

Computers and the Human Spirit

When I first turned to the study of computers and people in the late 1970s, I immersed myself in a world altogether strange to me. Trained as a humanist, I took a job at the Massachusetts Institute of Technology where I was surrounded by people who spoke about the mind in a language with which I was unfamiliar, a language of bits and bytes, registers and compilers. Many of them had strong, sometimes even passionate relationships with digital machines. Many claimed that working with computers changed the way they thought about the world, about their relationships with others, and most strikingly, about themselves. I first heard such extravagant sentiments expressed by computer enthusiasts within the academy, but as time went on, I came across them in personal computer clubs and grade school classrooms.¹ “When you program a computer, you put a little piece of your mind into the computer’s mind and you come to see yourself differently,” said Deborah, a sixth grade student in an elementary school that had recently introduced computer programming into its curriculum. By 1984, I had come to see the computer as a “second self.” At that time, the notion of mind as program was controversial. These days, the use of computational metaphors for mind has become banal. With the introduction of computers into mainstream culture in the late 1970’s and early 1980’s, large numbers of people began to describe human mental activity in computational terms. (“Excuse me, I need to clear my buffer; I won’t be happy until I debug this problem”). With an increasing acceptance of mind as mechanism, came an attendant question: If mind is program, where is free will? By the mid-1980s, the computer was clearly an evocative object, an object that provoked self-reflection.

Today, cognitive science has developed far more sophisticated computational models of mental processes than were dreamt of in the days of the nascent computer culture, and the Internet has made it possible for people to assume and explore multiple aspects of self in their online lives.² But with time grows familiarity and what was once exotic begins to seem “natural.” The computer is now so taken for granted that it has become cultural “background noise” and we may not notice its powerful effects on our thinking about self. We are on the verge of an era in which we feel ourselves in relationships of “mutual” affection with computational companions. These new relationships should not slip into our emotional lives as “background noise.” I revisit the recent history of how interacting with computers has affected our sense of self in the hope that “defamiliarizing” its effects will enhance the quality of our conversation about what comes next.

From Rorschach to Identity Workshop

When in the early 1980s I first called the computer a “second self” or a Rorschach, a projective screen, relationships with computers were usually one-to-one, a person alone with a machine. With the widespread use of the Internet, this was no longer the case. Virtual sociability changed the form of our communities and the expression of our sexuality. The Internet made it possible for users to cycle through different self-generated *personae* that could cut across “real life” distinctions of gender, race, class, and culture. On the Internet, the obese have a chance to be slender; for the beautiful, there is an opportunity to try out being plain. The fact that there is time to reflect upon and edit one’s self-composition makes it easier for the shy to be outgoing, the “nerdy” sophisticated. The relative anonymity of life on the screen—one has the choice of being known only by one’s

chosen “handle” or online name—gives people a chance to express unexplored aspects of their personalities. The same person can be known by several names. It would not be unusual for someone to be BroncoBill in one online context, ArmaniBoy in another, and MrSensitive in a third.

In the 1990s it became clear that cyberspace could serve as a kind of identity workshop.³ The people who make the most of online experiences are those who are capable of approaching them in a spirit of self-reflection. They ask: What does my behavior in cyberspace tell me about what I want, who I am, what I may not be getting in the rest of my life? Even the “windows” interface has become a potent metaphor for thinking about the self as a multiple, distributed, “time-sharing” system, suggesting a distributed self that exists in many worlds and plays many roles at the same time. To use the psychoanalyst Philip Bromberg’s language, online life facilitates a psychological culture in which one can “stand in the spaces between selves and still feel one, to see the multiplicity and still feel a unity.”⁴ To use the computer scientist Marvin Minsky’s language, it facilitates a culture where one can feel at ease cycling through one’s “society of mind.”⁵

Aliveness: From Motion to Emotion and Beyond ■■■■■■■■

When the Swiss psychologist Jean Piaget interviewed children in the 1920s and 1930s about which objects were “alive” and which were not, he found that children honed their definition of life by developing increasingly sophisticated notions about motion, the world of *physics*.⁶ In contrast, when I began to study the nascent computer culture in the late 1970s, children argued about whether a computer was alive through discussions about its *psychology*. Did the computer know things on its own or did it have to be programmed? Did it have intentions, consciousness, and feelings? Did it cheat? Did it know it was cheating? Although the presence of the first generation of computational toys (games like Merlin, Simon, and Speak and Spell) challenged the classical Piagetian story about children’s notions of aliveness, the story children were telling about such objects in the early 1980s had its own coherency. Faced with intelligent toys, children shifted from talking about the aliveness of an object in terms of motion to talking about it in terms of intentionality and cognition. They imposed a new conceptual order on a new world of objects.

