

Daniel Dennett, Interview 29.10.2017

Your new book *From Bacteria to Bach and Back* is one of striking comprehensiveness and breadth, it arrives at a theory of mind and consciousness, it takes in the evolution of life from its earliest beginnings and considers the origin of life itself, it also offers a theory of culture and cultural evolution. Could you tell us what *to you* are the most important claims in the book.

The reason it has to be so vast in temporal scope, going back to the origins of life, is that, as the British biologist D'Arcy Thompson once said, "everything is the way it is because it got that way", and if you don't think about how consciousness got that way--human consciousness in particular—it is just going to be mysterious. For many people that is just fine. But not for me. So a naturalistic theory of consciousness has to answer the question: how on earth did this ever arise? It is clearly an exquisitely efficient and breathtakingly powerful cognitive phenomenon. And you don't get those for free. They have to be designed. They are either designed by human designers, or by God, or by evolution. The claim that I am advancing is that evolution is what does the job, since there is no God, but there are two quite distinct processes of natural selection. First there is genetic natural selection, but that only takes us so far. Genetic natural selection gives us a primate brain, with many bells and whistles special to human beings (and not to chimps or bonobos for instance). But that in itself is nowhere near enough. A second process which is not just analogous to natural selection, but *is* natural selection, has to occur in the cultural realm, and so far it has occurred only in one species: us. It is responsible for the design of all of the thinking tools that we install in our primate brains to give us human minds. As my former student and colleague Bo Dahlbom once said, "you can't do much carpentry with your bare hands and you can't do much thinking with your bare brain". We have to equip our brains with the thinking tools of culture, and those thinking tools were not designed by any person. They were designed by cultural evolution: words, language primarily. No one invented language—not one in a hundred thousand words is a coined word introduced by some intelligent designer of words. And yet words are brilliantly designed data-structures that furnish our minds with capacities that our brains would not otherwise have.

There are those who think that consciousness is a big deal. You don't—or, at least not in the sense that many others think it is. Can you please elaborate on this.

I think that consciousness is the last truly mind-bothering puzzle. It is not a mystery. Or, if it is, nobody has given us good reason to think it is. There are some pretty strong causes of why we should think that consciousness was so spectacularly wonderful as to be beyond explanation. One is that we have no personal access to the medium in which our perception and thought happens. We don't know any more about what is happening in our brains than we know what is happening in our spleen or our lungs. We have a little bit of indirect information. And yet it seems that we have "infinite" information about what is going on in our minds. We do, but only of the content, not of the vehicles, not of the media. The media are invisible to us. We can't have access to them at all. And, people are scared of a naturalistic view of consciousness. They are afraid it is going to invade the last sanctuary of privacy and turn the scientists into evil, manipulating mind-readers. Well, that is not an entirely unfounded fear. But denying that the mind is explicable in terms of the brain is not going to make that fear go away and is not going to deal with the problem. So, I think that anybody who wants to put forward a firmly-grounded naturalistic theory of consciousness is going to face a fairly strong head-wind of anxiety and distrust. But let's get on with it and along the way we can learn some fascinating things.

Many people have worried that you somehow have denied the existence of consciousness. Is it correct to say that in *From Bacteria to Bach and Back* you provide your most explicit response to that criticism?

I suppose I do. I think, actually, my denial is right there at the beginning of *Consciousness Explained* where I go to some length to say that I am not denying that these phenomena exist, I am just saying that they aren't what you think they are. Now, if you think that consciousness occurs in a separate medium, in some dualistic ectoplasm or other realm, then I am denying that consciousness exists. That is not what consciousness is. I love to quote the philosopher-magician Lee Siegel, who has this wonderful passage in his book on magic where he says that he is writing a book on magic and people ask him "real magic?", and he says "no, conjuring tricks not real magic" and he goes on to say "real magic refers to the magic that is not real, while the magic that is real, that can actually be done, is not real magic". Many people think consciousness is in those terms "real magic" and if you deny that it is, you are denying that consciousness exists, by their lights. It's hard to get them to abandon that position. But I just keep trying and keep coming up with new ways of sneaking by their defences.

We have spoken of magic and mystery. But the so-called mysterians point to our epistemic boundedness which they argue result from internal limitations of

humans as biological systems. You have forcefully rejected this view in relation to the problem of consciousness. But do you recognise such a thing as human epistemic boundedness at all? If so, where might the bounds lie?

