We know that technology changes our lives –
but could it be changing our selves as well?

Technology and Human Vulnerability

A Conversation with MIT's Sherry Turkle

F OR MOST of the last 50 years, technology knew its place. We all spent a lot of time with technology – we drove to work, flew on airplanes, used telephones and computers, and cooked with microwaves. But even five years ago, technology seemed external, a servant. These days, what's so striking is not only technology's ubiquity but also its intimacy.

On the Internet, people create imaginary identities in virtual worlds and spend hours playing out parallel lives. Children bond with artificial pets that ask for their care and affection. A new generation contemplates a life of wearable computing, finding it natural to think of their eyeglasses as screen monitors, their bodies as elements of cyborg selves. Filmmakers reflect our anxieties about these developments, present and imminent. In Wim Wenders's *Until the End of the World,* human beings become addicted to a technology that shows video images of their dreams. In *The Matrix,* the Wachowski brothers paint a future in which people are plugged into a virtual reality game. In Steven Spielberg's *AI: Artificial Intelligence,* a woman struggles with her feelings for David, a robot child who has been programmed to love her.

Today, we are not yet faced with humanoid robots that demand our affection or with parallel universes as developed as the Matrix. Yet we're increasingly preoccupied with the virtual realities we now experience. People in chat rooms blur the boundaries between their online and off-line lives, and there is every indication that the future will include robots that seem to express feelings and moods. What will it mean to people when their primary daily companion is a robotic dog? Or to a hospital patient when her health care attendant is built in the form of a robot nurse? Both as consumers and as businesspeople, we need to take a closer look at the psychological effects of the technologies we're using today and of the innovations just around the corner.

Indeed, the smartest people in the field of technology are already doing just that. MIT and Caltech, providers of much of the intellectual capital for today's high-tech business, have been turning to research that examines what technology does *to* us as well as what it does for us. To probe these questions

further, HBR senior editor Diane L. Coutu met with Sherry Turkle, the Abby Rockefeller Mauzé Professor in the Program in Science, Technology, and Society at MIT. Turkle is widely considered one of the most distinguished scholars in the area of how technology influences human identity.

Few people are as well qualified as Turkle to understand what happens when mind meets machine. Trained as a sociologist and psychologist, she has spent more than 20 years closely observing how people interact with and relate to computers and other high-tech products. The author of two ground-breaking books on people's relationship to computers–*The Second Self: Computers and the Human Spirit* and *Life on the Screen: Identity in the Age of the Internet*–Turkle is currently working on the third book, with the working title *Intimate Machines,* in what she calls her "computational trilogy." At her home in Boston, she spoke with Coutu about the psychological dynamics between people and technology in an age when technology is increasingly redefining what it means to be human.

You're at the frontier of research being done on computers and their effects on society. What has changed in the past few decades?

To be in computing in 1980, you had to be a computer scientist. But if you're an architect now, you're in computing. Physicians are in computing. Business-people are certainly in computing. In a way, we're all in computing; that's just inevitable. And this means that the power of the computer–with its gifts of simulation and visualization–to change our habits of thought extends across the culture.

My most recent work reflects that transformation. I have turned my attention from computer scientists to builders, designers, physicians, executives, and to people, generally, in their everyday lives. Computer software changes how architects think about buildings, surgeons about bodies, and CEOs about businesses. It also changes how teachers think about teaching and how their

students think about learning. In all of these cases, the challenge is to deeply understand the personal effects of the technology in order to make it better serve our human purposes.

