Social Robotics as Moral Education? Fighting Discrimination Through the Design of Social Robots

Fabio FOSSA[a,1]
[a] *Department of Mechanical Engineering*
Politecnico di Milano
Italy

Abstract. Recent research in the field of social robotics has shed light on the considerable role played by biases in the design of social robots. Cues that trigger widespread biased expectations are implemented in the design of social robots to increase their familiarity and boost interaction quality. Ethical discussion has focused on the question concerning the permissibility of leveraging social biases to meet the design goals of social robotics. As a result, integrating ethically problematic social biases in the design of robots–such as, e.g., discriminatory gender stereotypes–has been opposed as morally unacceptable. Building on this debate, the present paper explores a related but different question: would it be permissible to design social robots in ways that intentionally challenge widespread discriminatory social biases, thus fostering moral education? The analysis shows that, while the potential benefits of such a design strategy could be significant, its practical endorsement raises important ethical issues. Hence, caution and further discussion are advised.

Keywords. Social Robotics; Design; Bias; Discrimination; Moral Education.

1. Introduction

The aim of this short paper is to discuss whether it would be permissible to design social robots in ways that intentionally challenge discriminatory social biases and, thus, foster moral education. I claim that the beneficial effects of such design strategy could be significant. However, its endorsement raises important ethical issues. For this reason, caution and further discussion are required.

---

[1] Fabio Fossa, Politecnico di Milano, Department of Mechanical Engineering, via privata Giuseppe La Masa 1, 20156 Milan, Italy. E-mail: fabio.fossa@polimi.it.

Here is a brief overview of my argument. Section 2 clarifies the reasons why social biases are relevant to social robotics. In a nutshell, since social robots are met within relational contexts where biased expectations influence interpersonal interactions, the same biases are likely to influence interactions between humans and robots. If robots are to be perceived as familiar and easy to interact with, it is therefore assumed, they must be designed in ways that align with such biased expectations.

Section 3 zooms in on leveraging the power of social biases to boost the quality of human-robot interaction. I call this design strategy the *bias alignment* approach. Even though in some cases this approach might not raise any ethical issue, I suggest that exploiting some types of social bias–such as, for instance, gender biases - to this end might lead to reinforcing the grip of discriminatory prejudices and hamper their eradication. Hence, leveraging problematic biases to maximise the quality of human-robot interaction is opposed as ethically unacceptable.

However, the integration of design cues stimulating biased reactions in users might perhaps be repurposed to pursue more ethical objectives. Section 4 explores the possibility of fostering moral education through social robotics by taking design choices that, instead of aligning to widespread discriminatory biases, intentionally challenge them. Interactions with robots designed to undermine biased expectations could indeed trigger moral self-reflection and help users free themselves from the unwelcome influence of unethical prejudices.

The prospect of promoting social inclusivity and fairness through social robotics, I believe, is promising. Nevertheless, important questions remain to be addressed concerning *who* should decide *which* biases can legitimately be opposed by design, and *how* this can be done without jeopardizing users' moral autonomy. Answering this question is difficult and bound to raise controversy among involved parties. In light of this, I conclude that the possibility of pursuing moral education through social robots should be further discussed. For the moment, caution is advised.

2. Social Bias and Human-Robot Interactions

As a first step, the relevance of social biases for the design of social robots must be clarified. Biases co-structure social relations and ease interactions between humans. Social robots are supposed to inhabit the space of human interactions as well. Therefore, they are likely to be met with the same biased expectations.
Social biases, then, set the conditions for human users to perceive and approach robots *as if they were* social actors. In other words, one might say that social biases could be construed as the implicit rules of the social game. Abiding by these rules is critical to be validated as a player–no matter if biological or robotic.

2.1. Interactions between Humans

Let us take a closer look at this point.
The role played by biases in structuring relations between humans is massive. The social world is messy and multifarious. Learning how to play the social game is hard. However, much depends on our ability to play it well enough. Interacting with other human beings is critical for accomplishing highly valued goals and enjoy a rich and satisfying life. Strategies to successfully navigate through the complexity of the social dimension are thus extremely useful.

