Available online at www.sciencedirect.com
ScienceDirect
Current Opinion in Psychology

Review

Prosocial behavior toward machines

Yngwie Asbjørn Nielsen[1], Stefan Pfattheicher[1] and Merel Keijsers[2]

Abstract
Building on the *computers are social actors* framework, we provide an overview of research demonstrating that humans behave prosocially toward machines. In doing so, we outline that similar motivational and cognitive processes play a role when people act in prosocial ways toward humans and machines. These include perceiving the machine as somewhat human, applying social categories to the machine, being socially influenced by the machine, and experiencing social emotions toward the machine. We conclude that studying prosocial behavior toward machines is important to facilitate proper functioning of human–machine interactions. We further argue that machines provide an interesting yet underutilized resource in the study of prosocial behavior because they are both highly controllable and humanlike.

Addresses
[1] Aarhus University, Denmark
[2] University of Canterbury, New Zealand

Corresponding author: Nielsen, Yngwie Asbjørn (yan@psy.au.dk)

Current Opinion in Psychology 2022, **43:**260–265

This review comes from a themed issue on **Prosociality (2022)**

Edited by **Stefan Pfattheicher** and **Isabel Thielmann**

For complete overview about the section, refer Prosociality (2022)

Available online 12 August 2021

https://doi.org/10.1016/j.copsyc.2021.08.004

Keywords
Anthropomorphism, Computers are social actors, Human–machine interaction, Prosocial behavior.

Introduction

Have you ever angrily blamed the computer for producing *yet another* error message during data analysis, knowing full well that it does not possess agency? Or have you ever apologized for getting in the way of a robot vacuum, despite knowing that robots do not have feelings? Humans interact with machines socially, even to the point of showing prosocial behavior, defined as “any type of voluntary behaviour that is intended to help others, while guaranteeing no rewards to the helper and requiring some kind of cost (e.g. money, time) to the person” [1•, p. 2]. The goal of this review is threefold. First, we will provide an overview of recent studies demonstrating prosocial responses to machines; second, we will highlight similarities in the motivational and cognitive processes underlying prosocial behavior toward humans and machines; and third, we will argue that research on human–machine interaction has interesting practical and methodological implications. To begin, we acquaint the reader with the *computers are social actors* framework [2–4] for understanding social behavior toward machines.

Computers are social actors

In the 1990s, Nass and colleagues conducted a series of experiments that would lay the foundation for the *computers are social actors* framework (hereinafter “CASA”). They followed a simple formula: take an experiment from social psychology, substitute the key people (i.e., those typically played by research confederates) with computers, and run the experiment to test if the original results replicated despite the substitution. Quite often, it would (see [3,4] for reviews). One paradigmatic example comes from a study on reciprocity [5]. Participants first used a computer to conduct a web search that yielded helpful results. Half were then moved to a different computer. Participants then completed a second task where they could spend any amount of time “helping” the computer generate color palettes. Participants seated at their original “helpful” computer spent more time “helping” their computer than those who were moved to a different computer prior to the second task. In other words, people reciprocated help received from computers.

What might explain this social behavior toward the machines? Nass and Moon [3] described social behavior toward machines as *mindless* (see also [4]). They posit that some cues activate social scripts (i.e., expectations for the sequence of events in a particular social setting) and automatic social behavior. Once the script is active, the search for more cues—ones that might reveal the social behavior is inappropriate or not applicable—is terminated. Thus, the automatic social responses occur despite the apparently nonsocial conditions. In short, the CASA framework regards social responses to machines as automatic and cued.

Although initially proposed within the realm of computers, CASA has also been applied in research into robots and virtual agents [6•]. Throughout the article, we will refer collectively to these as "machines." In the following, we review four psychological processes involved in prosocial behavior toward machines, all of which were identified in the literature on human–human interaction and subsequently extended to machines. As such, each process also illustrates a point of convergence between human-directed and machine-directed prosocial behavior. These processes are (i) perceiving humanness, (ii) social categorization, (iii) social influence, and (iv) affectively motivated prosociality.