In the 1990s, new computational objects that embodied principles of evolution (such as the Sim series of games) strained that order to the breaking point. Children still tried to impose order on these objects, but they did so in the manner of theoretical tinkerers or “bricoleurs,” constructing passing theories to fit prevailing circumstances. They “cycled through” various notions of what it took to be alive, saying for example that robots are in control but not alive, would be alive *if* they had bodies, are alive *because* they have bodies, would be alive if they had feelings, are alive the way insects are alive but not the way people are alive. They said that Sim creatures (for example in the game Sim City) are not alive but almost-alive, would be alive if they spoke, would be alive if they traveled, are alive but not “real,” are not alive because they don’t have bodies, are alive because they can have babies, would be alive if they could escape the game and “get out onto America Online.” In the presence of increasingly complex computational artifacts there had developed a radical heterogeneity of theory about how to speak about “aliveness.” This heterogeneity spilled over into children’s conversation when they were away from the computer. In the early 1990s, I observed a group of seven-year-olds playing with transformer toys that could take the shape of armored tanks, robots, or people. The transformers could also be put into intermediate states so that a “robot” arm could protrude from a human form or a human leg from a mechanical tank. Two of the children were play-

ing with the toys in these intermediate states [that is, in their intermediate states somewhere between being people, machines, and robots]. A third child insisted that this was not right. The toys, he said, should not be placed in hybrid states. “You should play them as all tank or all people.” He was getting upset because the other two children were making a point of ignoring him. An eight-year-old girl comforted the upset child. “It’s okay to play them when they are in between. It’s all the same stuff,” she said, “just yucky computer ‘cy-dough-plasm.’”

Today’s adults grew up in a psychological culture that equated the idea of a unitary self with psychological health and in a scientific culture that taught that when a discipline achieves maturity, it has a unifying theory. When adults find themselves cycling through varying perspectives on self (“I am my chemicals” to “I am my history” to “I am my genes”) they usually become uncomfortable.⁷ Such movement does not correspond to the unitary notion of self they were brought up to expect. But by the 1990s, children had learned a different lesson from their computational objects-to-think with. Having a range of ideas about mind and life may strike them as “just the way things are.” This is the lesson of the cy-dough-plasm: it is a lesson about fluid definitions of self and the discourse of aliveness.

Most recently, a new kind of evocative computational object has entered children’s lives. These include virtual creatures, digital dolls, robotic pets, humanoid robots, and software programs designed to monitor their users’ affect and show affect of their own. I call these relational artifacts—objects that present themselves as “affective” and “sociable.”

For the most part relational artifacts entered children’s lives with Tamagotchis, little screen creatures developed in Japan in the mid-1990s, that got bored and needed to be amused, got hungry and needed to be fed, got dirty and needed to be cleaned, got sick and needed to be nursed. Furbies, small furry owl-like creatures, the toy fad of 1998, shared many of the psychological properties that had animated the Tamagotchis. Most important, the Furbies demanded attention. They played games, “learned” to speak English, and said “I love you.” In 2000, My Real Baby, a robotic infant doll based on a prototype developed at the MIT AI Laboratory, appeared on the market. My Real Baby makes baby sounds and baby facial expressions, but more significant than its physical similarities to an infant, this computationally complex doll was designed to give the appearance of having baby “states of mind.” Bounce the doll when it is happy, and it gets happier. Bounce it when it is grumpy and it gets grumpier. Aibo, Sony’s entertainment robotic dog, develops different personalities depending on how it is treated. The newest models have facial and voice recognition software that enable Aibo to recognize its “primary caregiver.” These objects confront us with new questions: What kinds of relationships are appropriate, desirable, and imaginable with technology? What is a relationship?

These relational artifacts do not wait for children to “animate” them in the spirit of a Raggedy Anne doll or the Velveteen Rabbit, the stuffed animal who finally came alive because so many children had loved him. They present themselves as already animated and ready for relationship. I found that children describe these new toys as “sort of alive” not because of their cognitive capacities or seeming autonomy (as was the case for previous generations of computational objects) but because of the quality of their emotional attachments to the objects and the notion that the objects might be emotionally attached to them. For example, in my study of children and Furbies, when I asked the question, “Do you think the Furby is alive?” children answered not in terms of what the Furby could do, but rather in terms of how they felt about the Furby and of how, in their estimation, the Furby felt about them.

Ron (6): Well, the Furby is alive for a Furby. And you know, something this smart should have arms. It might want to pick up something or to hug me.

Katherine (5): Is it alive? Well, I love it. It's more alive than a Tamagotchi because it sleeps with me. It likes to sleep with me.

Jen (9): I really like to take care of it. So, I guess it is alive, but it doesn't need to really eat, so it is as alive as you can be if you don't eat. A Furby is like an owl. But it is more alive than an owl because it knows more and you can talk to it. But it needs batteries so it is not an animal. It's not like an animal kind of alive.

My study of children and relational artifacts is ongoing, but several things are already clear. Today's children are learning to distinguish between an "animal kind of alive" and a "Furby [or robot] kind of alive." The category of "sort of alive" is used with increasing frequency. Children already talk about an "animal kind of alive" and "a Furby kind of alive." Will they also come to talk about a "people kind of love" and a "computer kind of love?"