Well, obviously we have epistemic boundedness in one rather trivial sense. We are not eternal and there will always be unanswered questions, and the heat-death of the universe will overtake us, if we don't kill ourselves first, as we probably will at some time in the future (but maybe a thousand or a million years from now). Is there boundedness in any other sense? Maybe. But we seem to be dealing with it today by inventing, improving, refining and extending what might be called "distributed understanding". We have scientific articles published by the people at CERN with a thousand authors. No single one of them understands everything in that article, but we can chalk that one up as another thing that "we"—as a species, *Homo sapiens sapiens*—understand. Not all of us, and nobody understands it all—but it is a solved problem. Well suppose, following that model, we had a ten-volume explanation of consciousness, right down to the motor-proteins in the neurons, explaining all the social implications and extensions, and variations. And this ten-volume set which, of course, was authored by a thousand people, has become accepted—there are still some little pockets of controversy about some parts, but there it sits on the shelf. Do we understand consciousness now? Is it still a mystery? I think the answer to that would be no. That would be a solved problem. In fact, if we had that set of volumes, people would look elsewhere for really interesting career choices because the only thing to do in that area would be mopping up little details which are not really essential. Now, would any one person understand it all? No. Now one of the sad ironies of this move towards distributed understanding is that no sooner do we finally improve the social structure of science by giving women free-rein to rise to the top in every field, then we more or less abandon the ideal of the master-intelligent scientist. Aside from Marie Curie, whose name is going to be carved on the wall in the library, along with Copernicus and Einstein? The age of towering, giant figures in science may be over. Group understanding in philosophy is harder to get your head around, just because philosophers have tended to be solitary authors. Joint authorship is coming, but it is still fairly rare. I think that it is easy enough to imagine a more distributed form of expertise in philosophy where a group could have a division of labour and they all share some common theory—say naturalism—and some work on society, and others on religion, and others on epistemology, and others on the mind and so forth. It is already happening, but not in an organised way.

So, you think that the mysterians are clinging to an outdated model of philosophical and scientific work, where an isolated genius would have a full capacity to understand the whole thing?

Yes, I think there is an unnoticed equivocation in saying we will never understand consciousness. If by “we”, we mean as individuals no one individual will understand consciousness, well no one individual can understand the internal combustion engine. Pretty close, but not perfectly. So that is trivial. But the idea that there is something like the speed of light or the sound barrier that turns us back because we just don’t have the cognitive energy to go any faster--No, it is an interesting idea, but there is no particular reason to believe it. In particular, I find it curious to find those inspired by Noam Chomsky taking this very seriously. On the one hand they celebrate the breathtaking combinatorial powers of human language. But what they must mean when they say that consciousness is beyond all explanation is that in the Library of Babel there is no twenty-volume set of books, written in grammatical sentences which are available to every human being who speaks that language that answers the question. Well, that might be true, but they haven’t given us any reason to believe it.

We have been talking about consciousness, but your work actually reveals how much intelligence and design happens without conscious direction, you talk about “competence without comprehension” in the natural world generally, but more particularly in the human mind. Do you think philosophers have traditionally exaggerated the role of consciousness in cognition?

Yes I do, because most have adopted one variation or another of a Cartesian model of the mind, where first there is all the input-output machinery, that is the eyes and the ears and the optic nerve and the brain itself, and then the cognition happens at the summit—the understanding. Where does the understanding happen? It happens late in the inbound path and just before the decision, action, free-will and so forth happens. Well, that vision, isolating all the understanding in this special “consciousness-place”, is just hopeless! And it is provably false. The task of comprehension, of sorting out, interpreting, identifying, re-identifying, generating hypotheses, this all occurs over time, distributed around in the brain, and we have no direct access to it. So basically the machinery of understanding at every level is unconscious. We have an illusion of a space, the Cartesian theatre, where this happens, and that is a benign illusion, but it isn’t where the action is.

Previously you have talked about the mind as a mass of interacting homunculi, but in your latest book *From Bacteria to Bach and Back* your theory seems to

have become more explicitly neurobiological. Would it be fair to say that the homunculi have turned out to be neurons or coalitions of neurons?

Yes. Well, that was always my view. As I used to put it: Here is how you do cognitive science. You take one person, one whole cognitive agent, and you break that person down into sub-agencies, which are themselves rather homuncular, but they do not do the whole job. They have diminished understanding and are sort of myopic. Teams of those accomplish the highest level. But if you look at each one of those homunculi, they too are made up of more homunculi—it is like nested *matrioshka* dolls—and at the base you have neurons which can be basically replaced by machines. And then I realised that neurons are also little agents, and are more agent-like than I used to think. So, you have to add them to the homunculi and go to their inner-workings to find the parts that are replaced by the machine. Once you get down to the motor-proteins, those *are* robots. Those are machines. But every one of those 86 billion neurons in your brain is an agent with an agenda, and it is not easily captured as a machine.

