



Ethics of Corporeal, Co-present Robots as Agents of Influence: a Review

AJung Moon¹ · Shalaleh Rismani¹ · H. F. Machiel Van der Loos²

Accepted: 24 March 2021 / Published online: 14 April 2021
© The Author(s) 2021

Abstract

Purpose of Review To summarize the set of roboethics issues that uniquely arise due to the corporeality and physical interaction modalities afforded by robots, irrespective of the degree of artificial intelligence present in the system.

Recent Findings One of the recent trends in the discussion of ethics of emerging technologies has been the treatment of roboethics issues as those of “embodied AI,” a subset of AI ethics. In contrast to AI, however, robots leverage human’s natural tendency to be influenced by our physical environment. Recent work in human-robot interaction highlights the impact a robot’s presence, capacity to touch, and move in our physical environment has on people, and helping to articulate the ethical issues particular to the design of interactive robotic systems.

Summary The corporeality of interactive robots poses unique sets of ethical challenges. These issues should be considered in the design irrespective of and in addition to the ethics of artificial intelligence implemented in them.

Keywords Human-robot interaction · Robot ethics · AI ethics · Nonverbal communication · Robot presence · Influence

Introduction

In a 2017 report, a legal scholar, Radidja Nemar, presented how the majority of civilians living under the constant presence of US-led, remote-operated drones in Yemen suffer from the symptoms of post-traumatic stress disorder. She found the symptoms of mental illness irrespective of whether or not they have lost a loved one from a drone attack. A striking summary of her findings reads as follows:

“Our belief, which finds an empirical grounding in this study, is that the simple fact of living under drones has psychological consequences that are no different from those caused by the loss of a relative in a [drone] strike. ... The intensity of the suffering is such that we believe it amounts to cruel, inhumane, and degrading treatment of civilians.” (pg. 40 [1])

While there are many factors contributing to the issue of drone use in Yemen (e.g., geopolitical factors, exposure to drone use as weapons), one of the undeniable factors is the fact that the mere presence of robotic systems—albeit being remotely operated by humans—can profoundly impact people’s lives and pose serious harms and ethical issues [2].

At the time of writing, we are at an important juncture where the discussion of ethical, legal, and societal (ELS) issues of two highly related technologies are on the rise—robotics and artificial intelligence (AI). AI ethics in particular is a relatively new field that rapidly gained popularity over the last 5 years. The number of publications mentioning AI ethics saw an explosive, above 11-fold increase between 2015 and 2020 (see Fig. 1). This followed a number of high profile AI ethics issues published in the public media. A notable case is a 2015 ProPublica article highlighting the serious and real-life

This article is part of the Topical Collection on *Roboethics*

✉ AJung Moon
ajung.moon@mcgill.ca

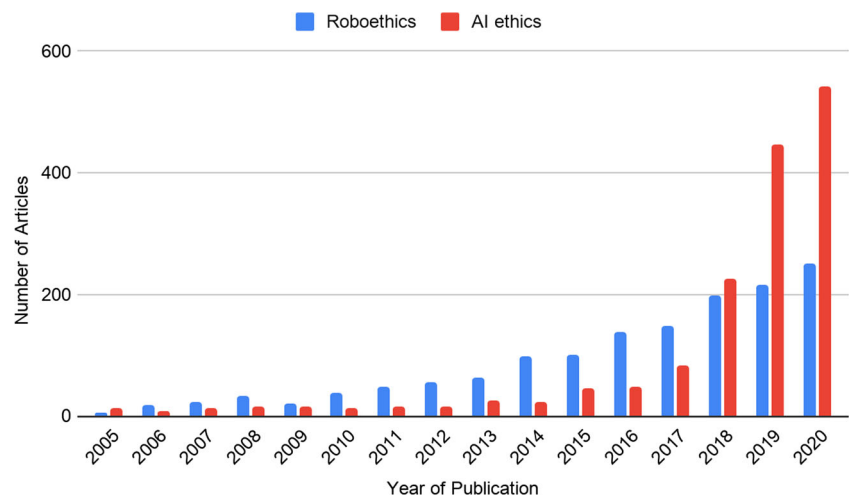
Shalaleh Rismani
shalaleh.rismani@mail.mcgill.ca

H. F. Machiel Van der Loos
vdl@mech.ubc.ca

¹ Department of Electrical and Computer Engineering, McGill University, Montreal, Canada

² Department of Mechanical Engineering, University of British Columbia, Vancouver, Canada

