general, highlighting the robot's status as a tool in the overall context of task fulfillment. The paragraph was, as all written information, linguistically adapted for both framing conditions. In the anthropomorphic condition the robot was referred to as "NAO" and "him" instead of "the robot" and "it".

Results

Perception of the robot. The results of study 2 are aligned with those of study one. First, all dimensions of the translated revised Godspeed questionnaire were rated nearly similar in the functionally and anthropomorphically framed conditions. Again, the highest ratings were given on the humanness scale with almost identical scores in the functional (*M*=3.75, *SD*=0.56) and anthropomorphic (*M*=3.9, *SD*=0.45) framing conditions. The global score as well showed a nearly equal perception of the robot in the functional (*M*=3.07, *SD*=0.47) and anthropomorphic framing (*M*=3.12, *SD*=0.4) condition. As in study 1, independent-samples t-test revealed neither significant differences between the framing conditions on the overall score ($t(38)=0.37$, $p=.36$) nor on the dimensions humanness ($t(38)=-0.1$, $p=.76$) eeriness ($t(38)=0.11$, $p=.2$) or attractiveness ($t(38)=0.93$, $p=.28$).

Willingness to donate. The willingness to donate was nearly equal in both framing conditions, with thirteen out of twenty in the functional framing condition and twelve out of twenty in the anthropomorphic framing condition (Figure. 3, left). No significant difference between framing and whether participants donated ($\chi^2(1,N=40)=0.11$, $p=.744$) was indicated by the results of a chi-square test.

Donation amount. The donation amount was analyzed for participants who were willing to donate money (*N*=15). The donated amount was higher in the functional framing condition

(M=3.62, SD=1.62) than in the anthropomorphic framing condition (M=2.96, SD=1.59; Figure 3, right). This difference between framing conditions was not statistically significant as an independent-samples t-test revealed ($t(23)$=1.02, p=.59).

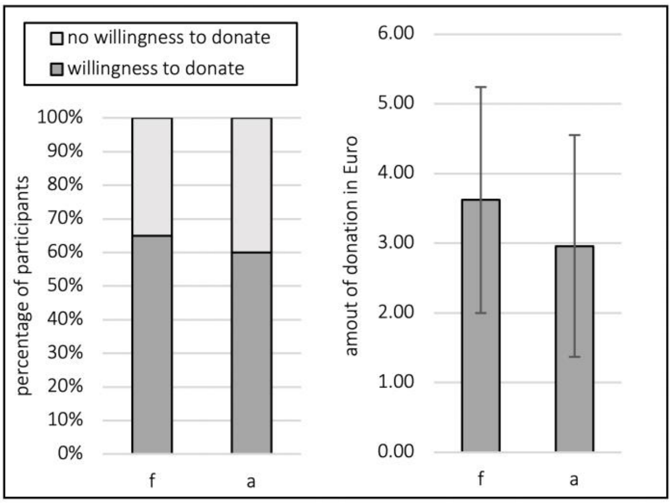

Figure 3. Results from study 2: Participant's willingness to donate (left); means and standard deviation of donation amount in Euro for functional (f) and anthropomorphic (a) framing condition (right).

Discussion

The second study empirically supported and extended the findings of Study 1 in several aspects. First, participants in both studies did not perceive the anthropomorphically framed robot as more humanlike than the functional framed one. The results of the two presented studies indicate that previous findings on non-humanoid robots (Darling, Nandy & Breazeal, 2015) might not be generalizable for humanoid robots. Further research is needed on the effectiveness of anthropomorphic framing regarding already humanoid robots, especially because of the already broad application of humanlike robots in social robotics (Duffy, 2003).

Second, the willingness to "save" the robot depended on the robot´s attributed value and not the robot´s framing. The negative effect of anthropomorphic framing found in Study 1 approximated to a likewise level if the value of the robot for research on human-robot interaction and task fulfillment was made explicit. This emphasizes the relevance for a functional awareness of technology and empirically underpins that "anthropomorphism and emotional bonding are undesirable, for example when this would diminish the function of the technology" (Darling, 2017, p.173).

