Diary

CONSIDER our moment: I take my 14-year-old daughter to the Darwin exhibition at the American Museum of Natural History. The exhibition documents Darwin's life and thought, and in a somewhat defensive tone presents the theory of evolution as the central truth that underpins contemporary biology. The exhibition wants to convince and it wants to please. At the entrance is a turtle from the Galapagos Islands. The turtle rests in its cage, utterly still. 'They could have used a robot,' my daughter remarks. She considers 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 am startled by her comments, both solicitous of the imprisoned turtle because it is alive and unconcerned about by its authenticity. The museum has been advertising these turtles as wonders, curiosities, marvels – among the plastic models of life at the museum, here is the life that Darwin saw. It is Thanksgiving weekend. The line is long, the crowd frozen in place. I begin to talk with some of the other parents and children. My question, 'Do you care that the turtle is alive?' is a 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 echo my daughter's sentiment that in this setting aliveness doesn't seem worth the trouble. A 12-year-old girl is adamant: 'For what the turtles do, you didn't have to have

the live ones.' Her father looks at her, uncomprehending: 'But the point is that they are real, that's the whole point.'

The Darwin exhibition gives authenticity major play: on display are the actual magnifying glass that Darwin used, the actual notebooks in which he recorded his observations, 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 has no place. I recall my daughter's reaction as a young child to a boat ride in the 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, a jellyfish! It looks so realistic!' 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' behaviour. The biological crocodiles, like the Galapagos turtle, pretty much kept to themselves.

I find the children's position unsettling. '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, say several of the children. Data on 'aliveness' can be shared on a need to know basis. But when do we need to know if something is alive?

Consider another moment: an older woman in a nursing home outside Boston

is sad. Her son has broken off his relationship with her. Her nursing home is part of a study I am conducting on robotics for the elderly. I am recording her reactions as she sits with the robot Paro, a seal-like creature, advertised as the first 'therapeutic robot' for its ostensibly positive effects on the ill, the elderly and the emotionally troubled. Paro is able to make eye contact by sensing the direction a human voice is coming from, it is sensitive to touch, and has 'states of mind' that are affected by how it is treated – for example, it can sense whether it is being stroked gently or more aggressively. In this session with Paro, the woman, depressed because of her son's abandonment, comes to believe that the robot is depressed as well. She turns to Paro, strokes him and says: 'Yes, you're sad, aren't you. It's tough out there. Yes, it's hard.' And then she pets the robot once again, attempting to provide it with comfort. And in so doing, she tries to comfort herself.

What are we to make of this transaction as it unfolds between a depressed woman and a robot? When I talk to others about it, their first associations are usually to their pets and the comfort they provide. I don't know whether a pet could feel or smell or intuit some understanding of what it might mean to be with an old woman whose son has chosen not to see her anymore. But I do know that Paro has understood nothing. The woman's sense of being understood is based on the ability of computational objects like Paro – 'relational artifacts' I call them – to convince their users that they are in a relationship by pushing certain 'Darwinian' buttons (making eye contact, for example) that cause people to respond as though they were in relationship. Relational artifacts are the new uncanny in our computer culture – as Freud put it, 'the long familiar taking a form that is strangely unfamiliar'.

Confrontation with the uncanny provokes new reflection. Do plans to provide relational robots to children and the elderly make us less likely to look for other solutions for their care? If our experience with relational artifacts is based on a fundamentally deceitful interchange can it be good for us? Or might it be good for us in the 'feel good' sense, but bad for us in a moral sense? What does it say about us, what kind of people are we becoming as we develop increasingly intimate relationships with machines?

For Winnicott objects such as a teddy bear or rag doll are mediators between the child's earliest bonds with its mother, who the infant experiences as inseparable from the self, and the child's growing capacity to develop relationships with other people who will be experienced as separate beings. The infant knows transitional objects as almost inseparable parts of the self and, at the same time, as the first not-me possessions. As the child grows, the objects are left behind. The lasting effects of early encounters with them, however, are manifest in the experience of a highly-charged intermediate space between the self and certain objects in later life.