Fig. 1 Web of Science literature search on AI ethics and roboethics between 2005 and 2020. Search queries used are (“AI” OR “artificial intelligence”) AND (“ethics” OR “ethical”) and (((“robot” OR “robotics” OR “robots” OR “robotic”) AND (“ethics” OR “ethical”)) OR “roboethics”), respectively. The term roboethics started being used since the First International Symposium on Roboethics in 2004



cases of bias in machine learning algorithms used in the US criminal court system [3]. This has resulted in the fields roboethics and AI ethics finding common ground and enabling efforts, such as sharing and shaping of ethical principles. These efforts are summarized in earlier publications of this journal (see [4, 5]). However, recent trends have also included the treatment of roboethics issues as a subset of issues pertaining to “embodied AI” (e.g., European Union’s report published March 2020 on *The ethics of artificial intelligence: Issues and initiatives* [6], and an influential review of AI ethics frameworks by Hagendorff [7]).

Both robotics and AI have multiple definitions. For the purposes of this paper, we borrow from Liao [8] and refer to AI as any machine that has cognitive abilities such as thinking and learning. AI ethicists have predominantly been focused on ethical implications of non-embodied, algorithmic systems (e.g., recommender systems) or virtual/digital agents (e.g., chatbots). An AI can be and often is embodied in physical or virtual systems (e.g., autonomous vehicles, virtual avatars, respectively) but embodiment or corporeality is not a necessary component of AI. In contrast, we refer to robots as systems with actuation and sensing capabilities where having a physical body (i.e., corporeal) and co-presence (i.e., being in the same physical space) in our physical environment are necessary conditions.

As the community of roboethicists and AI ethicists continue to cross-pollinate ideas and solutions to similar, and often interconnected challenges, we posit that articulating components of the technologies that give rise to a divergence of ethical issues is especially useful. In this paper, we aim to summarize the set of ethical issues that uniquely arise in the design and deployment of interactive robotic systems. In particular, we highlight how the corporeality, co-presence, and physical interaction modalities afforded by robots pose influence people in ways that pose unique sets of ethical challenges, irrespective of the degree of AI present in the system (see Fig. 2). To do so, we review recent findings in the field of

human-robot interaction (HRI) where the impact of interacting with co-present, corporeal robots is highlighted.

We first outline the two different beginnings that mark the launch of the domains roboethics and AI ethics. Subsequently, we highlight recent developments in HRI that articulate how a robot’s physical presence, movement, and capacity to touch pose unique ethical implications compared to our interaction with algorithmic systems. We then outline recent efforts to incorporate ethics in HRI design. We conclude with a summary of how the current community of researchers is framing the study of ethics within HRI.

Ethical Implications Arising from Interaction with Robots

As social animals, we are especially influenced by people around us both directly (e.g., one person influencing another) and indirectly (e.g., social norms, customs [9, 10]). We follow norms, the informal rules that shape and regulate the behaviors of members of a group [11]. Interactive technologies are increasingly part of what shape our behaviors. Notably, Reeves and Nass presented the media equation, illustrating how people treat and are affected by computers and other technological media as though they were people or places [12]. In particular, humans have evolved to be and are naturally influenced by our physical environment [13]. The corporeality and co-presence of robots allow them to leverage this human nature and influence people in a programmable, predictable, and scalable manner in ways distinguished from AI, virtual agents, or other digital technologies. These modalities include a robot’s physical presence and appearance, its capacity to move and perform nonverbal expressions in our physical space, and the capacity to come into physical contact with humans (touch).

As interaction with a variety of digital technologies have increasingly become an everyday part of our lives, decades of

