EXTENDING LEGAL PROTECTION TO SOCIAL ROBOTS

THE EFFECTS OF ANTHROPOMORPHISM, EMPATHY, AND VIOLENT BEHAVIOR TOWARDS ROBOTIC OBJECTS

Kate Darling[*]

"[H]e who is cruel to animals becomes hard also in his dealings with men." – Kant

INTRODUCTION

Robotic technology has been used in manufacturing for years. More recently, it's beginning to transform other areas of our world. From transportation to health care, the military to education, and elderly care to children's toys, robots are entering our lives in new ways, many of which raise social and ethical questions. In the face of increasingly autonomous technology, liability and privacy concerns have become prominent topics, but there is a lesser-discussed issue that is growing in significance: the emergence and effect of robots that are designed to interact with humans on a social level. Preliminary studies and anecdotal evidence show a human tendency to perceive robots differently

[*] Research Specialist, Massachusetts Institute of Technology (MIT) Media Lab, Fellow at Harvard University Berkman Center for Internet & Society and at the Yale University Information Society Project. kdarling@media.mit.edu

than we do other objects.[1] As more and more robotic companions enter into our lives and homes, and those robots are increasingly engineered to engage us socially, our inclination to project life-like qualities onto robots raises design and policy questions.

People are prone to anthropomorphism, that is, we project our own inherent qualities onto other entities to make them seem more human-like. Our well-documented inclination to anthropomorphically relate to animals translates remarkably well to robots.[2] A key characteristic of social robots is that they are specifically designed to elicit these projections. Studies with state-of-the-art technology indicate that humans already interact differently with social robots than they do with other objects.[3] But if we tend to perceive robots as lifelike things, should we be treating them more like devices or like creatures?

This chapter explores the near-future possibility of regulating people's behavior towards certain types of robots, given that such technology increasingly appeals to our anthropomorphic tendencies.[4] It offers some thoughts on why treating robots more like

[1] Heather Knight, *How Humans Respond to Robots: Building Public Policy through Good Design*, BROOKINGS REPORT (2014); Martin Saerbeck, Christoph Bartneck, *Attribution of Affect to Robot Motion*, PROCEEDINGS OF THE 5TH ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN-ROBOT INTERACTION (HRI2010), Osaka, 53-60 (2010); *see for example also* JULIE CARPENTER, THE QUIET PROFESSIONAL: AN INVESTIGATION OF U.S. MILITARY EXPLOSIVE ORDNANCE DISPOSAL PERSONNEL INTERACTIONS WITH EVERYDAY FIELD ROBOTS, Dissertation, University of Washington (2013).

[2] *See* Brian R. Duffy, *Anthropomorphism and the social robot*, 42 ROBOTICS AND AUTONOMOUS SYSTEMS 179-83 (2003).

[3] *See for example* Cynthia Breazeal, *Toward Sociable Robots*, 42 ROBOTICS AND AUTONOMOUS SYSTEMS, 167-75 (2003); Christoph Bartneck, Michel van der Hoek, Omar Mubin, Abdullah Al Mahmud, *"Daisy, Daisy, Give me your answer do!" – Switching off a robot,* PROCEEDINGS OF THE 2ND ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN-ROBOT INTERACTION, Washington DC, 217–22 (2007); Peter Kahn, Takayuki Kanda, Hiroshi Ishiguro, Nathan Freier, Rachel Severson, Brian Gill, Jolina Ruckert, Solace Shen, *'Robovie, you'll have to go into the closet now': Children's social and moral relationships with a humanoid robot*, DEVELOPMENTAL PSYCHOLOGY, 48(2), 303-14 (2012).

[4] This is not the same conversation as "rights for robots," which assumes a futuristic world of fully autonomous and highly sophisticated androids that are nearly indistinguishable from humans. While technological development may someday lead to such a Blade Runner-esque scenario, we have no idea what the technology or legal system will look like by that point, so most discussions of "robot rights" remain firmly

animals could help discourage human behavior that would be harmful in other contexts. Finally, it demonstrates how this regulation would fit into our current legal framework as analog to animal abuse laws.

WHAT IS A SOCIAL ROBOT?

A *social robot* is a physically embodied, autonomous[5] agent that communicates and interacts with humans on a social level.[6] For the purposes of this chapter, it is important to distinguish social robots from inanimate computers or software, as well as from industrial or service robots that are not designed to be social. Social robots communicate through social cues, display adaptive learning behavior, and mimic various emotional states.[7] Our interactions with them follow social behavior patterns, and are designed to encourage emotional relationships. Examples of early social robots include interactive robotic toys like Sony's AIBO dog and Innovo Labs' robotic dinosaur Pleo; robot companions such as Aldebaran's NAO and Pepper robots; medical and health monitoring devices like the theraputic Paro baby seal and Intuitive Automata's weight

in the philosophical rather than imminently pragmatic realm. This chapter suggests, however, that the development of social robots that interact with us on an emotional level could inspire a deliberation of "protecting" robots based on the societal implications of anthropomorphism.

[5] Autonomy in robotics can be understood to mean as little as the ability to perform tasks without continuous human input or control. For purposes of this chapter, autonomy is defined as the ability to "make (limited) decisions about what behaviors to execute based on perceptions and internal states, rather than following a pre-determined action sequence based on pre-programmed commands." *See* Matthias Scheutz & Charles Crowell, *The Burden of Embodied Autonomy: Some Reflections on the Social and Ethical Implications of Autonomous Robots*, WORKSHOP ON ROBOETHICS AT THE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION, 1 (2007).

