

Research Article

Jessica K. Barfield*

Discrimination against robots: Discussing the ethics of social interactions and who is harmed

<https://doi.org/10.1515/pjbr-2022-0113>
received October 17, 2022; accepted March 20, 2023

Abstract: This article discusses the topic of ethics and policy for human interaction with robots. The term “robot ethics” (or roboethics) is generally concerned with ethical problems that may occur when humans and robots interact in social situations or when robots make decisions which could impact human well-being. For example, whether robots pose a threat to humans in warfare, the use of robots as caregivers, or the use of robots which make decisions which could impact historically disadvantaged populations. In each case, the focus of the discussion is predominantly on how to design robots that act ethically toward humans (some refer to this issue as “machine ethics”). Alternatively, robot ethics could refer to the ethics associated with human behavior toward robots especially as robots become active members of society. It is this latter and relatively unexplored view of robot ethics that this article focuses on, and specifically whether robots will be the subject of discriminatory and biased responses from humans based on the robot’s perceived race, gender, or ethnicity. If so, the paper considers what issues are implicated, and how society might respond? From past research, preliminary evidence suggests that acts of discrimination which may be directed against people may also be expressed toward robots experienced in social contexts; therefore, discrimination against robots as a function of their physical design and behavior is an important and timely topic of discussion for robot ethics, human–robot interaction, and the design of social robots.

Keywords: ethics, robot discrimination, by-stander harm, human–robot interaction, roboethics, societal harm

1 Introduction

In the last few decades, robots have been widely used in industrial applications performing tasks such as part assembly, spot welding, material handling, and painting [1]. More recently, robots have entered other areas of the workforce, often coming into close contact with people and performing tasks which require a range of social skills. For example, in the retail and hospitality industries, robots may serve as customer greeters [2], shopping assistants [3], and guides [4]; in education, robots may serve as teachers and assistants [5]; and in entertainment, robots may serve as performers [6]. In addition, robots are currently entering society in the role of friend or companion, displaying ever increasing social skills [7].

There is widespread agreement among roboticists, philosophers, legal scholars, and legislators that as robots enter society it is desirable that they act ethically toward humans [1,8–13]. Generally, robot ethics is a topic within the emerging field of information ethics which focuses on the ethical standards and moral codes governing human conduct toward robots in society [14,15]. Tamburrini [16] described robot ethics as “a branch of applied ethics which endeavors to isolate and analyze ethical issues arising in connection with present and prospective uses of robots” (p. 12). Tamburrini also posed the question of whether we should “[...] regard robots, just like human beings, that is, as moral agents and bearers of fundamental rights?” (p. 12) (see [17]). Similarly, Asaro [8] commented that robot ethics should be concerned with the ethics associated with human behavior toward robots, a relatively unexplored topic which forms the focus of this article. Generally, ethics relates to societal standards of human conduct and consists of guidelines and principles that inform people about how to live in a society or how to behave in a particular situation [18]. This article takes the position that emerging social robots pose significant challenges to current views of ethics and that various actors may be harmed from a psychological and emotional perspective as a result of discriminatory, biased, and aggressive conduct directed against robots.

* Corresponding author: Jessica K. Barfield, School of Information Sciences, University of Tennessee-Knoxville, Knoxville, TN, 37966, USA, e-mail: jbarfiel@vols.utk.edu

In terms of how people interact with robots, emerging evidence suggests that people may discriminate against robots under various circumstances, for example, as a function of their perceived gender, ethnicity, or race [19–21]; yet policies to address robot discrimination are just beginning to be discussed [22,23]. In the US, legislation to address discrimination resulted in the passing of Title VII of the Civil Rights Act of 1964 which was enacted to prohibit racial, religious, gender, ethnic, and color discrimination in employment and housing (see also [24]). As mentioned, the Civil Rights Act identified ethnicity, gender, and color as categories which could result in discrimination toward individuals; interestingly these categories represent characteristics that may also be attributed to robots (see generally [25]). This raises the question of whether humans might engage in discriminatory behavior toward a robot designed with physical characteristics that in the past has led to discrimination among humans [19,26,27].

Under current anti-discrimination law, the aggrieved party is a natural person and not an artificial entity, thus, a question of interest for the robotics community is whether any party is harmed if humans discriminate against robots, and if so whether new policy, law, and ethical guidelines should be created to prohibit robot discrimination? Further, the possibility of discriminatory conduct directed against robots raises interesting ethical concerns regarding human treatment of robots, especially given that current robots are entities that are not considered to be independently morally considerable [28]. While rights for conscious robots is a fascinating albeit future-oriented topic, even so, if acts of discrimination against robots that lack self-awareness result in harm to human actors, then the ethical treatment of robots is of legitimate interest to society.

Perhaps the best-known discussion of robot ethics is Isaac Asimov's Three Laws of Robotics. The Three Laws are as follows: "a robot may not injure a human being or, through inaction, allow a human being to come to harm; a robot must obey the orders given it by human beings except where such orders would conflict with the First Law; and a robot must protect its own existence as long as such protection does not conflict with the First or Second Laws" [29]. Asimov's first two laws on robotics directly protect the rights of humans *vis-a-vis* robots, but do not provide rights for robots, that is, protection against discriminatory or other harmful acts directed against them. However, the third law provides some relief for the robot – it is allowed to protect its "own existence" (as long as that does not conflict with the first two laws), but this implies that the robot knows that it in fact exists,

and that it has rights that need protecting. That robots may eventually gain consciousness and act in such a way as to protect their rights is of course a controversial topic, but for the discussion within this article, robot consciousness is not considered a necessary condition for robot discrimination to have ethical and policy consequences for society in general and in some cases for specific individuals [30].

2 Current views of robot ethics

Generally, robot ethics is a broad topic that is of interest to roboticists, philosophers, and social scientists, and it is likely that a multitude of factors will influence whether or not humans are guided by ethical considerations when interacting with robots. Asaro [8], discussing what we might expect from robot ethics, identified three areas for consideration. The first is that a human (which could be the robot's operator) (Human₃, Figure 1) might act unethically toward another human (Human₂, Figure 1) through the use of a robot [30,31]. Second, according to Asaro [8], robots should be designed such that they do not act unethically themselves. Third, we might consider how people (Human₁ in Figure 1) should treat robots based on ethical standards of conduct [8]. A later section of the article focuses on the human actors shown in Figure 1.

