CrossMark

An Ethical Evaluation of Human–Robot Relationships

Maartje M. A. de Graaf[1]

Abstract When people interact with socially interactive robots on a regular basis, it could be that people start developing some kind of relationship with such robots. People are able to get attached to several objects in our everyday world. However, the relationships we build with regular objects differ significantly from those we may build with socially interactive robots. In contrast to common nonhuman objects, socially interactive robots act autonomously, which increases people's expectations about its capacities. Moreover, human–robot interactions are constructed according to the rules of human–human interaction, inviting users to interact socially with robots. This paper fosters a discussion on the ethical considerations of human–robot relationships and discusses whether these bonds between humans and robots could contribute to the good life. Research on people's interactions with and social reactions towards socially interactive robots is necessary to shape the ethical, societal and legal perspectives on these issues, and facilitates the design of responsible robotics and the successful introduction of robots into our society.

1 Introduction

The prices of robot manufacture are dropping, which facilitate the introduction of service robots into our lives in unprecedented numbers [66]. However, there is limited

✉ Maartje M. A. de Graaf
  maartje.degraaf@utwente.nl

[1] Department of Communication Science, University of Twente, Drienerlolaan 5, 7522 NB Enschede, The Netherlands

knowledge about the emotional effects robots could elicit from their users. Understanding about this topic is an essential requirement for discussions about the social and ethical issues regarding what roles we (want to) allow robots to fulfill in our lives. Research in the area of human–robot interaction provides evidence that people can establish some kind of emotional or social bond with socially interactive robots [33,39,41,72,78]. A socially interactive robot is a robot that elicits social responses from their human users because they follow the rules of behavior expected by these human users [31]. Robots designed for purposes in everyday environments must operate in spaces specifically designed for humans. To ease the communication with its users, robots are designed to evoke social interactions (or just reactions) following the rules of human social interaction behaviors. It is commonly assumed that people prefer to interact with machines in a similar manner they are used to with other human beings [26]. Ideally, a social robot is capable of communicating and interacting in such a sociable way that the robot allows its users to: (1) understand the robot in human social terms; (2) relate to the robot; and (3) to empathize with the robot [14]. Researchers in social robotics aim to develop such sociable machines by making use of models and techniques generally used in human–human communication.

People interact with robotic or computer interfaces in a similar way they do with other human beings [41,56]. Along with this human tendency to respond socially to nonhuman objects, it has been argued that the fundamental human motivation of the 'need to belong' [8,17] not only induces ones desire for meaningful and enduring relationships with other social beings, but also facilitates the likelihood people may form emotional attachments to artificial beings [43]. This issue of bonding with nonhuman objects is likely to be enlarged when these objects possess lifelike abilities and are

endowed with humanlike capacities, as is the case for socially interactive robots.

Relationships with others lie at the very core of human existence. Humans are conceived within relationships, born into relationships, and live their lives within relationships with others [11]. People are not only capable of forming relationships with other humans, they also have the ability to engage in social relationships with nonhumans. The interactions between humans and computer technologies are increasingly complex and humanlike, increasing the importance of the role psychological aspects of our relationships with those technologies take on [13]. In the future, robots are expected to serve humans in various social environments, such as the home, eldercare facilities, the workplace, and schools. In addition to their functional requirements, robots performing tasks in these social environments also require socially interactive components [24]. Besides performing their monitoring and assistive tasks, robots in social environments must engage in social interactions and create relationships with their users in order to achieve their goals. For example increasing an elderly person's health or successfully teaching new mathematical insights to scholars.

Currently, human–robot interactions are constructed along the rules of social behaviors in humanlike interpersonal interactions, which invites people to have meaningful social interactions with robots. With a minimal of social cues users tend to personify these socially interactive robots, which fosters a human willingness to form unidirectional emotional bonds with these robotic others [61]. Is there something morally wrong with deceiving humans into thinking they can foster meaningful interactions with a technological object? Or is this just a logical next step in our technological world? Would it be possible for people to become friends with robots? Are friendships between humans and robots desirable? What implications does this have on future generations, who will be growing up in the everyday presence of robots? How will this impact the formation of human friendships?

Regardless of the moral or ethical implications of the relationships between humans and robots, socially interactive robots will be entering our everyday lives as soon as their abilities are technically feasible for the application in real world contexts. This calls for an ethical evaluation of human–robot relationships. This paper aims to foster a discussion on the ethical considerations of human–robot relationships and discusses whether potential bonds between humans and robots could contribute to the good life. I start the ethical evaluation of human–robot relationships from the philosophical concepts of the good life, which is connected to the psychological concept of well-being. Consecutively, I will argue that the ethical evaluation of human–robot relationships needs to consider the roles nonhumans can play in our human social networks and that robots should be regarded as a new technological genre. Afterwards follows the ethical evaluation of

human–robot relationships, taking into account that people perceive and treat robots as social entities. The conclusion of this paper addresses some societal and ethical consequences of human–robot relationships, followed by an outline of some future directions for their ethical evaluation.