[6] There are various definitions of "social robot" in existing literature. *See for example* the compilation in Frank Hegel, Claudia Muhl, Britta Wrede, Martina Hielscher-Fastabend, Gerhard Sagerer, *Understanding Social Robots*, THE SECOND INTERNATIONAL CONFERENCES ON ADVANCES IN COMPUTER-HUMAN INTERACTIONS (ACHI), 169 – 74 (2009).

[7] *See also* CYNTHIA BREAZEAL, DESIGNING SOCIABLE ROBOTS, MIT Press, 1 (2004).

loss coach Autom; household robots like Jibo, and research robots like the Massachusetts Institute of Technology (MIT) robots Kismet, Cog, Leonardo, AIDA, Dragonbot, and Baxter.

In general, as more robots enter our lives and our homes, we are experiencing an increase in robots designed to engage us socially. This trend is not likely to slow. Social abilities in robots will continue to improve, and robotic companions will become more common as technology advances. Makers of toys, for example, have been working for decades to increase interactivity and engage children by creating the illusion of intentional behavior in robotic playthings.[8] This type of interactivity design is not restricted to children's markets, and has multiple potential uses, as explored below. But first, let us take a closer look at how this technology creates engagement, and why it is so effective.

ROBOTS VS. TOASTERS: PROJECTING OUR EMOTIONS

At first glance, it seems hard to justify differentiating between a social robot, such as a Pleo dinosaur toy, and a household appliance, such as a toaster. Both are man-made objects that can be purchased on Amazon and used as we please. Yet there is a difference in how we perceive these two artifacts. While toasters are designed to make toast, social robots are designed to act as our companions. Looking at state-of-the-art technology, our interactions with social robots are already strikingly full of projection. People will ascribe

[8] For instance, Pleo robot dinosaurs have been specifically marketed as a "truly autonomous Life Form capable of emotions that allow personal engagement." *See* Ryan Paul, *Robotic toy makes a big impression at the DEMO convention*, ARS TECHNICA (Feb 8, 2006), http://arstechnica.com/uncategorized/2006/02/6137-2/.

intent, states of mind, and feelings to robotic objects. Psychological research argues that we form emotional attachments to robots that are surprisingly strong.[9]

So what is the difference to other objects? It is true that humans have always been susceptible to forming attachments to things that are not alive, for example to their cars or stuffed animals. People will even become attached to virtual objects. In the video game *Portal*,[10] for example, when players are required to incinerate the companion cube that has accompanied them throughout the game, some will opt to sacrifice themselves rather than the cube, forfeiting their victory.[11] One factor that may play a significant role in the development of such unidirectional relationships to objects is a psychological caregiver effect. For example, Tom Hanks develops a relationship with a volleyball in the movie *Cast Away*.[12] The interesting aspect of his attachment is nicely demonstrated when he inadvertently lets the volleyball float out to sea. Realizing that he is unable to rescue his companion, he displays deep remorse for not taking better care. The focus thereby is not on his personal loss, but rather on his neglected responsibility toward the object: he calls out to it that he is sorry.[13]

But while it is clear that such attachments can exist outside of robotics, we have reason to believe that the relationships we form with social robots are significantly

[9] *See* Sherry Turkle, *In Good Company?: On the Threshold of Robotic Companions*, in CLOSE ENGAGEMENTS WITH ARTIFICIAL COMPANIONS: KEY SOCIAL, PSYCHOLOGICAL, ETHICAL AND DESIGN ISSUES, John Benjamins Publishing Company, 24 (Yorick Wilks ed. 2010).

[10] Portal is a single-player first-person puzzle-platform video game, released in 2007 by Valve Corporation.

[11] *See* "Portal Afterthoughts" on 1UP.COM http://www.1up.com/do/feature?pager.offset=2&cId=3165930. Thanks to Jennifer Berk for pointing out this example.

[12] Tom Hanks (Producer), & Robert Zemeckis, (Director), Cast Away (Motion picture), United States: Twentieth Century Fox Films (2000).

[13] *See* James Surowiecki, *Wilson the Volleyball, Reconsidered*, SLATE (March 23, 2001), http://www.slate.com/articles/arts/culturebox/2001/03/wilson_the_volleyball_reconsidered.html. Thanks to Chip Darling for sending this article.

stronger.[14] This effect is attributed to the interplay of three factors. The first is **physicality**. Humans are physical creatures and may be hardwired to respond differently to objects in their physical space as compared to, say, virtual objects on a screen. The second factor is **perceived autonomous movement**. Movements made by objects in our physical space that we can't quite anticipate will often lend themselves to projection of intent. Take, for example, the Roomba vacuum cleaner. The Roomba is a flat, round household robot that follows a simple algorithm to clean floors. It has no social skills whatsoever, but just the fact that it moves around on its own prompts people to name it, talk to it, and feel bad for it when it gets stuck under the couch.[15]

A more extreme example is military robots. When the United States military began testing a robot that defused landmines by stepping on them, the colonel in command ended up calling off the exercise. The robot was modeled after a stick insect with six legs. Every time it stepped on a mine, it lost one of its legs and continued on the remaining ones. According to journalist Joel Garreau, "[t]he colonel just could not stand the pathos of watching the burned, scarred and crippled machine drag itself forward on its last leg. This test, he charged, was inhumane."[16] Research by Julie Carpenter[17] and Michael Kolb [18] looks at other robots employed within military teams, finding that they evoke fondness and loyalty in their human teammates, who identify with the robots enough to name

[14] *For more examples, see* Matthias Scheutz, *The Inherent Dangers of Unidirectional Emotional Bonds between Humans and Social Robots*, in ROBOT ETHICS: THE ETHICAL AND SOCIAL IMPLICATIONS OF ROBOTICS, MIT Press, 205 (Patrick Lin, Keith Abney, George Bekey eds., 2012).