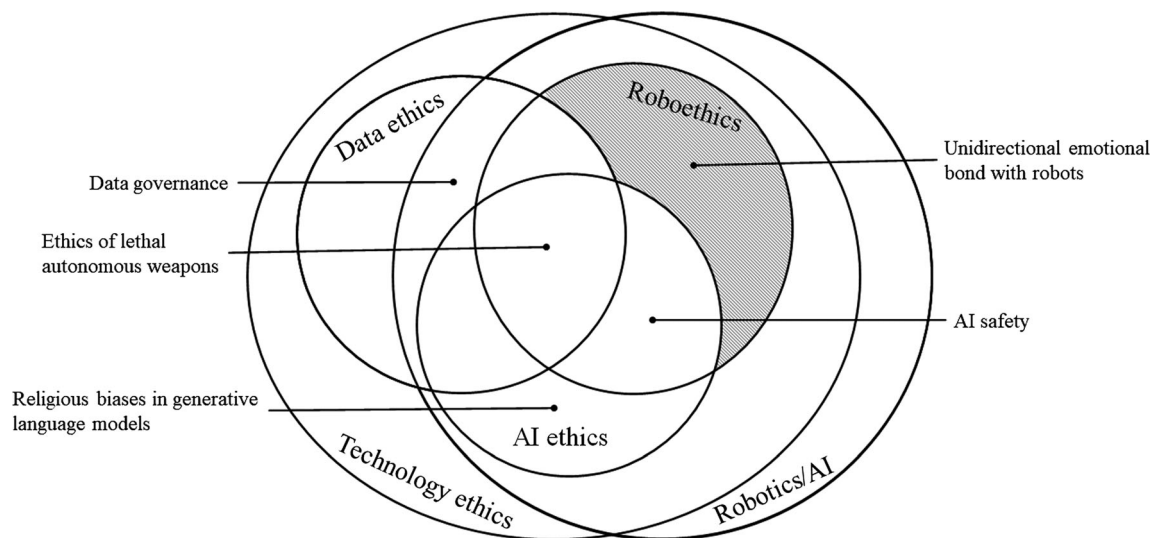


Fig. 2 An illustrated overlap of technology ethics issues and relevant examples. The dashed area of the diagram highlights the set of issues presented by the physicality of interactive robotic systems, a focus of this paper

research have been dedicated to investigating whether and how human perception and response to artificial agents differ across different embodiments of the agents (e.g., from virtual, telepresent, and physically co-present agents). Until recently, much of the research on human response to different morphology and corporeality of artificial agents have shown conflicting results (Li [14] provides a comprehensive overview on this topic). This has been in part, due to the variety of platforms, tasks, and measurements used in previous studies, making it difficult to compare stimuli in one study to another. Recent work by Hoffman et al. is making headway in providing a theoretical framework to disentangle these contradicting results [15]. In particular, they suggest that humans take their interaction with corporeal and co-present robots more seriously due to the possible capacity of the robots to physically interact and possibly harm people in contrast to virtual agents. We see evidence of this in the way humans respond to robot presence, touch, and movement in situ that present ethical implications. In this section, we highlight these particular modes of interaction and how they relate to ethical issues of interactive, corporeal robots.

Physical Presence and Appearance

As highlighted in the above example of drone use, a robot's presence in our physical space can have a profound impact on people. This is irrespective of a robot's degree of demonstrated intelligence or autonomous behavior.

Many studies in HRI have demonstrated that the presence of a robot in the room can have the same effect on humans as having another person in the room (i.e., social facilitation) in terms of task performance and elicitation of feelings [16, 17]. A series of recent experiments by Spatola and colleagues demonstrate how anthropomorphism and robot appearance play a

role in the influence robot co-presence has on people, and that such an effect may be tapping into the neurosocial mechanisms that are automatically triggered in our brains [17, 18]. One implication of these findings is that even before considering the design of the interactive behaviors for a robotic system, one must consider the ethics of putting a robot in a room in the first place.

The profound ways in which a robot's presence impacts human behavior over virtual agents are highlighted in the recent HRI experiments on human conformity to robots [16, 19•, 20]. In traditional experiments on conformity—originally devised by Asch [21]—individual participants, surrounded by a group of confederates, are asked to answer simple questions with one obvious right answer (e.g., “which of the following lines have the same length as one on the left?”). The experiments conducted with groups of humans consistently show that individuals easily resort to the same answers given by the group, even when the group is clearly wrong. HRI replication of the classic experiments suggests that humans also conform to a group of co-present robots (i.e., follow the robots' suggestions rather than standing by their own beliefs/decisions) [16, 19•, 20]. The results suggest that the co-present robots asserting peer pressure on humans are more effective than virtual agents trying to do the same, and that humans conform to robot suggestions significantly more when the task involves more uncertainties or ambiguities [22, 23].

One of the related factors to co-presence is a robot's appearance. Numerous studies illustrate that the appearance of a robot affects not only its user acceptance, but also how much users trust the system [24]. Inspired from the rising discussion of dark patterns in human-computer interaction (HCI)—referring to the design of user interfaces that mislead users into taking unwanted actions or decisions for the benefit of another

[25]—Lacey and Caudwell articulates that dark patterns exist in HRI designs as well [26]. They highlight that the recurring strategy to design consumer robots to be cute and baby-like in appearance constitutes a dark pattern, because the cute esthetics are used to manipulate and exploit the subsequently generated positive affective responses in users in order to attract user attention, continued interaction with the system, and enable further data collection.