A good example of such a challenge is the way we use PowerPoint presentation software, which was originally designed for business applications but which has become one of the most popular pieces of educational software. In my own observations of PowerPoint

in the classroom, I'm left with many positive impressions. Just as it does in business settings, it helps some students organize their thoughts more effectively and serves as an excellent note-taking device. But as a thinking technology for elementary school children, it has limitations. It doesn't encourage students to begin a conversation–rather, it encourages them to make points. It is designed to confer authority on the presenter, but giving a third or a fourth grader that sense of presumed authority is often counterproductive. The PowerPoint aesthetic of bullet points does not easily encourage the give-and-take of ideas, some of them messy and unformed. The opportunity here is to acknowledge that PowerPoint, like so many other computational technologies, is not just a tool but an evocative object that affects our habits of mind. We need to meet the challenge of using computers to develop the kinds of mind tools that will support the most appropriate and stimulating conversations possible in elementary and middle schools. But the simple importation of a technology perfectly designed for the sociology of the boardroom does not meet that challenge.

If a technology as simple as Power-Point can raise such difficult questions, how are people going to cope with the really complex issues waiting for us down the road – questions that go far more to the heart of what we consider

our specific rights and responsibilities as human beings? Would we want, for example, to replace a human being with a robot nanny? A robot nanny would be more interactive and stimulating than television, the technology that today serves as a caretaker stand-in for many children. Indeed, the robot nanny might be more interactive and stimulating than many human beings. Yet the idea of a child bonding with a robot that presents itself as a companion seems chilling.

> *"The machine might say, 'Mary, you are very tense this morning. It is not good for the organization for you to be doing X right now. Why don't you try Y?'"*

We are ill prepared for the new psychological world we are creating. We make objects that are emotionally powerful; at the same time, we say things such as "technology is just a tool" that deny the power of our creations both on us as individuals and on our culture. At MIT, I began the Initiative on Technology and Self, in which we look into the ways technologies change our human identities. One of our ongoing activities, called the Evocative Objects seminar, looks at the emotional, cognitive, and philosophical power of the "objects of our lives." Speakers present objects, often technical ones, with significant personal meaning. We have looked at manual typewriters, programming languages, hand pumps, e-mail, bicycle gears, software that morphs digital images, personal digital assistants–always focusing on what these objects have meant in people's lives. What most of these objects have in common is that their designers saw them as "just tools" but their users experience them as carriers of meanings and ideas, even extensions of themselves.

The image of the nanny robot raises a question: Is such a robot capable of loving us?

Let me turn that question around. In Spielberg's *AI,* 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 discussions that followed the release of the film, emphasis usually fell on the question of whether such a robot could *really* be developed. Was this technically feasible? And if it were 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. The question is not what computers can do or what computers will be like in the future, but rather, what *we* will be like. What we need to ask is not whether robots will be able to love us but rather why we might love robots.

Some things are already clear. We create robots in our own image, we connect with them easily, and then we become vulnerable to the emotional power of that connection. When I studied children and robots that were programmed to make eye contact and mimic body movements, the children's responses were striking: When the robot made eye contact with the children, followed their gaze, and gestured toward them, they responded to the robot as if it were a sentient, and even caring, being. This was not surprising; evolution has clearly programmed us to respond to creatures that have these capabilities as though they were sentient. But it was more surprising that children responded in that way to very simple robots – like Furby, the little owl-like toy that learned to speak "Furbish" and to play simple games with children. So, for example, when I asked the question, "Do you think the Furby is alive?" children answered not in terms of what the Furby could do but in terms of how they felt about the Furby and how it might feel about them.

Interestingly, the so-called theory of object relations in psychoanalysis has always been about the relationships that people – or objects – have with one another. So it is somewhat ironic that I'm now trying to use the psychodynamic object-relations tradition to write about the relationships people have with objects in the everyday sense of the word. Social critic Christopher Lasch wrote that we live in a "culture of nar-cissism." The narcissist's classic problem involves loneliness and fear of intimacy. From that point of view, in the computer we have created a very powerful object, an object that offers the illusion of companionship without the demands of intimacy, an object that allows you to be a loner and yet never be alone. In this sense, computers add a new dimension to the power of the traditional teddy bear or security blanket.

So how exactly do the robot toys that you are describing differ from traditional toys?