2 Friendship and the Good Life

The good life is a philosophical term for the most desirable life, often related to Aristotle's teaching in ethics, and refers to the physical and psychological well-being of people [15]. Aristotle held that Eudaimonia, often translated as happiness or personal flourishing, is the goal of human thought and action, and exhibited through the cultivation of human virtues in accordance with reason [3]. Although traditionally being a philosophical topic, in recent decades well-being has become an important concern in psychology [15]. The good life necessarily entails well-being, which denotes the desirable condition of our existence and the end state of our pursuit [83]. Positive psychology approaches to well-being focus on finding and nurturing talent and making normal life more fulfilling [64]. Research in positive psychology aims at developing positive practices that enhance human well-being, and focuses on supporting positive experiences, positive individual qualities, or positive social processes and institutions [15]. A Eudaimonia approach to positive psychology, which recognizes the moral and ethical underpinnings of well-being [83], denotes that happiness consists of objective factors and therefore correspond to objective list approaches in philosophy [63]. One part of this perfectionist approach to the good life is the engaged life, which involves, among other aspects, intimate relations or friendships.

Attachment is an essential part of the expression of virtue and having friendships forms a structural feature of the good life [68]. According to Aristotle [3], friendship holds the reciprocal and mutually acknowledged relation of goodwill and affection existing between individuals who jointly share an interest on the basis of virtue, pleasure or utility. The intrinsic worth of friendships lies in their ability to facilitate a person's realization of his or her virtue and the achievement of happiness. As Sherman [68] explains, 'to have intimate friends and family is to have interwoven in one's life, in an ubiquitous way, persons toward whom and with whom one can most fully and continuously express one's goodness' (p. 594). Postulating that individual happiness can flourish through friendships implies some notions of the extended self. People convey and extend their self-concepts from objects surrounding them, including persons, places and objects they feel attached to [9]. Objects bear our connection to others and help express our sense of self [58]. And in the case of intimate friendship, people are willing to give without even having to ask and often without a return expected [3],

because people perceive this as acting in the interests of this extended self [68]. The question remains how robots could fit in to these notions on friendship and the good life? Before we can adequately address this question, I would first like to call attention to the concept of technical mediation to frame the impact of technology on society, and the acknowledgement of robots as a new technological genre, which affects how we may apply ethics to human–robot relationships.

3 The Impact of Robot Technology on Society

A deterministic approach to technology perceives the role of technology in society from the concept of technology, which only allows for risk assessment, i.e., the ethical review of the normative aspects of the intended use purposes for which the technology is designed or the quality of the functionalities the technology represents [75]. The theory of scripts challenges this purely functional vision of technology by describing how the various roles technological artifacts play in everyday contexts and by illustrating how engineers' assumptions about user preferences and competences display themselves in the technical content of an object [2]. Designers anticipate on possible use scenarios for the products they are creating and, implicitly or explicitly, build prescriptions for these use scenarios into the materiality of that product [75]. Latour [45] expands this idea by showing how technologies steer human behaviors, morally and otherwise. For example, speed bumps make us slow down, and a door-spring makes us close the door. Thus, the scripts inherent to technology also help shape the behaviors of the user.

Following the line of thought from Akrich and Latour, Verbeek [75] demonstrates how our decision making is affected by technologies in such a way that moral decisions are a result of an intermediately relation between humans and technologies. When people start using technologies, these technologies shape and are shaped by their use context. Social norms, values and morals are both implicitly and explicitly intertwined with technologies, reinforcing or altering our beliefs and practices. Once a robot has entered a social environment, it will alter the distribution of responsibilities and roles within that environment as well as how people act in that use context or situation [84]. This shift in beliefs and practices is what Verbeek [76] calls technical mediation: "when technologies are used, they help to shape the context on which they fulfill their function, they help to shape human actions and perceptions, and create new practices and ways of living" (p. 92).

The notion of technical mediation is also presented in the sociological concept of social shaping of technology that, just like technical mediation, emerged from a long-standing critique on deterministic views on technology. The social shaping of technology perspective explaining that there are

choices, both consciously and subconsciously, inextricably linked to both the design of individual objects and in the direction or trajectory of innovation programs [82]. Hence, the meaning of a technology is not restricted to a technology's mechanisms, physical and technical properties, or actual capabilities. This view on technology also extends to how people believe they will or ought to interact with a technology, and the way it will or should be incorporated into and affect their lives. Similarly, the domestication theory denotes the complexity of everyday life and the role technology plays within its dynamics, rules, rituals, patterns and routines [69]. By investigating the impact technology has once it becomes an actor in a social network of humans and nonhumans in the home, the domestication theory relates to the concept of technical mediation. The domestication theory focuses on the impact a technology has in terms of its meaning in the social environment, how this meaning is established, and how the incorporation of the technology in daily routines propagates or alters existing norms and interpretation of values in the home.

Together, the above described theories or perspectives on technology (use) raise critical considerations for the roles nonhumans play in social interactions. By acknowledging technologies as more than just functional instruments, but without being withdrawn in classical images of technologies as a determining and threatening power, we can start investigating the impact of technologies in society. The concept of technology is largely defined by the way people and the societies in which they live in perceive, respond, and interact to these technologies. Consequently, technologies also have moral relevance given their role in mediating one's beliefs and practices [76]. Therefore, it can be concluded that technology has a socially-embedded meaning because the shaping process of people's understanding of a technology is intertwined with the evolving prevailing social norms. Recalling the subject of human–robot relationships, the remaining question is: how will we share our world with these new social technologies, and how will a future robot society change who we are, how we act and interact—not only with robots but also with each other?