In a sense, social biases offer an answer to this need. Basically, biases suggest associations between easily perceivable data or cues–such as ethnicity, age, gender, physical appearance, social roles–and other, less immediately accessible but highly valuable pieces of information–such as competence, trustworthiness, intelligence, kindness. In doing so, social biases concur to shaping expectations and subsequent social behaviour. Following their lead, we more or less adequately anticipate what might be about to happen and what we might have to do next.

Essentially, then, biases function as social scripts [1, 2] that help us cope with information incompleteness and complexity. As more or less implicit assumptions based on approximate generalizations, rules of thumb and long-lasting habits, social biases help us trace our path through the uncertain landscape of social interaction. They provide us with a simplified compass to chart courses of action that are supposed to work well enough, at least most of the times (fig. 1).
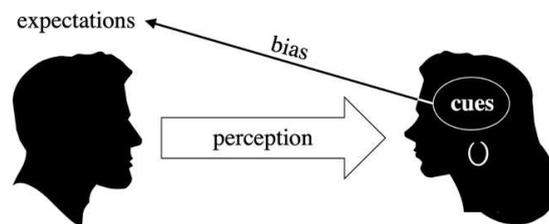

Figure 1: Bias dynamics in human-human interactions [3:3]

2.2. Human-Robot Interactions

If social biases were found to apply exclusively to interactions between humans, they would arguably be of little interest to social robotics. However, the case is precisely the opposite.

Social biases, stereotypes, and prejudiced expectations have indeed been found to transfer to interactions between humans and technological systems as well. This claim was firstly introduced and empirically confirmed in the human-computer interaction CASA approach studies [4, 5]. In more recent years, the same has been found to generally apply to human-robot interactions too [6, 7]. As such, it has grown to become a widely accepted notion in social robotics [8].

When social roles or tasks are automated, then, the way in which users respond to computer programs appear to be very similar to how they respond to social robots. In addition, interactions with both computer programs and robots are found to reproduce patterns that are typical of relationships between human beings. Hence, social biases co-shape interactions between humans and robots as well.

For example, users often tend to partially anthropomorphize social robots, projecting onto them typically human features such as gender, race, or social status. These projections, in turn, trigger biased information associations concerning competence, authority, trustworthiness, and other socially relevant features [9]. User dispositions, attitudes, and behaviours towards social robots cannot but be significantly shaped by these expectations. Ultimately, they decisively contribute to forming users' mental models of the technology (fig. 2).
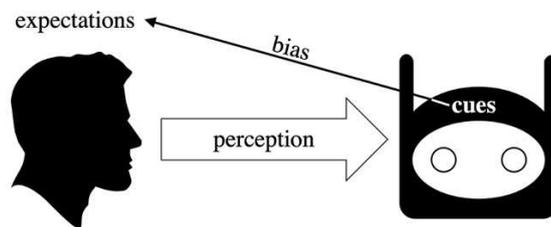
Figure 2: Bias dynamics in human-robot interaction [3: 4]

To a significant extent, then, what holds for human-human interactions also holds for human-robot interactions. The membrane separating interactions with other humans and with technological systems automating social roles is porous enough to let social biases apply to both cases.

In sum, when robots are deployed in social contexts, users appear to heavily rely on social biases during interaction. Interpretative schemes that firstly originated in human-human relations extends to interactions between human users and robots as well.

Given their relevance, the power of social biases must be acknowledged and coped with when designing social robots. The next section shows how this problem has been mostly tackled.

3. Bias, Design, and Discrimination

The fact that social biases are relevant for human-machine interactions as well is dense of design implications for social robotics. Indeed, social biases might be construed as forming a blueprint to specify robot features in ways that meet user expectations. Designing robots according to social biases, then, could significantly increase chances of acceptance, boost the perceived quality of interactions, and improve user experience.