Touch

Some of the ethical issues particular to interactive robotics arise from a robot's ability to touch and implicate the physical safety of humans. Recent studies indicate that humans seem to perceive a robot's physical touch in a wide variety of ways, and the notion of appropriateness of touch is starting to be discussed. Law et al., for example, conducted a series of experiments where participants watched a robot touch or not touch a human on the shoulder [27]. They found that while people perceive a robot in the touch condition more trustworthy and interpreted it as exhibiting comforting behaviors, they found it more inappropriate. To date, no unifying theory explains the variety of ways people interpret and respond to a robot touch behavior, and its impact on HRI is still in its infancy. However, appropriateness of robot use of touch as interaction modality with humans is tightly coupled with ethics in ways distinct from algorithmic systems. Development or use of sex robots, for instance, raises issues about acceptable touch behaviors between humans and robots independent of the degree of intelligence onboard the system (see [28] for a systematic review). Arnold and Scheutz articulate this as ethical issues arising from “the primacy and implicit dynamics of bodily perception” [29]. Given how the treatment of our bodies trigger moral psychological responses in humans, some make an explicit link between ethics of robot use to the religious ethics where the treatment of the human bodies is subject to religious moral rules [30].

Movements and Nonverbal Expressions

Other forms of robot influence on humans have non-obvious points of differentiations from algorithmic systems. Humans perceive the expressivity of nonverbal communication between a robot and a virtual representation of the robot to be similar [15]. Yet, movement in physical space can affect human behaviors distinct from motions depicted on a screen. Entrainment effects observed in HRI are an example. Entrainment is a phenomena where humans (agents) mirror or synchronize elements of behaviors of those around them such as unintentional synchronization of steps between a pair of people walking side by side [31]. In an imitation game involving a NAO robot, Ansermin et al. observed that humans entrain to the rhythmic behavior of the robot unidirectionally and unintentionally [32].

Ciardo et al. also documents similar effects on a robot teaching task, where a human teaching a robot attunes to and mirrors the pauses and delays exhibited in the learning robot behavior [33]. Such movement entrainment in HRI does not require social context as was used in Ansermin et al. [32]. An empirical experiment conducted by Lorenz et al., for example, involved a 7-DoF robotic manipulator that simply performs reach-retract motions on the shared tabletop workspace of a participant [34]. Results from Lorenz et al. [34] illustrate that humans synchronize their goal directed, arm motions to the movement of an articulated robot in a dyadic task, such as slowing down to follow the rhythm of the robot's motion. Humans also entrain to nonverbal expressions exhibited by robots. This includes affective and gesture-based expressions to protolinguistic behaviors. In 2002, Breazeal outlined how participants quickly mimic not only the body posture and facial expression of the robot, but also change their tone of voice (soothing, scolding, etc.) to regulate the robot's interactive response [35]. This is one way humans improve the efficacy of our communication with each other and co-present robots. Since then, we have found evidence that humans adopt prosodic behaviors (intonation and intensity of our speech) [27].

We have yet to come across real-life cases where human-robot entrainment effects have been abused in a morally harmful manner. However, it is not far-fetched to envision such cases. For example, collaborative robots in assembly lines could be programmed to speed up the pace of work unbeknownst to the human workers resulting in new workplace injuries, or programmed robot behaviors could inadvertently influence the norms of a particular task or place.

With its corporeality, social robots can also leverage its richer platform for nonverbal expressions to influence people. For instance, social robots can be designed to persuade people into making decisions ranging from food and entertainment choices to making donations [36]. In one study, a video recording of a person and a geminoid—a hyper-anthropomorphic robot doppelgänger of the person—was used to persuade people. In a product marketing task, the authors found that the geminoid was perceived to be just as persuasive as the video recording of a human [37]. Ghazali et al. demonstrated how social responses, such as compliance and reactance, affect user acceptance of robots' persuasion attempts [38].

The persuasive influence robots can have on people raises ethical issues due to its interference with human agency. Scheutz, for example, raised the concern that interactive robots can lead users to create emotional bonds that can be abused and are unidirectional [39]. In discussing ethics of robots that nudge people toward a choice, Borenstein and Arkin assert that the normative questions about the agency of users subjected to robotic nudge remain even if robots are designed to nudge them for their own good [40, 41]. They conclude by suggesting to consider questions such as whether the person subject to the nudge is aware of and has control

over it in our efforts to evaluate the ethics of a robot that nudges [41]. Although the context within which robots are used to persuade people remains limited today, the technology has the potential to serve as a programmed medium of influence in the future, thereby helping to establish new norms within society.

