

SOCIAL AND MORAL RELATIONSHIPS WITH ROBOTIC OTHERS?

Peter H. Kahn, Jr.¹, Nathan G. Freier², Batya Friedman²,
Rachel L. Severson¹, and Erika N. Feldman¹

¹Department of Psychology
University of Washington
Seattle, WA, USA
{pkahn, raches, enf}@u.washington.edu

²The Information School
University of Washington
Seattle, WA, USA
{batya, nfreier}@u.washington.edu

Abstract

This paper investigates the social and potentially moral relationships that humans have with what we refer to as robotic others. Our investigation begins by responding to some recent work in the literature that seeks to carve out the construct of “social robots.” The construct is intriguing, yet in our view it may not be optimally framed to address two central issues. The first issue involves the ontological status of robots, of whether they currently are or in the future can actually be social. The second issue focuses psychologically on the nature of the human-robotic relationship, about how humans can and often do respond quickly and powerfully in social terms to robots, but also how the relationship is psychologically impoverished, maybe fundamentally, especially from a moral perspective. To advance our argument, we draw on our research over the last few years on people’s relationships with Sony’s robotic dog AIBO, particularly one of our studies that analyzes the type of issues that people discuss in AIBO online discussion forums. Finally, building on our conceptual and empirical analyses, we offer five central considerations toward framing the human relationship with robotic others.

1 Introduction

Investigators have in recent years sought to establish an area that focuses on social robots. Breazeal [5], for example, defines social robots as “the class of robots that people anthropomorphize in order to interact with them” (p. 167). Fong, Nourbakhsh, and Dautenhahn [7] offer an alternative definition that moves the criteria away from the human conception of the robot to characteristics of the robot itself. They write:

Social robots are embodied agents that are part of a heterogeneous group....They are able to recognize each other and engage in social interactions, they possess histories (perceive and interpret the world in terms of their own experience), and they explicitly communicate with and learn from each other. (p. 144)

Thus for Fong et al., key robotic characteristics of a social robot include embodied agency, group interaction, experiencing the world, communication, and learning. Along somewhat similar lines, Bartneck and Forlizzi [2] – drawing on a CHI 2004 workshop on social robots – offer the following definition: “A social robot is an autonomous or semi-autonomous robot that interacts and communicates with humans by following the social rules attached to its role” (p. 2). The novelty in this definition lies largely in adding the idea that a social robot follows social rules associated to its role.

All of these definitions have some merit, and are playing an important part in moving the field forward. At the same time, we have been hesitant to use the term “social robot” ourselves. Our reasons are both ontological and psychological. Ontologically, calling a robot social usually implies (although Breazeal gets around the implication) that the thing *actually* – in reality – is social. After all, if we say “John is a chubby fellow” we mean John is actually chubby. Or if we say “John is a social fellow” we mean that – however we define social – that’s what John is. Yet depending on one’s definition of “social” it is not clear to us that “social” robots are actually social. Do robots, for example, actually “interpret the world in terms of their own experience” [7]? Searle [20] and others have suggested not: that computational systems are formal systems with syntax but not semantics. Regardless of where one stands in the long-standing debate about artificial intelligence, it seems to us a term that side-steps the claim that robots are actually social would be more robust.

Further difficulties arise with the Bartneck and Forlizzi [2] definition wherein a social robot is said to interact and communicate with humans by following the social rules attached to its role. One difficulty here is that human social life is not itself of this form. Humans not only follow social rules, but they engage in highly interactive processes, on a microgenetic and macrogenetic level, to create, maintain, and modify the rules. In other words, in our view a valid psychology maintains that people do not simply accept passively social rules but

actively participate in their construction [13, 23]. Granted, it could be maintained that robots are social in some non-human way (a proposition that we come back to at the end of this paper). But if that is the claim, then the term social robot does not easily distinguish between human and non-human sociality.