Well, if a child plays with a Raggedy Ann or a Barbie doll or a toy soldier, the child can use the doll to work through whatever is on his or her mind. Some days, the child might need the toy soldier to fight a battle; other days, the child might need the doll to sit quietly and serve as a confidante. Some days, Barbie gets to attend a tea party; other days, she needs to be punished. But even the relatively simple artificial creatures of today, such as Hasbro's My Real Baby or Sony's dog robot AIBO, give the appearance of having minds of their own, agendas of their own. You might say that they seem to have their own lives, psychologies, and needs. Indeed, for this reason, some children tire easily of the robots – they simply are not flexible enough to accommodate childhood fantasies. These children prefer to play with hand puppets and will choose simple robots over complicated ones. It was common for children to remark that they missed their Tamagotchis [a virtual pet circa 1997 that needed to be cleaned, fed, amused, and disciplined in order to grow] because although their more up-to-date robot toys were "smarter," their Tamagotchis "needed" them more.

If we can relate to machines as psychological beings, do we have a moral responsibility to them?

When people program a computer that develops some intelligence or social competency, they tend to feel as though they've nurtured it. And so, they often feel that they owe it something – some

Sherry Turkle

loyalty, some respect. Even when roboticists admit that they have not succeeded in building a machine that has consciousness, they can still feel that they don't want their robot to be mistreated or tossed in the dustheap as though it were just a machine. Some owners of robots do not want them shut off unceremoniously, without a ritualized "good night." Indeed, when given the chance, people wanted to "bury" their "dead" Tamagotchi in on-line Tamagotchi graveyards. So once again, I want to turn your question around. Instead of trying to get a "right" answer to the question of our moral responsibility to machines, we need to establish the boundaries at which our machines begin to have those competencies that allow them to tug at our emotions.

In this respect, I found one woman's comment on AIBO, Sony's dog robot, especially striking in terms of 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." The possibilities of engaging emotionally with creatures that will not die, whose loss we will never need to face, presents dramatic questions. The sight of children and the elderly exchanging tenderness with robotic pets brings philosophy down to earth. In the end, the question is not whether children will come to love their toy robots more than their parents, but what will loving itself come to mean?

What sort of relational technologies might a manager turn to?
We've already developed machines that can assess a person's emotional state. So for example, a machine could measure a corporate vice president's galvanic skin response, temperature, and degree of pupil dilation precisely and noninvasively. And then it might say, "Mary, you are very tense this morning. It is not good for the organization for you to be doing X right now. Why don't you try Y?" This is the kind of thing that we are going to see in the business world because machines are so good at measuring certain kinds of emotional states. Many people try to hide their emotions from other people, but machines can't be easily fooled by human dissembling.

So could machines take over specific managerial functions? For example, might it be better to be fired by a robot?
Well, we need to draw lines between different kinds of functions, and they won't be straight lines. We need to know what business functions can be better served by a machine. There are aspects of training that machines excel at–for example, providing information – but there are aspects of mentoring that are about encouragement and creating a relationship, so you might want to have another person in that role. Again, we learn about ourselves by thinking about where machines seem to fit and where they don't. Most people would not want a machine to notify them of a death; there is a universal sense that such a

moment is a sacred space that needs to be shared with another person who understands its meaning. Similarly, some people would argue that having a machine fire someone would show lack of respect. But others would argue that it might let the worker who is being fired save face.

Related to that, it's interesting to remember that in the mid-1960s computer scientist Joseph Weizenbaum wrote the ELIZA program, which was "taught" to speak English and "make conversation" by playing the role of a therapist. The computer's technique was mainly to mirror what its clients said to it. Thus, if the patient said, "I am having problems with my girlfriend," the computer program might respond, "I understand that you are having problems with your girlfriend." Weizenbaum's students and colleagues knew and understood the program's limitations, and yet many of these very sophisticated users related to