Incorporating Ethics into Interactive Robot Design

Philosophers have been identifying ways in which technology is used as a proxy for normative actions and decisions of humans, especially that of the designers [42]. For example, using a simple example of low-hanging bridges in Long Island designed to reinforce socioeconomic disparities, Langdon Winner famously proposed that artifacts (including technologies) have politics [43]. Unlike bridges, a robot poses a different set of ethics challenges owing to its ability to interact with people actively. It can also change its behavior over time, from one version of software upgrade to the next, so-to-speak. Due to the corporeality of robots discussed above, ethical issues that pertain to interactive robotic systems go beyond ensuring that a system is ethically used and an ethical decision be made by the system (issues that are very much shared with the community of AI ethicists), but also *how* an action is carried out. This implies that incorporating ethics into the design of an interactive robotic system includes the decision to place a robot in a room, what it should look like, and how it should communicate and interact with our physical bodies, objects, and the environment.

In one of the first attempts to tackle ethics issues unique to HRI, Riek and Howard developed a code of ethics for HRI practitioners [44]. In it, they point to some of the above-mentioned issues unique to HRI, such as the need to practice sensitivities in designing touch as interaction modality and decisions regarding robot morphology. The code of ethics states that “*Human frailty is always to be respected, both physical and psychological*” and to “[a]void racist, sexist, and ableist morphologies and behaviors in robot design.” Since the publication of the code of ethics, there active discussions that took place within the HRI community. This includes hosting of interdisciplinary workshops to bridge the gap between ethics and policy discussions in HRI [45] and to create community guidelines to address possible psychological manipulation and exploitation of humans by HRI [46].

New frameworks to better articulate the ELS issues particular to interactive robotics are being introduced as well. For instance, Payr [47] frames this line of inquiry as interaction ethics. Using a case study, she demonstrates the nature of ethical considerations involved in interaction design, including the detailed decisions in the implementation of turn-taking, topic change, and politeness behaviors on a robot. Webb et al. [48],

on the other hand, leverages Lucy Suchman’s pivotal work on situated action and plans to argue for creation of an ethical black box, a system that records HRI for the purposes of reviewing interaction failures and hazards by the system designer/user/operator and informs future design choices. Webb et al. assumes that interaction design for social robots has inherent ethical implications and that it is important to recognize these implications within their specific context.

These developments in roboethics highlight a perspective through which every detail of an interactive robot’s existence and its programmed behaviors can be said to have ELS implications. Noting the level of complexity involved considering ethics in interactive robot designs, some scholars are taking steps to tackle practical issues within the domain. One such development is the investigation of consent in HRI by Sarathy et al. [49]. In Sarathy et al. [49], the authors bridge the meaning of consent from legal scholarship to highlight the challenges of operationalizing consent in HRI and the need to implement consent for even normatively neutral contexts. However, much more remains to be done to guide the many ethics considerations involved in designing interactive robots that influence. This is not to mention the fact that ethics considerations become even more complex when interactive robots use AI—AI ethics issues pertaining to fairness, accountability, and transparency of the algorithm would need to be considered in addition to the above-mentioned issues unique to interactive robotics.

Conclusion

In this paper, we outlined how corporeality and co-presence of robots marks a point of divergence for roboethics from AI ethics. The latest findings in HRI and the growing discussions from the community highlight that the treatment of roboethics issues as a subset of AI ethics has the potential to undermine these unique issues present in the design of interactive robots. To date, among numerous published ethics principles, frameworks, and guidelines discussing AI or robotics, there are only two published standards on ethical design of robotic devices (*BS 8611:2016—Robots and robotic devices: Guide to the ethical design and application of robots and robotic systems* [50] and *IEEE Recommended Practice for Assessing the Impact of Autonomous and Intelligent Systems on Human Well-Being* [51]). Much more work is needed to navigate the ethics issues particular to the design of interactive robots to help bridge ethics thinking into robot engineering design, technology governance, and coordinated research toward the development of robots with ethics in mind.

While this paper emphasized the differentiating factors between roboethics and AI ethics issues, this is by no means to be taken as a discouragement for cross-pollination of ideas between the two communities. Interdisciplinary conferences

such as We Robot and AAAI/ACM Conference on Artificial Intelligence, Ethics, and Society have been a catalyst and inspiration for urgent ethics discussions in the roboethics and AI ethics community. Underlying all of the ELS issues across the corporeality-algorithmic spectrum is the fundamental question about what kind of society we would like to envision for our increasingly technological society. Synergy between the communities is essential for accelerated advancement in the respective domains and to catalyze coordinated action toward responsible engineering practices across the corporeality-algorithmic spectrum.