4 Applying Ethics to a New Technological Genre

The application of robot ethics is strongly related to the way robots are perceived. Acknowledging robots strictly as machines implies that they do not possess any hierarchically higher characteristics, nor should they be provided with consciousness, free will, or a level of autonomy superior to what is technically embedded by the designer [77]. Yet, people's interaction experiences with robots fundamentally differ from people's experiences interacting with most other technologies due to a strong social or emotional com-

ponent in human–robot interactions. For instance, people act upon or talk to some robots as if the robot has sensations and emotions, even though existing robots are not sentient and lack feelings. Indeed, people's social responses to robots have been reported in many human–robot interaction studies, e.g., [6,39,41,47,72]. A lot of robot ethics discussions focus on the technological abilities of robots, i.e., what is in the 'mind' of the robot [8]. It has been argued that the emphasis of robot ethics should shift to the robot's outside, i.e., what robots do to us [20]. Although most people would reasonably agree that robots are programmed machines that only simulate social behavior, the same people seem to 'forget' this while interacting with these sociable machines [32].

4.1 Physical Presence and Environmental Awareness

Robots differentiate from the most other technologies in that they are embodied, autonomous, and mobile technologies capable of sensing our (social) environments. Although new media is also advancing the integration of their communication modes, e.g., with images, sounds, text and data in a single medium [23], robotic devices hold even richer communication modes since they add nonverbal communication and higher social presence to their interaction capabilities. Because robots are embodied, we share our physical worlds with these socially interactive technologies [14]. Robots can produce a physical impact on our worlds by holdings hands, cleaning our home and tripping over objects. Having a physical body enables robots to perturb and be perturbed by their environment, which creates a stronger social presence [46, 48], and affects the users' perceptions and interactions with robots [14,80,81]. Specifically, socially interactive robots are different from other technologies because they must be able to perform their actions in our social environments as well. And our human world is not easily simplified without the implementation of potentially unacceptable restrictions for their users [14]. Robots are able to perceive naturally offered cues of their users using cameras or microphones. Although certain other interactive media have the same affordance, this is clearly more limited for a screen character than for a robot whose sensors can move with it and can remain compelling eye contact. Due to the strong social and emotional component caused by their physical presence in our personal spaces and their ability to sense our social environments, people's interacting experiences with robots appears to be different from people's interaction experiences with most other technologies.

4.2 Social Responses to Robots

With a minimum of social cues, people tend to evaluate technological objects as social entities. This phenomenon is known as the media equation [56], which has also been suc-

cessfully applied to the field of social robotics, e.g., [39,41]. People have a tendency to assign a level of intelligence and sociability to robots which affects the perceptions of those people of the way interactions should proceed. When robots are capable of natural language dialog, the users expectations are not only raised with respect to the natural language, but also regarding the robot's intentionality of both verbal and nonverbal behaviors, its autonomy, and its awareness of the sociocultural context [70]. Robots embedded with sociable interaction features, such as familiar humanlike gestures or facial expressions in their designs, are likely to further encourage people to interact socially with those robots in a fundamentally unique way. For human users, interaction with robots is, in a sense, more as if one is interacting with an animal or another person rather than interacting with a technology.

Furthermore, the increased autonomy of robot behavior extends the likelihood that users will regard such robots as humanlike social entities [42,62]. People tend to associate a robot's autonomous behavior with intentionality, which prompts and reinforces people's perception of agency in robots. Agency entails the capacity to act, and transfers the notion of intentionality [22]. Robots with physically embodiments capable of sociable interactions elicit a unique and affect-charged sense of active agency, which people experience as similar to that of living entities [86]. This ensures that human–robot interaction, in a sense, is perceived as if one is interacting with an animal or another person, and not as if one is interacting with a technological artifact.

4.3 A New Ontological Category

With robots recognizing our faces, making eye contact, and responding socially, they are pushing our Darwinian buttons by displaying behaviors that we associate with sentience, intentions and emotions [72]. Social robots could be interpreted as both animate and inanimate. Animate, because they move and talk. On the other hand they are inanimate because they are programmed machines. So it seems like a new ontological category is about to emerge through the creation of socially interactive robots, and this process magnifies when robots become increasingly sophisticated in their social abilities [39]. This line of work suggests that socially interactive robots should be perceived as a new technological genre, which has been argued by other researchers in social robotics as well [85]. A first evidence for this new technological genre can be found in the notion that young children cannot easily distinguish robots from alive or not alive [12,36]. This despite the fact that children of the same age consistently group living and nonliving entities as distinct classifications in terms of biological, psychological and perceptual characteristics from the age of five [35]. So it seems that children

do not attribute aliveness to robots in similar ways as they do with other canonical entities.

4.4 Moral Agency for Robots

A specific characteristic of robots points towards an extremely relevant notion of how everyday life is changing: "robots are a new form of living glue between our physical world and the digital universe we have created" (p. XV) [55]. Together with the statements from the previous paragraphs, this denotes that we need to perceive robots as an evolution of a new species. Consequently, this requires us to consider robots to have autonomy and consciousness, and that robots need to be created with moral and intellectual features [77]. Current robot owners conceptualize their robotic pets in terms of essence, agency, and social standing, but seldom attribute moral standing [39]. Thus, although people are willing to treat a robot as a companion according to these findings, they are not ready to endow it with human rights. Yet, there is evidence that autonomous robots can be a target for moral blame. People expect robots to act upon moral decisions and more so blame them for failing to act, whereas this is the exact opposite for humans in a similar situation [52]. Moreover, some researchers expect that the manifold ways in which we may treat robots will be similar to how that we treat other humans, and therefore postulate that ethical behavior towards robots is merely an extension of such treatment [49]. These arguments obliges us to develop robots that are embedded with morality and the ability to make sound decisions.