[15] *See* Scheutz, *supra* note 14; *see also* Ja-Young Sung, Lan Guo, Rebecca Grinter, Henrik Christensen, *"My Roomba Is Rambo": Intimate Home Appliances*, 9TH INTERNATIONAL CONFERENCE ON UBIQUITOUS COMPUTING, 145-62 (2007).

[16] Joel Garreau, *Bots on The Ground in the Field of Battle (Or Even Above It), Robots Are a Soldier's Best Friend*, WASHINGTON POST (May 6, 2007) http://www.washingtonpost.com/wp-dyn/content/article/2007/05/05/AR2007050501009.html.

[17] CARPENTER, *supra* note 1.

[18] MICHAEL KOLB, SOLDIER AND ROBOT INTERACTION IN COMBAT ENVIRONMENTS, Dissertation, University of Oklahoma (2012).

them, award them battlefield promotions and "purple hearts", introduce them to their families, and become very upset when they "die".

While none of the above-described robots are designed to display emotional cues, their autonomous behavior already makes them appear lifelike enough to generate emotional projection. This inclination to anthropomorphize objects that act autonomously is further magnified when we introduce the third factor: **social behavior**. Cleverly designed social robots are not only made to look adorable, they are also able to mimic cues that we automatically, even subconsciously associate with certain states of mind or feelings.[19] Even in today's primitive form, devices that encompass all three factors are able to elicit emotional reactions from people that are similar, for instance, to how we react to animals and to each other. From being reluctant to switch off robots that give the appearance of animacy,[20] to ascribing mental states to AIBO dogs,[21] we respond to the cues given to us by lifelike machines, even if we know that they are not "real".

The projection of lifelike qualities begins with a general tendency to over-ascribe autonomy and intelligence to the way that things behave, even if they are just following a simple algorithm.[22] But not only are we prone to ascribing more agency than is actually present, we also project intent and sentiments (such as joy, pain, or confusion) onto other entities. When robots are able to mimic lifelike behavior, react to social gestures,

[19] *See for example* Turkle, *supra* note 9.

[20] *See* Bartneck et al., *supra* note 3 at 217-22.

[21] *See* Batya Friedman, Peter Kahn, Jennifer Hagman, *Hardware Companions? – What Online AIBO Discussion Forums Reveal about the Human-Robotic Relationship*, PROCEEDINGS OF THE SIGCHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS, 273–80 (2003).

[22] For starters, how many of us have been caught in the assumption that the shuffle function on our music players follows more elaborate and intricate rules than merely selecting songs at random? *See also* Michael Shermer, *Patternicity: Finding Meaningful Patterns in Meaningless Noise: Why the brain believes something is real when it is not*, SCIENTIFIC AMERICAN 299, 48 (2008).

and use sounds, movement, and facial expressions to signal emotions in a way that we immediately recognize, this may target our involuntary biological responses.[23] Owners of Sony AIBO dogs in the 1990s, while fully aware that it was a robot, would regularly ascribe lifelike essences and mental states to their artificial companion.[24] The robotic seal Paro, currently used as a therapeutic device in nursing homes, reacts to touches and words. It conveys a sense of animacy by exhibiting emotional states, responding to people's actions, and learning individual voices. Most of the patients (and other people) who work with Paro treat it as if it were alive.[25] The emotional effect of social robots has the potential to strongly supersede the "accidental" projection invoked by non-social robots, because it is intentional. In fact, it is often their main function.

Returning to the concept of the "caregiver effect," psychologist Sherry Turkle explains that this effect is particularly strong with social robots that are designed to evoke feelings of reciprocity. "Nurturing a machine that presents itself as dependent creates significant social attachments."[26] She finds that there is a difference between the type of projection that people have traditionally engaged in with objects, such as small children comforting their dolls, and the psychology of engagement that comes from interacting with social robots, which create an effective illusion of mutual relating.[27] While a child is aware of the projection onto an inanimate toy and can engage or not engage in it at will, a robot

[23] *See* Turkle, *supra* note 9, at 24.

[24] *See* Friedman/Kahn/Hagman, *supra* note 21.

[25] *See* Christopher Calo, Nicholas Hunt-Bull, Lundy Lewis, Ted Metzler, *Ethical Implications of Using the Paro Robot, with a Focus on Dementia Patient Care*, WORKSHOP AT THE TWENTY-FIFTH AAAI CONFERENCE ON ARTIFICIAL INTELLIGENCE, 21 (2011).

[26] *See* Sherry Turkle, *A Nascent Robotics Culture: New Complicities for Companionship*, TECH. REP., AAAI, (2006).

[27] *See id*. at 2.

that demands attention by playing off of our natural responses may cause a subconscious engagement that is less voluntary.