Compliance with Ethical Standards

Conflict of Interest The authors declare no competing interests.

Human and Animal Rights and Informed Consent This article does not contain any studies with human or animal subjects performed by any of the authors.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

Papers of particular interest, published recently, have been highlighted as:

- Of importance
1. The humanitarian impact of drones, edited by Ray Acheson, Matthew Bolton, Elizabeth Minor, and Allison Pytlak, Women's International League for Peace and Freedom (WILPF), 2017
 2. van Wynsberghe A, Soesilo D, Thomassen K, Sharkey N. Drones in the service of society. 2018. Retrieved from Foundation for Responsible Robotics website: <https://responsiblerobotics.org/2018/06/05/report-drones-in-the-service-of-society/>.
 3. Angwin J, Larson J, Mattu S, Kirchner L. Machine bias. ProPublica. 2016. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.
 4. Studley M, Winfield A. ELSA in industrial robotics. *Curr Robot Rep*. 2020;1:179–86. <https://doi.org/10.1007/s43154-020-00027-0>.
 5. Wullenkord R, Eyssel F. Societal and ethical issues in HRI. *Curr Robot Rep*. 2020;1:85–96. <https://doi.org/10.1007/s43154-020-00010-9>.
 6. The ethics of artificial intelligence: issues and initiatives, European Union, Brussels, 2020. Available: [https://www.europarl.europa.eu/RegData/etudes/STUD/2020/634452/EPRS_STU\(2020\)634452_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2020/634452/EPRS_STU(2020)634452_EN.pdf)
 7. Hagedorff T. The ethics of AI ethics: an evaluation of guidelines. *Mind Mach*. 2020;30:99–120. <https://doi.org/10.1007/s11023-020-09517-8>.
 8. Liao M. Ethics of artificial intelligence. New York: Oxford University Press; 2020.
 9. Sechrist GB, Stangor C. Perceived consensus influences intergroup behavior and stereotype accessibility. *J Pers Soc Psychol*. 2001;80: 645–54.
 10. Terry DJ, Hogg MA. Group norms and the attitude-behavior relationship: a role for group identification. *Personal Soc Psychol Bull*. 1996;22:776–93. <https://doi.org/10.1177/0146167296228002>.
 11. Feldman DC. The development and enforcement of group norms. *Acad Manag Rev*. 1984;9:47–53. <https://doi.org/10.2307/258231>.
 12. Reeves B, Nass C. The media equation: how people treat computers, television, and new media like real people and places. Cambridge: Cambridge University Press; 1996.
 13. Forgas JP, Williams KD (eds) (2001) Social influence: direct and indirect processes. *Soc Infl*. <https://doi.org/10.4324/9781315783031>
 14. Li J. The benefit of being physically present: a survey of experimental works comparing copresent robots, telepresence robots, and virtual agents. *Int J Human-Comput Stud*. 2015;77(2015):23–37. <https://doi.org/10.1016/j.ijhcs.2015.01.001>.
 15. Hoffmann L, Bock N, Astrid M, Pütten R vd. The peculiarities of robot embodiment (EmCorp-Scale): development, validation and initial test of the embodiment and corporeality of artificial agents scale. In: Proceedings of the 2018 ACM/IEEE international conference on human-robot interaction (HRI '18). New York: Association for Computing Machinery; 2018. p. 370–8. <https://doi.org/10.1145/3171221.3171242>.
 16. Riether N, Hegel F, Wrede B, Horstmann G. Social facilitation with social robots? In: Proceedings of the seventh annual ACM/IEEE international conference on human-robot interaction(HRI'12). New York: Association for Computing Machinery; 2012. p. 41–8. <https://doi.org/10.1145/2157689.2157697>.
 17. Spatola N, Belletier C, Chausse P, Augustinova M, Normand A, Barra V, et al. Improved cognitive control in presence of anthropomorphized robots. *Int J Soc Robot*. 2019;11(3):463–76. <https://doi.org/10.1007/s12369-018-00511-w>.
 18. Spatola N, Monceau S, Ferrand L. Cognitive impact of social robots: how anthropomorphism boosts performances. *IEEE Robot Auto Magaz*. 2020;27(3):73–83. <https://doi.org/10.1109/MRA.2019.2928823>.
 19. Salomons N, Van Der Linden M, Strohkorb Sebo S, Scassellati B. Humans conform to robots: disambiguating trust, truth, and conformity, ACM/IEEE int conf human-robot interact; 2018. p. 187–95. <https://doi.org/10.1145/3171221.3171282>. **This paper provides an empirical evidence linking the rate of human conformity to a robot with human trust in the robot**
 20. Woods S, Dautenhahn K, Kaouri C Is someone watching me? - consideration of social facilitation effects in human-robot interaction experiments. In 2005 int symp on comp int in robotics and automation. Espoo, Finland; 2005. p. 53–60. <https://doi.org/10.1109/CIRA.2005.1554254>.
 21. Asch SE. Effects of group pressure upon the modification and distortion of judgments. In: Groups, leadersh. men; res. hum. relations. Oxford: Carnegie Press; 1951. p. 177–90.
 22. Hertz N, Wiese E. Influence of agent type and task ambiguity on conformity in social decision making. *Proc Hum Factors Ergon Soc*. 2016:313–7. <https://doi.org/10.1177/1541931213601071>.
 23. Fuse Y, Tokumaru M. An investigation of social influence of group norms on human in human-robot groups. In proc of the IEEE symp series on comp int. Xiamen, China; 2019. p. 1407–1414. <https://doi.org/10.1109/SSCI44817.2019.9002796>.