So the neuron is by no means the ultimate unit and the homunculi occur within the neuron?

As usual we get a gradualism. I just mentioned motor proteins. If you see the highly detailed animations of motor proteins that have been made, trudging along on their micro-tubule highways inside a cell, carrying goods and services inside an individual cell, you can appreciate that they are truly minimal agents. But they have jobs to do and are quite robotic. They are like the brooms of the sorcerer's apprentice. For some purposes it is useful to call them agents.

You describe neurons as being like agents, engaging in competitive and collaborative behaviour that is often reminiscent of human individuals. This invites the thought that human minds might interact to produce collective minds, just as neurons give rise to ours. Do think this kind of speculation is worth exploring?

It is worth exploring to see its limits. If you were to take a bowl of 86 billion live neurons and set it loose it would not form a mind. They would all die before anything remotely like a mind would be formed because the organisation, and its very specific details, is so important to the brain. Could human beings set out deliberately to design a structure of human beings that could mimic the social structure of the human brain? They could in principle. It would be a sad day because it would mean relegating people to slavery, and I don't think any of us really want that. Ned Block famously

had a thought experiment about enslaving the Chinese nation and having them simulate a brain, and he declared that it would not be conscious. I think that is just a failure of imagination. If it really did duplicate the structure, then there would be an individual completely distinct from all those Chinese slaves. Who knows what language it would speak, if any? And no one of the Chinese participants would have the faintest idea what it was doing, what its agenda was, what it was talking about. But of course, my neurons don't know what I am doing. It doesn't stop me from doing these things. So I think that Block's thought experiment is a classic case of a failure of imagination.

In your evolutionary story, you speak about the acquisition of language and how crucial it is in enabling the emergence of consciousness and reflective experience. We therefore no longer have the bare primate brain. But it is not always clear where this leaves animal sentience. Animals obviously lack the tools of linguistic expression so they won't have conscious self-monitoring, introspection, or the "user-illusion". But would it still not be reasonable to attribute to other mammals the conscious experience of physical pain on the basis of their behaviour?

Yes, it would. And they can have a user-illusion, it is just not as elaborate as ours, because it doesn't have so many thinking tools. In fact, it has almost none, aside from what they acquire in the course of their own individual experience. If you want to understand consciousness in any organism, or in any entity, you have to ask: And then what happens? What does consciousness enable the organism to do? And if we look at a wolf we see that its cognitive system enables it to do a lot of very cunning and insightful things. And it is the most natural thing in the world to imagine that it is running through its plans in its head, thinking ahead, reflecting on whether it did the wrong thing the last time, and so on. It is very hard to find any evidence for that, in spite of the adroitness of wolves. In an English cliché we say "well, it seemed like a good idea at the time", usually regarded as a rueful acknowledgement of stupidity. On the contrary, I say that any agent that can think to itself, whether it has language or not, "well, it seemed like a good idea at the time", has got some of the fundamental capacities of intelligence, and an animal that can reflect thus on its recent experience and use that reflection to shape its next action, is a very smart animal, and I would call that a kind of consciousness. But if you cannot do that, then, indeed, injuries will create pain-states which are distressing and reliable dissuaders from courses of action—you can train a dog by punishing it and rewarding it, so of course in that sense they have feelings. But they don't, so far as we can tell, have a capacity to reflect on their feelings, their urges, the way even small children can. That is a huge

difference. But still, it is not as if at some point you get an agent that can think it seemed a good idea at the time and *boom!* something wonderful happens and they have just kindled the great fire of consciousness. There is no magic point like that.

You said earlier that people are worried that scientific research into consciousness might endanger the precious sanctuary and intimate space that we have. In *Breaking the Spell*, your book on religion, this worry seems to be tackled in a different form. Here again the spell may be broken. Is there a deep connection here?

That is a good comparison. This ties in well with my theme about competence without comprehension too. I think we have elaborate alarm-systems and danger-systems in our brains and when we perceive something as a threat we don't have to understand why it is a threat, and we may be wrong that it is a threat. But it still has a potent effect upon subsequent action and decision. Sometimes this is a wonderful thing. You definitely want to be on your guard. You want to teach children "look out for that too-friendly stranger". You want them to perceive the threat and respond to it appropriately, and run and yell if necessary. That is a good thing, but like everything else good it can misfire and malfunction. I think that the hostility towards atheism is fuelled by that anxiety. There is lots of research by social psychologists and neuroscientists on the effects of fear, anxiety and expectation on people's thinking. When it gets to the theories, it often leads them to refuse to consider the relevant hypotheses. They just don't want their minds to go there for fear of what they may find. I found this again and again in my work on consciousness. People even admonish their readers: "don't let Dennett adopt the third-person point of view, you have got to stick to the first-person point of view!" That is, in a raw and naked form, this fear.