This anthropomorphism is especially plausible when people have little sense of how a complex robot works, and so are especially inclined to assign autonomy, intent, or feelings to actions that result from algorithms they do not understand. Small children are regularly confused when asked whether the social robots they interact with experience pain or other sentiments.[28] Elderly people unfamiliar with modern technology struggle with the difference between robotic companions and live animals.[29] But the effect of projection holds even for those who are perfectly informed as to the exact, detailed functionality of the robots with which they interact. For example, AIBO owners reported that they would remove their AIBO from the room while changing, so that they would not be "watched",[30] or that they experienced feelings of guilt when putting the device back in its box.[31] Students in MIT's Media Lab would often put up a curtain between themselves and Kismet, a social robot that simulates emotion through facial expressions, because the lifelike behavior of the face distracted them.[32] And Cynthia Breazeal,

[28] *See* podcast documenting the confusion of 7-year-olds in differentiating between the „pain" felt by a Furby (social robot technology from 1999) and „real" pain: *Furbidden Knowledge*, RADIOLAB (May 31, 2011) http://www.radiolab.org/2011/may/31/furbidden-knowledge/; *see also* recent study of children of various ages interacting with a social robot: Kahn et al., *supra* note 3.

[29] *See* Will Taggart, Sherry Turkle, Cory Kidd, *An Interactive Robot in a Nursing Home: Preliminary Remarks*, PROCEEDINGS OF COGSCI WORKSHOP ON ANDROID SCIENCE, 56-61 (2005); Anne Tergesen, Miho Inada, *Paro the Robo-Seal Aims to Comfort Elderly, but Is It Ethical?* WALL STREET JOURNAL (June 21, 2010) http://online.wsj.com/article/SB10001424052748704463504575301051844937276.html.

[30] Albeit admitting that they feel slightly silly for doing so. *See* Friedman/Kahn/Hagman, *supra* note 21, at 276.

[31] *See id*. at 277.

[32] *See* Garreau, *supra* note 16; *see also* Scheutz, *supra* note 14.

Kismet's developer, reports experiencing "a sharp sense of loss" when she parted ways with her own creation at the end of her dissertation.[33]

In summary, it appears that social robots elicit behavior in us that is significantly different from what we exhibit towards other devices, like toasters. While people have for decades named their cars and developed emotions towards other inanimate objects, the effect of robots that actively and intentionally engage our ingrained anthropomorphic responses is considerably stronger. We are already disposed towards social engagement with the robotic companions available to us today, and we can only imagine what the technological developments of the next decade will be able to effect. As we move within the spectrum between treating social robots like toasters and treating them more like our cats, there may be some ethical issues to consider.

THE ISSUES AROUND EMOTIONAL ATTACHMENTS TO ROBOTS

Some of the researchers who have studied our tendencies for human-robot projection and emotional attachment have concluded that this effect is undesirable.[34] There are multiple concerns. One is that the development and dissemination of such technology encourages a society that no longer differentiates between real and fake,[35] thereby undermining values of authenticity we want to preserve.[36]

Another concern is the use of social robots to replace human interaction. For example, if robots are used in elderly care, childcare, heath care, and educational contexts, how are

[33] *See* Turkle, *supra* note 26, at 9.

[34] *See* Turkle, *supra* note 9; Scheutz, *supra* note 14.

[35] *See* Turkle, *supra* note 26, at 3, lamenting the loss of „authenticity" as a value.

[36] This seems unconvincing. It is also possible that ubiquitous robotic technology could enhance the value of biologically "authentic" entities and relationships.

they being used? Are they assisting teachers and care personnel in their interactions with patients and students? Or are they being used where we previously provided inter-human communication and care? If social robots are used as a helpful supplement, or in places where human care is not possible, then allowing them seems desirable.[37] But if robots start replacing human companionship, this might warrant careful consideration of whether and what aspects get lost in the process.

Another concern is the danger of manipulation through social robots.[38] If people develop attachments to their robotic companions, can the companies who control the hard- and software exploit this attachment? Should companies be allowed to make use of the fact that individuals might have a massively increased willingness to pay for technology upgrades or repairs? What about sneaking advertisements for products into children's dialogs with their robotic toys? Furthermore, not only do household robots introduce new ways of data collection and retention, but people might also be more inclined to inadvertently reveal personal information to their robotic companions than they otherwise would ever willingly enter into a database.[39] Is the data from the cameras and microphones uploaded to the cloud?[40] Is private information secure? Under what

[37] For example, the Paro seal robot can be used in nursing homes for pet therapy when allowing live pets would be unhygienic or a health hazard to the patients. In this case, having the robot seal is likely better than having nothing.

[38] *See* Ian Kerr, *Bots, Babes and the Californication of Commerce*, 1 U OTTAWA L & TECH J 284-324 (2004) Ryan Calo, *People Can Be So Fake: A New Dimension to Privacy and Technology Scholarship*, 114 PENN ST L REV, 809-55 (2010); Scheutz, *supra* note 14.

[39] Ryan Calo, *Robots and Privacy*, in: ROBOT ETHICS: THE ETHICAL AND SOCIAL IMPLICATIONS OF ROBOTICS, Patrick Lin, George Bekey, and Keith Abney, eds. (Cambridge: MIT Press, 2011); Kristen Thomasen, *Liar, Liar, Pants on Fire! Examining the Constitutionality of Enhanced Robo-Interrogation*, paper presented at We Robot 2012, University of Miami (April 21, 2012).