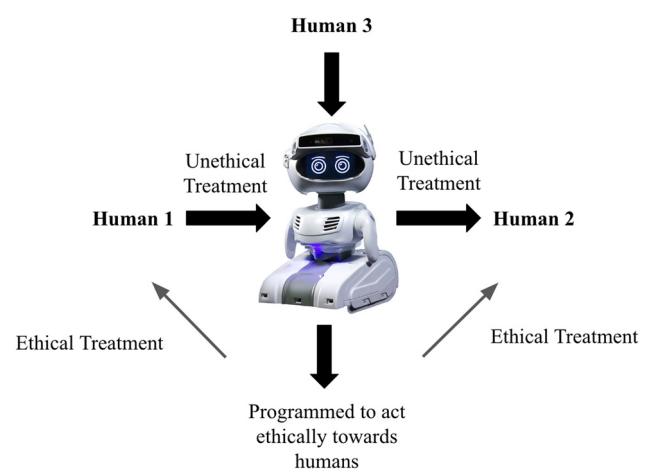


Figure 1: Different scenarios in which robots are implicated by ethical considerations. The robot is an actor which can engage in discriminatory conduct against a human with or without human control (right side); the robot can be programmed to act ethically toward humans; and humans may act unethically toward robots (left side). The arrows can represent an observation of the robot or interaction with the robot depending on the circumstances.

Furthermore, Abney [32] described robot ethics as having different meanings: (i) the professional ethics of roboticists who design robots; (ii) the moral code programmed into robots (sometimes referred to as machine ethics); or (iii) the conscious ability of a robot to engage in ethical reasoning. Extending Abney's categorization and Asaro's [8] view, for robots entering society and experienced in social contexts, I propose that robot ethics should be concerned with the ethical treatment that robots receive by those who interact with them. And further that a robot need not be consciously aware of discriminatory conduct directed against it for the topic of robot ethics to have relevance for the field of ethics and for human–robot interaction (HRI). However, extending the main focus of this article on ethical treatment of robots which lack awareness of an act of discrimination directed against them (the situation which exists now), I briefly touch upon a future in which robots gain consciousness and are aware that it was the subject of discrimination; in that case, an important issue would be to determine what ethical rules and policy should be applied to guide HRIs when a conscious artificial agent is involved [33–35].

2.1 Ethics, robots, and discrimination

Robots that interact with people do so in a multitude of ways; they speak to us, obey our commands, and may examine our facial features to discern our intentions or even our emotional state [36]. In response, according to the computers as social actors (CASA) paradigm, humans tend to interact with robots in a comparable manner as they do with people [37]. On this point, Eysel and Kuchenbrandt [25] observed that people draw on their own self-knowledge, or their knowledge about other people, when judging unfamiliar non-human entities (see [38]). Similarly, Kang *et al.* [39] summarizing decades of research on discrimination have shown that social cognitions influence our evaluation and behavior toward individuals who are placed into a different group. It should be noted that the CASA paradigm does not predict that interactions between humans and robots will be positive (or in some way ethical), just that the interaction will be similar to how humans interact with each other. On this point, the results of prior research suggest that as people interact with robots, they may express biases toward the robot not unlike the biases that people of color, or of certain ethnic groups, or of a particular gender currently receive (see generally [21,40,41]). For this reason, as robots become more and more prevalent throughout society, ethical issues associated with HRI have become an

important topic of discussion. Additionally, discrimination against robots raises a number of questions which need to be addressed sooner rather than later as robots become the subject of human attention. Briefly, some issues of ethics and policy for robots are as follows: (i) Can personhood status be granted to a robot such that it can claim rights? (ii) Would it be a violation of current ethical standards to discriminate against robots? and (iii) Is the society-at-large or people who witness acts of discrimination directed against robots harmed by viewing the act of discrimination?

The question of whether robots may be the subject of discrimination, or in some cases even extreme animus, is motivated by examples on the current treatment of robots and from historical events. A recent example of hostility directed against a robot occurred in 2019 when a person kicked and knocked over a 400-pound security-guard robot that was patrolling a parking-garage structure [42]. In another example, hitchBot, a robot which had successfully hitchhiked across Canada and parts of Europe, was destroyed in Philadelphia by unknown assailants [43]. Also, from a historical perspective as far back as the eighteenth century, there was a rebellion by textile workers in Great Britain who thought that the introduction of technology into their workplace posed a threat to their livelihood; as a result, the “Luddites” destroyed the equipment [44].

While the aforementioned examples show an extreme animus toward machines, as discussed by the CASA theory, people often react to robots in far more subtle or stereotypical ways such as attributing human characteristics to the robot based on its physical appearance or behavior [45,46]. For robots, Sparrow [19] commented that people attributed racial and/or ethnic identities to robots and suggested that research which showed that robots are placed into racial categories pose unique ethical and political challenges to building humanoid robots. Sparrow [19,27] further argued that the design of humanoid robots that may be perceived as representing a particular race presents a difficult ethics problem for robot designers [19]. Barfield [26] found that the surface color of a robot (black, brown, white) could influence the task the robot would be selected to perform suggesting that perceived robot race (as a function of the robots' surface color) may influence the user's evaluation of the robot.

That people anthropomorphize robots has been shown in numerous studies and suggests that if a person interacts with a robot that has attributes which differ from those of the observer, the person may react in a stereotypical or discriminatory manner toward the robot [25,38,45,47–50]. In fact, a number of social psychology studies have concluded that in-group or out-group bias, and thus preferential or non-preferential treatment, can be triggered by markers of physical similarity, such as skin color; or in the case of robots, the

color of the material used to design the surface of the robot [26]. Illustrating this point, Bartneck *et al.* [51] showed that robots designed with different colors were “racialized” by observers. Further, Eyssel and Kuchenbrandt [25] found that observers exhibited a biased reaction toward robots which were thought to represent a member of a different ethnic group. What the aforementioned studies highlight is that robots perceived as a member of the observer’s “out-group” are judged less favorably than robots perceived as members of the observer’s “in-group” – a reaction which researchers have shown can lead to discrimination against out-group members under various contexts (see generally [21]).

From a theoretical perspective, as predicted by the CASA paradigm, people treat robots as social actors and under some circumstances, as if they have a race [21,50]. Bartneck *et al.* [51] and Piske [52] commented that when forming opinions about other people we often rely on social cues such as age, gender, and race (see also [53]). Based on the results of past research studies, the human tendency to identify and stereotype along racial lines is also prevalent in human interaction with robots [21,41]. For example, Eyssel and Hegel [20] have shown that people use a variety of cues to categorize non-human entities and there is also a tendency to place robots into racial and ethnic categories. Eyssel and Kuchenbrandt [25], studying biases directed against robots, used German citizens to evaluate robots that were given a Turkish or German identity. The results of their study indicated that the robot introduced to the German citizens as a Turkish product compared to the same robot introduced to subjects as an in-group, or German product, received less preference among German citizens. Further, De Angeli *et al.* [53] found that people are more likely to engage in antisocial behaviors when interacting with technology designed with humanlike and engendered forms, and that robots designed with female gender cues could be the subject of unintended sexual attention and harassment (see [54]). Additionally, Bartneck *et al.* [51] noted that manipulations of robot shape and hairstyle often elicited gender stereotypical responses. Given that robots may be perceived as representing a particular gender, these findings raise interesting questions for HRI: as an example, whether it is ethical to design female-appearing robots that are expected to serve gender roles in society [46]. Summarizing the earlier discussion, previous studies have shown that individuals evaluate robots as an in-group or out-group member along the dimensions of perceived ethnicity, race, and gender. A question raised in this article is whether discrimination occurs for robots perceived to be a member of an out-group, and if so, which actors are harmed, and what is the nature of the harm?