***Breaking the Spell* of 2006 made you part of the so-called "new atheism" movement, along with other famous figures—particularly Richard Dawkins, Christopher Hitchens, and Sam Harris. Can you explain your distinct perspective on atheism, in particular how it differed, or still differs, from the views of those other authors?**

I think the difference is one of emphasis. For the first time in my life I get to be the white sheep rather than the black sheep, I get to be good cop rather than bad cop. Of the Four Horsemen, I am the gentlest, the most sympathetic to religion, and I think that is partly due to personal histories. Hitch saw first-hand, and very dangerously, some of the most violent evils of religion, which I have read all about and I know about, but I never had to face them in a situation of jeopardy, and he did. Richard had

some extremely anxious times with his daughter, who was obliged by her mother to attend a Roman Catholic school at a very tender age. She is now fine, a chip off the old block, like her Dad. But he was deeply concerned with that, and that was one of the personal motivations for him. I have known many people who are mild, sane, helpful, just what Christians are supposed to be, and for whom religion has been a lifeboat, a bulwark, a source of community and love when nothing else was available. That's the role of religion that I will regret losing when religion withers away, as I think it is doing right now. I think we cannot expect the state to do the comforting, nurturing, the protecting, and to provide the asylum that some people just need. If we can have secular organisations, charities etc., that can pick up that role, that will be just wonderful. In fact, what may well happen is that religions will evolve into that role, and the creed will become less and less important and you will still have the allegiance and the team-spirit and community, outreach and organisation, but without the irrationality of superstition.

We have seen internal conflict in the New Atheism movement. So is the movement still going on and what are its accomplishments, if any?

Well, it was never an organised movement. The so-called Four Horsemen got together in Hitch's apartment in Washington DC to film that interview which certainly has had a big impact. But that is the only time that the four of us were in the same room at the same time. It wasn't all planned out, but was quite spur of the moment. It is true that we have all been in communication since (though of course Hitch has now died). We don't have any agenda or progress reports. But we all follow the latest polls and research (done by the Pew Institute and others), and we take heart from the fact that the fastest growing category in the world is *no religion at all*. I am inclined to think that the noise and amplified outcries today are the death cries of people who see the handwriting on the wall, not the triumphant cries of the new and rising. I am not worried about a take-over by religion, although I am afraid we are going to see a lot of suffering still and, who knows, a nuclear catastrophe or some religious nut releasing some deadly toxin or something. These are dangerous times. But I do not think that the hold of religion over people has a future because the transparency of the world doesn't leave room for it, although the fake news phenomenon is worrying. How successful it might be in isolating large groups of people and having them believe in an alternative reality--I am a little worried about that power.

In your book on religion, you spoke of some religious people, mostly perhaps sophisticated Westerners, who, though they reject the claims of traditional

religion, still profess a “belief in belief.” Could that perhaps be the future of religion, at least in the West?

There are unrecognised victims of that. Linda LaScola and I did a book about closeted atheist clergy, *Caught in the Pulpit*. (It has now been made into a play, and we are working on mounting a production in New York.) These folks are trapped. They are good people who went into religion as sincere believers. Ironically, some of them lost their faith when they decided that they should study the enemy—the books by Dawkins and me and others—read the books and were convinced by them, and now they are stuck in the pulpit, preaching week after week, and they don’t believe it. Their lives are extremely lonely. They are trying to do the right thing and they have made so many commitments that it is very hard for them. Some of them are still in the ministry. Out of that book came the organisation The Clergy Project, which is a website that has over 800 members. These are practising clergy who are atheists, or who are formerly practising clergy who are atheists. They are secret, they are confidential, but their ranks grow and so there are a lot of religious leaders who are already struggling with this hypocrisy that lies right at the heart of organised religion. I don’t know if there is a good solution to it. I have had some fantasies about possible developments. Many people love the ceremony, the ritual, the drama, and I love it too. It can make the hair stand up on the back of my neck.

Richard Dawkins says that he loves that too.