[40] Andrew Proia, Drew Simshaw, and Kris Hauser, *Consumer Cloud Robotics and the Fair Information Practice Principles: Recognizing the Challenges and Opportunities Ahead*, 16 MINN. J. L. SCI. & TECH. 145-214 (2015).

circumstances can the government or a private actor access personal data collected through a toy or household robot?[41]

But while all of the above are important conversations to have, they do not necessarily mean that emotional attachments to robotic objects are inherently negative, or even generally undesirable. We are already seeing some highly positive use cases of engaging people with social robot technology. The Paro seal, used as a medical device, has proven effective in treating dementia patients.[42] The NAO Next Generation robot is successfully being used to help engage children with autism.[43] Preliminary studies show that even simple robotic companionship can motivate people to reach goals, for example to lose weight twice as effectively as with other methods.[44] The possibilities for health, education, and other areas are endless.

This technology is getting smarter and we are likely to see an increase in social robots designed to "manipulate" us on various levels. Not only are humans inclined to anthropomorphize the robots they interact with, they also actively enjoy doing so. Universities, toy companies, and other private actors will therefore continue to search for ways to make robots attractive to us and to engage people's emotions with technology. It will ultimately be a balancing act to find ways to deal with the concerns without overly restricting research and development, and without giving up too much of the potential this technology has to offer.

---

[41] *See* Kerr, *supra* note 38, Calo, *supra* note 38, Thomasen, *supra* note 39.

[42] Kaoru Inoue, Kazuyoshi Wada, and Reona Uehara, *How Effective Is Robot Therapy?: PARO and People with Dementia,* 37 5TH EUROPEAN CONFERENCE OF THE INTERNATIONAL FEDERATION FOR MEDICAL AND BIOLOGICAL ENGINEERING IFMBE, 784-787 (2012).

[43] *See* Syamimi Shamsuddina, Hanafiah Yussof, Luthffi Ismail, Salina Mohamed, Fazah Hanapiah, Nur Zahari, *Initial Response in HRI- a Case Study on Evaluation of Child with Autism Spectrum Disorders Interacting with a Humanoid Robot NAO,* 41 IRIS 1448–55 (2012).

[44] CORY KIDD, DESIGNING FOR LONG-TERM HUMAN-ROBOT INTERACTION AND APPLICATION TO WEIGHT LOSS, Dissertation, Massachusetts Institute of Technology (2008).

Given our inclination to become emotionally involved with robotic companions, there is another question to explore. Even with primitive, state-of-the-art technology, some people have visceral feelings about the treatment of robots. For example, I conducted a workshop with my colleague Hannes Gassert at the LIFT13 conference in Geneva, Switzerland. In the workshop, groups of participants were given Pleos—cute robotic dinosaurs that are roughly the size of small cats.[45] After interacting with the robots and performing various tasks with them, the groups were asked to tie up, strike, and "kill" their Pleos. Drama ensued, with many of the participants refusing to "hurt" the robots, and even physically protecting them from being struck by fellow group members. One of the participants removed her Pleo's battery, later sheepishly admitting that she had instinctively wanted to "spare it the pain." Although the groups knew we had purchased the robots to be destroyed, we could only persuade them to sacrifice one Pleo in the end. While everyone in the room was fully aware that the robot was just simulating its pain, most participants giggled nervously and felt a distinct sense of discomfort when it whimpered while it was being broken.

Similarly, not long after the Pleo became commercially available in 2007, videos of Pleo "torture" began to circulate online. The comments left by viewers are strikingly polarized – while some derive amusement from the videos, others appear considerably upset, going so far as to verbally attack the originators and accuse them of cruelty.[46] An older review of Sony AIBO online message boards reveals that people were dismayed to

[45] For more information on the Pleo, *see* Mattias Jacobsson, *Play, belief and stories about robots: A case study of a pleo blogging community*, THE 18TH IEEE INTERNATIONAL SYMPOSIUM ON ROBOT AND HUMAN INTERACTIVE COMMUNICATION, 232-37, 233 (2009).

[46] *See, for example* "Pleo vs. ComBot", http://youtu.be/95MAyWE0ERk.

witness the story of an AIBO being tossed into a garbage can.[47] A recent neurological study shows that people in fact seem wired to feel more empathy towards robots than towards other objects.[48]

What we see from these examples and from our currently ongoing studies is that violent behavior towards robotic objects *feels* wrong to many of us, even if we know that the "abused" object does not experience anything. We may be hardwired to respond instinctively to cues like simulated pain or need. If this discomfort becomes stronger with increasingly sophisticated technology, how should we deal with it? Should we encourage or discourage it? Might we even reach a point where we want rules in place governing the treatment of our robot companions? It seems far-fetched at first, but there are a few reasons why regulation of this sort might make sense.

One reason why people might want to prevent the "abuse" of robotic companions is the protection of societal values. Parents of small children with a robotic pet in their household are likely to energetically intervene when their toddler kicks or otherwise physically abuses the robot. Of course, one reason for doing so is to protect the (usually expensive) object from breaking, but another is to discourage the child from engaging in types of conduct that could be harmful in other contexts. Given the lifelike behavior of the robot, a child could easily equate kicking it with kicking a living thing, such as a dog or another child. As it becomes increasingly difficult for small children to fully grasp the difference between live pets and lifelike robots, we may want to teach them to act equally

[47] With comments along the lines of "[T]hat poor puppy," and "That is so sick to me!", *see* Friedman/Kahn/Hagman, *supra* note 21, at 277.