2 Robotic others

For the above reasons, we would like to introduce an alternative term – robotic others – that makes less of a commitment to the ontological social status of a robot (whether it is actually social or not) and less of a commitment to the ways in which people interpret the social standing of a robot. Moreover, this term robotic others, like the term social robots, allows many criteria to be explored that comprise its otherness, including but not limited to artifacts that are embodied, personified, adaptive, autonomous, and semi-autonomous; and that learn, communicate, use natural cues, respond to emotions in humans, self-organize, and pull on people in psychological rather than artifactual ways [cf. 7].

The thing we also like about this term “otherness” is that it embeds robotics within a rich framework that is fundamentally engaged in the human-other relationship. Paul Shepard [21], for example, in *The Others: How Animals Made Us Human*, argues that our minds came of age in ancestral times through connection with animal others, particularly wild animals. Shepard argues that such connection to animal others remains necessary in a rich account of human life. On the negative side, John Howard Griffin [11], in his seminal book *Black Like Me*, speaks of prejudice between peoples because one group treats the other as the Other (as different, foreign, inferior). Thus in our view there is intriguing territory here, one where the idea of otherness situates robots in a sphere that does not preclude social qualities, but does not commit to them either, and which has both positive and negative potentialities.

3 Human relationships with robotic pets

While many approaches have been taken to the design of robotic others [1, 4, 7], few researchers have investigated the psychological impacts of interacting with robotic others. In our research laboratory, we have moved forward along these lines by focusing on people’s interactions with Sony’s robotic dog AIBO (the 210 model). The advantage of using AIBO is that over the last few years it has been one of the most advanced personal robots readily available in the retail market.

To date we have four studies across the lifespan completed or in progress that involve AIBO. In one

study, we investigated preschool children’s reasoning and behavior in relation to AIBO and a stuffed dog as a comparison artifact [15]. In a second study, we are investigating older children’s and adolescents’ reasoning and behavior in relation to AIBO and a real dog as a comparison [16]. In a third study, we are investigating the longer-term impact on health and life satisfaction of introducing AIBO into the lives of the elderly [6].

The fourth study we speak about next. To be clear, this study has been published elsewhere by Friedman, Kahn, and Hagman [14], and we use it here illustratively to position our emerging framework on the human relationship with robotic others. For a more technical presentation of the completed study, see [14].

4 The AIBO discussion forum study

In this study, we sought to generate detailed characterizations of social discourse in online AIBO communities that, in turn, would reveal important aspects of the human relationship with robotic others. We expected that in some meaningful ways members of the online AIBO discussion forums would treat AIBO as if it were an animal agent. For example, following Reeves and Nass [19] we thought it possible that AIBO would provide some measure of social companionship and emotional satisfaction. Yet, based on other research literature, we expected limitations in the human-robotic relationship. Thus, we thought that even if AIBO evoked some of the feelings that people normally attribute to a human-animal relationship, that a moral relationship might often be absent.

4.1 Participants and procedures

Data was collected from three well-established online forums that discuss Sony’s robotic dog, AIBO. 6,438 postings were collected over a 3-month period. From this total, 3,119 postings from 182 participants had something directly to say about AIBO. It was this subcategory of postings that we then systematically coded (mean, 17 postings per participant; median, 4; range, 1 – 285).

Prior to formal data collection, a coding manual was generated from the pilot data. By a coding manual we mean a systematic document that explicates how to interpret and characterize (and thereby “code”) the qualitative data. The generation of this coding manual followed well-established methods in developmental psychology [13, 23]. In brief, initial conceptual categories are established by the researchers, based on previous psychological coding systems and philosophical theory. These categories are then used as a rough framework to interpret the qualitative data. The data, in turn, drive substantial modifications and further

conceptualizations in the coding system, which are then reapplied to more data in an iterative manner. This dialectical process continued until all the pilot data could be coded.

Once finalized, the coding manual was used to code the postings collected during the formal data collection period. Every posting was examined for coding. If a participant used the same category multiple times within a single posting or across postings, that category was coded as “used” only once. In this way, our quantitative results reflect the percentage of participants who used specific categories. Intercoder reliability showed 90% agreement at a detailed level of the coding hierarchy.