Machines are perceived as (somewhat) human

Perceived humanness is a matter of degree, that is, perceptual phenomena are not just categorized as either human or not, but as *more or less* human. In turn, perceived humanness has a major influence on behavior; people who are *dehumanized* are more likely to be aggressed against, whereas, conversely, people who are perceived as more human (e.g., capable of feeling complex emotions) are more likely to receive help [7]. Perceiving humanness also extends to machines as is apparent from the reaction of a colonel viewing a test run of a land mine detection robot: "The colonel just could not stand the pathos of watching the burned, scarred and crippled machine drag itself forward on its last leg. This test, he charged, was *inhumane*" [8, emphases added]. In scientific terms, we *anthropomorphize* machines, that is we "imbue the real or imagined behavior of nonhuman agents with humanlike characteristics, motivations, intentions, or emotions" [9, p. 864]. Anthropomorphism functions as a *metaphor*; we apply our vast understanding of humans in order to comprehend a non-human agent "as if" it were a human [9,10]. This is especially true when the agent has a high number of humanlike features or traits or when one has an increased need for either company or for effectively interacting with one's environment [9]. Anthropomorphism is not necessarily conscious; instead, Kim and Sundar [11] argue that it occurs mindlessly and is thus consistent with CASA.

Anthropomorphism is linked to both prosocial and antisocial behavior toward machines. Recently, Keijsers and Bartneck [12] found a negative relationship between mind attribution—a central part of anthropomorphism—and verbally bullying a robot in a video game. Furthermore, Bartneck et al. [13] found that people hesitate longer when switching off an intelligent and agreeable robot compared to an unintelligent and disagreeable one. However, physical manipulations of anthropomorphism such as robot shape and interactivity have received mixed support (see [1•] for a review), possibly because the relation between physical human-likeness and perceived human-likeness is rather complex; for instance, conflicting machinelike and humanlike cues can make a machine seem eerie or "uncanny" [14]. In sum, to the degree that machines are ascribed humanlike characteristics, they elicit humanlike treatment.

Machines are socially categorized

Humans belong to a variety of different *social categories* (e.g., gender, ethnicity, age group). Whenever these are salient, we use them to guide our judgments, attitudes, and prosocial behavior [15]. Once again, the same is true of machines [16•,17••]. For example, people attribute gender to machines: Siegel et al. [18] showed that men donate more money to a female-voiced robot versus a male-voiced robot. Along the same lines, people also apply gender stereotypes. In a study by Eyssel and Hegel [19], participants would rather cooperate on a math task—versus a verbal task—when paired with a male (i.e., short-haired) robot, but showed no preference for task type when paired with a female (i.e., long-haired) robot (see also [20] for a recent conceptual replication). Accordingly, people conform more to the advice of a male compared to a female computer when quizzed on a stereotypically masculine topic [21] (but see [22]).

Besides social categorization, machines may be subject to social identification and the accompanying intergroup biases. Indeed, Fraune [23] demonstrated ingroup favoritism in minimal groups: participants were given a red or blue wristband and assigned to a team of humans and robots with matching colors. Next, they played a game against a differently colored team of humans and robots in which they had to assign noise blasts to all players (including their own teammates). Strikingly, participants administered louder noise blasts to an outgroup human compared to an ingroup robot, indicating that ingroup favoritism extends to robots (see also [17••] for a state-of-the-art review). In addition, Jong et al. [24] recently demonstrated that team membership increased elicited empathy toward a robot in an online game. Finally, multiple studies have manipulated ethnic identification, for example, by informing participants that a machine had a foreign versus a familiar name or country of origin [25]. Such studies have found that machines belonging to an ethnic ingroup are anthropomorphized more [25,26] and induce more cooperative behavior [27,28]. In sum, the prosocial treatment of machines depends, in part, on which social categories they are assigned to.

Machines exert social influence

If machines are perceived as somewhat human and are categorized socially, it should come as no surprise that they also elicit *social influence* (i.e., induce behavior change by virtue of their social presence). Of interest, responses to machines follow the norm of reciprocity (cf. the prior example from [5]). Lee and Liang [29,30] found that people are more willing to help a machine if it provided helpful (versus unhelpful) hints in a prior trivia game. Relatedly, Sandoval et al. [31] found that humans reciprocate the decisions of robots in economic games.