With robots becoming more humanoid in appearance, intelligent, and social in behavior, there is a growing effort to investigate whether people discriminate against robots and respond to them based on gender, ethnicity, and racial stereotypes. For example, Eyssel and Hegel [20] investigating the effect of facial cues on the perception of robot gender asked whether a robot designed as gendered female would be stereotyped in user interactions as female, and whether a robot designed as gendered male would be stereotyped as male. The findings indicated that the same gender stereotypes which bias social perceptions of humans are also applied to robots. For example, “male-appearing” robots were ascribed more agency-related traits, and “female-appearing” robots were ascribed more communal traits [20]. More recently, Otterbacher and Talias [55] also found that people responded using stereotypical responses to robots thought to represent a particular gender. Using videos of male- or female-appearing robots, participant’s evaluations of the female-gendered robots were categorized as being emotionally warm and the male-gendered robots as being more agentic [55].

In addition to gender, there are other factors which influence how an individual evaluates another person and whether they exhibit a bias or discriminatory response against that person. For example, as alluded to earlier, it has been shown that people are evaluated more positively if they are perceived as a member of the evaluator’s “in-group” and negatively if not [49]. This in-group bias effect was tested in the domain of social robots by Eyssel and Longham [21]. In their study, participants who identified as Caucasian were asked to rate robots whose appearance resembled the participants’ in-group or resembled a social out-group designed with “Afrocentric features.” Contrary to expectations, more agency was attributed to the out-group robots, and, as an explanation, Eyssel and Longham [21] attributed the unexpected finding to the participants’ desire to appear egalitarian and unprejudiced. The extent to which this represents a consistent and valid response to social robots remains to be determined in future studies. Eyssel and Kuchenbrandt [25] also investigated the effect of social category membership on the evaluation of humanoid robots. Their results showed that subjects who rated a robot which was designed to represent either their in-group or a national out-group not only rated the in-group robot more favorably but used social categorization processes and differential social evaluations when viewing the robots [25].

On the topic of robot discrimination and the stereotyping of robots, Louine *et al.* [41] investigated the perception of robots of certain colors. Respondents were presented with pictures of black-, yellow-, and neutral-colored robots and were asked to indicate their evaluation of the robot on

a number of dimensions. The results suggested that black-colored robots were viewed as significantly stronger than yellow robots, that respondents were more likely to move away from black-colored robots, and that yellow robots were viewed as more affable than black robots. Further, using the shooter-bias paradigm, Bartneck *et al.* [51] provided strong evidence that people discriminate against robots thought to represent a different race than the observer. In their study, subjects viewed white- or black-colored robots that either had a gun or other object in their hand. Bartneck and colleagues found that people shot quicker if the robot was darker than if it was lighter. In a related study, Barfield [26] showed that robot color can evoke an emotional response from people in various situations. Varying the surface color of robots, she found that participants thought society would discriminate against a black- or rainbow-colored robot more so than a robot colored as white. Further, a black-colored robot was thought to be stronger than a white- or yellow-colored robot and participants indicated that a red- and black-colored robot would be selected more often to commit an assault than the other robots.

One overarching conclusion from the aforementioned studies is that human observers are able to place robots into different social categories based on their appearance. Remarkably, this is a human skill developed at an early age. For example, Matsuda *et al.* [56] used robots to examine infant discrimination ability between humans, an android robot, and a more “mechanical-looking” robot. Subjects consisted of three groups, 6- to 8-month, 9- to 11-month, and 12- to 14-month-old infants. In the study, a human and robot image, or the two robot images, were presented side-by-side in the visual field facing the infants and using eye-tracking equipment the infant’s time spent focusing on each of three areas of the images were recorded. The results provided by Matsuda and colleagues showed that infants that were within the 6- to 14-month age range were able to discriminate a mechanical-appearing robot from a human but were not able to distinguish between an android robot and human [56]. This result is particularly interesting for HRI given that psychologists have concluded from extensive research that people learn to discriminate against other people early in their lives [57,58].

3 Robot discrimination: the actors and harms involved

An important question for robot ethics and policy relating to the design and use of social robots is who is harmed if a robot is the subject of discrimination [59]? As a starting

point to the discussion, we know that current robots are not conscious agents and thus have no awareness of whether or not they are the subject of discrimination; therefore, as a threshold question, does discriminatory acts directed against a robot result in any harm or ethical consideration that needs to be addressed [60]? Further, under current legal doctrine we know that robots are not considered legal persons; therefore, even if their physical appearance or behavior elicits a negative reaction, robots have no individual rights that they can pursue [61]. So, lacking consciousness and legal rights, if a robot is the subject of discrimination, are any parties harmed? Put another way, what harms could be involved based on discrimination against robots?

In general, Lippert-Rasmussen [60] has argued that multiple actors may suffer adverse consequences if acts of discrimination occur within society. One harm which may occur is what I term “witness harm,” which could manifest itself if a third-party witnessed robot discrimination and experienced adverse consequences such as depression and stress [62]. Another harm which may occur is more general, but I view it as an important harm to “society-at-large” [63]. Generally, acts of robot discrimination could lead to an unjust society and lack of trust within society. Further, if robot discrimination is tolerated, current gender and race discrimination prevalent within society could be exacerbated [62]. In addition, the operator controlling the robot could also suffer emotional harm if the robot controlled by the operator was discriminated against or treated in an unfair manner, especially if the robot was anthropomorphized by the operator. Finally, the individual discriminating against the robot could be harmed. For example, Coeckelbergh [22,23] described virtue ethics as being concerned with a person’s moral character when they engage in discriminatory behavior, stating that fundamentally what is at issue is the moral character of the individual, more so than the object of discrimination. Similarly, Mamak [64] commented that discriminating against another “[...] damages the kindly and human qualities in himself, which he ought to exercise in virtue of his duties to mankind” (p. 6). My conjecture that the aforementioned harmful outcomes could occur when humans interact with robots is supported by past research (reviewed in an earlier section of this article) which has shown that a robot’s physical design, and particularly its color, perceived gender, and ethnicity may lead to people engaging in discriminatory or stereotypical reactions toward the robot [19,21,26,27,45].