[48] Astrid Marieke Rosenthal-von der Pütten, Frank Schulte, Sabrina Eimler, Laura Hoffmann, Sabrina Sobieraj, Stefan Maderwald, Nicole Krämer, and Matthias Brand, *Neural Correlates of Empathy Towards Robots*, HRI '13 PROCEEDINGS OF THE 8TH ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN-ROBOT INTERACTION, 215-216 (2013). We are currently in the process of conducting further controlled studies exploring empathy towards robots in situations of physical violence at MIT.

considerately towards both. While this is easily done when a parent has control over both the robot and the child, protecting social robots more generally would ensure that societal standards are set and prevent children from witnessing or adopting undesirable behavior elsewhere. It could protect children from potentially traumatizing experiences, for instance seeing older children or adults "torture" a robotic toy, the likes of which the child has developed an emotional relationship to at home.[49] Finally, if children face no societal repercussions for behaving violently towards a robot with lifelike reactions, they might become generally desensitized to this type of conduct.

Even for fully informed adults, the difference between alive and lifelike may be muddled enough in our subconscious to warrant adopting the same attitudes toward robotic companions that we carry towards our pets. Given that many people already feel strongly about state-of-the-art robot "abuse", it may gradually become more widely perceived as out of line with our social values to treat robotic companions in a way that we would not treat our pets.

A related aspect that our society has strong sentiments about is the question of sexual behavior. In the near future, we may have to consider whether or not to permit sexual practices between humans and social robots that we currently do not permit with live humans or animals. Bestiality, rape, and sexual acts with under-age children are condemned in our culture and heavily governed by our legal system. It is thinkable that the desire to protect our current social values could cause people to demand laws that prohibit the sexual "abuse" of social robots.

[49] If children have trouble distinguishing between robots and live pets, they will have a similar reaction in this scenario as to the same involving a kitten or a puppy.

Most importantly, if more research shows that the line between lifelike and alive is muddled in our subconscious when interacting with something physically, then certain behavior towards robots could either traumatize or desensitize us. In particular, there is concern that mistreating an object that reacts in a lifelike way could impact the general feeling of empathy we experience when interacting with other entities. We already see some correlation between non-empathetic behaviors. For example, abuse reporting laws in many States in the U.S., recognize that animal abuse and child abuse are frequently linked,.[50] On a more positive note, studies have shown links between pet ownership as a child and general empathetic tendencies towards others.[51] Currently, our research is exploring whether human-robot-interaction fits this model.[52] Further research may dovetail with existing literature on the societal effects of violence in video games. But compared to virtual characters on a screen, there may be a significant difference in translation of behavior when it comes to physically embodied objects and physical actions.[53]

---

[50] *See, for example*, Code of the District of Columbia, § 4-1321.02 "Persons required to make reports; procedure"; 510 Illinois Compiled Statutes 70/18, Sec. 18, "Cross-reporting"; Louisiana State Legislature, RS 14 §403.6 "Reporting of neglect or abuse of animals". Other States with cross-reporting of animal and child abuse include California, Colorado, Connecticut, Maine, Massachusetts, Nebraska, Ohio, Tennessee, Virginia, and West Virginia, *see* American Veterinary Medical Association, *Cross-Reporting of Animal and Child Abuse*, State Summary Report, available at https://www.avma.org/Advocacy/StateAndLocal/Pages/sr-animal-abuse-cross-reporting.aspx (last updated June 2014).

[51] Beth Daly & L.L. Morton, *An Investigation of Human–Animal Interactions and Empathy as Related to Pet Preference, Ownership, Attachment, and Attitudes in Children*, 19 ANTHROZOOS 113-27 (2006). Conversely, urban legends of SS training purport that soldiers were given a puppy to raise, and later ordered to kill it to harden them (thanks to Nathana O'Brian for this example). Stories such as this indicate that our society views empathy as a dynamic tendency that can be influenced through our behavior.

[52] Kate Darling, Palash Nandy & Cynthia Breazeal, *Empathic Concern and the Effect of Stories in Human-Robot Interaction*, PROCEEDINGS OF THE IEEE INTERNATIONAL WORKSHOP ON ROBOT AND HUMAN COMMUNICATION (ROMAN), (2015).

[53] Stela H. Seo, Denise Geiskkovitch, Masayuki Nakane, Corey King & James E. Young, *Poor Thing! Would You Feel Sorry for a Simulated Robot? A comparison of empathy toward a physical and a*

If research establishes a connection between empathetic or violent behavior and our interactions with robots, this raises the question of regulation. In fact, even without scientific evidence, society might at some point begin to push for rules governing violent or sexual behavior towards certain types of robots. The following section offers some thoughts on how the legal system could accommodate this.

EXTENDING LEGAL PROTECTION TO ROBOTIC OBJECTS

This section proposes that abuse protection for social robots could follow the analogy of our animal abuse protection laws. Despite the fact that the exact underpinnings of animal abuse protection are contested and many do not match the reasons we might protect robots, there are both psychological and philosophical parallels.

Philosophical arguments for animal rights range from moral obligations to prevent pain and suffering in sentient beings,[54] to an abstract recognition of certain animals' inherent dignity.[55] Many such positions use factors like cognitive abilities or sentience to differentiate between the moral treatments of various types of life forms. Societal discussion surrounding the prevention of animal abuse centers on the fact that animals

simulated robot, PROCEEDINGS OF THE TENTH ANNUAL ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN-ROBOT INTERACTION, 125-132 (2015).