Beyond simple reciprocity, machines exert social influence in a variety of ways. Recently, Lee and Liang [32] demonstrated the foot-in-the-door technique, that is, robots that made an initial small request had greater success at persuading people to commit to a large request. Under conditions of high ambiguity, people feel pressured to conform—and, indeed, do conform—to groups of robots in Asch-like scenarios [33]. There is also preliminary evidence of robots inducing pluralistic ignorance: Connolly et al. [34] recently found that a mistreated robot was more likely to receive help if bystander robots displayed a sad versus a neutral expression. Finally, Riether et al. [35] demonstrated a social facilitation effect, whereby humans perform better in easy tasks and worse in complex tasks if a robot is present versus absent. In sum, machines can induce prosocial treatment via social influence.

Machines evoke emotions

When it comes to machines, most people recognize their capability to arouse emotions, as the opening example of the anger-inducing error messages shows. Interestingly, emotions are complicit in many forms of prosocial behavior—even those directed toward machines. For instance, de Melo et al. [36] demonstrated that compared to humans, people exhibit equally envious behavior toward a computer with a better payoff in an economic game but appear to be less guilt-ridden by exploiting the computer in said game.

The empathy-altruism hypothesis, which implies that empathic concern elicits a motivation to help others [37], has also been replicated with machines. When people are exposed to videos of robot mistreatment (as compared to a neutral video of the robot), they report feeling more pity toward the robot, as well as higher negative affect and lower positive affect [38]. Furthermore, a study using encephalographic measures indicates that a similar empathic response is activated by watching painful videos of human hands compared to robot hands [39]. Crucially, empathic concern toward machines leads to downstream behavior consequences. For example, knowing the background story of a bug robot makes people more hesitant to strike it with a mallet, likely due to increased anthropomorphism and empathic concern [40]. In another example, Briggs and Scheutz [41] had participants watch a robot proudly build a tower before commanding the robot to knock it over. If the robot started sobbing and pleading to keep the tower, participants felt more discomfort and were more hesitant to insist on their command. Overall, machines and humans alike elicit prosocial behavior through empathic concern.

Implications for future research

Researching human–machine interaction has both practical and methodological implications. In the applied domain, roboticists have already integrated robots into areas such as transportation, health care, and education—a trend that will likely persist (see [42••] for a recent review). However, for machines to properly fulfill their function, they must be treated prosocially. For instance, delivery robots in hospitals rely on nurses and housekeepers for loading and unloading [43]; robot vacuums need help getting un-stuck from furniture [44]; and autonomous vehicles need to be respected in traffic, despite the fact that they brake automatically for jaywalkers [45]. Thus, it is important to understand and foster prosocial behavior toward machines. Indeed, encouraging any kind of prosocial behavior may be an end in itself [1•].

As for methodological implications, Wykowska [46•] recently argued for the use of robots in social cognition research, stating,

> Robots can be informative with respect to understanding the human mind because they can be tools to study social cognition through their embodied physical (and perhaps even social) presence, which provides higher ecological validity than screen-based stimuli and better experimental control than human–human interaction. (p. 1).

The present review suggests that similar processes underlie human-directed and machine-directed prosocial behavior. As such, it seems reasonable to suggest that machines can be applied in the service of understanding human-directed prosociality (see [47] for similar reasoning). Like Wykowska [46•], we argue that machine behavior is more controlled than research confederate behavior and arguably more ecologically valid than primitive stimuli. In addition, using machines as "research confederates" eliminates the need for blinding procedures, possibly reduces the need for deception (see [48]), and makes more experimental procedures morally permissible, such as Milgram's obedience paradigm [49] (but see [50]). In the future, machines might even allow for reliable manipulation of complex variables such as human-likeness and mind attribution, which are central in areas such as social cognition research or moral psychology. Finally, the CASA framework asserts that behavior toward machines is mindless; thus, applying machines might allow the researcher to focus on the automatic aspects of prosocial behavior. In sum, applying machines has the potential to enrich the "tool kit" of prosociality researchers.

