Consciousness Revisited

In the following pages we return to a topic first addressed in FREE INQUIRY *in its Fall 1994 issue ("New Conceptions of the Mind"). Some of the foremost researchers in the field discuss what is known about the nature of consciousness and the direction further examinations should take.—*EDS.

FI Interview

A Conversation with Daniel C. Dennett

Daniel C. Dennett is Director of the Center for Cognitive Studies at Tufts University, and author of such highly praised books as Consciousness Explained, Elbow Room, *and* Brainstorms. *His latest book is* Darwin's Dangerous Idea: Evolution and the Meanings of Life *(Simon & Schuster, 1995). The following interview was conducted by* FREE INQUIRY *editors Tom Flynn and Tim Madigan.—*EDS.

FREE INQUIRY: Synthesizing the work of many thinkers, yourself included, Adam Carley wrote a recent article in FREE INQUIRY (Fall 1994) in which he described "consciousness" as an illusion arising from the operation of short-term memory. In his view, consciousness is an evolutionarily useful fiction that enables the intelligent brain to reflect on its own decision-making processes and learn from its experiences more effectively, but it is altogether fictional in any naïve sense of reality. What do you think of that model, and how would you go about testing it?

DANIEL DENNETT: I think that's partly right, but it's very misleading in some regards. The way I would put it is that the brain needs what's called a "user illusion" the same as when we use a word processor. You can use a word processor without knowing what's actually going on inside the computer. You need to have some useful metaphors though, and that's the user illusion. But notice, it's a benign and extremely useful illusion.

Well, the brain works similarly—the brain of a human being, not of any other species. It is equipped with a very powerful user illusion. It sounds like a trick with mirrors, I know, but I'm saying that the brain is both the user and the provider of the user illusion. There are various agencies in the brain that require information from other agencies within the brain, and

Photo by Susan Dennett

this is provided in limited useful form by the way the brain is organized. This gives rise to the illusory sense that there is this one place—which I call "the Cartesian theater"—where it all comes together: the subject, the ego, the "I." There's no denying that that's the way it *seems*. But that is *just* the way it seems. That is the benign illusion of consciousness.

FI: If we dismantle the Cartesian theater, if we abandon the idea of an "I" who makes our decisions, what are the implications for moral philosophy? Is there any point in talking about moral responsibility if the moral agent cannot be conceived as a unitary entity?

DENNETT: The moral agent *can* be conceived as a unitary entity, but not as a sort of organ of the brain, or a particular subsection of the brain. It is, if you like, an artifact of the organization of the brain. The unity is imperfect and changing and problematic, and it is something that is *achieved*—it's not something one is born with. The idea that one has, for instance, an immortal soul, which is the locus of responsibility—we replace that idea with one of a developing, virtual center of control, what I call "the center of narrative gravity." Once it develops, and once it is trained by enculturation and learning, then the body that is thus organized has a self, is a person, and is a locus of responsibility.

But we are all familiar with the ways in which people move the boundaries of the self. There is nothing more familiar than to hear somebody say, after having uttered something terrible, "That wasn't me speaking, no, no, I disown the recent motion of my own body. The real me would never have said anything like that."

FI: "It was the booze talking."

DENNETT: Yes. "The devil made me do it." Or "That wasn't me, that was just the darn urge that I have somewhere that sometimes takes control." This is a very natural way of talking and there's an element of truth in it. That is to say, a person is in fact a rather delicately poised committee—or not even something as well organized as a committee—sort of a crowd of sometimes cooperating, sometimes competing agencies, and one is just the sum total of them. And in fact, when

people say, "That wasn't me," we very often disallow that, and say "Now, take responsibility." As I put it in *Elbow Room,* if you make yourself really small you can externalize just about everything.

FI: If you do away with the idea of the soul as an immortal entity and a moral agent, what implications does this have for religion, in particular for Christianity? Do you see Christianity and other Western religions as being opposed to your ideas, in the way that many are opposed to the theory of evolution?

DENNETT: It continues to fascinate me how ambivalent Christians are about the relationship between morality and part of the Judeo-Christian tradition. Nietzsche was wonderfully scornful of the vision of Christianity that made morality depend upon pie in the sky; that is, be a good boy or girl now and in heaven you'll get your reward. On the face of it, this is an ignoble foundation for morality. It concedes the selfishness of the agent. It abandons the hope of an agent's conceiving of his or her acts as worthy in their own right, independent of any reward. The idea that a reward in heaven or a punishment in hell is a necessary foundation for morality is a deeply pessimistic, almost nihilistic, idea. And yet, of course, there it is, enshrined in many aspects of the Judeo-Christian and Islamic traditions. I think that most religious thinkers recognize this, and, if pressed, would disavow the importance of heaven and hell and even of immortality of the soul as a foundation of ethics. We can have ethics; we can have responsibility without pie in the sky and without the tortures of hell.