In Steven Spielberg's movie, *AI: Artificial Intelligence*, scientists build a humanoid robot, David, who is programmed to love. David expresses his love to a woman who has adopted him as her child. In the discussion that followed the release of the film, much conversation centered on the question of whether such a robot could *really* be developed. Was this technically feasible? And if it was feasible, how long would we have to wait for it? People thereby passed over another question, one that historically has contributed to our fascination with the computer's burgeoning capabilities. That question concerns not what computers can do or what computers will be like in the future, but rather, what we will be like. *What kinds of people are we becoming as we develop increasingly intimate relationships with our machines?*

We are in a different world than that in which the old AI debates about whether machines could be "really" intelligent were conducted. The old debates were about the machines themselves, about what they could and could not do. New debates – debates that will have an increasingly high cultural profile – will rather concern the impact that these objects are having on *us*. When an object invites us to care for it, when the cared-for object thrives under our care, we experience that object as intelligent (whether or not we are justified in so doing). More important, we feel a connection to it. So the question for the future is not whether relational artifacts "really" have emotions, but rather what these objects evoke in their users.

In this context, the pressing issue in Spielberg's *A.I.* is not the potential "reality" of a robot that loves, but rather the conflicts faced by its adoptive mother—a human being whose response to a machine that asks for nurturance is the desire to nurture it; whose response to a non-biological creature who reaches out to her is attachment, love, horror, and confusion.

Today, we are faced with relational artifacts that elicit responses from their users / owners that have much in common with those of the mother in *A.I.* These artifacts are not perfect human replicas like the imaginary David, but they are able to push certain emotional buttons (think of them perhaps as evolutionary buttons). To take the simplest example: When a robotic creature makes eye contact, follows your gaze, and gestures towards you, you are provoked to respond to that creature as a sentient and even caring other.

I have most recently been studying children playing with virtual pets and digital dolls, and the elderly to whom robotic companions are starting to be aggressively marketed.

How will interacting with relational artifacts affect people's way of thinking about themselves, their sense of human identity, of what makes people (and pets) special? Children

have traditionally defined what makes people special in terms of a theory of “nearest neighbors.” So, when the nearest neighbors (in children’s eyes) were their pet dogs and cats, people were special because they had reason. The Aristotelian definition of man as a rational animal made sense even for the youngest children. But when, in the 1980s, it seemed to be the computers who were the nearest neighbors, children’s approach to the problem changed. Children still used the “nearest neighbors” methodology. But now, people were special not because they were rational animals but because of their differences from the rational computers: people were emotional machines. So, in 1983, a ten-year-old told me: “When there are the robots that are as smart as the people, the people will still run the restaurants, cook the food, have the families. I guess they’ll still be the only ones who’ll go to Church.” Today, speaking about robot pets, one hears echoes of this “romantic reaction.” Some children say that the robots could be friends, but not “best friends” because they are “too perfect,” and people are not. Others, as for example, one eleven-year-old girl, are more concrete: “They can’t be friends because you can’t take them to lunch.” And yet there is movement in another direction. In Ray Bradbury’s story, “I sing the body electric,” a robotic, electronic grandmother is unable to win the trust of the girl in the family, Agatha, until the girl learns that the grandmother, unlike her recently deceased mother, cannot die.⁸ In many ways throughout the story we learn that the grandmother is actually better than a human caretaker—more able to attend to each family member’s needs, less needy, with perfect memory and inscrutable skills—and most importantly – not mortal. One woman’s comment on AIBO, Sony’s household entertainment robot, startles us with what it might augur for the future of person-machine relationships: “[AIBO] 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.”

Mortality has traditionally defined the human condition; a shared sense of mortality has been the basis for feeling a commonality with other human beings, a sense of going through the same life cycle, a sense of the preciousness of time and life, of its fragility. Loss (of parents, of friends, of family) is part of the way we understand how human beings grow and develop and bring the qualities of other people within themselves.

The question, “What kinds of relationships is it appropriate to have with machines?” has been explored in science fiction and in technophilosophy. But the sight of children and the elderly exchanging tenderness with robotic pets brings science fiction into everyday life and technophilosophy down to earth. In the end, the question is not just whether our children will come to love their toy robots more than their parents, but what will *loving* itself come to mean?

- 1 See Turkle, Sherry. *The Second Self: Computers and the Human Spirit*, Simon and Schuster, New York, 1984
- 2 See Turkle, Sherry. *Life on the Screen: Identity in the Age of the Internet*, Simon and Schuster, New York, 1995
- 3 This felicitous phrase was coined by my then student, Amy Bruckman.
- 4 Bromberg, Philip. “Speak that I May See You: Some Reflections on Dissociation, Reality, and Psychoanalytic Listening,” in *Psychoanalytic Dialogues*. 4 (4). pp.517–547. 1994
- 5 Minsky, Marvin. *The Society of Mind*, Simon and Schuster, New York, 1987
- 6 Piaget, Jean. *The Child’s Conception of the World* (trans. by Joan and Andrew Tomlinson Totowa) Adams, Littlefield N. J., 1960
- 7 Kramer, Peter. *Listening to Prozac: A Psychiatrist Explores Antidepressant Drugs and the Remaking of the Self*, p. Xii–xiii. Viking, New York, 1993
- 8 Bradbury, Ray. *I Sing the Body Electric and Other Stories*, Avon Books, New York, 1998 [1946]