ELIZA as though it were a person. With full knowledge that the program could not empathize with them, they confided in it and wanted to be alone with it. ELIZA was not a sophisticated program, but people's experiences with it foreshadowed something important. Although computer programs today are no more able to understand or empathize with human problems than they were 40 years ago, attitudes toward talking things over with a machine have gotten more and more positive. The idea of the nonjudgmental computer, a confidential "ear" and information resource, seems increasingly appealing. Indeed, if people are turning toward robots to take roles that were once the sole domain of people, I think it is fair to read this as a criticism of our society. So when I ask people why they like robot therapists, I find it's because they see human ones as pill pushers or potentially abusive. When I've found sympathy for the

"Are you really you if you have a baboon's heart inside, had your face resculpted by Brazil's finest plastic surgeons, and are taking Zoloft to give you a competitive edge at work?"

idea of computer judges, it is usually because people fear that human judges are biased along lines of gender, race, or class. Clearly, it will be awhile before people say they prefer to be given job counseling or to be fired by a robot, but it's not a hard stretch for the imagination.

The story of people wanting to spend time with ELIZA brings me to what some have termed "computer addiction." Is it unhealthy for people to spend too much time with a computer?
Usually, the fear of addiction comes up in terms of the Internet. In my own studies of Internet social experience, I have found that the people who make the most of their "lives on the screen" are those who approach on-line life in a spirit of self-reflection. They look at what they are doing with their virtual selves and ask what these actions say about their desires, perhaps unmet, as well as their need for social connection,

perhaps unfilled. If we stigmatize the medium as "addictive" (and try to strictly control it as if it were a drug), we will not learn how to more widely nurture this discipline of self-reflection. The computer can in fact serve as a kind of mirror. A 13-year-old boy once said to me that when you are with a computer, "you take a little piece of your mind and put it into the computer's mind…and you start to see yourself differently." This sense of the computer as second self is magnified in cyberspace.

For some people, cyberspace is a place to act out unresolved conflicts, to play and replay personal difficulties on a new and exotic stage. For others, it provides an opportunity to work through significant problems, to use the new materials of "cybersociality" to reach for new resolutions. These more positive identity effects follow from the fact that for some, cyberspace provides what psychologist Erik Erikson would have called a "psychosocial moratorium," a central element in how Erikson thought about identity development in adolescence. Today, the idea of the college years as a consequence-free time-out seems of another era. But if our culture no longer offers an adolescent time-out, virtual communities often do. It is part of what makes them seem so attractive. Time in cyberspace reworks the notion of the moratorium because it may now exist on an always-available window.

A parent whose child is on heroin needs to get the child off the drug. A parent whose child spends a great deal of time on the Internet needs, first and foremost, to be curious about what the child is doing there. Does the child's life on the screen point to things that might be missing in the rest of his or her life? When contemplating a person's computer habits, it is more constructive to think of the Internet as a Rorschach than as a narcotic. In on-line life, people are engaged in identity play, but it is very serious identity play.

Isn't there a risk that we'll start to confuse simulation with reality?
Yes, there certainly is. When my daughter was seven years old, I took her on a

vacation in Italy. We took a boat ride in the postcard-blue Mediterranean. 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 to a research scientist at Walt Disney, he responded by describing the reaction of visitors to Animal Kingdom, Disney's newest theme park in Orlando, populated by "real," that is, biological, animals. He told me that the first visitors to the park expressed disappointment that the biological animals were not

ple to take care of them in order to function well and thrive, they present themselves *as if* they had emotions. As a consequence, for many people I interview, feelings begin to seem less special, less specifically human. I am hearing people begin to describe humans and robots as though they somehow shared emotional lives.

If emotions are not what set us apart from machines, then people search for what does, and they come up with the biological. What makes human beings

gramming?" In terms of our identities, we're getting squeezed in every direction as new technologies provoke us to rethink what it means to be authentically human.