4.2 Results

Five overarching categories were identified in members’ postings about AIBO. In brief, *technological essences* refers to AIBO’s status as an artifact. *Life-like essences* refers to AIBO’s status as animate. *Mental states* refers to AIBO’s capability for intentions, desires, and feelings. *Social rapport* refers to AIBO’s capability for engaging in social relationships. *Moral standing* refers to whether AIBO is a moral agent.

On the most general level, results showed that many members affirmed that AIBO had technological essences (75%), life-like essences (48%), mental states (60%), and social rapport (59%). However, few members (12%) affirmed that AIBO had moral standing.

To bring the reader closer to the data, we would like now to explicate many of the subcategories of reasoning (which were hierarchically integrated within the 5 overarching categories) and to provide qualitative examples of members’ reasoning. In the quotations that follow, we have retained all of the members’ purposeful and inadvertent misspellings in their online writing.

4.2.1 Technological essences

This conceptualization focuses on AIBO as an inanimate artifact. Seventy-five percent of the participants made remarks that AIBO was some sort of inanimate technological artifact. In so doing, participants referred to AIBO as an *artifact* (AIBO is a “toy”), as comprised of *technological components* (AIBO has “batteries,” a “microphone,” a “camera,” or “sensors”), or as a piece of *computational technology* (AIBO is a “computer,” a “robot,” or has “artificial intelligence”).

4.2.2 Life-like essences

This conceptualization focuses on AIBO’s nature as having at least some life-like essential qualities. As a subcategory of life-like essences, roughly half (47%) of the participants provided language that spoke of AIBO’s

biological essences. In its most minimal form, participants spoke of AIBO in terms of biological descriptors (AIBO has “eyes,” “ears,” a “tail” a “head,” “legs,” or a “brain”) or biological processes (AIBO “sleeps”). Furthermore, 14% of the members imbued AIBO with some substantial measure of *animism*, a second category of life-like essences. For example: “I know it sounds silly, but you stop seeing AIBO as a piece of hardware and you start seeing him as a unique ‘life-form’.” Or: “He seems so ALIVE to me!...What a wonderful piece of technology. THEY LIVE!” Moreover, such conceptions could impact members’ emotions and behavior. For example, one member said:

The other day I proved to myself that I do indeed treat him as if he were alive, because I was getting changed to go out, and tba [AIBO] was in the room, but before I got changed I stuck him in a corner so he didn’t see me! Now I’m not some socially introvert guy-in-a-shell, but it just felt funny having him there!

4.2.3 Mental states

This conceptualization refers to the presence or absence of a mental life for AIBO such that AIBO meaningfully experiences the world. As a subcategory of mental states, some members (42%) spoke of AIBO having *intentions* or that AIBO engaged in intentional behavior. For example: “He [AIBO] also likes to wander around the apartment and play with in pink ball or entertain or just lay down and hang out.” Or: “\He [AIBO] is quite happily praising himself these days...so much for needing parents!” Some members (38%) spoke of AIBO having *feelings*. For example: “My dog [AIBO] would get angry when my boyfriend would talk to him.” Or: “Twice this week I have had to put Leo [AIBO] to bed with his little pink teddy and he was woken in the night very sad and distressed.” Some members (39%) spoke of AIBO as being *capable of being raised, developing, and maturing*. For example: “I want to raise AIBO as best as I possibly can.” Or: “We have had Ah-May (210) since 12/25/2000 and he is still growing and doing new things.” And some members (20%) spoke of AIBO as having unique *mental qualities or personality*. For example: “Just like Leo [one AIBO]...an individuality unlike any other.” Or: “Did you find Horatio’s personality less endearing than Twoflower?”