In the past, the power of objects to play this transitional role has been tied to the ways in which they enabled the child to

project meanings onto them. The doll or the teddy bear didn't change, didn't do anything. Relational artifacts are decidedly more active. With them, children's expectations that their dolls want to be hugged, dressed, or lulled to sleep don't only come from the child's projection of fantasy or desire onto inert playthings, but from such things as the digital dolls' crying inconsolably or even saying: 'Hug me!' or 'It's time for me to get dressed for school!' In the move from traditional transitional objects to contemporary relational artifacts, projection gives way to engagement.

The psychoanalyst Heinz Kohut talks about the way some people shore up a fragile sense of self by turning another person into a 'self object'. In her role of self object, the other is experienced as part of the self, a kind of spare part, in perfect tune with the fragile individual's inner state. Disappointments inevitably follow. Relational artifacts (not as they exist now but as their designers promise they soon will be) clearly present themselves as candidates for this kind of role. If they can give the appearance of aliveness and yet not disappoint, they may have an advantage over human beings, and open new possibilities for narcissistic experience with machines. From this point of view, relational artifacts make a certain amount of sense as successors to the always more resistant human material.

In *Computer Power and Human Reason*, Joseph Weizenbaum wrote about his experiences with his invention, ELIZA, a computer program that seemed to serve as self object as it engaged people in a dialogue similar to that of a Rogerian psychotherapist. It mirrored one's thoughts; it was always supportive. To the comment: 'My mother is making me angry,' the program might respond: 'Tell me more about your mother,' or 'Why do you feel so negatively about your mother.' Weizenbaum was disturbed that his students, knowing quite well that they were talking with a computer program, wanted to chat with it, indeed, wanted to be alone with it. Weizenbaum was my colleague at MIT at the time; we taught courses together on computers and society. And when his book came out, I wanted to reassure him. ELIZA seemed to me like a Rorschach test through which people expressed themselves. They became involved with ELIZA, but in a spirit of 'as if'. They thought: 'I will talk to this program "as if" it were a person; I will vent, I will rage, I will get things off my chest.' The gap between person and program was vast, and at the time ELIZA seemed to me no more threatening than an interactive diary. Thirty years later, I wonder if I underestimated the quality of the connection. A newer technology has created computational creatures that evoke a sense of mutual relating. The people who meet relational artifacts feel a desire to take care of them. And with that comes the fantasy of reciprocation. They want the creatures to care about them in return. Very little about these relationships seemed to be experienced 'as if'. The story of computers and their evocation of life had come to a new place.

We already know that the 'intimate machines' of the computer culture have shift-

ed how children talk about what is and isn't alive. For example, children use different categories to talk about the aliveness of 'traditional' objects from the language they use when confronted with computational games and toys. A traditional wind-up toy was considered 'not alive' when children realised that it did not move of its own accord. Here, the criterion for aliveness was autonomous motion. Faced with computational media, children's way of talking about aliveness changed. From the late 1970s and the days of the electronic toys Merlin, Simon, and Speak and Spell, children classified computational objects as alive if they could think on their own. Faced with a computer toy that could play noughts and crosses, what counted to a child was not the object's physical but its psychological autonomy.

Children of the early 1980s came to define what made people special in opposition to computers, which they saw as our 'nearest neighbours'. Computers, the children reasoned, are rational machines; people are special because they are emotional – 'emotional machines'. In 1984, when I completed my study of a first generation of children who grew up with electronic toys and games, I thought that children might take the intelligence of artifacts for granted, understand how it was created, and be less inclined to give it importance. I did not anticipate how quickly robotic creatures that presented themselves as having both feelings and needs would enter mainstream American culture. By the mid-1990s, as emotional machines, people were not alone.