As a summary to the earlier discussion, Figure 2 identifies several parties that could suffer emotional or psychological harm as a result of discrimination directed

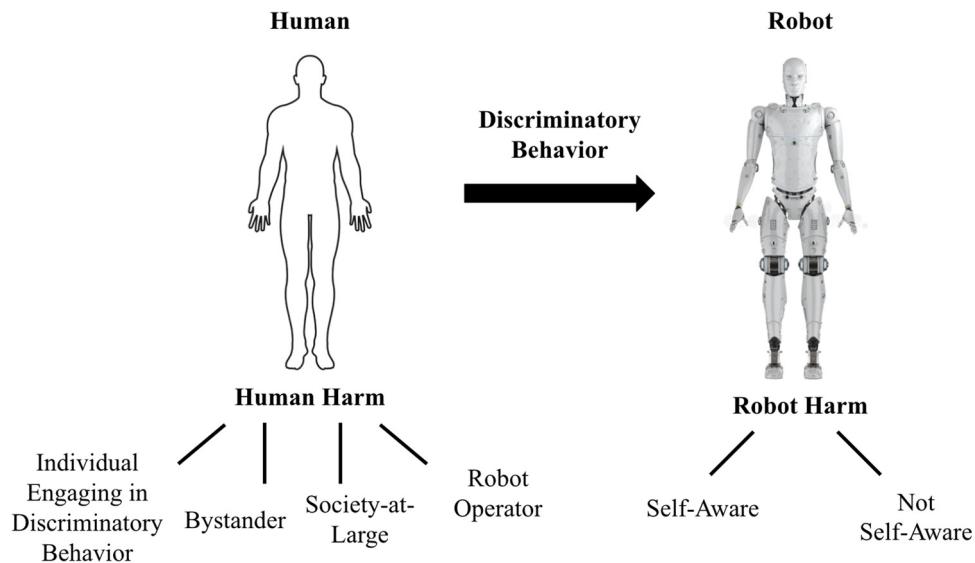


Figure 2: Discrimination against robots as a function of robot appearance and which actor could experience harm by the discriminatory acts directed against robots.

against a robot. Specifically, the left side of the figure indicates harm to members of society, and, for completeness, the right side of the figure identifies harm which could occur if a future robot were aware (or not) of the discriminatory acts directed against it. However, in this article I focus predominately on the left side of Figure 2, which addresses current issues which relate to harm resulting from discriminatory acts directed against non-conscious robots (although the same analysis presented here would apply to conscious robots and I briefly address that point in the following).

Considering robot discrimination, Figure 2 summarizes the various actors which could experience harm (resulting from mental to physical harm). First, the “harm to society-at-large” argument states that irrespective of whether the robot has self-knowledge of discriminatory acts directed against it, society-at-large may still experience harm due to the consequences of discriminatory acts occurring within society. A related topic may help illustrate this point. Consider the discussion among media scholars and legislators which concerns the playing of video games that contain violent content [65]. While the virtual avatars that are killed, maimed, and assaulted in such games neither actually suffer any physical harm nor are they consciously aware of being harmed, there is concern that society itself could suffer negative consequences based on a game player becoming desensitized to violence [66]. On this point, Ryan *et al.* [67] commented that moral choices such as those involved in game play could affect the ethical values of the player and that unethical behavior learned in game play could “leak out” into the world outside the video

game. Therefore, as with the ethical issues associated with game play, even though current robots are not aware of discriminatory acts directed against them, society may experience the harmful effects of robot discrimination which may also “leak out” and influence human interactions with other individuals. Illustrating this, while discussing the ethical treatment of robots, Sparrow [68] argued that it would be unethical to design robots that were programmed to explicitly refuse sexual advances in order to facilitate the rape fantasy of some individuals. As Sparrow noted, such acts could symbolically represent the rape of a real woman, show disrespect for women, and represent and exploit a significant character defect in the individual. As with video games, the concern is that behavior learned or practiced with robots could manifest itself in society. On this point, consider the Kantian view on animal cruelty which holds that our actions toward animals reflect our morality [69]; therefore, by extension, if we treat robots in inhumane ways, we become inhumane persons, and as a result society could be negatively impacted.

Additionally, a harm to humans shown on the left side of Figure 2 is “by-stander” harm (see generally, [70]), which I describe as the harm which could be experienced by individuals witnessing acts of discrimination directed against robots. A study by Connolly *et al.* [71] illustrates that if a robot were abused, a bystander would be more likely to prosocially intervene when the robot expressed sadness in response to the abuse compared to a condition in which the robot ignored the abuse. Given that people anthropomorphize robots, I propose that a bystander may experience emotional

reactions triggered by witnessing robot discrimination, and thus they may become indirect victims of discrimination. On this point, Wofford *et al.* [72] found that people witnessing discrimination may experience depression and other health-related effects.

Another harm that I propose may occur when acts of discrimination are directed toward a robot is the harm to the individual engaging in discriminatory conduct. While not the victim of discrimination, still, the person is engaging in negative behavior toward a robot and thus may experience the negative consequences of inappropriate behavior based on normative societal standards. On this point, Mamak [64] asks what the mistreatment of robots informs us about a person's character. Sparrow [68] commented that even if the mistreatment of robots does not predict a person's future behavior toward other people, it may reveal something about the person's character in general, offering society reason to be concerned with their behavior.

Given that robots may be the subject of discrimination as they interact with people, it is worth briefly discussing the right side of Figure 2, and whether it would be ethical to create robots that were consciously self-aware and thus able to experience the deleterious effects of discrimination that may accompany robots as they enter society. Basl [73] discussed several ideas concerning the ethics of creating robots exhibiting artificial consciousness. His main argument is that creating artificial entities that are conscious might be unethical on the grounds that humans would likely wrong such robots. On this point, note the results of the aforementioned literature review and real-world examples where robots have been wronged in various ways. In Basl's [73] view to determine whether it is wrong to harm an entity that has artificial consciousness, it is necessary to discuss whether such entities have moral status. Basl [73] concludes that to have moral status is to have moral significance, meaning in some contexts a moral human agent will be responsive to a robot because it has moral status. Refraining from performing discriminatory acts directed against robots considered to have moral status would be an example. Further, Basl [73] uses the term "moral patient" as a particular form of moral status, which means that an entity's interests will be taken into account by human agents in their moral deliberations. Basl [73] proposed that it would be acceptable for scientists to create an entity with artificial consciousness, but only if the entity was treated with the status of a moral patient. However, according to Basl if there was any reason to expect that such a being would be treated in a way not commensurate with the status of a moral patient, then we should not create the artificially

conscious entity, at least without adequate protection. Similarly, Sparrow [74] when discussing the ethical treatment of robots focused on the obligations that people should have toward robots that gain human levels of intelligence. Asaro [8] also commented that creating conscious robots would lead to the question of what ethics are required for HRI. Finally, in a discussion of robot ethics, the *Montréal Declaration for Responsible Development of Artificial Intelligence* from the Forum on the Socially Responsible Development of AI, it was emphasized that there should not be cruel behavior toward robots that take on the appearance of human beings and act as humans do in society (see generally [75]).