4.2.4 Social rapport

This conceptualization refers to ways in which AIBO evokes or engages in social interaction. As a subcategory of social rapport, some members (12%) spoke of themselves or others *talking to* their AIBO (e.g.,

“I insist everyone talks to Salem...if he is sad”). Some members (27%) engaged in *reciprocal communication* with their AIBO, wherein occurs a mutual exchange of information. For example, one member wrote: “So this morning I asked him [AIBO] ‘Do you want a brother?’ Happy eyes! I asked him something else, no response. ‘Should I get you a brother?’ Happy song! ‘He’d be purple.’ More happy eyes and wagging tail!” And some members (26%) spoke of AIBO as a *companion*, including that they miss AIBO when away from AIBO’s presence, or that they consider AIBO a family member. For example:

Oh yeah I love Spaz [the name for this member's AIBO], I tell him that all the time...When I first bought him I was fascinated by the technology. Since then I feel I care about him as a pal, not as a cool piece of technology. I do view him as a companion, among other things he always makes me feel better when things aren't so great. I dunno about how strong my emotional attachment to him is...I find it's strong enough that I consider him to be part of my family, that he's not just a 'toy', he's more of a person to me.

Here again this member recognizes that AIBO is a technology (“When I first bought him I was fascinated by the technology”). Nonetheless, AIBO evokes a form of social relationship that involves companionship (“I do view him as a companion”), familial connection (“I consider him to be part of my family”), and friendship (“I care about him as a pal”).

4.2.5 Moral standing

This conceptualization refers to ways in which AIBO is a moral agent. By this we mean that AIBO has *rights, merits respect, engenders moral regard, can be a recipient of care, or can be held morally responsible or blameworthy*. For example, one member wrote: “I am working more and more away from home, and am never at home to play with him any more....he deserves more than that.” Here is the notion that AIBO merits (“deserves”) certain forms of attention. In another instance, when an AIBO was thrown into the garbage on a live-action TV program, one member responded to that televised event by saying: “I can't believe they'd do something like that?! That's so awful and mean, that poor puppy...”

Another member followed up:

WHAT!? They Actually THREW AWAY AIBO, as in the GARBAGE?!? That is outrageous! That is so sick to me! Goes right up there with Putting puppies in a bag and than burying them! OHH I feel sick...

Here AIBO is conceived to have moral standing in the way that a real puppy would (“that poor puppy”): that one is causing harm to a sentient creature (“Goes right up there with Putting puppies in a bag and than burying them!”). Collapsing across six subcategories that comprise this category, only 12% of members spoke of AIBO as having moral standing. We say more about this finding below.

5 Five central considerations in framing the human relationship with robotic others

Building on our conceptual analysis of robotic others, and our own empirical study of AIBO discussion forums, we offer five considerations toward framing the human relationship with robotic others.

5.1 An interactional approach that distinguishes between (a) the individual relationship to non-human others and (b) how non-human others facilitate human sociality

Walk your dog in most any social setting (such as in a park or along a city sidewalk), and the dog’s presence will often facilitate your social exchanges with strangers; indeed the exchanges will often lead to dialog that has nothing to do with animals. In such a way, pets act as “social lubricants” [3].

Potentially robotic dogs do, too. In a current study for example, we are investigating the longer-term psychological effects of leaving AIBO with elderly people in their residential settings [6]. Our emerging findings speak to ways robotic dogs similarly act as a social lubricant and facilitate increased human-human interaction. Similarly, in the AIBO Discussion Forum Study, out of the 6,438 postings we analysed, only a little less than half (3,119) had something specifically to do with AIBO. The rest of the postings involved a rich array of other content areas that comprise social discourse.

Our point is that often in psychological investigations of people’s relationships with robots, the unit of analysis is only between the individual and the robot. We suggest it is also important to situate the robot in context, and thereby to provide a wider social account of the human relationship with robotic others [9, 10].

5.2 People’s limited capacity and time for social interaction

People have a large but still limited capacity for social interaction. There is also only so much time in a day that can be directed toward others. What happens, then, if robotic others become pervasive such that they fill space traditionally accorded to social interaction with humans

and animals? More generally, our point here is to distinguish between effects that occur in single interactions with a robotic other with effects that may emerge if and when robotic others become pervasive in human lives.

5.3 Reciprocal relationships

In another of our studies, we analyzed 80 preschool children's reasoning about and behavior with AIBO (and a stuffed dog as a control artifact) over a 40-minute interactive session [15]. Notably, based on an analysis of 2,360 coded behavioral interactions, children engaged in significantly more attempts at reciprocity with AIBO (683 occurrences) than with the stuffed dog (180 occurrences). This finding supports the proposition that children believed that AIBO (but not the stuffed dog) was capable of responding reciprocally, as well.