Independent of the discussion of moral agency for robots, the fact remains that people perceive socially interactive robots as social entities at a much higher level than justified by their technological capacities, even by humans who are fully aware of a robot's technological limitations. For social interactions to occur, people only need to assume the social potential of others [18]. Endowing robots with human capacities allows us to warrant nonhumans as viable others in social interactions. Therefore, I postulate that, when dealing with ethics, we should primarily part from this perception of robots as social entities when examining and assessing human–robot relationships.

5 Evaluating Human–Robot Relationships

To begin the ethical evaluation of human–robot relationships, I embark on a conceptual investigation of the concept of friendship. Aristotle [3] argues that so called 'perfect friendship' is the most worthy form of relationship between people, which requires that people desire their friend for the sake of the other and desires the good for the other [5]. This implies the existence of a relationship between moral equals, in which we acknowledge and enjoy virtue in the other. However, this notion of 'perfect friendship' by Aristotle would only be appropriate for the discussion of human–robot relationships if we view these interactive technologies as morally equal. And the above discussion on morality for robots remained inconclusive. Therefore, it is suggested to focus on the concept of companionship, which realizes only the social end of humans [21]. Companionship contains a similar amount of warmth but is less demanding as it doesn't acquire moral symmetry, and it doesn't entail the need to desire for the other. Desire for company is sufficient.

The current definition of companionship can be applied to human relationships with nonhumans, such as robots, and these relationships can be interpreted as unidirectional emotional bonds initiated from the human user [61]. As stated earlier, people do not only acknowledge the useful aspects of robots, but they also experience robots as social entities. And this challenge traditional ontological categories between animate and inanimate. Then if people treat robots as more than lifeless objects, is it then possible to push this argument into the possibility of moral symmetry? Looking back at the Aristotelian definition of friendship [3], from a user's perspective, this means that the user only has to perceive that the robot initiates friendship by showing good will and affection, and that the robot allows its human user to realize his or her virtue and achievement of happiness.

Research in human–robot interaction have already shown that human users establish feelings of reciprocity and mutuality in their interactions with robots [30,33,44], as well as present findings that human users flourish in the presence of robots [16,78]. If we accept the Aristotelian definition of friendship from a user's perspective, we would ascribe virtual virtue to robots, whatever their designed and intended functions may be. This enables us to let robots contribute to the good life of people. People tend to ascribe human virtue on the bases of consistently good behavior, so this qualification would probably be acknowledged to robots with 'good habits' [21]. Thus robots should no longer be perceived as just machines that follow algorithms, which demands from the designers to create the illusion of virtue in terms of the robot's behavior and motivations. The question is whether or not this is deception, and whether or not this is morally wrong?

5.1 Deception and Delusion?

Ethical concerns related to socially interactive robots, especially those developed for care settings, are increasingly gaining attention [67,71]. However, a lot of the robotics research efforts focus on emotional recognition, interpretation and expression [8]. There are already robots that express emotions through facial expression, but their actual interaction and communication abilities are still very limited [25]. The perception of life largely be determined by

the observation of intelligent behavior, and observed behavior of greater intelligence is granted by people with more rights [6]. People in general, thus, may succumb to accepting robotic companionship without the moral responsibilities that real, reciprocal relationships involve [39]. It is therefore argued that any benefits gained from interactions with robots are the consequences of deceiving people into thinking they could establish a relationship with robots over time.

People can feel happy when interacting with robots and forming relationships with them. However, some argue that this is a delusion because those people mistakenly believe that these robots have properties which they do not, and the failure to apprehend the world accurately is a moral failure [71]. This concern is shared by Turkle [72] who states that we need to think about degree of authenticy we acquire from our technology. Some researchers even claim that already a robot detecting human social cues and responding with humanlike cues should be regarded as a form of deception [79]. However, Coeckelbergh [21] denotes that healthy people, when interacting with virtual agents in virtual worlds, are aware of the virtual others while being aware that this other is not real. People can decide on acting as if something is real even though they understand it is not [8]. Indeed, people imagine about others all the time, as portrayed in the theory of mind [34,37,40]. Severson and Carlson [65] advocate that it is still an open question whether people genuinely perceive robots as social entities with internal states of mind and perceptual experiences, or whether such ascriptions are merely a product of imagination. For example, children often play make-believe or let's pretend games, they knowingly play as-if their dolls are alive [19]. Yet Severson and Carlson acknowledge that this dividing line between 'as' and 'as-if' may further fade away when robots become increasingly more sophisticated and socially interactive. Nonetheless, we should assume that people can enjoy interacting with a robot companion without believing they are interacting with a full-fledged social other.