Depending on the ethical profile of the leveraged bias, however, this strategy might arguably lead to reinforcing discriminatory stereotypes. As such, the integration of social biases by design has mostly been addressed in negative terms: ethical reflection is needed to avoid spreading discrimination through social robots.

In what follows, I offer a brief sketch of how biases can be leveraged as design tools to boost the quality of human-robot interactions. Afterwards, I argue why I believe that this design strategy could lead to reinforcing unacceptable discrimination.

3.1. Bias Alignment

Social biases have appeared to many in the social robotics community as conditions for designing successful interactions–which is arguably the central design goal in this field. While ignoring their power would probably lead to interaction failures, harnessing it through the intelligent insertion of the right design cues might be the most successful strategy for optimizing interactions [10].

Perceivable traits commonly associated with ethnicity, gender, age, physical beauty, and so on become tools in the designers' hands. Through the adoption of various cues–round or squared body shapes, low- or high-pitched voices, lighter or darker artificial skin tones, gendered names, etc.–designers can influence the formation of users' mental models of the technology and steer them towards desired outcomes, such as maximizing

the feeling of trustworthiness suggested by a healthcare robot or the impression of competence produced by a smart assistant. Aligning the design of social robots to users' expectations, however biased they might be, can smooth interactions and boost acceptance.

Since this design strategy aims at aligning social robots to the expectations of probable users, let us called this approach *bias alignment* [3]. Similarly, [7, 8] describe the tendency of adapting the design of social robots to anticipated biased expectations held by potential users as *stereotyping*. Even though there is no full agreement on the actual effectiveness of this strategy, design choices inspired by the approach are not uncommon. For the purpose of this paper, let us set aside issues concerning the effectiveness of bias alignment and focus on its ethical assessment.

3.2. Potential Discriminatory Outcomes

Depending on the moral profile of the bias that is leveraged, troubling ethical issues might arise.

Sure enough, not every instance of bias alignment is necessarily problematic. Suppose that, according to empirical data, blue robots would elicit calmer and more collaborative user behaviour than red robots. No ethical risk would arise in producing blue robots instead of red ones.

Leveraging more delicate biases, such as those associated with gender, might also be ethically unproblematic in some circumstances. Consider the case of a social robot handling consumer service tasks in a gynaecological clinic. Suppose that gendered robots were found to be significantly easier to interact with compared to robots lacking gendered design cues. In this instance, there would arguably be no ethical caveat in endowing the social robot with cues stimulating users to project the female gender, rather than the male gender. Analogously, the choice of andrology clinics opting for a "male" robot would be equally understandable.

Things, however, are substantially different when the leveraged biases convey ethically unacceptable beliefs. Consider, for instance, the case of social biases associating professional qualities to gender.

Suppose that empirical research would show that potential users of a medical robot strongly associate expertise and competence to the male gender. Endowing the robot with male cues would likely increase the perceived trustworthiness of its behaviour. However, this would feed into discriminatory expectations suggesting that male doctors are more competent and reliable than female doctors just because of their gender. It would validate and strengthen the discriminatory prejudice, ultimately making it more difficult to eradicate.

Similarly, biases associating gender to social roles or working positions could also lead to consolidating discriminatory prejudices. For example, the bias alignment approach might induce to endow robotic secretaries with bodily traits proper to young, obliging women. This would perhaps require less cognitive efforts on the part of users, which would be presented with what they more or less implicitly expect. However, it would validate the discriminatory stereotype according to which female workers are better hired as secretaries rather than sales associates, team managers, and so on, thus reinforcing harmful discrimination.

As the two examples show, leveraging discriminatory biases to boost the quality of interactions is arguably inadequate from an ethical point of view. While the stimulation of some biases through dedicated design cues might not raise any ethical worry, aligning

the design of social robots to unethical stereotypes and prejudices would arguably yield unacceptable results. If, of course, the presupposition of the present argument is accepted.