[54] *See, for example*, PETER SINGER, ANIMAL LIBERATION: A NEW ETHICS FOR OUR TREATMENT OF ANIMALS, New York Review (1975); JEREMY BENTHAM, AN INTRODUCTION TO THE PRINCIPLES OF MORALS AND LEGISLATION, Chapter xvii (1789) "[T]he question is not, Can they reason?, nor Can they talk? but, Can they suffer?"

[55] *See, for example*, TOM REGAN, THE CASE FOR ANIMAL RIGHTS, University of California Press (1983).

experience pain. Because many people in our society agree that hurting certain animals is wrong, laws have been put in place to prevent their mistreatment.[56]

But while animal rights philosophy and discussion revolve around concepts like pain, our laws actually indicate that these concerns are secondary when it comes to legal protection. Many successful societal pushes for animal abuse laws have followed popular sentiment rather than consistent biological criteria. Our apparent desire to protect those animals to which we more easily relate indicates that we may care more about our own emotional state than any objective biological criteria. For example, laws banning the slaughter of horses for meat in the United States have been enacted because of the general sentiment that such treatment of horses is offensive.[57] In contrast to many Europeans, a large part of the United States population seems strongly opposed to the idea of horses being killed and eaten.[58] This is not justified by any biological differences between horses and cows. Similarly, very few people were interested in early campaigns to save the whales until the latter half of the twentieth century when the beautiful sounds of whale songs were first recorded. Support for the cause rose dramatically, as people discovered whales to be creatures they could relate to.[59] All of this indicates that we may care more about our own feelings than those of other entities.

Our desire to protect animals from abuse may be based on our relationship to the animals, as well as on a projection of ourselves. A lot of people do not like to see kittens

[56] *For example, see* The European Union Council Directive 1999/74/EC; United States Animal Welfare Act, 7 U.S.C. § 2131 et seq. The United States has individual State legislation governing animal cruelty, *see for example* Massachusetts General Laws MA ST 272 § 77 - 95; MA ST 272 § 34; MA ST 22C § 57.

[57] *See* Brian Palmer, *The Delicious Mr. Ed - Why Don't Americans Eat Horse Meat?* SLATE (Oct 24, 2011) http://www.slate.com/articles/health_and_science/explainer/2011/10/slaughtering_horses_for_meat_is_banned_in_the_u_s_why_.html.

[58] In contrast, horsemeat is a popular delicacy in many European countries. Dogs and cats are also considered acceptable to eat in other parts of the world.

[59] Thanks to Ryan Calo for this example.

be held by the tail. It is certainly possible that we feel so strongly about this because of the specific details of kittens' inherent biological pain. But it is also possible that it simply causes us discomfort to see a reaction that we associate with suffering. Our emotional relationship to kittens, plus the strong response of the kitten to being held by the tail, may trigger protective feelings in us that have more to do with anthropomorphism than moral obligation.

However, even if there is general agreement that projecting our emotions onto other things is part of why we protect animals from abusive behavior, many will argue that we should draw the line at something that does not actually "suffer". After all, despite the behavior we display towards them, most of us know that robots are not alive. And while we find differential treatment of animals in the law, the actual discussions surrounding their moral inclusion do not usually consider anthropomorphism to be a justification.[60] Rather, they invoke the experience of pain, or concepts of sentience, consciousness, and the general question of what exactly sets humans apart in a way that is relevant to a morally distinct treatment. Even amongst those who see no difference to humans that is sufficiently relevant to deny animals certain rights, there are likely to be a number of people who draw a moral line at biological criteria.[61] For instance, people in favor of protecting creatures that experience biological pain from abuse may have no moral reason to extend it to anything beyond that. But those unconvinced by popular sentiment towards robots might nevertheless agree with the more practical concept of protecting

[60] Even if it were, one might still oppose the idea that law be based on social sentiment rather than morally consistent criteria.

[61] While it is true that someday there may be less of a biological difference between humans and robots, opening rich avenues for debate, this argument does not seem timely for purposes of this discussion.

robots to protect ourselves. As discussed above, we might want to demand protection for social robots in order to discourage behavior that would be harmful in other contexts.

The Kantian philosophical argument for preventing cruelty to animals is that our actions towards non-humans reflect our morality — if we treat animals in inhumane ways, we become inhumane persons.[62] This logically extends to the treatment of robotic companions. Granting them protection may reinforce behavior in ourselves that we generally regard as morally correct, or at least behavior that makes our cohabitation more agreeable. It may also prevent desensitization towards actual living creatures and protect the empathy we have for each other. It is true that this reason may not map to most people's current understanding of animal abuse laws, but the analogy to Kant's reasoning makes sense.

Whether for psychological reasons of projection onto other entities, moral philosophical concepts, or practical concerns, the regulation of violent behavior towards robotic objects could be structured similarly to the animal abuse laws we already have in place. In terms of practicality, one challenge lies in establishing limiting factors.