FI: At what point do you think your explanation of consciousness will be generally accepted? Do you think there may be some physical discovery that will lead to this, in the same way that evolution is now generally accepted?

DENNETT: Ah! If only that were true. The pockets of resistance to the theory of evolution are not just confined to fundamentalist religious groups and scientifically uneducated people. So widespread is the antagonism to or fear of Darwinian thinking in science that I decided I had to write a book about it, which has just come out, *Darwin's Dangerous Idea.* The point of that book is to show that Darwinian thinking cuts a lot deeper than a lot of apologists have been willing to admit, and the residual antagonism to the idea has to be exposed, analyzed, and finally dismantled.

I think eventually—I'm not sure how long it will take—people will calm down about this. It's been hundreds of years since Copernicus and Galileo overthrew the idea that the Earth is the center of the universe, and now every school child learns that without tears or terror. This is *not* a deeply troubling idea, it turns out, although there was a time when it caused a lot of anxiety. In due course, every child will learn about the four-billion-year history of evolution on this planet, and how from self-replicating macromolecules every living thing on this planet is descended, without divine intervention at any point. But the date when calm acceptance of the truth is widespread is still a ways off. I'm very actively advocating that the time has come to say to the creationists and the "intelligent design" theorists, "Come on, give us a break. This is a losing battle. You're only upsetting your children and prolonging the agony. Eventually, and perhaps quite soon, they're going to know that you tried to misinform them when they were children, and that's a really bad thing. *That's* an unforgivable sin."

I'm reminded of an occasion in Sweden, in 1967. On one shocking and wonderful weekend an amazing thing happened. Drivers were instructed to switch from driving on the left to driving on the right side of the road. The change was planned very well and carried out with scarcely a hitch. There were almost no accidents. It would have been completely foolish to have tried to do that gradually. Some things you want to do in one big gulp, with everybody holding hands. I think that's exactly what we have to do with evolution in this country. We have to say "Come on, everybody. Let's drop this silly business about creationism and accept that evolutionary biology is good science, on a par with everything in chemistry and physics. It explains so many things so well that it doesn't have any plausible rivals."

FI: It may be that taking away people's belief in a soul and immortality may strike much deeper existentially than Galileo's discovery. Perhaps most people can't accept this.

DENNETT: British opera, theater, and television director and polymath Jonathan Miller, the Chicago neuropsychologist Jerre Levy, and I were invited to a weekend retreat for ABC television executives. When we got there we found that our role was to be the heavies—to debate with believers in the paranormal and other sorts of West Coast holistier-than-thou mystics. We tore them to shreds. The discussion wasn't ferocious or embittered—we just made nice mincemeat of them. At the end of the workshop, Jonathan taught me a lesson. He asked the assembled group of several hundred people—highly educated people, movers and shakers in the very competitive world of television—how many of them, before the weekend, had thought there was something to the paranormal. Perhaps 20 to 30 percent put up their hands. "Now, after this weekend," he asked, "How many of you think there's something to the paranormal?" Maybe 50 or 60 percent of the hands went up. This was really disturbing to me. "Well, Dan," he said, "you've got to understand, the way these people are thinking is, if these three industrial-strength academics work this hard to combat it, there *must* be something to it." That's a depressing lesson.

I think that people are almost immune to rational persuasion about some of the things that they really want to believe. They are sometimes candid and self-reflective enough to say, "I don't care about your arguments, I don't care if the arguments on my side are invalid or question-begging. I like the conclusion."

FI: You are quite interested in the area of artificial intelligence (AI). Can you describe the Turing test?

DENNETT: Alan Turing, who is as deserving as anybody to be called the inventor of the computer, proposed this test back in 1949, in an article in the journal *Mind,* as a way of stopping a futile debate over the question, Can machines think? He said, "Let's set that question aside. I propose a test, and why don't we just ask, could a machine pass this test?" He deliberately chose a test that he thought was so difficult that nobody would quibble about whether anything that passed the test could think.

The test that he chose was in fact inspired by Descartes. Nobody knows

whether Turing read Descartes, but in *The Discourse on Method,* written in the mid-seventeenth century, Rene Descartes proposed essentially the same test. He said, "Oh, you could make a machine that could utter a few sentences. If you pushed it here it might say, 'That hurts, stop.'" But one couldn't make a machine, he said, that could hold a conversation, that could respond appropriately to the essentially infinite variety of topics that any normal human being can converse about.