Although Sparrow and Sparrow [71] rightfully point to the subconscious processes involved in human–robot interactions, these subconscious processes also occur in interpersonal interactions between humans. We can be deceived about other people and even by other people, which is actually more malicious because intended deception is morally worse than unintended deception [21]. As humans, most of the time we play a certain role to some extend and it is difficult to distinguish between what appears to be and reality in human social life. Retrospectively, we consider it to be acceptable that people mask their own emotions to achieve ulterior goals [28,38], such as protecting someone else's feelings and maintaining harmonious relationships or gaining advantages, avoiding negative consequences and preserving self-esteem. Thus putting all emphasis on deception when evaluating human–robot relationships seems misplaced. As long as the user perceives to be served well by a robot and is satisfied with the behavior of that robot, there should not be a problem in this account. Thus the question whether or not human–robot relationships can contribute to the good life and enhance well-being, at least for now, needs to be separated from the discussion on deception. However, one may argue that these counterfeit emotions and behaviors could elicit misplaced trust, which is yet another discussion.

6 Conclusion

This paper aimed to foster a discussion on the ethical considerations of human–robot relationships and discusses whether unidirectional emotional bonds between humans and robots could contribute to the good life. As shown in the paper, we can presume that people may establish some relationship with robotic others [39], and that people could even benefit from these relationships under particular circumstances [16,78]. Therefore, it seems nothing is intrinsically wrong with human–robot relationships as long as we can develop robotic systems that effectively deliver what user's believe to be appropriate care behavior [7,21]. Considering socially interactive robots unethical because their effectiveness relies on deception just oversimplifies the issue [66]. The currency of all human social relationships is performance [29], and rather than it being a bad thing, this is simply how human social interactions work. People have always been performing for other people and now the robots too will perform. When people interact with socially interactive robots, they go beyond the psychology of projection to that of engagement. Although we talk to our pets as we do to socially interactive robots, we do not expect them to respond with a validation of our ideas and there is no one suggesting that pets are similar to humans or on their way becoming like us [72]. So, could a robot really be your friend if it acts like it? As discussed in this paper, robots cannot be our Aristotelian friends since they genuinely lack mutuality and reciprocity [21], even if this is perceived as such by human users. Thus, we need to ensure that human–robot relationships will not replace their human counterparts in social relations, as Sparrow and Sparrow [71] rightfully fear.

Yet, this does not discard the fact that people willingly initiate different types of relationships with socially interactive robots. If people share feelings with robotic others, they may become accustomed to the reduced emotional range that these machines can offer [72]. We know that, especially young, people will carry over their ways of treating one ontological entity over to another [4], for example from animals to humans. What will be the impact of human–robot relationships on the relationships we have with other humans? Will

it change our (moral) standards of friendship and ultimately lower them? Hence, another concern here is that if we come to accept these unidirectional emotional bonds with robotic others, will this degrade our relationships with other humans? There is indeed a concern whether robot technology might replace human contact all together [67]. For example, the application of multiple technological systems that provide eldercare would cease the need for care visits from human caregivers. This is troublesome not only since many elderly already feel socially isolated, but also because social isolation has frequently been associated with negative health outcomes and potential risk factors [54]. Having interactions with significant others, not only for the elderly but for all people, give meaning and guidance to behavior and the development of the social self.

Social interactions are necessary for normal personality development and to appropriate social conduct. Social interaction produces the social self, which is the part of one own personality that connects the individual to society, and is considered an important intervening aspect in human behavior [53]. The interactions we have with both the humans and nonhumans surrounding us serve as parts of ourselves [9]. And, as discussed in the paper, both humans and nonhumans should be considered as actors in a social network of our daily environment [2,69,76,82]. With the increase of digital media use, the possibilities for self-extension have tremendously increased, which is fundamentally changing human behavior [10]. For example, when playing video games, avatars are often seen as an extension of the self, with which we strongly identify and which is transferred into our online behavior and sense of self [74]. And indeed, the effects of the digital age we are living in are already shown in our decreasing social abilities. Adolescents with higher rates of internet use indicated to have significantly worse relationships with their mothers and friends [60]. Additionally, there is evidence that social media and smartphone use makes us narcissistic, selfish, deceitful, dishonest, compulsive and vicious [1,73]. In the prospect of the robotic revolution, some researchers fear the effect frequent close interactions with robots will damage their emotional and social development and may ultimate lead to attachment problems [67]. These concerns are linked to the notion of dehumanization which happens when rationalization [57] exceeds its mark leading to the creation of anti-human technological systems [59] and ultimately the de-socialization of our society when people start striving for 'perfect' relationships [72]. These concerns emphasize that a future robot society could result not only in a possible decrease of our social skills but also a reduced willingness to deal with the complexity of real human relationships. Ultimately, we need to consider not only how everyday encounters and social interactions with robots will change us, but also what implications these changes could have on our society.

7 Future Directions for the Ethical Evaluation of Human–Robot Relationships

Despite its relevance, an integrative approach to address the consequences of innovations lacks attention in the literature. One reason for this gap might be that the sponsors of innovative research are often the companies supplying the innovations who silently assume that their innovations only result in positive consequences. Another reason for the underexposure of the consequences of innovations in general is that the common research method involves questionnaires, which are less appropriate to investigate the impact of innovations. The ideal method for studying the impact of innovations involves multiple observations over extended periods of times. A final reason for the research gap comprises the difficulty of measuring consequences. People are mostly unaware of the possible consequences associated with the introduction of an innovation, resulting in incomplete and misleading conclusions when only the opinions of the general public about possible consequences are included in a study. Addressing the potential consequences of human–robot relationships within our future robot society requires the multidisciplinary insights from the HRI researcher community, as well as the necessity to design future robots with the inclusion of understandings from philosophers, legal scholars and policy makers in the deliberation process.