4 Toward a theoretical framework for ethics and HRI

Based on the earlier discussion, I propose that a theoretical framework be developed to guide efforts on establishing ethical and policy guidelines for HRI given that robots may be the subject of discrimination and animus in society. Such a framework will help facilitate interdisciplinary collaboration between researchers across disciplines and provide a structure in which to evaluate HRI in the context of ethical treatment of robots. Based on a literature review of models used to investigate HRI, I summarize five different approaches which I propose can be used to conceptualize research on the ethical treatment of robots. While there is overlap between the five areas, still, I believe the categories reflect a formal way to discuss theories which could be applied to the ethical treatment of robots. The proposed theoretical frameworks are as follows:

- 1) *Theories relating to robot anthropomorphism:* Several theories have attempted to explain how individuals interact with social robots based on anthropomorphizing robots [76]. Generally, such theories are based on the perceptual and cognitive processes that are used to interpret the physical features and behavior of a robot in human terms. One example of such a theory is the Sociality, Effectance, and Elicited Agent Knowledge (SEEK) theory which was proposed by Epley *et al.* [47]. Just as discrimination can result from an automatic process [77], SEEK states that anthropomorphism occurs subconsciously within moments of interacting with a non-human agent. Additionally, one aspect of interacting with robots that are anthropomorphized is that they elicit an emotional response from users, and the Cognitive Appraisal Theory proposed by Lazarus [78] discusses

how arousal provides the basis for any emotion (here arousal could be a function of the robot's appearance or behavior). So how arousal leads to emotions and how emotion triggers ethical considerations for HRI are topics in need of further exploration.

- 2) *Theories based on the Uncanny Valley effect:* The Uncanny Valley effect has been replicated in numerous studies and represents the eerie reaction to robots that approach human likeness, but not quite achieving that goal. Given that people may discriminate against robots based on their appearance, the Uncanny Valley effect is useful for explaining why robots of a certain appearance could elicit a discriminatory response from those who interact with them. A related theory proposed by Gray and Wegner [79] indicates that humans may feel threatened by humanoid robots based on their appearance which could lead to discriminatory responses. Similarly, under Self-Completion theory, people whose self is threatened, for example, by the physical appearance of a robot, are motivated to acquire symbols to offset the threat [80]; the extent to which people accumulate symbols as a response to social interactions with a robot could potentially represent a metric to measure the Uncanny Valley effect.
- 3) *Gender and ethnic stereotypes:* Several HRI studies have been carried out to determine if the design of robots results in triggering stereotypes based on the robot's appearance and behavior. An extension of the CASA paradigm states that humans mindlessly apply the same social heuristics used for human interactions to robots because robots call to mind similar social attributes as humans [50]. Thus, robots that are genderized may elicit gender stereotype responses [81] and those appearing as a certain ethnicity may elicit discriminatory responses based on ethnicity.
- 4) *Role theory:* As robots perform social tasks, they take on social roles which could implicate ethical considerations. Role theory examines robots as a function of the roles they take within society. Under role theory, a role is thought to be a cluster of functional, social, and cultural norms that dictate how interacting parties should act in a given situation [82]. Within this general genre of theories, the Social Identity Theory discussed by Turner and Oakes [83] says that social identity is the portion of an individual's self-concept derived from perceived membership in a relevant social group and has been used to explain why people discriminate against robots perceived as an out-group member.
- 5) *Information processing (IP) theories:* Several theories have focused on how individuals process information

in response to interactions with robots. Such studies include reaction time measures and memory questions. The Social Exchange Theory discusses social behavior (and thus involves IP) for the interaction of two parties and implements a cost–benefit analysis to determine risks and benefits of the interactions [84]. By extension, it would be of interest to determine if a cost–benefit analysis was done when individuals discriminate against robots in social contexts.

5 Addressing robot discrimination and concluding remarks

Summarizing the earlier discussion on ethics, discrimination, and HRI, research has shown that people categorize robots as a function of the robot's appearance, and in many circumstances react to robots in a positive or negative manner as they do with other people [37]. Further, Coghlan *et al.* [75] commented that there appears to be a link between robots and human character suggesting that our treatment of robots could have societal consequences. Past research has shown that people may react in a hostile or discriminatory manner toward robots, especially if the robot's design represents the appearance of a different group than the observer [9,21,26,51].

Discussing the ethical treatment of robots, according to Darling [85], robots designed with social skills by nature of their form and behavior will elicit virtue-promoting or vice-promoting effects which may impact their treatment and the treatment of people. On this point, some philosophers have asked whether treating social robots kindly would make people kinder, and conversely, if treating robots with cruelty, could make people more callous [75]? There is evidence to support both of these conjectures. For example, Darling [85] reported the results of a study done by the US military that involved crippling a six-legged robot and stated that the study was halted by the subjects for being inhumane. In a workshop run by Darling, she noted that after an hour allowing people to socialize with a robot, people refused to hurt the robot. In contrast, another study found that children in Japan would, in the absence of their parents, verbally abuse, kick, and punch a service robot in a shopping mall [86]. Specifically, Nomura *et al.* [86] found that children verbally abused a robot, repeatedly obstructed its path, and sometimes even kicked and punched the robot. Based on a process of interviews, the majority of the children indicated that they engaged in abuse because they wanted

to test the robot's reactions, actually enjoyed abusing it, considered the robot human-like, and actually thought the robot was capable of perceiving their abusive behaviors [86].

Interestingly, as robots continue to enter society with the likelihood of taking on different physical appearances, human reactions might mirror the “Uncanny Valley” effect in which images that approach humanness, but not quite so, may elicit an eerie feeling among observers [87]. In terms of robot's physical appearance, one may be surprised to learn that, as late as the 1970s, there were so-called “ugly laws” operating in the US (within local municipalities) which were designed to discourage people with deformities from appearing in public [88]. Such laws made it illegal for any person, who is diseased, maimed, mutilated, or deformed in any way, so as to be an unsightly or disgusting, to expose himself or herself to public view [89]. Given the Uncanny Valley effect for robots, society may express a “humancentric bias” against robots and judge them as members of out-groups. As a response, could “ugly laws” be enacted to regulate the appearance of robots in society? If so, then perhaps robots whose appearance placed them in the dip of the Uncanny Valley curve would not be allowed to expose themselves in public; they would only be allowed in manufacturing facilities or other non-public spaces. However, I do not advocate for such a law or think such laws are a likely outcome in HRI, but discuss the possibility based on the extensive research which has shown that the Uncanny Valley effect applies to humanoid robots [79,90–92] and that “ugly laws” were enacted in the past for humans.

