

# Close Engagements with Artificial Companions

Key social, psychological, ethical and design issues

*Edited by*

Yorick Wilks

University of Oxford

2010

John Benjamins Publishing Company

Amsterdam / Philadelphia

## In good company?

On the threshold of robotic companions

Sherry Turkle

Most contributors to this volume believe that only technical matters stand between where we are now and a time when robots will be our companions and teachers. In this view, robots need to expand their domains of understanding, and if those domains should be emotional, well, that will be a technical matter as well. So while this volume expresses designers' enthusiasm about robots as technical objects, it challenges us to see robots as something more, as evocative objects. What are we thinking about when we are thinking about robots? We are thinking about aliveness and authenticity, love and spirituality. We are thinking about what it means to build a psychology. We are thinking about what makes people special. Or perhaps that they are not so special after all.

If one thinks of a classic "upstairs/downstairs" scenario, it is no longer clear where the robots will be lodging. For some years, robot vacuum cleaners have staked their claim as household helpers, but more recently, roboticists would expand the domain of the robotic from doing chores to offering companionship and care. But to make this step, one needs a new kind of robot, one equipped to be good company. These are sociable robots, robots programmed to have mental states that shift as they interact with people. To take those already on the scene, consider, for example, the robotic doll My Real Baby. Bounce the doll and it gets happy; bounce it too much and it might get ornery. Or Pleo, a small dinosaur robot advertised for its psychological '*autonomy*'; or Paro, a baby seal robot designed as a therapeutic companion. Like the others, Paro has "states of mind", even if primitive ones. Interacting with Paro requires understanding its states, a demand for vigilance that, in the end, contributes to its appeal. Such digital creatures are the shock troops of a cultural moment when figuring out how a robot is "feeling" in order to get along with it begins to seem a natural thing.

When robots make eye contact, recognize faces, mirror human gestures, they push our Darwinian buttons, exhibiting the kinds of behavior people associate

with sentience, intentions, and emotions. Once people see robots as creatures, people feel a desire to nurture them. With this feeling comes the fantasy of reciprocation: as we begin to care for robots, we want them to care about us. In our nascent robotics culture, nurturance turns out to be a “killer app”. Eleven-year-old Fara reacts to a play session with Cog, a humanoid robot at MIT that can meet her eyes, follow her position, and imitate her movements, by saying that she could never get tired of the robot because “it’s not like a toy because you can’t teach a toy; it’s like something that’s part of you, you know, something you love, kind of like another person, like a baby.”

What seems natural to the contributors to this volume – that robots are part of our relational futures – is now a widely shared idea, and not just among experts. Indeed, it is so widely shared that it is easy to forget that, until recently, it was hardly shared at all. In the 1980s, one common response to the computer presence could be summed up as a “romantic reaction”.<sup>1</sup> It accepted that people might well be a kind of computer (a major concession to models of mind as program) but stressed the soul and the spirit in the human machine: simulated thinking might be thinking but simulated feeling is never feeling, simulated love is never love. Computers were fine, more than fine, if they were doing instrumental jobs, or jobs that could be neatly compartmentalized in the ‘*thinking*’ category. But computational objects – robots included – should not be allowed into the realm of human relationships.<sup>2</sup> So, when in 1984 I called the computer a “second self” and demonstrated people’s strong personal connection to them, many objected by insisting on the truism: “the computer is just a tool.”

Over the next decade, opinions shifted.<sup>3</sup> Computers became everyday objects and it became commonplace to see one’s laptop as an extension of self. “What is on your PowerBook?” was a good advertising slogan in the mid-1990s because it acknowledged the degree to which a computer desktop reflected personal as well as intellectual commitments. People still saw computers as tools but recognized that they were tools with a difference. They had special vocations: not-yet-minds, but on the boundaries of mind, they were close to being minds whose opinions counted.<sup>4</sup>

1. Turkle, *The Second Self*.

2. Ibid.

3. Turkle, *Life on the Screen*. There I discuss the case of computer psychotherapy to illustrate a turning away from the romantic reaction as a response to programs as appropriate dialogue partners in the realm of the personal.