Conclusion

For thirty years, researchers have used evidence from (social) psychology as a means to improve the design of robots, computers, and more. Despite this, going in the opposite direction and applying machines (or findings from human—machine interaction) to the study of human behavior seems to be a rare occurrence. This is unfortunate, as there are clear similarities between how people interact with other humans and how they interact with machines. With regard to prosocial behavior, these similarities include diverse topics such as dehumanization, stereotypes, intergroup biases, norms, social influence, and empathy. At the same time, applying machines in research allows for highly controllable, human-like, and more ethically viable research confederates. Although machines have already been applied to the study of social cognition [46•], this conclusion leaves plenty of unexplored territory for psychologists to map out. In fact, we contend that studying prosocial behavior toward machines has a unique and untapped potential for testing theories. For instance, in the context of empathy, one can experimentally vary stereotypical appearance and behavior of a human-like robot along central dimensions (e.g., warmth, competence) and test under which conditions people behave most empathically toward the robot. Similarly, in the context of dehumanization, one can study which features (e.g., appearance, agency) are conducive to the attribution of humanness [46•]. Finally, in the context of aggression, one can study highly harmful aggressive behavior, as seen in Milgram's obedience paradigm [49] (harming machines, however, introduces new ethical concerns because doing so is often unpleasant to the actor [50]).

With these opportunities in mind, we wish to identify some challenges for future research. First, not all interactions with machines are mindless (e.g., [51]). Similarly, some prosocial behavior (e.g., teaching a child how to tie their shoes) might only occur mindfully. Thus, understanding the boundary conditions of how prosocial behavior is related to the CASA framework is necessary. Second, many studies within the field of human—machine interaction apply small sample sizes [1•], and future research should seek to apply larger samples and to preregister confirmatory research. Third, we have stressed the similarities, but it is equally important to recognize the differences between human-directed and machine-directed prosociality. There is plenty of evidence that people discriminate between humans and machines (e.g., [36,49]), and the researcher can expect smaller effect sizes when applying machines instead of humans—machines are perceived as no more than *somewhat* human. Finally, as technology evolves, the CASA framework faces new challenges. Compared to at the time of its conception, the CASA of today must account for longer, more complex interactions between ever more experienced machine users and increasingly anthropomorphic machines (see [6•]). Surely, interesting times lie ahead for the study of prosocial behavior toward machines.

Conflict of interest statement

Nothing declared.

Acknowledgements

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Papers of particular interest, published within the period of review, have been highlighted as:

* of special interest
** of outstanding interest

1. • Oliveira R, Arriaga P, Santos FP, Mascarenhas S, Paiva A: **Towards prosocial design: a scoping review of the use of robots and virtual agents to trigger prosocial behaviour**. *Comput Hum Behav* 2021, **114**, https://doi.org/10.1016/j.chb.2020.106547. Article 106547.
This systematic review (k = 19) is the first of its kind to address research on robots and virtual agents as elicitors of prosocial behavior. The authors highlight the diversity in extant research methods, areas with mixed support, and—crucially—the problems of low statistical power as well as conceptual disagreements within the field.

2. Nass CI, Steuer J, Tauber ER: **Computers are social actors**. In *Proc. SIGCHI Conf. Hum. Factors Comput. Syst.* New York, NY, USA: Association for Computing Machinery; 1994:72–78, https://doi.org/10.1145/191666.191703.

3. Nass CI, Moon Y: **Machines and mindlessness: social responses to computers**. *J Soc Issues* 2000, **56**:81–103, https://doi.org/10.1111/0022-4537.00153.

4. Reeves B, Nass CI: *The media equation: how people treat computers, television, and new media like real people and places.* New York, NY, US: Cambridge University Press; 1996.

5. Fogg B, Nass C: **How users reciprocate to computers: an experiment that demonstrates behavior change**. In *CHI 97 ext. Abstr. Hum. Factors Comput. Syst.* New York, NY, USA: Association for Computing Machinery; 1997:331–332, https://doi.org/10.1145/1120212.1120419.

6. • Gambino A, Fox J, Ratan R: **Building a stronger CASA: extending the computers are social actors paradigm**. *Hum-Mach Commun* 2020, **1**, https://doi.org/10.30658/hmc.1.5.
This theoretical article argues that the basic CASA paradigm should be extended to include human–machine scripts (as opposed to *just* human–human scripts) in order to keep up with the technological advances of the past 30 years.

7. Haslam N, Loughnan S: **Dehumanization and infrahumanization**. *Annu Rev Psychol* 2014, **65**:399–423, https://doi.org/10.1146/annurev-psych-010213-115045.