In this article, the results of several studies showed that robots may be the subject of discriminatory acts, particularly based on their physical appearance. I also discussed how society-at-large and third-party individuals may experience harm as a result of robot discrimination. I view the discussion presented in this article as highlighting important issues for human treatment of current versions of non-sentient robots, and I recognize that significant other issues of ethics and policy will result if future robots gain consciousness and thus are aware of how they are treated by their human companions. If robots do gain consciousness (at least at some rudimentary level), then the issue of legal personhood for robots becomes relevant, and whether such robots should be considered to have moral status [93–95]. Looking to the future, as Basl [73] notes, for any debate on the moral and legal status of robots, a better understanding of artificial consciousness, artificial rationality, artificial sentience, and similar concepts is needed.

Acknowledgement: The author thanks the School of Information Sciences at the University of Tennessee-Knoxville for support during the writing of this manuscript.

Funding information: The author states no funding to declare.

Conflict of interest: The author states no conflict of interest.

Ethical approval: The conducted research is not related to either human or animal use, thus ethical approval for the work is not applicable.

Informed consent: The article does not include any study; thus, no informed consent was obtained during the research work.

Data availability statement: Data sharing is not applicable to this article, as no datasets were generated during the research work.

References

- [1] M. Miller, *Robots & robotics: Principles, systems, and industrial applications*, McGraw-Hill Education, 2017.
- [2] S. Moore, S. Bulmer, and J. Elms, “The social significance of AI in retail on consumer experience and shopping practices,” *J. Retail. Consum. Serv.*, vol. 64, p. Article 102755, 2022.
- [3] V. Flores, J. F. Villa, M. A. Porta, and J. G. Jaguey, “Shopping market assistant robot,” *IEEE Trans. Lat. Am. Trans.*, vol. 13, no. 7, pp. 2559–2566, 2015.
- [4] J. H. Lim and H. I. Kim, “Development of an autonomous guide robot for campus tour,” *Trans. Korean Soc. Mech. Eng.*, vol. 41, no. 6, pp. 543–551, 2017.
- [5] Z. Sun, Z. Li, and T. Nishimori, “Development and assessment of robot teaching assistant in facilitating learning,” *6th International Conference of Educational Innovation Through Technology (EITT)*, 2017, pp. 165–169.
- [6] J. Hudson, *The robot revolution: Understanding the social and economic impacts*, Edward Elgar Publisher, 2019.
- [7] H. Jiang, S. Y. Lin, V. Prabakaran, M. R. Elara, and L. Y. Sun, “A survey of users' expectations towards on-body companion robots,” *ACM Designing Interactive Systems Conference (DIS)*, 2019, pp. 621–632.
- [8] P. M. Asaro, “What should we want from a robot ethic?,” *Int. Rev. Inf. Ethics*, vol. 6, no. 12, pp. 9–16, 2006.
- [9] P. M. Asaro, A body to kick, but still no soul to damn: Legal perspective on robotics, In: *Robot ethics: The ethical and social implications of robotics*, P. Lin, K. Abney, G. A. Bekey, (eds), MIT Press, 2012, 169–186.
- [10] A. Jori, *Principi di roboetica filosofia practica e intelligenza artificiale, Paternò, Nuova Ipsa*, 2019.
- [11] D. Leben, *Ethics for robots: How to design a moral algorithm*, Routledge Press, London, 2019.