Robotic technologies are growing faster, increasingly finding their way into different spheres of everyday life, ranging from the work space to private homes, and from transport sector to environmental monitoring, crowd controlling and warfare. The upcoming development of technologies brings along countless social and ethical issues. One serious societal consequence of the unidirectional bonds between users and socially interactive robots is that they provoke psychological dependencies which are likely to be exploited by the companies who created these robots [61]. Moreover, because different cultures and religions have different 'virtues' and 'vices', exhibiting from different worldviews, leading to different results on the same questions [51], it is necessary to conduct research on ethics and rights for robots in different cultural settings and contexts.

In the future, when robot will gradually become more like us and people will increasingly interact with them in an ever more 'natural' social way, the need to address legal, societal and ethical issues becomes increasingly pressing. Since humans perceive and treat robots as social actors in our physical environments, robots should be acknowledged as such and we should apply robot ethics from that perspective. For future generations, interactions with robotic others will be part of every ordinary day. The evermore evaporating line between online and offline will make it easier for the children raised in a future robotics society to socialize with and through these robotic others. Robotics researchers need to

heighten the awareness of the relevant legal, societal and ethical issues before robots become ubiquitous in society. The focus should not only be on preventing 'bad' things from happening, but there is also the necessity to explore the social roles robots can (not) or should (not) perform in the future. When autonomous technologies such as robots will be used in everyday environments, researchers need to map a wide range of possible use scenarios including associated potential consequences for both individual users and society as a whole. If the dissemination of robots within society replicates that of personal computers a few decades ago, the rise of important legal, societal and ethical issues should be expected for our future robot society as well. Therefore, robotics researchers need to deal with these issues if we want to anticipate the potential (negative) consequences of the ubiquitous use of robots in our society.

Recognizing the notion of technical mediation not only fosters value-centered design approaches that consciously incorporate social and cultural meaning into robotic design but also suggests a need to involve users early in the design process. Therefore, addressing these legal, societal and ethical issues cannot be done thoroughly without empirical data on the relationship between human users and robots. Observations of real human–robot interactions in people's natural environments are necessary to reveal the consequences of such interactions both on an individual and societal level. The common presumption is that robots will become ubiquitous in our societies and that it is inevitable that everyone will be using a social robot in their own homes within the not-so-distant future [27,50]. Research on people's interactions with and social reactions towards socially interactive robots is necessary to shape the ethical, societal and legal perspectives on these issues. By understanding people's acceptance of robots and the consequences of potential roles for robots in society, robotics researchers can more effectively address their research efforts towards the design of responsible robotics and the successful introduction of robots into our society.

References

1. Aboujaoude E (2011) Virtually you: the dangerous powers of the e-personality. Norton, New York
2. Akrich M (1992) The description of technological objects. In: Bijker W, Law J (eds) Shaping technology. MIT Press, Cambridge, MA, USA
3. Aristotle: 384–322 BC, Ethica Nichomachea
4. Ascione FR, Arkow P (1999) Child abuse, domestic violence, and animal abuse: linking the circles of compassion for prevention and intervention. Purdue University Press, West Lafayette
5. Badhwar NK (1987) Friends as ends in themselves. Philos Phenomenol Res 48:1–23
6. Bartneck C, van der Hoek M, Mubin O, Al Mahmud A (2007) "Daisy, Daisy, Give me your answer do!": switching off a robot. In: International conference on human–robot interaction, Arlington, VA, USA
7. Baumeister RF, Leary MR (1995) The need to belong: desire for interpersonal attachments as a fundamental human motivation. Psychol Bull 117:497–529
8. Baumgaertner B, Weiss A (2014) Do emotions matter in the ethics of human–robot interaction? Artificial empathy and companion robots. In: International symposium on new frontiers in human–robot interaction, London, UK
9. Belk RW (1988) Possessions and the extended self. J Consum Res 15:139–168
10. Belk RW (2013) Extended self in a digital world. J Consum Behav 40:477–500
11. Berscheid E, Peplau LA (1983) The emerging science of relationships. In: Kelly HH et al (eds) Close relationships. WH Freeman, New York
12. Bernstein D, Crowley K (2008) Searching for signs of intelligent life: an investigation of young children's beliefs about robot intelligence. J Learn Sci 17:225–247
13. Bickmore TW (1998) Friendship and intimacy in the digital age. Online report retrieved from http://www.media.mit.edu/Bbickmore/Mas714/finalReport.html. Accessed 6 May 2011
14. Breazeal CL (2004) Designing sociable robots. MIT Press, Cambridge
15. Brey P (2012) Well-being in philosophy, psychology, and economics. In: Brey P, Briggle A, Spence E (eds) The good life in a technological age. Routledge, New York, pp 15–34
16. Broadbent E, Stafford R, MacDonald B (2009) Acceptance of healthcare robots for the older population: review and future directions. Int J Soc Robot 1:319–330
17. Cacioppo JT, Patrick B (2008) Loneliness: human nature and the need for social connection. Norton, New York
18. Cerulo KA (2009) Nonhumans in social interaction. Annu Rev Sociol 35(531):552
19. Cayton H (2006) From childhood to childhood? Autonomy and dependence through the ages of life. In: Hughes JC, Louw SJ, Sabat SR (eds) Dementia: mind, meaning, and the person. Oxford University Press, Oxford
20. Coeckelbergh M (2009) Personal robots, appearance, and human good: a methodological reflection on roboethics. Int J Soc Robot 1:217–221
21. Coeckelbergh M (2012) Care robots, virtual virtue, and the best possible life. In: Brey P, Briggle A, Spence E (eds) The good life in a technological age. Routledge, New York
22. Dewey J (1980) Art as experience. Perigee Books, New York
23. van Dijk JAGM (2012) The network society, 3rd edn. Sage, London
24. Feil-Seifer D, Mataric MJ (2011) Socially assistive robotics. IEEE Robot Autom Mag 18:24–31
25. Fischinger D, Einramhof P, Wohlkinger W, Papoutsakis K, Mayer P, Panek P, Koertner T, Hofmann S, Argyros A, Vince M, Weiss A, Gisinger C (2013) Hobbit: the mutual care (workshop paper). In: International conference on intelligent robots and systems, Tokyo, Japan
26. Fong T, Nourbakhsh I, Dautenhahn K (2003) A survey of socially interactive robots. Robot Auton Syst 42:143–166
27. Gates B (2007) A robot in every home. Sci Am 296:58–65
28. Gnepp J, Hess DLR (1986) Children's understanding of verbal and facial display rules. Dev Psychol 22:103–108