3.3. Feedforward, Feedback

Ultimately, the ethical issue discussed in the previous section rests on the hypothesis that stimulating discriminatory biases through social robots will feedback negatively onto corresponding human relations as well, thus reinforcing unethical prejudices and exacerbating the conditions of already discriminated social groups.

There is no agreement among commentators, however, on the actual consistency of this worry. The idea that interactions with robots would exert an influence on corresponding interactions with humans can indeed be opposed. A bright line could be drawn between the two dimensions. Interactions with robots, it could be argued, are different enough from interactions with human beings to avert any feedback. What happens with robots, stays with robots.

I believe, however, that the abovementioned ethical preoccupation deserves to be taken seriously. Conceptually speaking, the possibility of biased beliefs feeding back from robots to humans sounds reasonable. If social biases feed forward from interactions among humans to corresponding interactions with robots, it should at least be theoretically assumed that biases triggered by robots could feed back onto corresponding human relations as well [11, 12, 13, 14, 15]. If the membrane is porous, it is reasonable to contemplate the possibility that it could be crossed both ways.

According to what discussed so far, social robotics needs an ethical methodology to sort through the social biases the power of which can be safely leveraged. Ethics is needed to make sure that no discriminatory belief is validated and strengthen by-design. In other words, it is necessary to guarantee that social robotics do not contribute to the moral corruption of society.

In this sense, the moral profile of social robotics has been traced mostly in negative terms. Yet, a more proactive possibility still awaits to be examined. Could social robotics do something to help us free ourselves from discriminatory biases? Could it be used as a means to moral education?

4. A Proactive Perspective: Fighting Bias Through Design

Let us now explore the possibility of approaching social robotics as a means to the moral education of society.

This possibility stems from the very same presupposition that has just been discussed–i.e., that the membrane separating interactions among humans and with robots is porous and can be crossed both ways. If this is the case, avoiding unethical outcomes would represent just one side of the matter. Indeed, discriminatory social biases could be not just eschewed, but also intentionally challenged, subverted, and fought by-design.

In the following lines it is discussed why, in my opinion, the prospect of such a proactive approach is definitely worth examining. However, the presence of considerable obstacles should curb any easy enthusiasm. In the end, caution and further research are advised to responsibly weigh potential opportunities against possible risks.

4.1. Social Robotics as a Means to Moral Education

If it is reasonable to hypothesise that interactions between humans and robots would affect corresponding interactions among humans, social robotics might become a means to the moral education of society. In other words, social robots could serve as tools for fighting discriminatory biases.

Instead of aligning robot design to pre-existing discriminatory biases, misalignment could be intentionally pursued. As [16] suggests, unethical social prejudices could be actively opposed through the design of *counterstereotypical machines*. By incorporating cues intentionally designed to challenge user expectations based on discriminatory biases, social roboticists might promote moral awareness, kickstart self-reflection processes, and support users in the ethical effort of getting rid of discriminatory and unfair prejudices.

According to this ethically proactive approach, robots would be designed to interact effectively with humans while at the same time explicitly challenging discriminatory user expectations. This would lead to mitigate discriminatory attitudes towards the human subjects whose place is taken by robots. For example, explicitly designing robots in ways that subvert prejudices against women would hopefully mitigate discriminatory attitudes towards real women and generate a fairer vision of women in society [17].

To further clarify the logic behind the approach, let us briefly go back to our previous examples. In the case of the medical robot, opting for design cues that challenge biased associations between age, gender, ethnicity on the one hand and commitment, competence, proficiency on the other hand might lead users to becoming aware of discriminatory expectations and support processes of moral growth. Endowing secretary robots of cues that do not prompt the projection of the female gender might help defuse discriminatory associations between this working position and a given social category. Moreover, it might induce self-reflective reconsiderations of stereotypes in users. Their moral efforts aimed at silencing discriminatory biases would thus be supported.