A recent *New Yorker* cartoon summed up these recent anxieties: Two grown-ups face a child in a wall of solidarity, explaining, "We're neither software nor hardware. We're your parents." This cartoon reminds me of a statement someone I interviewed once made about simulation and authenticity: "Simulated thinking can be thinking, but simulated feeling can never be feeling. Simulated love is never love." The more we manipulate ourselves and the more our artifacts seek pride of place beside us as social and psychological equals, the more we find the issue of authenticity confronting us. Authenticity is becoming to us what sex was to the Victorians – an object of threat and obsession, of taboo and fascination.

> *"One woman, a successful journalist, described the experience of losing the contents of her PDA: 'When my Palm crashed, it was like a death. More than I could handle. I had lost my mind.'"*

realistic enough. They did not exhibit the lifelike behavior of the more active robotic animals at Disney World, only a few miles away. What is the gold standard here? For me, this story is a cautionary tale. It means that in some way the essence of a crocodile has become not an actual living crocodile but its simulation. In business, one is tempted to sell the simulation if that is what people have come to expect. But how far should you go in selling the simulation by marketing it as authentic?

You've said that computers change the way we think about ourselves. How so?
People tend to define what is special about being human by comparing themselves to their "nearest neighbors," so when our nearest neighbors were pets, people were special because of their intellects. When computers were primitive machines and began to be analogized to people, people were superior because of their superior intellects. As the computers became smarter, the emphasis shifted to the soul and the spirit in the human machine. When Gary Kasparov lost his match against IBM's chess computer, "Deep Blue," he declared that at least he had feelings about losing. In other words, people were declared unique because they were authentically emotional. But when robot cats and dogs present themselves as needing peo-

special in this new environment is the fact that we are biological beings rather than mechanical ones. In the language of children, the robot is smart and can be a friend but doesn't have "a real heart or blood." An adult confronting an "affective" computer program designed to function as a psychotherapist says, "Why would I want to talk about sibling rivalry to something that was never born?" It would be too simple to say that our feelings are devalued; it would be closer to the mark to say that they no longer seem equal to the task of putting enough distance between ourselves and the robots we have created in our image. Our bodies, our sexuality, our sensuality do a better job.

Of course, defining people in biological terms creates its own problems. For one thing, we are already blurring the distinction between people and machines by making machines out of biological materials and using machine parts within the human body. And we are treating our bodies as things–in our investigations of our genetic code, in the way we implant pumps and defibrillators in our flesh, in our digitizing of our bodies for education, research, and therapeutic purposes. Additionally, a psychopharmacologist might well say, "Excuse me, sir, but have you noticed that you are taking ten psychotropic medications to alter your mental pro-

Could you expand on that?
In many intellectual circles, notions of traditional, unitary identity have long been exiled as passé – identity is fluid and multiple. In a way, the experience of the Internet with its multiple windows and multiple identities brings that philosophy down to earth. But human beings are complex, and with fluidity comes a search for what seems solid. Our experiences with today's technologies pose questions about authenticity in new, urgent ways. Are you really you if you have a baboon's heart inside, had your face resculpted by Brazil's finest plastic surgeons, and are taking Zoloft to give you a competitive edge at work? Clearly, identity comes to be seen as malleable when the distinction between the real and the artificial fades. Personally, I find it amazing how in less than one generation people have gotten used to the idea of giving their children Ritalin – not because the children are hyperactive but because it will enhance their performance in school. Who are you, anyway – your unmedicated self or your Ritalin self? For a lot of peo-

ple, it has become unproblematic that their self is their self with Ritalin or their self with the addition of a Web connection as an extension of mind. As one student with a wearable computer with a 24-hour Internet connection put it, "I become my computer. It's not just that I remember people or know more. I feel invincible, sociable, better prepared. I am naked without it. With it, I'm a better person."