CONCLUSION

Overall, both studies do not support the conventional wisdom that anthropomorphic framing facilitates human-robot interaction on principle. The results suggest a negative effect of anthropomorphic framing, if the robot´s functional value for task fulfillment is not explicitly emphasized.

Framing might not have a strong effect on perceived anthropomorphism when interacting with humanoid robots. In this case, the robot's morphology dominates the perception of human likeness. However, the results clearly support the relevance of emphasizing the robot´s value for a cooperative task. Study 1 showed that with functional framing, more people were willing to resign from maximum benefit to ensure the proper functioning of the robot. Study 2 revealed that an emphasis of the functional value of the robot is a successful countermeasure for possible negative effects of anthropomorphism. In this case, framing did not have an impact on participants' behavior. The explication of the functional robot's value led people to donate even higher amounts of money (in both conditions) compared to findings of study 1. Future research should therefore investigate the relevance of framing effects in human-robot interaction with already highly functional robots, like industrial robots. In these cases, where the robot's functional value is clearly apparent, framing might not have negative effects on people's behavior. However, in terms of non-humanoid robots, framing could still be beneficial for the perception of the robot in terms of higher acceptance and likeability.

Besides, the studies additionally contribute to the body of research using realistic scenarios and implementing actual and embodied interactivity between human and robot instead of applying hypothetical and very artificial HRI setups. So far, many studies have used pictures or videos to evaluate human response to different robots. Nonetheless, two-dimensional illustrations are not able to depict the complex three-dimensional appearance, movements and sounds of social human-robot interaction (e.g. Wainer et al., 2007). Future research needs to take the embodied nature of human-robot interaction into account to achieve a sound transferability from the lab into the real world.

References

Bartneck, C., Verbunt, M., Mubin, O., & Al Mahmud, A. (2007). To kill a mockingbird robot. In *Proceedings of the ACM/IEEE International Conference on Human-robot Interaction* (81–87). New York, NY, USA: ACM.

Darling, K. (2017). "Who's Johnny?" Anthropomorphic framing in human-robot: Interaction, integration, and policy. In *Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence* (173–188).

Darling, K., Nandy, P., & Breazeal, C. (2015). Empathic concern and the effect of stories in human-robot interaction. In *2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)* (770–775).

Duffy, B. R. (2003). Anthropomorphism and the social robot. *Robotics and Autonomous Systems*, *42*(3–4), 177–190.

Ho, C. C., & MacDorman, K. F. (2010). Revisiting the uncanny valley theory: Developing and validating an alternative to the Godspeed indices. *Computers in Human Behavior*, *26*(6), 1508-1518.

Keay, A., & Graduand, M. (2011). Emergent phenomena of robot competitions: Robot identity construction and naming. In *Advanced Robotics and its Social Impacts* (12–15).

Kiesler, S., & Hinds, P. (2004). Introduction to this special issue on human-robot interaction. *Human - Computer Interaction :* Philadelphia, PA :

Nijssen, S. R. R., Müller, B. C. N. ,van Baaren, R. B., & Paulus, M.(2019). Saving the robot or the human? Robots who feel deserve moral care. *Social Cognition*, *37*(1), 41–56.

Riek, L. D., Rabinowitch, T.-C., Chakrabarti, B., & Robinson, P. (2009). How anthropomorphism affects empathy toward robots. In *Proceedings of the 4th ACM/IEEE International Conference on Human Robot Interaction* (245–246). New York, NY, USA: ACM.

Sheridan, T. B. (2016). Human-robot interaction: Status and challenges. *Human Factors*, *58*(4), 525–532.

Wainer, J., Feil-Seifer, D. J., Shell, D. A., & Mataric, M. J. (2007). Embodiment and human-robot interaction: A task-based perspective. In *RO-MAN 2007-The 16th IEEE International Symposium on Robot and Human Interactive Communication* (872-877). IEEE.

Waytz, A., Cacioppo, J., & Epley, N. (2010). Who sees human? The stability and importance of individual differences in anthropomorphism. *Perspectives on Psychological Science : A Journal of the Association for Psychological Science*, *5*(3), 219–232.