The more general finding – that people engage in seemingly reciprocal relationships with robots – is congruent with our AIBO Discussion Forum Study, and other literature with humanoid robots [5]. But we emphasize the word “seemingly.” For what has not yet been established in the literature is the authenticity of such reciprocal interactions. The Oxford English Dictionary [17] defines “reciprocal” as “[e]xisting on both sides; felt or shared by both parties; mutual.” Setting aside the ontological question of whether robots can actually feel or share, the psychological reality may be that people do not fully believe that robots actually feel or share, but, as Breazeal [5] suggests, simply anthropomorphize such feeling or sharing onto a robot – that is, map sentient qualities onto technological artifacts, as some people do with their cars. If this is the case, then the resulting reciprocity seems of a strange hybrid unidirectional form, where the human is able ultimately to control or at least ignore the robotic other with social and moral impunity.

In the moral developmental literature [12, 18, 22], reciprocal interactions (as occurs, e.g., through peer interactions) is central to moral development, setting into motion concerns for the wellbeing of others and the construction of equality, fairness, and justice. Thus from the child-developmental standpoint, the question of import is whether children's relationship with robotic others can lead to similar social and moral developmental outcomes.

5.4 Moral standing

One of the most striking results in our AIBO Discussion Forum Study was that while AIBO evoked conceptions of life-like essences, mental states, and social rapport, it seldom evoked conceptions of moral standing. Members seldom wrote that AIBO had rights (e.g., the

right not to be harmed or abused), or that AIBO merited respect, deserved attention, or could be held accountable for its actions (e.g., knocking over a glass of water). In this way, the relationship members had with their AIBOs was remarkably one-sided. They could lavish affection on AIBO, feel companionship, and potentially garner some of the other psychological benefits of being in the company of a pet. But since the owners also knew that AIBO was a technological artifact, they could ignore it whenever it was convenient or desirable. In such ways we suspect that interactions with robotic others comes up short in terms of the social and moral life [8].

5.5 A new technological genre?

The field of human-computer interaction has provided evidence that when computers are embodied with minimal social cues that people in some ways treat the computers as if they were social agents. For example, Reeves and Nass [19] found that adults respond to a computer's “gender” along stereotypical lines (e.g., male voice interfaces are believed to be more knowledgeable about technical topics, and female voice interfaces more knowledgeable about topics like love and relationships); that adults respond to multiple voices from a single computer as though they were separate entities; and that adults are less likely to criticize a computer directly (i.e. if the computer itself asks for an evaluation) than if a third party (a human or different computer) asks for the evaluation. The results from the AIBO Discussion Forum Study extended this body of work by showing that when an animal persona is embedded in the computation that adults can readily accept that the resulting technological artifact pulls for their social responses.

Thus a new technological genre may be emerging that challenges traditional ontological categories (e.g., between animate and inanimate). If we are correct, then it may be that the English language is not yet well equipped to characterize or talk about this genre. As an analogy, we do not normally present people with an orange object and ask, “Is this object red or yellow?” It is something of both, and we call it orange. Similarly, it may not be the best approach to keep asking if this emerging technological genre is, for example, “alive” or “not alive” if from the person's experience of the subject-object interaction, the object is alive in some respects and not alive in other respects, and is experienced not simply as a combination of such qualities (in the way one can inspect a tossed salad and analytically distinguish, for example, between the green leaf lettuce and the red leaf lettuce) but as a novel entity.

If it is true that a new technological genre is emerging, questions remain that have been a focus of this paper: Are pervasive interactions with a wide array of

instantiations of this genre a good thing for human beings? If so, in what ways? In what contexts? And where might such interactions impoverish the human experience? These are important questions, and in our view warrant substantial psychological research in the immediate years ahead.

Acknowledgments

This research was supported in part by NSF Awards IIS-0102558 and IIS-0325035.