Traditionally, artificial intelligence concentrated on building engineering systems that impressed by their rationality and cognitive competence – whether in playing chess or giving 'expert' advice. Relational artifacts by contrast are designed to impress not so much through their 'smarts' as through their sociability. Tamagotchis, virtual creatures that live on tiny LCD screen, itself housed in a small plastic egg, were the first relational artifacts to enter the American marketplace. A toy fad of the 1997 holiday season, they were presented as creatures from another planet that needs both physical and emotional nurturing. An individual Tamagotchi grows from child to healthy adult if it is cleaned when dirty, nursed when sick, and fed when hungry. It communicates its needs through a screen display. If its needs are not met, it dies. Owners of Tamagotchis became responsible parents; they enjoyed watching their Tamagotchis thrive and did not want them to die. During school hours, parents were enlisted to take care of the creatures; beeping Tamagotchis became background noise during business meetings. Although primitive as relational artifacts, the Tamagotchis illustrated a consistent element of a new human/machine psychology; when it comes to bonding with computers, nurturing is the 'killer app'. When people are asked to care for a computational creature, they become attached, feel connection, and as we have seen with the old woman and her Paro, sometimes they feel much more.

The Tamagotchis asked for attention in order to thrive. And when these first virtu-

al creatures prospered under their care, what children said about aliveness shifted: they no longer discussed a robot's 'aliveness' in terms of its motion or cognitive abilities. They came to describe robots as alive or 'sort of alive' not because of what the robots could do (physically or cognitively) but because of their own emotional connection to the robots and their fantasies about how the robots might be feeling about them. The focus of discussion about whether robots might be alive moved from robotic competence to robot connection.

A five-year-old talks about her Furby, a robotic creature that resembles an owl and appears to learn English under the child's tutelage, as alive 'because it might want to hug me.' A six-year-old declares his Furby 'more alive than a Tamagotchi because it likes to sleep with me'. A nine-year-old is convinced that her Furby is alive because she 'likes to take care of it'. She immediately amends her comment to acknowledge her new pet's limitations. 'It's a Furby kind of alive, not an animal kind of alive.' Children 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 'robot kind of love'?

In the early 1980s, I met a 13-year-old, Deborah, who responded to the experience of computer programming by speaking about the pleasures of putting 'a piece of your mind into the computer's mind and coming to see yourself differently'. Twenty years later, 11-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.' The contrast between the two responses reveals a shift from projection onto an object to engagement with a subject.

In the presence of relational artifacts, people feel attachment and loss; they want to reminisce and feel loved. In a year-long study of human-robot bonding, a 74-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.' In my study of robots in Massachusetts nursing homes, 74-year-old Jonathan responds to his robot baby 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 criticise me.' Andy, also 74, says that the My Real Baby robotic infant doll, which like Paro responds to caretaking by developing 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's – 'over there, it makes me feel so nice, like somebody is watching over me.'

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 and what they could and couldn't do. Current debates about relational and sociable machines – debates that will have an increasingly high profile in mainstream culture – are not about the machines' capabilities but about our vulnerabilities. When we are asked to care for an object, when that object thrives and offers us its attention and concern, we feel a new level of connection to it. The new questions have to do with what relational artifacts evoke in their users.

Science fiction has long presented robots as objects-to-think-with when we are considering who we are as people. The classic

In Philip K. Dick's classic story, 'Do Androids Dream of Electric Sheep', the story that became the movie *Blade Runner*, androids begin to act like people when they learn that they have a predetermined lifespan, and in the case of one android, Rachel, when they are programmed with memories of a childhood. Mortality and a sense of a life cycle are offered as the qualities that make the robots more than machines. *Blade Runner*'s hero, Decker, is able to tell humans from robots on the basis of their reactions to emotionally-charged images. The rotting carcass of a dead animal should cause no reaction in an android, but should repel a human, causing a change in pupil dilation. What does it take, asks the film, for a simulation to become indistinguishable from the reality? Decker, as the film progresses, falls in love with the near-perfect simulation, the android Rachel. Implanted with memories of a hu-