- [12] B. Malle, “Integrating robot ethics and machine morality the study and design of moral competence in robots,” *Ethics Inf. Technol.*, vol. 18, no. 4, pp. 243–256, 2016.
- [13] W. Wallach, C. Allen, *Moral machines: Teaching robots right from wrong*, Oxford University Press, New York, 2009.
- [14] P. Lin, Introduction to robot ethics, In: *Robot ethics: The Ethical and Social Implications of Robotics*, P. Lin, K. Abney, G. A. Bekey, (eds), MIT Press, Cambridge, MA, 2012.
- [15] V. K. Suraj, *Encyclopaedic dictionary of library and information science*, Isha Book, Gyan Publishing House, Delhi, India, 2005.
- [16] H. Tamburini, “Robot ethics: A view from the philosophy of science,” *Ethics Robot*, pp. 11–22, 2009.
- [17] J. Robertson, “Human rights vs. robot rights: Forecasts from Japan,” *Crit. Asian Stud.*, vol. 46, no. 4, pp. 571–598, 2014.
- [18] G. Reynolds, *Ethics in information technology*, 6th edn, Cengage Learning Press, Boston, MA, 2018.
- [19] R. Sparrow, “Do robots have race? Race, social construction, and HRI,” *IEEE Robot. Autom. Mag.*, vol. 27, pp. 144–150, 2020.
- [20] F. Eyssel and F. Hegel, “(S)he’s got the look: Gender stereotyping of robots,” *J. Appl. Soc. Psychol.*, vol. 42, no. 9, pp. 2213–2230, 2012.
- [21] F. Eyssel and S. Loughnan, It don’t matter if you’re black or white? Effects of robot appearance and user prejudice on evaluations of a newly developed robot companion,” In: *Social Robotics, ICSR 2013, Lecture Notes in Computer Science*, G. Herrmann, M. J. Pearson, A. Lenz, P. Bremner, A. Spiers, U. Leonards, (eds), vol. 8239, Cham, Springer, 2013, pp. 422–433.
- [22] M. Coeckelbergh, “Why care about robots? Empathy, moral standing, and the language of suffering,” *J. Philos. Sci.*, vol. 20, pp. 141–158, 2018.
- [23] M. Coeckelbergh, “How to use virtue ethics for thinking about the moral standing of social robots: A relational interpretation in terms of practices, habits, and performance,” *Int. J. Soc. Robot.*, vol. 13, pp. 31–40, 2021.
- [24] J. J. Ramsey, “Basics of employment law: Understanding and dealing with adverse employment actions and discrimination in the workplace,” *J. Extra-corporeal Technol.*, vol. 37, no. 3, pp. 253–255, 2005.
- [25] F. Eyssel and D. Kuchenbrandt, “Social categorization of social robots: Anthropomorphism as a function of robot group membership,” *Br. J. Soc. Psychol.*, vol. 51, pp. 724–731, 2012.
- [26] J. K. Barfield, “Discrimination and stereotypical responses to robots as a function of robot colorization,” *Adjunct Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization*, 2021, pp. 109–114.
- [27] R. Sparrow, “Robotics has a race problem,” *Sci. Technol. Hum. Values*, vol. 45, no. 3, pp. 538–560, 2020.
- [28] M. Coeckelbergh, “Robot rights? Towards a social-relational justification of moral consideration,” *Ethics Inf. Technol.*, vol. 12, pp. 209–221, 2010.
- [29] I. Asimov, Runaround, *Astounding Science Fiction (magazine)*, 1942.
- [30] D. Levy, “The ethical treatment of artificially conscious robots,” *Int. J. Soc. Robot.*, vol. 1, pp. 209–216, 2009.
- [31] S. Gless, E. Silverman, and T. Weigend, “I robots cause harm, who is to blame? Self-driving cars and criminal liability,” *N. Crim. Law Review: An. Int. Interdiscip. J.*, vol. 19, no. 3, pp. 412–436, 2016.
- [32] K. Abney, Robotics, ethical theory, and metaethics: A guide for the perplexed, In: *Robot ethics: The ethical and social implications of robotics*, P. Lin, K. Abney, G. A. Bekey, (eds), MIT Press, 2012, pp. 35–52.
- [33] T. Kitamura, T. Tahara, and K. Asami, “How can a robot have consciousness? *Adv. Robot.*, vol. 14, no. 4, pp. 263–275, 2000.
- [34] B. J. MacLennan, *Consciousness in robots: The hard problem and some less hard problems*, 14th IEEE Workshop on Robot and Human Interactive Communication (RO-MAN), 2005, pp. 434–439.
- [35] J. Waskan, Robot consciousness, *Routledge handbook of consciousness*, London, UK, 2018, pp. 408–419.
- [36] D. McColl, A. Hong, N. Hatakeyama, G. Nejat, and B. Benhabib, “A survey of autonomous human affect detection methods for social robots engaged in natural HRI,” *J. Intell. Robot. Syst.*, vol. 82, pp. 101–133, 2016.
- [37] C. Nass, J. Steuer, and E. R. Tauber, “Computers as social actors,” *CHI’95 Conference*, 1994, pp. 72–78.
- [38] N. Epley, A. Waytz, S. Alkalis, and J. T. Cacioppo, “When we need a human: Motivational determinants of anthropomorphism,” *Soc. Cognit.*, vol. 26, pp. 143–155, 2008.
- [39] J. Kang, N. Dasgupta, K. Yogeeswaran, and G. Blasi, “Are ideal litigators white? Measuring the myth of colorblindness,” *J. Empir. Leg. Stud.*, vol. 7, no. 4, pp. 886–915, 2010.
- [40] M. Keijsers and C. Bartneck, “Mindless robots get bullied,” *Proceedings of the ACM/IEEE International Conference on Human Robot Interaction*, Chicago, 2018, pp. 205–214.
- [41] J. Louine, D. C. May, D. W. Carruth, C. L. Bethel, L. Strawderman, and J. M. Usher, “Are black robots like black people? Examining how negative stigmas about race are applied to colored robots,” *Social. Inq.*, vol. 88, no. 4, pp. 626–648, 2018.
- [42] A. Torresz and P. D. Hayward, Searching for person who kicked, damaged robot security guard, *KTVU FOX 2 News*, 2019. <https://www.ktvu.com/news/hayward-p-d-searching-for-person-who-kicked-damaged-robot-security-guard>.
- [43] D. H. Smith and F. Zeller, “The Death and lives of hitchBOT: The design and implementation of a hitchhiking robot,” *Leonardo*, vol. 50, no. 1, pp. 77–78, 2017.
- [44] S. W. Elliott, “Anticipating a luddite revival,” *Issues Sci. Technol.*, vol. 30, no. 3, pp. 27–36, 2014.
- [45] C. Bartneck, E. Croft, D. Kulic, and S. Zoghbi, “Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots,” *Int. J. Soc. Robot.*, vol. 1, pp. 71–81, 2009.
- [46] J. Bernotat, F. Eyssel, and J. Sachse, “Shape it – The influence of robot body shape on gender perception in robots,” *International Conference on Social Robotics (ICSR)*, 2017, pp. 75–84.
- [47] N. Epley, A. Waytz, and J. T. Cacioppo, “On seeing human: A three-factor theory of anthropomorphism,” *Psychol. Rev.*, vol. 114, pp. 864–886, 2007.
- [48] H. Kamide, F. Eyssel, T. Arai, Psychological anthropomorphism of robots, In: *Social robotics, ICSR 201, Lecture Notes in Computer Science*, G. Herrmann, M. J. Pearson, A. Lenz, P. Bremner, A. Spiers, U. Leonards, (eds), vol. 8239, Cham, Springer, 2013.
- [49] J. C. Turner, Social comparison, similarity and ingroup favoritism, In: *Differentiation between social groups: Studies in the social psychology of intergroup relations*, H. Tajfel, (ed.), Academic Press, Cambridge, MA, 1978.

[50] C. Nass and S. Brave, *Wired for speech*, MIT Press, Cambridge, MA, 2005.

[51] C. Bartneck, K. Yogeeswarm, Q. M. Ser, G. Woodward, S. Wang, R. Sparrow, et al., "Robots and racism," *ACM/IEEE International Conference*, 2018, pp. 1–9.

[52] S. T. Piske, *Stereotyping, prejudice, and discrimination*, Boston, MA, McGraw Hill, 1998.

[53] A. De Angeli, S. Bahnman, P. Wallis, and A. Dix, "Misuse and use of interactive technologies," *CHI'06 Ext. Abstr. Hum. Factors Comput. Syst.*, pp. 1647–1650, 2006.

[54] S. Bahnman and A. De Angeli, "Gender affordances in conversational agents," *Interact. Comput.*, vol. 24, no. 3, pp. 139–153, 2012.

[55] J. Otterbacher, M. Talias, "S/he's too warm/agentic! The influence of gender on uncanny reactions to robots," *HRI'17 Conference*, 2017, pp. 214–223.

[56] G. Matsuda, H. Ishiguro, and K. Hiraki, "Infant discrimination of humanoid robots," *Front. Psychol.*, vol. 6, pp. 1–7, 2015.

[57] J. Bolgatz, "Revolutionary talk: Elementary teacher and students discuss race in a social studies class," *Soc. Stud.*, vol. 96, no. 6, pp. 259–264, 2005.

[58] A. Skinner, How do children acquire prejudices? *Psychology Today*, 2019. <https://www.psychologytoday.com/us/blog/catching-bias/201911/how-do-children-acquire-prejudices>.

[59] K. B. Rasmussen, "Harm and discrimination," *Ethic Theory Moral. Pract.*, vol. 22, pp. 873–891, 2019.

[60] K. Lippert-Rasmussen, (ed.), *The routledge handbook of the ethics of discrimination*, New York, Routledge, 2018.

[61] D. J. Gunkel, "Robots can have rights; Robots should have rights," *Robot. Rights*, pp. 79–116, 2018.

[62] S. Samuel, *Humans keep directing abuse – Even racism – At robots*, 2019. <https://www.vox.com/future-perfect/2019/8/2/20746236/ai-robot-empathy-ethics-racism-gender-bias>.