References

- [1] Aylett, R. 2002. *Robots: Bringing Intelligent Machines to Life?* Hauppauge, NY: Barron.
- [2] Bartneck, C., & Forlizzi, J. 2004. *A design-centered framework for social human-robot interaction*. Unpublished manuscript, Eindhoven University of Technology, the Netherlands & Carnegie Mellon University, Pittsburgh, PA.
- [3] Beck, A., & Katcher, A. 1996. *Between Pets and People*. West Lafayette, IN: Purdue University Press.
- [4] Breazeal, C. L. 2002. *Designing Sociable Robots: Intelligent Robotics and Autonomous Agents*. Cambridge, MA: MIT Press.
- [5] Breazeal, C. 2003. Toward sociable robots. *Robotics and Autonomous Systems*, 42, 167-175.
- [6] Edwards, N., Beck, A., Kahn, P. H., Jr., & Friedman, B. in progress. [Robotic pets and the elderly].
- [7] Fong, T., Nourbakhsh, I., & Dautenhahn, K. 2003. A survey of socially interactive robots. *Robotics and Autonomous Systems*, 42, 143-166.
- [8] Friedman, B., & Kahn, P. H., Jr. 1992. Human agency and responsible computing: Implications for computer system design. *Journal of Systems Software*, 17, 7-14.
- [9] Friedman, B., & Kahn, P. H., Jr. 2003. Human values, ethics, and design. In J. A. Jacko and A. Sears (Eds.), *The Human-Computer Interaction Handbook* (pp. 1177-1201). Mahwah, NJ: Erlbaum.
- [10] Friedman, B., Kahn, P. H., Jr., & Borning, A. in press. Value sensitive design and information systems. To appear in D. Galletta & P. Zhang (Eds.), *Human-Computer Interaction in Management Information Systems*. Armonk, NY: Sharpe.
- [11] Griffin, J. H. 1960. *Black Like Me*. New York: Signet.
- [12] Kahn, P. H., Jr. 1992. Children's obligatory and discretionary moral judgments. *Child Development*, 63, 416-430.
- [13] Kahn, P. H., Jr. 1999. *The Human Relationship with Nature: Development and Culture*. Cambridge, MA: MIT Press.
- [14] Friedman, B., Kahn, P. H., Jr. & Hagman, J. 2003. Hardware companions? – What online AIBO discussion forums reveal about the human-robotic relationship. *Conference Proceedings of CHI 2003* (pp. 273-280). New York, NY: Association for Computing Machinery.
- [15] Kahn, P. H., Jr., Friedman, B., Perez-Granados, D. R., Freier, N. G. 2004, April. Robotic pets in the lives of preschool children. *Extended Abstracts of CHI 2004*. New York, NY: Association for Computing Machinery Press. [Includes video figure.]
- [16] Melson, G. F., Kahn, P. H., Jr., Beck, A., & Friedman, B. in progress [Robotic pets and children: A developmental study].
- [17] Oxford English Dictionary. 2004. Oxford: Oxford University Press. Retrieved May 24, 2004 from http://dictionary.oed.com/cgi/entry/00199182?single=1&query_type=word&queryword=reciprocal&editon=2e&first=1&max_to_show=10
- [18] Piaget, J. 1969. *The Moral Judgment of the Child*. Glencoe, IL: Free Press. (Original work published 1932)
- [19] Reeves, B., & Nass, C. 1996. *The Media Equation. How People Treat Computers, Television, and New Media Like Real People and Places*. New York and Stanford, CA: Cambridge University Press and Center for the Study of Language and Information Publications.
- [20] Searle, J. R. 1990. Is the brain's mind a computer program? *Scientific American*, 262, 26-31.
- [21] Shepard, P. 1996. *The others: How animals made us human*. Washington, DC: Island Press.
- [22] Turiel, E. 1983. *The development of social knowledge*. Cambridge, England: Cambridge University Press.
- [23] Turiel, E. 1998. Moral development. In W. Damon (Ed.) *Handbook of Child Psychology* (5th ed.). Vol. 3: N. Eisenberg (Ed.), *Social, Emotional, and Personality Development* (pp. 863-932). New York: Wiley.