Of course, designing social robots according to this perspective would expose the products to the risk of being misaligned to user expectations and, therefore, rejected. The contrast between efficiency and moral needs becomes utterly evident when expectations are explicitly challenged in order to trigger processes of moral self-reflection. However, one cannot exclude that a balance can be successfully struck between opposing unethical biases and boosting acceptability [18, 19]. Moreover, the idea of promoting fairness and non-discrimination through social robots cannot but be *prima facie* extremely attractive from an ethical and social standpoint.

In light of the above, it might seem reasonable to support this position and advocate for its endorsement throughout the social robotic community. However, a closer look raises several important questions that require attention.

4.2. Challenges: What, Who, How

The prospect of fighting discrimination through robotics is enticing, but not void of risks. Three main ethical challenges stand out suggesting caution against too enthusiastic endorsements of the proactive approach. They concern the *what*, the *who*, and the *how* of the design strategy.

Let us begin with the *what*. The approach presupposes that discriminatory biases can easily be identified and defined. To an extent, this presupposition is sound. Some forms of discrimination are already acknowledged as utterly unethical. Their eradication is a social objective enshrined in laws and pursued through institutional efforts.

For instance, Article 2 of the Universal Declaration of Human Rights [20] states that human rights and freedoms are entitled to everyone "without distinction of any kind, such as race, colour, sex, language, religion, political or other opinion, social origin, property, birth, or other status". Biases conveying discriminatory attitudes based on these social features could then be taken as starting points to determine which beliefs could be legitimately opposed through social robotics.

Apart from the most evident cases, however, specifying these general principles in order for them to apply to specific circumstances is bound to generate controversy. Social disagreement is to be expected concerning which biases should be legitimately challenged through social robots and which, on the contrary, should not. The risk of influencing ethical beliefs that are not generally perceived as discriminatory might, in fact, open dangerous avenues to manipulation and propaganda.

These considerations directly lead to doubts regarding the *who*. Who should be put in the position of deciding which ethical beliefs to oppose or promote through the design of social robots?

It would likely be inadequate to entrust this task to robotic companies or to any other group of private citizens. Since the moral fibre of society is at stake, the decision is arguably a collective one. Democratic societies would require general agreement among all relevant stakeholders on the identification of discriminatory biases to eradicate. How to reach agreement on such sensitive and divisive matters remains an open question.

Even if agreement on at least some social biases were achievable, problems would not end there. For example, who should supervise and control that marketed social robots comply with bias-related policies? Again, relying on companies to provide ethical assessments of their own products through internal ethics committees does not seem too promising a strategy. Institutional bodies should then perhaps be established both to supervise and enforce policies on bias. However, this might significantly slow down product development and deployment, impacting negatively on the competitiveness of companies in such a frantic, ever-evolving, and global sector. Moreover, cultural specificities would likely lead to different positions on this matter, which would add yet another degree of complexity to the picture.

Finally, let us considering issues regarding the *how*.

In this sense, threats to users' moral autonomy must be carefully discussed. The risk must at least be considered that nudging people's ethical beliefs through social robotics would amount to moral paternalism [21, 22, 23, 24].

The relationship between moral paternalism and social efforts aimed at eradicating discriminatory biases is complex. One might doubt that the accusation of moral paternalism rightfully applies to this case. Is it truly a form of paternalism to implement cues in technological products which are supposed to limit the spread of discriminatory behavioural patterns? The weaker the agreement concerning the moral status of a bias, and the less the public involvement in the decision-making process, the stronger the charge of paternalism would become.

Let us assume for now that the charge is valid. In this regard, respecting some conditions could perhaps help circumvent the problem [25]. For example, proactive design cues might be required to be always made transparent and not too intrusive, so to respect user moral autonomy. Doing this would attenuate the accusation of moral paternalism and concur to involving users as voluntary participants in a conscious social effort towards the realization of a fairer society through responsible technological innovation.

Nevertheless, the matter is so delicate that caution cannot but be warmly advised. The line that separates ethically commendable or acceptable influence from intrusion and manipulation is too thin and difficult to trace not to ring countless alarm bells.