[63] R. A. Lengardt, "Understanding the mark: Race, stigma, and equality in context," *N. Y. Univ. Law Rev.*, vol. 79, no. 3, pp. 803–931, 2004.

[64] K. Mamak, "Should violence against robots be banned? *Int. J. Soc. Robot.*, vol. 14, no. 1, pp. 1–50, 2022.

[65] A. Suziedelyte, "Is it only a game? Video games and violence," *J. Econ. Behav. Organ.*, vol. 188, pp. 105–125, 2021.

[66] A. T. Prescott, J. D. Sargent, and G. H. Hull, "Metanalysis of the relationship between violent video game play and physical aggression over time," *Natl Acad. Sci.*, vol. 115, no. 40, pp. 9882–9888, 2018.

[67] M. Ryan, P. Formosa, P. Howarth, and D. Staines, "Measuring morality in videogames research," *Ethics Inf. Technol.*, vol. 22, pp. 55–68, 2020.

[68] R. Sparrow, "Robots, rape, and representation," *Int. J. Soc. Robot.*, vol. 889, pp. 465–477, 2017.

[69] L. Fedi, "Mercy for animals: A lesson of secular morality and its philosophical history," *ROMANTISME*, vol. 142, pp. 1–25, 2008.

[70] S. Iguchi, H. Takenouchi, M. Tokumaru, "Sympathy expression model for the bystander robot in group communication," *7th International Conference on Humanoid Nanotechnology, Information Technology, Communication and Control, Environment and Management*, 2014, pp. 1–6.

[71] I. Connolly, V. Mocz, N. Salomins, J. Valdez, N. Tsoi, B. Scassellati, et al., "Prompting prosocial human intervention in response to robot mistreatment," *ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2020, pp. 211–220.

[72] N. Wofford, A. M. Defever, and W. J. Chopik, "The vicarious effects of discrimination: How partner experiences of discrimination affect individual health," *Soc. Psychol. Personal. Sci.*, vol. 10, no. 1, pp. 121–130, 2019.

[73] J. Basl, "The ethics of creating artificial consciousness," *APA Newsl. Philos. Comput.*, vol. 13, no. 1, pp. 23–29, 2013.

[74] R. Sparrow, Can Machines be People? Reflections on the Turing Triage Test, In: *Robot Ethics: The Ethical and Social Implications of Robotics*, P. Lin, K. Abney, G. A. Bekey, (eds), MIT Press, Cambridge, MA, 2012, pp. 301–315.

[75] S. Coghlan, F. Venere, J. Waycott, and B. B. Neves, "Could social robots make us kinder or crueler to humans and animals? *Int. J. Soc. Robot.*, vol. 11, pp. 741–751, 2019.

[76] A. Waytz, J. Cacioppo, and N. Epley, "Who sees human? The stability and importance of individual differences in anthropomorphism," *Perspect. Psychol. Sci.*, vol. 5, no. 3, pp. 219–232, 2010.

[77] R. M. Blank, M. Dabady, and C. F. Citro, (eds), *Measuring robot discrimination*, National Academy of Sciences, Washington, D.C., 2013.

[78] R. S. Lazarus, *Emotion and adaptation*, Oxford, Oxford University Press, 1991.

[79] K. Gray and D. M. Wegner, "Feeling robots and human zombies: Mind perception and the uncanny valley," *Cognition*, vol. 125, no. 1, pp. 125–130, 2012.

[80] A. T. Higgins, "Self-discrepancy: A theory relating self and affect," *Psychol. Rev.*, vol. 94, no. 3, pp. 319–340, 1987.

[81] A. M. Koenig and A. H. Eagly, "Evidence for the social role theory of stereotype content: observations of groups' roles shape stereotypes," *J. Personality Soc. Psychol.*, vol. 107, no. 3, pp. 371–392, 2014.

[82] M. Soloman, C. Suprenant, J. Czepial, and E. Gutman, "The role theory perspective on dyadic interactions: The service encounter," *J. Mark.*, vol. 40, no. 1, pp. 99–111, 1985.

[83] J. Turner and P. Oakes, "The significance of the social identity concept for social psychology with reference to individualism, interactionism and social influence," *Br. J. Soc. Psychol.*, vol. 25, no. 3, pp. 237–252, 1986.

[84] R. Cropanzano and M. S. Mitchell, "Social exchange theory: An interdisciplinary review," *J. Manag.*, vol. 31, no. 6, pp. 874–900, 2005.

[85] K. Darling, Extending legal protection to social robots: The effects of anthropomorphism, empathy, and violent behavior towards robotic objects, In: *We Robot Conference*, University of Miami; Published in Robot Law, R. A. Calo, M. Froomkin, I. Kerr, (eds), Edward Elgar, 2012.

[86] T. Nomura, T. Kanda, H. Kidokoro, Y. Suehiro, and S. Yamada, "Why do children abuse robots? *Interact. Stud.*, vol. 17, no. 3, pp. 347–369, 2016.

[87] M. Mori, K. F. MacDorman, and N. Kageki, "The uncanny valley," *IEEE Robot. Autom.*, vol. 19, no. 2, pp. 98–100, 2012.

[88] S. M. Schweik, *The ugly laws: Disability in public*, NYU Press, New York, NY, 2010.

[89] P. Burgdorf and R. Burgdorf, Jr., "A history of unequal treatment: The qualifications of handicapped persons as a suspect class under the equal protection clause," *St. Clara Lawyer*, vol. 15, no. 4, pp. 855–910, 1975.

- [90] H. Brenton, M. Gillies, D. Ballin, and D. Chattin, *The Uncanny Valley: Does it Exist?*, 2005. <http://www.davidchatting.com/research/uncanny-valley-hci2005.pdf>.
- [91] K. Dautenhahn, Robots we like to live with? —A developmental perspective on a personalized, life-long robot companion? *Proceedings of the 13th IEEE International Workshop on Robot and Human Interactive Communication (RO-MAN)*, 2004, pp. 17–22.
- [92] K. Dautenhahn, S. N. Woods, C. Kaouri, M. L. Walters, K. L. Koay, I. Werry, “What is a robot companion—friend, assistant or butler?” *Proceedings of IEEE RSJ International Conference on Intelligent Robot Systems (IROS’05)*, 2005, pp. 1488–1493.
- [93] J. J. Bryson, M. E. Diamantis, and T. D. Grant, “Of, for and by the people: The legal lacuna of synthetic persons,” *Artif. Intell. Law*, vol. 25, pp. 273–291, 2017.
- [94] S. M. Solaiman, “Legal personality of robots, corporations, idols and chimpanzees: A quest for legitimacy,” *Artif. Intell. Law*, vol. 25, pp. 155–179, 2017.
- [95] L. B. Solum, “Legal personhood for artificial intelligences”, 70 *N.C. L. Rev*, vol. 1231, 1992.