Hearing one of the choirs at Oxford at evensong and listening to the King James version being read in those Oxonian tones is thrilling. I have imagined setting up a theatre which week in and week out conducts absolutely authentic ceremonies—wedding ceremonies, funeral ceremonies, high masses, gospel meetings, Jewish weddings, you name it, and people come and sit politely and are moved or not. Maybe they ask the Baptist minister in town to come in and be their guest star this week. If something like that were managed, it would help to demystify and, of course, to desanctify the roles. Nobody accuses actors of hypocrisy. If people would just get in the habit of realising the priest or the preacher who is up there is basically an actor, then maybe it would be easier on those actors, on those who excel at that kind of work.

In Denmark and the Netherlands it is apparently fine for Lutheran pastors to come out and say that they do not believe in God, and for many people in their parish it is fine. So maybe that transition is happening.

That happens. Famously in the United States reform rabbis are most of them atheists. I certainly know a number of them and they are quite candid about this, and with their congregations too. And also, Episcopalians of one sort (there is a big split there) and people in the United Church of Christ, they downplay creed and make a big thing of ritual and allegiance. The Episcopalian ceremony, which is high on ritual and low on dogma, is sometimes called “smells and bells”.

In addition to mind/consciousness and religion, your third big topic has been free will, to which you have devoted two monographs—*Elbow Room* (1984) and *Freedom Evolves* (2003). In these books you have defended a compatibilist position. Yet in a 2013 interview you suggested that the concept of free will is too folksy and unclear to be useful in philosophy and science. Could this be interpreted as your confession of hard determinism, instead of compatibilism?

I can see how what I had said in that interview could be interpreted like that, but it would be a misunderstanding. I have not really changed my opinion on the issue of free will recently. My view is that free will has nothing to do with determinism. I wish people would finally start to see this. Instead, free will has everything to do with an engineering context. We should speak of different ways of being free. Take Cog, the robot at MIT who was stationary, but could move his head and his arms, to some extent. We could say that Cog, with its eyes, head, arms and fingers, had maybe 60 degrees of freedom. The ideal of freedom is having indefinitely many degrees of freedom. But nobody has indefinitely many degrees of freedom -- it's a fantasy. Is it even desirable? That is why I talk of the kind of free will worth having, and I argue that we do, indeed, have it. In particular, it's worth being free of manipulation without one's knowledge. The whole idea of determinism not being compatible with free will is a mistake that goes back to Democritus, and we should not make that mistake anymore.

You have commented on many occasions, orally and in print, on what you see as the ills of contemporary philosophy. On the one hand, the discipline has reached unseen levels of sophistication, but, on the other hand, it has been insulated from many of the vital issues that face humanity. What are some of the vital issues that philosophers should be tackling?

I think the one that concerns me the most and the one that I have not been able to find a crack to drive a wedge into, is the issue of the bounty of opportunities that we now have and our inability to prioritise them. If there is one thing that philosophers agree on (there aren't many), it is that “ought implies can”, that you are not obliged to

do something that is outside your powers. The thing is that our powers have grown exponentially in recent decades. Right now we can pull out our cell-phones and donate a \$100 to Oxfam or another cause of choice. We can adopt an orphan. We can provide monetary or technical support to a thousand different causes. We don't. We've got our own lives. We do some of that and we worry and are oppressed by the plethora of opportunities to do good and to contribute to the resistance of evil. Our ethical theory so far does not have any very good perspective on that phenomenon that I can see. That strikes me as a pressing moral concern, and I wish that people in ethics were thinking more practically about the implications and presuppositions of action in the world. I am also very interested in the development of a better theory of blame and punishment. It is not that we don't have free-will—we do have free-will in the ways that are worth wanting. I have written a lot about that. But the implications of that for how we should treat those that we convict of crimes is appalling. I am happy to say that my colleague Erin Kelly has a wonderful book coming out on this which I think is the skeleton of a major recasting of our perspective. Retributivism is, and should be, largely discarded, but not by replacing it with a bland, straightforward medicalisation of wrongdoing. That is not the way to go. I think a fundamental misunderstanding is that the only way that punishment can be justified is with a retributivist view. I don't think that is the case. I think there are indirect and compelling consequentialist justifications of punishment. I have sketched them out a bit and I would like to see that pursued much more vigorously because, at least in the United States, our penal system, our prison system, is obscene. It is a disgrace, and I would like to see philosophers unite in condemning the brutal and inhuman and immoral treatment to which we subject people in our prisons and jails.

You speak of a public role for philosophy. Does this not connect you to the American tradition of pragmatism, personified by John Dewey.

Yes, I consider myself to be in the American pragmatist tradition along with my mentor Quine. Absolutely.

Thank you very much.