4. Ibid.

These days, in studying reactions to sociable robots, robots that do such things as look you in the eye, remember your name, and track your motion, I find numbers of people who consider such objects as potential friends, confidants, and (as they imagine technical improvements) even lovers. I listen for what stands behind this new attitude and I hear three things. There is openness to seeing computational objects as “other minds”; there is willingness to consider what a computer and human mind have in common; and, in a different register, there is evidence of a certain fatigue with the difficulties of dealing with people. A female graduate student comes up to me after a lecture and tells me she would gladly trade in her boyfriend for a sophisticated humanoid robot as long as the robot could produce what she calls “caring behavior”. She tells me that “I need the feeling of civility in the house and I don’t want to be alone”. She says: “If the robot could provide a civil environment, I would be happy to help produce the illusion that there is somebody really with me.” What she is looking for, she says, is a no-risk relationship that will stave off loneliness; a responsive robot, even if it were just exhibiting scripted behavior, seems better to her than a demanding boyfriend.

The distance travelled over the past twenty years has been impressive. Bruce, a thirteen-year-old I interviewed in 1983, thinks of computers and robots as “perfect” and therefore different from flawed and frail people. Robots do everything right; people do the best they know how. But for Bruce, it is human imperfection that makes for the ties that bind. Specifically, his own limitations make him feel close to his father (“I have a lot in common with my father ... we both have chaos”). Perfect robots could not understand this very important relationship. Twenty years later, Howard, fourteen, compares his father to the idea of a robot counsellor and the human does not fare as well. Howard thinks that the robot would be better able to grasp the intricacies in the day of a high school student. He comments on what the robot would bring to the table: “Its database would be larger than Dad’s. Dad has knowledge of basic things, but not enough of high school. Robots can be made to understand things like “feelings”.

With the belief that robots can understand them, people are, of course, more likely to warm to their company. In the presence of sociable robots people feel attachment and loss; they want to reminisce and feel loved. In a year-long study of human-robot bonding, one seventy-four-year-old Japanese participant said of her Wandukun, a furry robot creature designed to resemble a koala bear: “When I looked into his large, brown eyes, I feel in love after years of being quite lonely ... I swore to protect and care for the little animal.”<sup>5</sup> In my study

5. Suvendi Kakushi, “Robot Lovin” Asia Week Magazine Online, November 9, 2001. <http://www.asiaweek.com/asiaweek/magazine/life/0,8782,182326,00html>. Accessed on 5/9/05.

of robots in Massachusetts nursing homes, seventy-four-year-old Jonathan responds to his My Real Baby robot doll by wishing it were a bit smarter because he would prefer to talk to a robot about his problems than to a person. "The robot wouldn't criticize me." Andy, also seventy-four, says that the My Real Baby, which responds to caretaking by exhibiting different states of mind, bears a resemblance to his ex-wife Rose, "something in the eyes". He likes chatting with the robot about events of the day. "When I wake up in the morning and see her face [the robot] over there, it makes me feel so nice, like somebody is watching over me."

I recently had an exchange with colleagues who wrote about the 'I-Thou' dyad of people and robots and I could only see Martin Buber spinning in his grave.<sup>6</sup> The *I* was the person in the relationship, but how could the robot be the *Thou*? In the past, I might have focused on how my colleagues' projected feelings on a robot that definitionally could not have them. But I had taken that position when I interpreted attitudes toward robots as a kind of Rorschach for better understanding people's hopes and frustrations. Now, there was a new earnestness, a new literal-mindedness to the consideration of a robot companion. My colleagues saw the robot in the wings and were eager to welcome it onstage.

It seemed no time at all that a book came out, *Love and Sex with Robots* and a reporter from *Scientific American* was interviewing me about the psychology of robot marriage.<sup>7</sup> I found the conversation memorable. I was asked if my opposition to people marrying robots didn't put me in the same camp as those who oppose the marriage of lesbians or gay men. I tried to explain that just because I didn't think people could marry machines didn't mean that any mix of people with people was fair play. The reporter accused me of species chauvinism. Wasn't this the kind of talk that homophobes once used, not considering gays as "real" people? Our culture had clearly come to a new place.