29. Goffman E (1956) The presentation of self in everyday life. Doubleday, New York
30. de Graaf MMA, Ben Allouch S, van Dijk JAGM (forthcoming) Long-term evaluation of a social robot in real homes. In: Interaction studies special issue: new frontiers in human–robot interaction
31. de Graaf MMA, Ben Allouch S, van Dijk JAGM (2016) Long-term acceptance of social robots in domestic environments: insights from a user's perspective. In: AAAI 2016 Spring symposium on "Enabling computing research in socially intelligent human–robot interaction: a community-driven modular research platform", Palo Alto, CA, USA
32. de Graaf MMA, Ben Allouch S, van Dijk JAGM (2015) What makes robots social? A user's perspective on characteristicsfor social human–robot interaction. In: International conference on social robotics, Paris, France
33. de Graaf MMA, Ben Allouch S, Klamer T (2015) Sharing a life with Harvey: exploring the acceptance of and relationship building with a social robot. Comput Hum Behav 43:1–14
34. Heider F (1958) The psychology of interpersonal relations. Wiley, New York
35. Inagaki K, Hatano G (2006) Young children's conception of the biological world. Curr Dir Psychol Sci 15:177–181
36. Jipson J, Gelman SA (2007) Robots and rodents: children's inferences about living and nonliving kinds. Child Dev 78:1675–1688
37. Jones EE, Davis KE (1965) From acts to dispositions: the attribution process in person perception. In: Berkowitz L (ed) Advances in experimental social psychology. Erlbaum, Hillsdale, NJ
38. Josephs E (1994) Display rule behavior and understanding in preschool children. J Nonverbal Behav 18:301–326
39. Kahn PH, Gary HE, Shen S (2013) Children's social relationships with current and near-future robots. Child Dev Perspect 7:32–37
40. Kelley HH (1967) Attribution theory in social psychology. In: Levine D (ed) Nebraska symposium on motivation. University of Nebraska Press, Lincoln
41. Kerepesi A, Kubinyi E, Jonsson GK, Magnusson MS, Miklosi A (2006) Behavioural comparison of human–animal (dog) and human–robot (AIBO) interactions. Behav Process 73:92–99
42. Kiesler S, Hinds P (2004) Introduction to the special issue on human–robot interaction. Hum Comput Interact 19:1–8
43. Krämer NC, Eimler SN, Pütten AM, Payr S (2011) Theory of companion: what can theoretical model contribute to applications and understanding of human–robot interaction? Appl Artif Intell 25:474–502
44. Lammer L, Huber A, Weiss A, Vincze M (2014) Mutual care: how older adults react when they should helptheir care robot. In: AISB workshop on new frontier in human–robot interaction, London, UK
45. Latour B (1994) On technical mediation: philosophy, sociology, genealogy. Common Knowl 3:29–64
46. Lee KM (2004) Presence, explicated. Commun Theory 14:27–50
47. Lee K, Park N, Song H (2005) Can a robot be perceived as a developing creature? Effects of a robot's long-term cognitive developments on its social presence and people's social responses toward it. Hum Commun Res 31:538–563
48. Lee KM, Jung Y, Kim J, Kim SR (2006) Are physically embodied social agents better than disembodied social agents? The effects of physical embodiment, tactile interaction, and people's loneliness in human–robot interaction. Int J Hum Comput Stud 64:962–973
49. Levy D (2009) The ethical treatment of artificial conscious robots. Int J Soc Robot 1:209–216
50. Lin P, Abney K, Bekey G (2011) Robot ethics: mapping the issues for a mechanized world. Artif Intell 175:942–949
51. MacDorman KF, Cowley SJ (2006) Long-term relationships as a benchmark for robotpersonhood. In: International symposium on robot and human interactive communication, Hatfield, UK
52. Malle BF, Scheutz M, Arnold T, Voiklis J, Cusimano C (2015) Sacrifice one for the good of many? People apply different moral norms to human and robot agents. In: International conference on human–robot interaction, Portland, OR, USA
53. Mead GH (1934) Mind, self, an society. University of Chicago Press, Chicago
54. Nicholson NR (2012) A review of social isolation: an important but underassessed condition in older adults. J Prim Prev 33:137–152
55. Nourbakhsh IR (2013) Robot futures. MIT Press, Cambridge
56. Reeves B, Nass C (1996) The media equation: how people treat computers, television, and new media like real people and places. CSLI, New York
57. Ritzer G (1983) The McDonaldization of society. J Am Cult 6:100–107
58. Rook D (1985) The ritual dimensions of consumer behavior. J Consum Res 12:231–246
59. Royakkers L, van Est R (2015) A literature review on new robotics: automation from love to war. J Soc Robot 7:549–570
60. Sanders CE, Field TM, Diego M, Kaplam M (2000) The relationship of internet use to depression and social isolation among adolescents. Adolescence 35:2387–242
61. Scheutz M (2012) The inherent dangers of unidirectional emotional bonds between humans and socially interactive robots. In: Lin P, Abney K, Bekey GA (eds) Robot ethics: the ethical and social implications of robotics. MIT Press, Cambridge
62. Scheutz M, Schermerhorn J, Kramer J, Anderson D (2007) First steps toward natural human-like HRI. Auton Robots 22:411–423
63. Seligman M (2002) Authentic happiness: using the new positive psychology to realize your potential for lasing fulfillment. Free Press, New York
64. Seligman M, Csikszentmihalyi M (2000) Positive psychology: an introduction. Am Psychol 55:5–14
65. Severson RL, Carlson SM (2010) Behaving as or behaving as if? Children's conceptions of personified robots and the emergence of a new ontological category. Neural Netw 23:1099–1103
66. Sharkey NE (2008) Computer science: the ethical frontiers of robotics. Science 322:1800–1801
67. Sharkey AJC, Sharkey NE (2010) The crying shame of robot nannies: an ethical appraisal. Interact Stud 11:161–190
68. Sherman N (1987) Aristotle on friendship and the good life. Philos Phenomenol Res 47:589–613
69. Silverstone R, Haddon L (1996) Design and the domestication of ICTs: technical change and everyday life. In: Silverstone R, Mansell R (eds) Communication by design. The politics of information and communication technologies. Oxford Press, Oxford, pp 44–74
70. Simmons R, Makatchev M, Kirby R, Lee MK, Fanaswala I, Browning B, Forlizzi J, Sakr M (2011) Believable robot characters. AI Mag 32:39–52
71. Sparrow R, Sparrow L (2006) In the hands of machines? The future of aged care. Mind Mach 16:141–161
72. Turkle S (2011) Alone together: why we expect more from technology and less from each other. Basic Books, New York
73. Twenge J (2009) The narcissism epidemic: living in the age of entitlement. Free Press, New York
74. Vasalou A, Joinson AN (2009) Me, myself and I: the role of interactional context on self-presentation through avatars. Comput Hum Behav 25:510–520
75. Verbeek PPCC (2006) Materializing morality: design ethics and technological mediation. Sci Technol Hum Values 31:361–380
76. Verbeek PPCC (2008) Morality in design: design ethics and technological mediation. In: Vermaas P, Kroes P, Light A, Moore S (eds) Philosophy and design: from engineering to architecture. Springer, Berlin, pp 91–102
77. Veruggio G, Operto F (2006) Roboethics: a bottom-up interdisciplinary discourse in the field of applied ethics in robotics. In: Capurro R, Nagenborg M (eds) Ethics and robotics. OIS Press, Amsterdam, pp 3–8