In our culture, technology has moved from being a tool to a prosthetic to becoming part of our cyborg selves. And as a culture, we've become more comfortable with these closer bonds through our increasingly intimate connections with the technologies that we have allowed onto and into our person. For most people, it hasn't been through technologies as exotic as a wearable computer. It's been through technologies as banal as a Palm Pilot (which, of course, when you think about it, is a wearable computer). In the Evocative Objects seminar at the Initiative on Technology and Self, one woman, a successful journalist, described the experience of losing the contents of her PDA: "When my Palm crashed, it was like a death. More than I could handle. I had lost my mind." Such objects are intimate machines because we experience them as extensions of self.

Do you think that kind of dependence is dangerous?

Not necessarily. Nursing homes in Japan increasingly make use of robots that give elders their medicine, take their blood pressure, and serve as companions. The Japanese are committed to this form of care for their elders; some say that they see it as more respectful than bringing in foreigners from different cultural backgrounds. When I first heard about this trend toward the use of robotics for elder care, I felt troubled. I feared that in our country there might be a danger that the widespread use of robotics would be used to legitimate social policy that does not make elder care a priority and does not set aside the resources, both in time and money, to have people there for the elderly. How-

ever, I have been doing fieldwork with robots for the elderly in local nursing homes. My project is to introduce simple robotic creatures – for example, robotic dogs and robotic baby dolls – in nursing homes and see what kinds of relationships the elderly form with these robots. Of course, when you look at particular institutions, families, and individuals, the question of the humane use of robotics for elder care is in fact quite complex.

At one nursing home, for example, the nursing staff has just gone out and bought five robot baby dolls with their own funds. The nurses are not doing this so that each elderly person can go to his or her room with a robot baby. They are doing this because it gives the elders something to talk about and share together, a community use of the robots that was totally unexpected when I began the project and which is quite promising.

One goal of my work is to help designers, businesspeople, and consumers keep human purposes in mind as they design and deploy technology and then choose how to make it part of daily life. For me, authenticity in relationships is a human purpose. So, from that point of view, the fact that our parents and grandparents might say "I love you" to a robot, who will say "I love you" in return, does not feel completely comfortable to me and raises, as I have said, questions about what kind of authenticity we require of our technology. We should not have robots saying things that they could not possibly "mean." Robots do not love. They might, by giving timely reminders to take medication or call a nurse, show a kind of caretaking that is appropriate to what they are, but it's not quite as simple as that. Elders come to love the robots that care for them, and it may be too frustrating if the robot does not say the words "I love you" back to the older person, just as I can already see that it is extremely frustrating if the robot is not programmed to say the elderly person's name. These are the kinds of things we need to investigate, with the goal of having the robots serve our human purposes.

How can we make sure that happens?

It's my hope that as we become more sophisticated consumers of computational technology – and realize how much it is changing the way we see our world and the quality of our relationships – we will become more discerning producers and consumers. We need to fully discuss human purposes and our options in technical design before a technology becomes widely available and standardized. Let me give you an example. Many hospitals have robots that help health care workers lift patients. The robots can be used to help turn paralyzed or weak patients over in bed, to clean them, bathe them, or prevent bedsores. Basically, they're like an exoskeleton with hydraulic arms that are directly controlled by the human's lifting movements.

Now, there are two ways of looking at this technology. It can be designed, built, and marketed in ways that emphasize its identity as a mechanical "flipper." With this approach, it will tend to be seen as yet another sterile, dehumanizing machine in an increasingly cold health care environment. Alternatively, we can step back and imagine this machine as a technological extension of the body of one human being trying to care for another. Seen in the first light, one might argue that the robot exoskeleton comes between human beings, that it eliminates human contact. Seen in the second light, this machine can be designed, built, and marketed in ways that emphasize its role as an extension of a person in a loving role.

During one seminar at the Initiative on Technology and Self in which we were discussing this robotic technology, a woman whose mother had just died spoke about how much she would have loved to have had robot arms such as these to help her lift her mother when she was ill. Relatively small changes in how we imagine our technologies can have very large consequences on our experiences with them. Are the robot arms industrial "flippers" or extensions of a daughter's touch?