5. Conclusions

This paper set out to discuss whether social robotics could be framed as a means to the moral education of society. To answer this question properly, it has been necessary to shed light on the relation between social biases, robotic design, and discrimination. The ensuing analysis has reached the following results.

First, the relevance of social biases for the design of social robots has been clarified. Design strategies aimed at leveraging the power of biases in order to enhance the quality of human-robot interactions have been introduced and examined. Moreover, ethical risks consisting in the validation and strengthening of discriminatory social biases have been evaluated.

What makes of social robots potential means for moral corruption, however, also allows for them to be envisioned as possible tools for fighting social discrimination. Given the importance of the matter, the prospect thus revealed is definitely worth exploring–both theoretically and practically.

However, considerable obstacles stand in the way of its pursuit. Problems regarding *what* social beliefs to challenge through social robotics; *who* should be delegated with the task of determining and enforcing them; and, finally, *how* to oppose discrimination without illegitimately infringing on human autonomy have been preliminary sketched. Ultimately, the outlook of social robotics as a means to the moral education of society depends on how thoroughly these issues will be clarified, assessed, and coped with in the future.

References

[1] De Angeli A, Brahnam S. Sex Stereotypes and Conversational Agents. In: Paper Presented at the AVI 2006 Workshop Gender and Interaction: Real and Virtual Women in a Male World, Venice, Italy. 2006. http://brahnam.info/papers/EN2033.pdf
[2] Brahnam S, De Angeli A. Gender affordances of conversational agents. Interacting with Computers 2012 May;24(3):139-153. https://doi.org/10.1016/j.intcom.2012.05.001
[3] Fossa F, Sucameli I. Gender Bias and Conversational Agents: an Ethical Perspective on Social Robotics. Science and Engineering Ethics 2022;28:1-23. https://doi.org/10.1007/s11948-022-00376-3
[4] Nass C, Steuer J, Tauber ER. Computers are social actors. In: Proceedings of the SIGCHI conference on Human factors in computing systems celebrating interdependence - CHI '94; New York: ACM Press; 1994. p. 72-78.
[5] Nass C, Moon Y, Green N. Are machines gender neutral? Gender-stereotypic responses to computers with voices. Journal of Applied Social Psychology 1997;27(10):864-876. https://doi.org/10.1111/j.1559-1816.1997.tb00275.x
[6] Weßel M, Ellerich-Groppe N, Schweda M. Stereotyping of social robots in eldercare: An explorative analysis of ethical problems and possible solutions. In: Nørskov M, Seibt J, Quick O, editors. Culturally Sustainable Social Robotics-Proceedings of Robophilosophy'2020. Frontiers in Artificial Intelligence and Applications; Amsterdam: IOS Press; 2020. p. 239-246.
[7] Weßel M, Ellerich-Groppe N, Schweda M. Gender Stereotyping of Robotic Systems in Eldercare: An Exploratory Analysis of Ethical Problems and Possible Solutions. International Journal of Social Robotics 2021:1-14. https://doi.org/10.1007/s12369-021-00854-x
[8] McDonnell M, Baxter D. Chatbots and Gender Stereotyping. Interacting with Computers 2019;31(2):116-121. https://doi.org/10.1093/iwc/iwz007

[9] Tay B, Jung Y, Park T. When stereotypes meet robots: The double-edge sword of robot gender and personality in human-robot interaction. Computers in Human Behaviour 2014;38:75-84. https://doi.org/10.1016/j.chb.2014.05.014

[10] Siegel M, Breazeal C, Norton MI. Persuasive robotics: The influence of robot gender on human behaviour. In: 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems; 2009. p. 2563-2568. https://doi.org/10.1109/IROS.2009.5354116

[11] Carpenter J, Davis JM, Erwin-Stewart N, Lee TR, Bransford JD, Vye N. (2009). Gender representation and humanoid robots designed for domestic use. International Journal of Social Robotics 2009;1(3):261-265. http://dx.doi.org/10.1007/s12369-009-0016-4