To me, robots represent the new uncanny in our culture of simulation. Here I refer to the uncanny in Freud's sense – something known of old and long familiar, yet now made strangely unfamiliar.<sup>8</sup> They are not like the dolls of the past that offered possibilities for pure projection. The new relational robots are built with psychologies and needs of their own. As uncanny objects, objects on the bound-

aries of categories, robotic creatures provoke us to ask questions about traditional categories, questions such as, *What kinds of relationships are appropriate to have with machines?* And more generally, *What is a relationship?* The question is not whether children will grow up to love their robots more than other toys, or indeed, their parents, but what will loving come to mean?

The psychoanalyst Heinz Kohut describes how some people may temporarily strengthen their fragile sense of self by turning another person into a "self-object".<sup>9</sup> In the role of self-object, the other is experienced as part of the self, thus in perfect tune with the fragile individual's inner state. Disappointments inevitably follow. The relationships people form on social networking sites (whether in MySpace or Facebook chat or in virtual worlds) are excellent contenders for the role of self-object. There, we come to know people through a curious half-light, where people can be imagined to be what the fragile self needs them to be.<sup>10</sup>

Similarly, robots that look into your eyes, trace your movements, perhaps say your name, clearly present themselves as candidates for the role of self-object. If they can give the appearance of aliveness and yet not disappoint, they may even have a "comparative advantage" over people for this job, and thus open new possibilities for narcissistic experience with machines. Why do I say, "comparative advantage?" When people turn other people into self-objects, one might say they are making an effort to turn a person into a kind of "spare part" rather than taking him or her as an autonomous, individual personality. The artificial companion or robot is, of course, already a spare part.

The seductions of the robotic provide a window onto how much people are tempted to sidestep encounters with friends and family. Over-stressed, over-worked, people claim exhaustion and overload. Loneliness is failed solitude. Are cyber-connections paving the way to considering robotic companions as sufficient unto the day? These days, people readily admit that they would rather leave a voice mail or send an email than talk face-to-face. And from there, they say: "I'd rather talk to the robot. Friends can be exhausting. The robot will always be there for me. And whenever I'm done, I can walk away." Or as one woman said about AIBO, Sony's household entertainment robot, "It is better than a real dog ... It won't do dangerous things, and it won't betray you ... Also, it won't die suddenly and make you feel very sad." The romantic reaction to the computer presence stressed that simulation had no place in matters of love. These days, people are

6. *Interaction Studies*, Ibid.

7. David Levy, *Love and Sex With Robots: The Evolution of Human-Robot Relationships* (New York: HarperCollins, 2007).

8. Sigmund Freud, "The Uncanny," In *The Standard Edition of the Complete Works of Sigmund Freud*, edited by James Strachey, et al. London: The Hogarth Press and The Institute of Psychoanalysis, 1953–74.

9. P. H. Ornstein, Ed., *The search for the self: Selected writings of Heinz Kohut: 1950–1978*, (Vol. 2). New York: International Universities Press, Inc., 1978.

10. On this, see Sherry Turkle, "Whither Psychoanalysis in Computer Culture," *Psychoanalytic Psychology*, Vol. 21, 2004.

likely to speculate on the possibility that humans, like robots, get by on simulation. One thirty-year-old puts it this way: "How do I know that my lover is not just simulating everything he says he feels?" A fifty-nine-year-old man said: "My first wife faked her orgasms for twenty-five years ... so what would make a robot inauthentic? She set the bar pretty low." Or recall the graduate student who was willing to "help produce the illusion that there is somebody really with me" if a robot would just provide some background.

In the 1980s, people insisted that the bedrock of human uniqueness was what computers could not do or be, placing special value on the importance of the human life cycle in defining what was essential about being a person. One man, considering the possibility of confiding in a computer psychotherapist put it this way: "How can I talk about sibling rivalry to something that never had a mother?" It put a premium on the idea that only people can give each other understanding and empathy. It was invested in the idea that there is something essential about the human spirit, and that this essential quality resides in human inner states. Now this essentialist assumption is challenged. Today one does not linger over inner states. The new focus is on behavior. What matters is how the robots perform and how we perform for each other – the essence, after all, of life in virtual communities where we create an avatar and put it on a self-built stage. With the focus on behavior rather than inner states, a creature that behaves appropriately is an appropriate creature.