In order to pass protective laws, we would ideally come up with a good definition of "social robot". In other words, we would want to delineate normal objects or robots from those that the majority of people feel sufficiently anthropomorphic towards to warrant

[62] „If a man shoots his dog because the animal is no longer capable of service, he does not fail in his duty to the dog, for the dog cannot judge, but his act is inhuman and damages in himself that humanity which it is his duty to show towards mankind. If he is not to stifle his human feelings, he must practice kindness towards animals, for he who is cruel to animals becomes hard also in his dealings with men." *See* IMMANUEL KANT, LECTURES ON ETHICS, Cambridge University Press, 240 (J.B. Scheewind, ed., 1997).

special treatment.[63] Depending on relevant factors shown through future research, this definition would be something along the lines of (1) an embodied object with (2) a defined degree of autonomous behavior that is (3) specifically designed to interact with humans on a social level and respond to mistreatment in a lifelike way. This definition will not cover all of the robotic objects that people project onto (for instance the robots that evoke emotional engagement by accident, like the above-mentioned military robots). And it may prove to be overly broad and require tailoring to specific types of programming[64] or specific social interactions.[65]

We would also have to clearly determine the extent of protection, including what constitutes "mistreatment". Many issues could be resolved analogous to animal abuse laws, but there may be some edge cases, especially in light of changing technology. For example, people's hesitation to switch off a robot[66] may not be a good analogy to terminating an animal's life, or a good enough reason to regulate behavior. The circumstances under which such behavior would cross a line might take into consideration the capabilities of the technology (such as irreversible data loss), or programmed reactions of the robot that would make switching it off emotionally traumatizing to humans. But it probably wouldn't need to fall under a restriction at all. Animals themselves are not protected from being put down, but rather only when ending their lives is deemed cruel and unnecessary given the method or circumstances. Similarly, it would make little sense to give robots a "right to life."

[63] Another option could be to have a law that is triggered by a certain type of behavior or specific level of emotion, rather than the nature of the entity harmed. But to avoid overly broad legislation, it seems helpful to have a guiding definition of social robot for practical purposes of setting a first threshhold.

[64] For example, some social robots are specifically programmed to respond with lifelike pain cues when treated violently, like the Pleo dinosaur.

[65] For example, types of sexual behavior.

[66] *See* Bartneck et al, *supra* note 3.

The delineations of mistreatment will require consideration, and they may need to reflect both human sentiment and technological capability in specific contexts. They do not, however, constitute preventive hurdles. While it may take some effort to find the right balance, the challenge of drawing such lines is not new to the law. Another thing that is not new to the law is the impact of this legislation on property rights, as explored below.

A NOTE ON PROPERTY RIGHTS

When people care about protecting something, there are different ways that the law can address this. One way is by maintaining protection through property rights that are inherent to an owner. For example, the owner of a dog can take legal action against somebody who infringes on the owner's property by abusing the dog. But it is not uncommon for society to enact laws that go beyond personal property rights. Although individual dog owners may be able to protect their dogs from being abused by others, we may want to ensure the abuse protection of all dogs, whether we own them or not. Granting protection to certain animals from being mistreated, and thereby governing the actions of their owners, goes beyond property rights. It goes so far as to restrict other people's property rights, for example by preventing farmers from treating their chickens poorly. As such, creating a legal rule to extend such protection to certain robots warrants an analysis of potential costs.[67] But despite historical discussion over the sanctity of

[67] Restricting property rights can have economic effects, such as reducing the value of the property. In the case of robots, this could influence research & development, or drive it in directions that circumvent the legal rule.

property rights,[68] protective laws that go beyond property rights are fairly common in this day and age. Examples of wider-reaching laws include environmental protection, endangered species preservation, minimum wages, or punishment for child abuse. These laws tend to be enacted when society cares deeply enough about an issue or its effects. The next section explores when this could and should apply to robots.

WHEN SHOULD WE START PROTECTING ROBOTS?

Whether out of sentiment or to promote socially desirable behavior, some parts of society may sooner or later begin to ask that legal protections be extended to robotic companions. If this happens, politicians and lawmakers will need to deliberate whether it would make sense to accommodate this societal preference.

One view supports granting legal protection to social robots as soon as there is sufficient societal demand. Assuming that our society wants to protect animals regardless of their capacities, because of our personal attachments to them, society may someday also want to protect social robots regardless of their capacities. Humans' moral consideration of robots may simply depend more on our own feelings than on any societal effects or inherent qualities built into robots. Catering to this preference views the purpose of law as a social contract. We construct behavioral rules that most of us agree on, and we hold everyone to the agreement. In theory, the interest of the majority prevails in democratic societies, and the law is tailored to reflect social norms and preferences. If this is the purpose of the legal system, then societal desire for robot

[68] *See, for example,* Gerald Friedman, *The Sanctity of Property in American History*, PERI WORKING PAPER NO. 14 (2001).

protection should be taken into account and translated to law as soon as the majority calls for it. Whether or not one believes that the majority always makes the best decisions for society, and even if one believes in a natural rights theory of higher truth, there are pragmatic reasons to accommodate societal preferences. Legislatively ignoring that people feel strongly about an issue can lead to discontent and even a lack of compliance with the law.

Ideally, however, social sentiment and political push would be based on an understanding of the actual effects of anthropomorphism and driven by well-founded concern for societal welfare. Law can also be used to govern behavior for the greater good of society and have a positive influence on people's preferences, rather than the other way around. In this case, the question of when we should extend legal protections to social robots depends on whether we find evidence that our behavior towards robots translates to other contexts. As mentioned above, if lifelike and alive is subconsciously muddled, then treating certain robots in a violent way could desensitize actors towards treating living things similarly. If research conclusively establishes such effects, then this chapter advocates consideration of extending animal abuse laws to certain types of robots.

While it seems likely that people will increasingly develop strong attachments to robotic companions, the question of whether we should legally protect them is by no means simple. Further research will be important in helping to confirm or debunk our hunches. But in any case, as technology widens the gap between social robots and toasters, it seems timely to begin thinking about the societal implications of anthropomorphism.

***