[12] Robertson J. Gendering Humanoid Robots: Robo-Sexism in Japan. Body & Society 2010;16(2):1-36. https://doi.org/10.1177/1357034X10364767

[13] Sparrow R. Robots, Rape, and Representation. International Journal of Social Robotics 2017;9(4):465-477. https://doi.org/10.1007/s12369-017-0413-z

[14] Bisconti P. Will sexual robots modify human relationships? A psychological approach to reframe the symbolic argument. Advanced Robotics 2021;35(9):561-571. https://doi.org/10.1080/01691864.2021.1886167

[15] Flensborg Damholdt M, Vestergaard C, Seibt J. Ascribing Gender to a Social Robot. In: Nørskov M, Seibt J, Quick OS, editors. Culturally Sustainable Social Robotics- Proceedings of Robophilosophy'2020. Frontiers in Artificial Intelligence and Applications, vol. 335. Amsterdam: IOS Press; 2020. p. 247-256. https://doi.org/10.3233/FAIA200921

[16] Eyssel F, Hegel F. (S)he's Got the Look: Gender Stereotyping of Robots. Journal of Applied Social Psychology 2012;42(9): 2213-2230. https://doi.org/10.1111/j.1559-1816.2012.00937.x

[17] Reich-Stiebert N, Eyssel F. (Ir)relevance of Gender?: On the Influence of Gender Stereotypes on Learning with a Robot. In: 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI'2017), 2017. p. 166-176. https://ieeexplore.ieee.org/document/8534896

[18] Brahnam S, Weaver M. Re/Framing Virtual Conversational Partners: A Feminist Critique and Tentative Move Towards a New Design Paradigm. In: Marcus A, editor. Design, User Experience, and Usability: Users and Interactions. DUXU 2015. Lecture Notes in Computer Science, vol. 9187; Cham: Springer; 2015. p. 172-183. https://doi.org/10.1007/978-3-319-20898-5_17

[19] Shaw-Garlock G. Gendered by Design: Gender Codes in Social Robotics. In: Seibt J, Hakli R, Nørskov M, editors. Sociable Robots and the Future of Social Relations-Proceedings of Robophilosophy'2014. Frontiers in Artificial Intelligence and Applications, vol. 273; Amsterdam: IOS Press; 2014. p. 309-317. https://doi.org/10.3233/978-1-61499-480-0-309

[20] Universal Declaration of Human Rights. https://www.un.org/en/about-us/universal-declaration-of-human-rights

[21] Borenstein J, Arkin R. Robotic Nudges: The Ethics of Engineering a More Socially Just Human Being. Science and Engineering Ethics 2016;22: 31-46. https://doi.org/10.1007/s11948-015-9636-2

[22] Rodogno R. Nudging by Social Robots. In: Nørskov M, Seibt J, Quick OS, editors. Culturally Sustainable Social Robotics- Proceedings of Robophilosophy'2020. Frontiers in Artificial Intelligence and Applications, vol. 335. Amsterdam: IOS Press; 2020. p. 337-345. https://doi.org.10.3233/FAIA200930

[23] Sullins JP, Dougherty S. Ethical Nudging of Users While They Interact with Robots. In: Nørskov M, Seibt J, Quick OS, editors. Culturally Sustainable Social Robotics- Proceedings of Robophilosophy'2020. Frontiers in Artificial Intelligence and Applications, vol. 335. Amsterdam: IOS Press; 2020. p. 346-358. https://doi.org/10.3233/FAIA200931

[24] Westberg M, Jingar M. Preserving Personal Perspectives in Coaching Technology. In: Nørskov M, Seibt J, Quick OS, editors. Culturally Sustainable Social Robotics- Proceedings of Robophilosophy'2020. Frontiers in Artificial Intelligence and Applications, vol. 335. Amsterdam: IOS Press; 2020. p. 359-369. https://doi.org/10.3233/FAIA200932

[25] Sunstein C. The Ethics of Nudging. Yale Journal on Regulation 2015;32(2):413-450. https://www.yalejreg.com/print/the-ethics-of-nudging/