24. Natarajan M, Gombolay M. Effects of anthropo-morphism and accountability on trust in human robot interaction. In: Proceedings of the 2020 ACM/IEEE international conference on human-robot interaction (HRI '20), March 23–26, 2020, Cambridge, United Kingdom. New York: ACM; 2020. 10 pages. <https://doi.org/10.1145/3319502.3374839>.
25. Gray C, Kou Y, Battles B, Hoggatt J, Toombs AL. The dark (patterns) side of UX design. In: Proceedings of the 2018 CHI conference on human factors in computing systems (CHI '18) Montreal, QC, April 21–26. New York, ACM Paper 534, 14 pages; 2018. <https://doi.org/10.1145/3173574.3174108>.
26. Lacey C, Caudwell C. Cuteness as a 'dark pattern' in home robots. In: Proceedings of the 14th ACM/IEEE international conference on human-robot interaction (HRI '19). New York: IEEE Press; 2019. p. 374–81. <https://doi.org/10.1109/HRI.2019.8673274>.
27. Law T, Malle BF, Scheutz M. A touching connection: how observing robotic touch can affect human trust in a robot. *Int J Soc Robot*. 2021;20210105. <https://doi.org/10.1007/s12369-020-00729-7>.
28. González-González CS, Gil-Iranzo RM, Paderewski-Rodríguez P. Human-robot interaction and sextbots: a systematic literature review. *Sensors*. 2021;21(1):216. <https://doi.org/10.3390/s21010216>.
29. Arnold T, Scheutz M. Beyond moral dilemmas: exploring the ethical landscape in HRI. In: Proceedings of the 2017 ACM/IEEE international conference on human-robot interaction (HRI '17). New York: Association for Computing Machinery; 2017. p. 445–52. <https://doi.org/10.1145/2909824.3020255>. **Articulates the particulars of physically present robots and the unique ethical challenges posed by physical HRI.**
30. de Graaf MMA. An ethical evaluation of human–robot relationships. *Int J Soc Robot*. 2016;8:589–98. <https://doi.org/10.1007/s12369-016-0368-5>.
31. Mörtl A, Lorenz T, Vlaskamp BNS, Gusrialdi A, Schubö A, Hirche S. Modeling inter-human movement coordination: synchronization governs joint task dynamics. *Biol Cybern*. 2012;106:241–59. <https://doi.org/10.1007/s00422-012-0492-8>.
32. Ansermin E, Mostafaoui G, Sargentini X, Gaussier P. Unintentional entrainment effect in a context of human robot interaction: an experimental study. In: *Proc int symp on robot and human interact com (RO-MAN'17)*. Lisbon; IEEE; 2017. p. 1108–14. <https://doi.org/10.1109/ROMAN.2017.8172442>.
33. Ciardo F, De Tommaso D, Wykowska A. Humans socially attune to their “follower” robot. In: Proceedings of the 14th ACM/IEEE international conference on human-robot interaction (HRI '19). New York: IEEE Press; 2019. p. 538–9. <https://doi.org/10.1109/HRI.2019.8673262>.
34. Lorenz T, Mortl A, Vlaskamp B, Schubo A, Hirche S (2011) Synchronization in a goal-directed task: human movement coordination with each other and robotic partners. In: *Proc int symp on robot and hum interact commun (RO-MAN'11)*. Atlanta; IEEE; 2011. p. 198–203. <https://doi.org/10.1109/ROMAN.2011.6005253>.
35. Breazeal C. Regulation and entrainment in human - robot interaction. *Int J Robot Res*. 2002;21(10–11) October–November 2002: 883–902. <https://doi.org/10.1177/0278364902021010096>.
36. Herse S, Vitale J, Ebrahimi D, et al. Bon appetit! Robot persuasion for food recommendation. In: *ACM/IEEE int. conf. human-robot interact*. New York: IEEE Computer Society; 2018. p. 125–6. <https://doi.org/10.1145/3173386.3177028>.