In the 1980s, debates in Artificial Intelligence centred on the question of whether machines could "really" be intelligent. These debates were about the objects themselves, what they could and could not do. Our new debates about relational and sociable machines – debates that will have an increasingly high profile in mainstream culture – are not only about the machines' capabilities but about our vulnerabilities, both to machines that push our Darwinian buttons and to the promise of relationship, any relationship, in a world where humans so often seem to disappoint. For many who are lonely yet fearful of intimacy, a robotic companion offers the illusion of companionship without the demands of sustained, intimate friendship. One can be a loner yet never alone. It is a small step from this position to the question, one implicit in so many of the contributions in this volume: *What are the purposes of living things?*

From the perspective of today's young people, the answers may not be obvious. Recall fourteen-year-old Howard who thinks that robots might be better than people when it comes to understanding the intricacies of high school and eleven-year-old Fara who looks forward to nurturing a baby-like robot. As for me, the question has been raised very close to home. I took my daughter, then fourteen, to visit the Darwin exhibit at the American Museum of Natural History. The exhibit documented Darwin's life and thought, and with a somewhat

defensive tone (in light of current challenges to evolution by proponents of intelligent design), presented the theory of evolution as the central truth that underpins contemporary biology. The Darwin exhibit wanted to convince and it wanted to please. At its entrance was a turtle from the Galapagos Islands, a seminal object in the development of evolutionary theory. The turtle rested in its cage, utterly still. "They could have used a robot," commented my daughter. She considered it a shame to bring the turtle all this way and put it in a cage for a performance that draws so little on the turtle's "aliveness." I was startled by her comments, both solicitous of the imprisoned turtle because it is alive and unconcerned about its authenticity. The museum had been advertising these turtles as wonders, curiosities, and marvels – among the plastic models of life at the museum, here is the life that Darwin saw.

I began to talk with others at the exhibit, parents and children. It was Thanksgiving weekend. The line was long, the crowd frozen in place. My question, "Do you care that the turtle is alive?" was welcome diversion. A ten-year-old girl would prefer a robot turtle because aliveness comes with aesthetic inconvenience: "its water looks dirty. Gross." More usually, votes for the robots echoed my daughter's sentiment that in this setting, aliveness doesn't seem worth the trouble. A twelve-year-old girl opined: "For what the turtles do, you didn't have to have the live ones." Her father looked at her, uncomprehending: "But the point is that they are real, that's the whole point."

The Darwin exhibit gave authenticity major play: on display were the actual magnifying glass that Darwin used, the actual notebooks in which he recorded his observations, indeed, the very notebook in which he wrote the famous sentences that first described his theory of evolution. But, in the children's reactions to the inert but alive Galapagos turtle, the idea of the original was in crisis. I recall my daughter's reaction when she was seven to a boat ride in the postcard blue Mediterranean. Already an expert in the world of simulated fish tanks, she saw a creature in the water, pointed to it excitedly and said: "Look mommy, a jellyfish! It looks so realistic!" When I told this story to a friend who was a research scientist at the Walt Disney Company, he was not surprised. When Animal Kingdom opened in Orlando, populated by "real", that is, biological animals, its first visitors complained that these animals were not as "realistic" as the animatronic creatures in Disneyworld, just across the road. The robotic crocodiles slapped their tails, rolled their eyes, in sum, displayed "essence of crocodile" behavior. The biological crocodiles, like the Galapagos turtle, pretty much kept to themselves. What is the gold standard here?

I have long believed that in our culture of simulation, the notion of authenticity is for us what sex was to the Victorians – threat and obsession, taboo and fascination. I have lived with this idea for many years, yet at the museum, I found

the children's position strangely unsettling. For them, in this context, aliveness seemed to have no intrinsic value. Rather, it was useful only if needed for a specific purpose. "If you put in a robot instead of the live turtle, do you think people should be told that the turtle is not alive?" I ask. "Not really", said several of the children. Data on "aliveness" can be shared on a "need to know" basis, for a purpose. This volume with its enthusiastic embrace of machines as kin both provokes and challenges: What indeed are the purposes of living things?