8. Garreau J: *Bots on the ground, Wash.* Post; 2007. http://www.washingtonpost.com/wp-dyn/content/article/2007/05/05/AR2007050501009.html. Accessed 10 April 2021.

9. Epley N, Waytz A, Cacioppo JT: **On seeing human: a three-factor theory of anthropomorphism**. *Psychol Rev* 2007, **114**: 864–886, https://doi.org/10.1037/0033-295X.114.4.864.

10. Duffy BR: **Anthropomorphism and the social robot**. *Robot Autonom Syst* 2003, **42**:177–190, https://doi.org/10.1016/S0921-8890(02)00374-3.

11. Kim Y, Sundar SS: **Anthropomorphism of computers: is it mindful or mindless?** *Comput Hum Behav* 2012, **28**:241–250, https://doi.org/10.1016/j.chb.2011.09.006.

12. Keijsers M, Bartneck C: **Mindless robots get bullied**. In *Proc. 2018 ACMIEEE Int. Conf. Hum.-Robot interact.* New York, NY, USA: Association for Computing Machinery; 2018:205–214, https://doi.org/10.1145/3171221.3171266.

13. Bartneck C, van der Hoek M, Mubin O, Al Mahmud A: **"Daisy, daisy, give me your answer do!" Switching off a robot**. In *2007 2nd ACMIEEE Int. Conf. Hum.-Robot interact. HRI*; 2007: 217–222, https://doi.org/10.1145/1228716.1228746.

14. Kätsyri J, Förger K, Mäkäräinen M, Takala T: **A review of empirical evidence on different uncanny valley hypotheses: support for perceptual mismatch as one road to the valley of eeriness**. *Front Psychol* 2015, **6**, https://doi.org/10.3389/fpsyg.2015.00390. Article 390.

15. Dovidio JF, Gaertner SL: **Intergroup bias**. In *Handb. Soc. Psychol.*. Edited by Fiske ST, Gilbert DT, Lindzey G. 5th ed., Hoboken, NJ, US: John Wiley & Sons; 2010:1084–1121, https://doi.org/10.1002/9780470561119.socpsy002029.

16. • Vanman EJ, Kappas A: **"Danger, Will Robinson!" the challenges of social robots for intergroup relations**. *Soc Personal Psychol Compass* 2019, **13**, https://doi.org/10.1111/spc3.12489. Article e12489.

This theoretical article posits that similar mechanisms underlie both the experience of empathy towards robots and the perception of robots as an existential threat to humans. The authors discuss this "paradox" in terms of potential future challenges for human–robot intergroup relations. Of interest, the authors' paradox provides a novel explanation for the mixed effect of anthropomorphism manipulations.

17. •• Smith ER, Šabanović S, Fraune MR: **Human-robot interaction through the lens of social psychological theories of intergroup behavior**. *Technol Mind Behav* 2021, **1**, https://doi.org/10.1037/tmb0000002.

This review article on human–robot intergroup relations covers research into perceptions of robot groups and interventions aimed at reducing prejudices towards robots. Notably, the paper includes 12 previously unpublished null results from the US and Japan.

18. Siegel M, Breazeal C, Norton MI: **Persuasive robotics: the influence of robot gender on human behavior**. In *2009 IEEERSJ Int. Conf. Intell. Robots syst.*; 2009:2563–2568, https://doi.org/10.1109/IROS.2009.5354116.

19. Eyssel F, Hegel F: **(S)he's got the look: gender stereotyping of robots**. *J Appl Soc Psychol* 2012, **42**:2213–2230, https://doi.org/10.1111/j.1559-1816.2012.00937.x.

20. Bernotat J, Eyssel F, Sachse J: **The (fe)male robot: how robot body shape impacts first impressions and trust towards robots**. *Int J Soc Robot* 2021, **13**:477–489, https://doi.org/10.1007/s12369-019-00562-7.

21. Lee E-J: **Effects of "gender" of the computer on informational social influence: the moderating role of task type**. *Int J Hum-Comput Stud* 2003, **58**:347–362, https://doi.org/10.1016/S1071-5819(03)00009-0.