78. Wada K, Shibata T (2007) Living with seal robots: its sociopsychological and physiological influences on the elderly at a care house. IEEE Trans Robot 23:972–980
79. Wallach W, Allen C (2008) Moral machines: teaching robots right from wrong. Oxford University Press, Oxford
80. Walters ML, Dautenhahn K, Woods SN, Koay KL (2007) Robot etiquette. In: International conference of human–robot interaction, Arlington, VA, USA
81. Wang E, Lignos C, Vatsal A, Scassellati B (2006) Effects head movement on perceptions of humanoid robot behavior. In: International conference of human–robot interaction, Salt Lake City, UT, USA
82. Williams R, Edge D (1996) The social shaping of technology. Res Policy 25:865–899
83. Wong PTP (2011) Positive psychology 2.0: towards a balanced interactive model of the good life. Can Psychol 52:69–81
84. van Wynsberghe AM (2012) Designing robots for care: care centered value-sensitive design. Sci Eng Ethics 19:40–433
85. Young JE, Hawkins R, Sharlin E, Igarashi T (2007) Towards acceptable domestic robots: applying insights from social psychology. Int J Soc Robot 1:95–108
86. Young JE, Sung JY, Voida A, Sharlin E, Igarashi T, Christensen HI, Grinter RE (2011) Evaluating human–robot interaction. Int J Soc Robot 3:53–67

Maartje M. A. de Graaf is currently a postdoctoral researcher at the Department of Communication Science, University of Twente. Her Ph.D. research on the long-term user acceptance of social robots in domestic environments indicated a strong role of people's anthropomorphic responses to robots in the process of long-term acceptance. Envisioning a future in which the social abilities of robots wil only increase, her current research interest focuses on the user's social, emotional and cognitive responses to robots and the societal and ethical consequences of such responses.