man past and a belief that she will inevitably die make her 'human'. By the end of the film we wonder whether Decker himself may be an android, unaware of his status. Unable as viewers to resolve this question, we are left cheering for our hero and heroine as they escape to whatever time they have remaining – in other words, to the human condition. And perhaps by the time we have to face the reality of computational devices passing the Turing test – i.e. becoming indistinguishable through their behaviour from human beings – we will no longer care about the test at all. By that point, people will love their machines and be more concerned about their machines' happiness than their test scores. This conviction is the theme of 'Supertoys Last All Summer Long', the short story by Brian Aldiss that was made into the Steven Spielberg film, *A.I.: Artificial Intelligence*.

In *A.I.*, scientists build a humanoid robot, David, who is programmed to love. David expresses his love to a woman, Monica, who has adopted him as her child. Current experience suggests that the pressing issue raised by the film is not the potential reality of a robot who 'loves', but the feelings of the adoptive mother: a human being whose response to a machine that asks to be looked after is a desire to look after it and whose response to a non-biological creature who reaches out to her is to feel attachment and horror, love and confusion. Even today we are faced with relational artifacts that elicit human responses that are in some ways not unlike those of the mother in *A.I.* Decisions about the role of robots in the lives of children

and old people cannot turn simply on whether children and the elderly 'like' the robots. We need to think about the kinds of relationships it is appropriate to have with machines.

My work in robotics laboratories has offered some images of how future relationships with machines may look, appropriate or not. Cynthia Breazeal was the leader on the design team for Kismet, the robotic head that was designed to interact with humans 'sociably,' much as a two-year-old child would. Breazeal was its chief programmer, tutor and companion. Kismet needed Breazeal to become as 'intelligent' as it did and then Kismet became a creature Breazeal and others could interact with. Breazeal experienced what might be called a maternal connection to Kismet; she certainly describes a sense of connection with it as more than a connection with a 'mere' machine. When she graduated from MIT and left the AI Laboratory where she had done her doctoral research, the tradition of academic property rights required that Kismet be left behind in the laboratory that had paid for its development. What she left behind was the robot 'head' and its attendant software. Breazeal described a sharp sense of loss. Building a new Kismet would not be the same.

In the summer of 2001, I studied children interacting with robots, including Kismet, at the MIT AI Laboratory. It was the last time Breazeal would have access to Kismet. It's not surprising that separation from Kismet was not easy for Breazeal, but more striking was how hard it was for the rest of us to imagine Kismet without her.

'But Cynthia is Kismet's mother,' one ten-year-old objected.

Comparing Breazeal's situation to that of Monica, the mother in A.I. might seem facile, but in her separation from Kismet, Breazeal is one of the first people to have had the experiences described in A.I. In a very limited sense, Breazeal 'brought up' Kismet. But even this very limited experience evoked strong emotions. My experiences watching people connect with relational artifacts, from primitive Tamagotchis to the sophisticated Kismet and Paro, suggests that being asked to look after a machine that presents itself as a young creature of any kind, constructs us as dedicated cyber-caretakers. When psychoanalysts talk about object relations, the objects they have in mind are usually people. The new objects of our lives (sociable, 'affective', relational) demand an object relations psychology that will help us navigate our relationships with material culture in its new, animated manifestations.

How will interacting with relational artifacts affect people's way of thinking about what, if anything, makes people special? The sight of children and the elderly exchanging tendernesses with robotic pets brings science fiction into everyday life and techno-philosophy down to earth. The question here is not whether children will love their robotic pets more than their real life pets or even their parents, but rather, what will 'loving' come to mean?

One woman's comment on AIBO, Sony's household entertainment robot is startling in what it suggests about 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.' Relationships with computational creatures may be deeply compelling, perhaps educational. But they don't teach us what we need to know about empathy, ambivalence, and life lived in shades of grey. To say all of this about our love of our robots does not diminish their interest or importance. It only puts them in their place.

Sherry Turkle