22. Kuchenbrandt D, Häring M, Eichberg J, Eyssel F, André E: **Keep an eye on the task! How gender typicality of tasks influence human–robot interactions**. *Int J Soc Robot* 2014, **6**:417–427, https://doi.org/10.1007/s12369-014-0244-0.

23. Fraune MR: **Our robots, our team: robot anthropomorphism moderates group effects in human–robot teams**. *Front Psychol* 2020, **11**, https://doi.org/10.3389/fpsyg.2020.01275. Article 1275.

24. de Jong D, Hortensius R, Hsieh T-Y, Cross ES: *Empathy and schadenfreude in human–robot teams*. 2020, https://doi.org/10.31234/osf.io/2wyxp.

25. Eyssel F, Kuchenbrandt D: **Social categorization of social robots: anthropomorphism as a function of robot group membership**. *Br J Soc Psychol* 2012, **51**:724–731, https://doi.org/10.1111/j.2044-8309.2011.02082.x.

26. Eyssel F, Loughnan S: **"It don't matter if you're black or white"?** In *Soc. Robot.*. Edited by Herrmann G, Pearson MJ, Lenz A, Bremner P, Spiers A, Leonards U, Cham: Springer International Publishing; 2013:422–431, https://doi.org/10.1007/978-3-319-02675-6_42.

27. Häring M, Kuchenbrandt D, André E: **Would you like to play with me? How robots' group membership and task features influence human–robot interaction**. In *2014 9th ACMIEEE Int. Conf. Hum.-Robot interact. HRI*; 2014:9–16, https://doi.org/10.1145/2559636.2559673.

28. Pratt JA, Hauser K, Ugray Z, Patterson O: **Looking at human–computer interface design: effects of ethnicity in computer agents**. *Interact Comput* 2007, **19**:512–523, https://doi.org/10.1016/j.intcom.2007.02.003.

29. Lee SA, Liang Y: **Reciprocity in computer–human interaction: source-based, norm-based, and affect-based explanations**. *Cyberpsychol, Behav Soc Netw* 2015, **18**, https://doi.org/10.1089/cyber.2014.0458.

30. Lee SA, Liang Y: **The role of reciprocity in verbally persuasive robots**. *Cyberpsychol, Behav Soc Netw* 2016, **19**:524–527, https://doi.org/10.1089/cyber.2016.0124.

31. Sandoval EB, Brandstetter J, Obaid M, Bartneck C: **Reciprocity in human-robot interaction: a quantitative approach through the prisoner's dilemma and the ultimatum game**. *Int J Soc Robot* 2016, **8**:303–317, https://doi.org/10.1007/s12369-015-0323-x.

32. Lee SA, Liang Y: **Robotic foot-in-the-door: using sequential-request persuasive strategies in human-robot interaction, Comput**. *Hum Behav* 2019, **90**:351–356, https://doi.org/10.1016/j.chb.2018.08.026.

33. Salomons N, Sebo SS, Qin M, Scassellati B: **A minority of one against a majority of robots: robots cause normative and informational conformity**. *ACM Trans Hum-Robot Interact* 2021, **10**:1–22, https://doi.org/10.1145/3442627.

34. Connolly J, Mocz V, Salomons N, Valdez J, Tsoi N, Scassellati B, Vázquez M: **Prompting prosocial human interventions in response to robot mistreatment**. In *Proc. 2020 ACMIEEE Int. Conf. Hum.-Robot interact.* New York, NY, USA: Association for Computing Machinery; 2020:211–220, https://doi.org/10.1145/3319502.3374781.

35. Riether N, Hegel F, Wrede B, Horstmann G: **Social facilitation with social robots?**. In *Proc. Seventh Annu. ACMIEEE Int. Conf. Hum.-Robot interact.* New York, NY, USA: Association for Computing Machinery; 2012:41–48, https://doi.org/10.1145/2157689.2157697.

36. de Melo CM, Marsella S, Gratch J: **People do not feel guilty about exploiting machines**. *ACM Trans Comput-Hum Interact* 2016, **23**:1–8, https://doi.org/10.1145/2890495. 17.