37. Ogawa K, Bartneck C, Sakamoto D, Kanda T, Ono T, Ishiguro H. Can an android persuade you? In: *Geminoid stud*. Singapore: Springer Singapore; 2018. p. 235–47. https://doi.org/10.1007/978-981-10-8702-8_14.
38. Ghazali AS, Ham J, Barakova E, Markopoulos P. Persuasive robots acceptance model (PRAM): roles of social responses within the acceptance model of persuasive robots. *Int J Soc Robot*. 2020;12: 1075–92. <https://doi.org/10.1007/s12369-019-00611-1>.
39. Scheutz M. The inherent dangers of unidirectional emotional bonds between humans and social robots. In: Lin P, Abney K, Bekey G, editors. *Robot ethics ethical soc. implic. robot*. Cambridge: MIT Press; 2012. p. 205–21.
40. Borenstein J, Arkin R. Robotic nudges: the ethics of engineering a more socially just human being. *Sci Eng Ethics*. 2016;22:31–46. <https://doi.org/10.1007/s11948-015-9636-2>.
41. Borenstein J, Arkin RC. Nudging for good: robots and the ethical appropriateness of nurturing empathy and charitable behavior. *AI & Soc*. 2017;32:499–507. <https://doi.org/10.1007/s00146-016-0684-1>. **This work bridges the discussions on the ethics of nudges, a behavioural economics and policy concept, to the discussion of robots as a mechanism that nudges.**
42. Millar J. Technology as moral proxy: autonomy and paternalism by design. *IEEE Technol Soc Mag*. 2015;34:47–55. <https://doi.org/10.1109/ETHICS.2014.6893388>.
43. Winner L. Do artifacts have politics? *Daedalus*. 1980;109:121–36.
44. Riek, Laurel and Howard, Don, A code of ethics for the human-robot interaction profession (April 4, 2014). Proceedings of we robot, 2014, Available at SSRN: <https://ssrn.com/abstract=2757805>
45. Riek LD, Hartzog W, Howard DA, Moon AJ, Calo R. The emerging policy and ethics of human robot interaction. In: Proceedings of the tenth annual ACM/IEEE international conference on human-robot interaction extended abstracts (HRI '15 extended abstracts). New York: Association for Computing Machinery; 2015. p. 247–8. <https://doi.org/10.1145/2701973.2714393>.
46. Haring KS, Novitzky MM, Robinette P, de Visser EJ, Wagner A, Williams T. The dark side of human-robot interaction: ethical considerations and community guidelines for the field of HRI, 2019 14th ACM/IEEE international conference on human-robot interaction (HRI), Daegu, Korea (South); 2019. p. 689–90. <https://doi.org/10.1109/HRI.2019.8673184>.
47. Payr S. Towards human–robot interaction ethics. In: Trapp R, editor. *A construction manual for robots' ethical systems*, Cognitive technologies. Cham: Springer; 2015. https://doi.org/10.1007/978-3-319-21548-8_3.
48. Webb H, Jirotko M, Winfield AFT, Winkle K. Human-robot relationships and the development of responsible social robots. *ACM Int Conf Proc Ser*. 2019;1–10. <https://doi.org/10.1145/3363384.3363396>.
49. Sarathy V, Arnold T, Scheutz M. When exceptions are the norm: exploring the role of consent in HRI. *J Hum-Robot Interact*. 2019;8(3) Article 14 (August 2019), 21 pages:1–21. <https://doi.org/10.1145/3341166>.
50. BSI Group (2016) BS 8611:2016 - robots and robotic devices: guide to the ethical design and application of robots and robotic.
51. (2020) 7010-2020 - IEEE recommended practice for assessing the impact of autonomous and intelligent systems on human well-being.