Now, Turing said, let's set up a test where we put a human being behind one screen and a computer behind another screen. Have each contestant interrogated by a vigorous and skeptical judge. The judge's task is to figure out which one is the human being. If, more often than not, the machine can fool the human judge into declaring that it is the human being, then it is manifestly very intelligent. He didn't propose this test as a research strategy for AI, but just as a thought experiment that was supposed to stop a certain sort of discussion. But in fact, brilliant as the test was, it completely backfired in one of its aims, because untold gallons of ink have been spilled by philosophers arguing about the validity or probability of the Turing test. I think in fact it is quite clear that the test is extraordinarily difficult, if properly conceived, and indeed that Turing was right. There is just no serious possibility at all that something could pass that test and not be in every important sense a thinker.

People have fantasized about how one can pass the test with one cheap bucket of tricks or another. The most well known of these fantasies is by the philosopher Ned Block: There is a giant lookup table, in which all possible intelligent conversations in English are stored in alphabetical order. When the judge comes up and says, "Hi there, what's your name?" it goes to "H," and looks under "Hi there, what's your name?" and then it simply gives the canned response. Block says, quite plausibly, that this imaginary computer could pass the Turing test. There's no conversation that is unavailable to it, and manifestly it *wouldn't* be thinking.

The problem with this thought experiment is that this giant lookup table isn't just giant, it is so much larger than the universe that if we harnessed every electron in the universe we would still not be within a trillionth of 1 percent of the computer we would need to pass the test, and unless we violate the speed of light we aren't going to be able to answer these questions in real time. Of course, Turing didn't even bother to rule this possibility out, but if need be we can add a proviso—"It has to pass the Turing test, and it has to be smaller than the Pentagon, and we have to hold that the rules of physics are intact."

Although it's very hard to say what structural or internal requirements we really have as thinkers, at a minimum we suppose that the work gets done locally; that is, it's a response in real time that takes the current input and does some sort of an analysis of it, that determines its meaning, and then, on the basis of that meaning, generates a response. How many different sorts of systems could there be that might actually meet those results? Well, who knows? Those are very abstract specifications. Turing saw an elegant way of in effect operationalizing them.

FI: Adam Carley suggests that, in the future development of AI, the Turing test essentially has it backward. He says that the most productive way to go is not toward working to fool a human observer, but rather toward working to fool the machine, to create for the machine an illusion of self much like your description of what goes on in the human person.

DENNETT: In some ways I agree with that, but then the Turing test was never a realistic goal. That was never anybody's idea of how to constrain one's researches into AI. Regarding better methodologies or research strategies, indeed the better goal of AI is to build robots that need to develop a sense of self in order to get through the day. In fact, we're working on just such a program now at the Massachusetts Institute of Technology. Rodney Brooks and Lynn Andrea Stein are the directors of the COG Project, and I'm part of it. We're building the most humanoid robot that has ever been attempted by a very wide margin.

It is now not quite intact. It's adult human size, but it doesn't have legs, because if it walked it would have to have a huge umbilical cord to carry its power supply and its input/output to its brain, which is about the size of a telephone booth. So it is fixed on a pedestal, but then from there it is very human. It can bend and turn at the waist. It has two arms, two hands, and a head with 3 degrees of freedom of motion. It can track with its eyes and ears and react to sudden motions and sounds. It has pain sensors everywhere and touch sensitive skin on its fingertips.

Now, our humanoid robot is going to have an infancy. It won't grow in size, but it's going to start out like a human infant, with a lot of hardwired autonomic underlying capacities. It's going to have to learn hand-eye coordination, a sense of its own body, and its place in the world.

This is an enormously ambitious project, of which a tremendous amount has already been completed. At this time, its left arm and left hand are essentially complete, and its head and eyes and torso are intact and working. Its eyes have pretty good capacities to notice and track things. It has color, three-dimensional vision. It doesn't have any language capacity at birth but this is a dream of the future.

FI: Does it have a soul?

DENNETT: Not yet. But I must say, one of my favorite sociological predictions is already being borne out. Long before we have robots who have moral status, we're going to have onlookers who are fervently convinced that this is a suitable moral patient, whose rights have to be protected. And sure enough, both the graduate students working on this project and a lot of people who've learned about it have begun to be really quite concerned about what their moral responsibilities are to COG itself, not just to the COG team.

FI: People for the Ethical Treatment of Robots?

DENNETT: Absolutely. I don't know if a committee has been formed yet, but some of my students are in effect mobilizing to address the questions of just when COG develops the moral status of an animal. And I quite agree with their initiative. To give you an instance of this, when you walk by COG and its eyes flick over and lock on you and then track you across the room, there's an overpowering sense that you are being watched by another conscious agent. Now, those of us who are working on the project know that these are still basically blind eyes that have just a few autonomic reflexes, but already the sense of presence is very strong, and it's going to get stronger and stronger. •