37. Batson CD, Lishner DA, Stocks EL: **The empathy–altruism hypothesis**. In *Oxf. Handb. Prosocial Behav.*. Edited by Schroeder DA, Graziano WG, New York, NY, US: Oxford University Press; 2015:259–281, https://doi.org/10.1093/oxfordhb/9780195399813.013.023.

38. Rosenthal-von der Pütten AM, Krämer NC, Hoffmann L, Sobieraj S, Eimler SC: **An experimental study on emotional reactions towards a robot**. *Int J Soc Robot* 2013, **5**:17–34, https://doi.org/10.1007/s12369-012-0173-8.

39. Suzuki Y, Galli L, Ikeda A, Itakura S, Kitazaki M: **Measuring empathy for human and robot hand pain using electroencephalography**. *Sci Rep* 2015, **5**, https://doi.org/10.1038/srep15924. Article 15924.

40. Darling K, Nandy P, Breazeal C: **Empathic concern and the effect of stories in human-robot interaction**. In *2015 24th IEEE Int. Symp. Robot Hum. Interact. Commun. RO-MAN*; 2015: 770–775, https://doi.org/10.1109/ROMAN.2015.7333675.

41. Briggs G, Scheutz M: **How robots can affect human behavior: investigating the effects of robotic displays of protest and distress**. *Int J Soc Robot* 2014, **6**:343–355, https://doi.org/10.1007/s12369-014-0235-1.

42. •• Broadbent E: **Interactions with robots: the truths we reveal about ourselves**. *Annu Rev Psychol* 2017, **68**:627–652, https://doi.org/10.1146/annurev-psych-010416-043958.

In this thorough review article, the author covers the practical applications of robots, the challenges faced in current applied research, and the major psychological theories that have spawned within the field of human–robot interaction.

43. Mutlu B, Forlizzi J: **Robots in organizations: the role of workflow, social, and environmental factors in human-robot interaction**. In *Proc. 3rd Int. Conf. Hum. Robot interact. - HRI 08*. Amsterdam, The Netherlands: ACM Press; 2008:287–294, https://doi.org/10.1145/1349822.1349860.

44. Sung J-Y, Guo L, Grinter RE, Christensen HI: **“My roomba is rambo”: intimate home appliances**. In *UbiComp 2007 Ubiquitous Comput.*. Edited by Krumm J, Abowd GD, Seneviratne A, Strang T, Berlin, Heidelberg: Springer Berlin Heidelberg; 2007:145–162, https://doi.org/10.1007/978-3-540-74853-3_9.

45. Moore D, Currano R, Shanks M, Sirkin D: **Defense against the dark cars: design principles for griefing of autonomous vehicles**. In *Proc. 2020 ACMIEEE Int. Conf. Hum.-Robot interact.* New York, NY, USA: Association for Computing Machinery; 2020: 201–209, https://doi.org/10.1145/3319502.3374796.

46. • Wykowska A: **Robots as mirrors of the human mind**. *Curr Dir Psychol Sci* 2021, **30**:34–40, https://doi.org/10.1177/0963721420978609.

This theoretical article highlights the potential opportunities in using robots as a tool to research social cognition. For instance, the author argues that robots are uniquely suited to address the question, “what constitutes humanness?”

47. Eyssel F: **An experimental psychological perspective on social robotics**. *Robot Autonom Syst* 2017, **87**:363–371, https://doi.org/10.1016/j.robot.2016.08.029.

48. Hilbig BE, Thielmann I, Böhm R: **Bending our ethics code**. *Eur Psychol* 2021:1–9, https://doi.org/10.1027/1016-9040/a000431.

49. Bartneck C, Rosalia C, Menges R, Deckers I: **Robot Abuse - a limitation of the media equation**. In *Proc. Interact 2005 Workshop Abuse*; 2005. http://hdl.handle.net/10092/16925.

50. Darling K: **Extending legal protection to social robots: the effects of anthropomorphism, empathy, and violent behavior towards robotic objects**. In *Robot Law*. Edited by Calo R, Froomkin AM, Kerr I; 2016:213–232, https://doi.org/10.4337/9781783476732.

51. Fischer K: **Interpersonal variation in understanding robots as social actors**. In *2011 6th ACMIEEE Int. Conf. Hum.-Robot interact. HRI*; 2011:53–60, https://doi.org/10.1145/1957656.1957672.