

Consciousness Explained

D.Dennett

December 5 - 10, 2005

In the eighties the subject of consciousness, which until then had been more or less of a taboo, became an almost respectable and legitimate subject for scientific inquiry. Traditionally it was the domain exclusively that of philosophers, with Descartes naturally cast in the role of the founding father. The rise of experimental psychology in the 19th century expanded the field. W. James monumental work, the Principles of Psychology I-II, is essentially a philosophical contemplation of the nature of consciousness¹ refining the philosophical tradition of say Locke, Berkeley, Hume and Kant, mainly using introspection as the main probing tool (and thus placing the reader on the same level as the author, inviting him or her to check and confirm). James admittedly also included some experimental work of the quantitative kind, yet more as 'tag-ons' than to centrally buttress arguments or initiate new avenues of inquiry. The more scientific aspects of his work have consequently dated, while the philosophical approach remains as fresh as ever, and his prose is a delight to read, so different from the jocular jargon of so many contemporary writers. The musings of Freud a few decades later had a tremendous impact (comparable to the theory of relativity) on fashionable intellectual thought in the first part of the 20th century, turning consciousness into a thin film of the surface of the subconscious² The ideas of Freud eventually fell out of fashion, and as a reaction psychology became dominated by the behaviourism of Skinner, taking to the extreme the notion that only what is objectively observable is a proper study of science. In addition to the traditional fields of philosophy and its off-spring psychology³(or more precisely cognitive psychology or science), neurobiology and the burgeoning field of computer science, entered the fray, with totally new approaches⁴. So in recent decades the attack on the problem of consciousness has been effected by four different armies, but as Dennett notes, each of them not thinking very highly of the others. This may explain, the author points out tongue in cheek, why so little progress has been made, when everyone involved seems, at least based on the verdicts of the others, to be an idiot⁵. Dennett himself, on the other hand, is filled

¹ James was the one to coin the notion of a stream of consciousness

² His disciple, ally and eventual rival, Jung, carried matters even further, speculating about a collective unconsciousness (a preposterous idea, which can be given a more respectable interpretation in terms of preset structures evolved in the brain) and bringing in religion as an integral part in the psychic make-up of man.

³ As everyone knows, Russell thought of the progress of philosophy as being effected by successive amputations, as each sub-field reached scientific maturity it could be set free from the merely speculative considerations of philosophy proper.

⁴ Neurobiology was present already at the time of James, be it in a rudimentary form, and commented on by him in his preliminary chapters on the effects of excisions of parts of the brain.

⁵ The interdisciplinary warfare extends to the individual fields themselves. McGinn personally informed

with respect, finding the scientists blessed with impressive intelligence (with the arrogance and impatience that invariably goes with it). The role of Dennett is ostensibly that of the philosopher, although he does interpret that role as some kind of high-level journalist, whose task it is to listen and to synthesize and to present to a wider lay-audience.

Dennett's basic attitude, unlike that of Nagel, McGinn and to some extent that of Penrose, is that 'consciousness' can be scientifically understood, just as we have come to understand life and its evolution by doing away with notions like vital forces and intelligent design. To claim the opposite is just defeatism and sentimentalism. But what should really count as scientific understanding? We all know this is based on reductionism, but as every child realises, this leads to an infinite regress, unless we think of explanations or rather their ambitions, not as absolutes, but wedded to particular situations and problems. The basic problem of 'consciousness' is its inherently subjective character, and science, if anything concerns itself with the objective, but a science of the mind that ignores the subjective runs the risk of ending into the sterility of behaviourism, leaving out everything that is 'really' of interest. Thus any scientific inquiry of consciousness really boils down to finding more and more objective features of the phenomenon, and its success depends on how much intriguing territory it is thus ultimately able to uncover, but there will always be a residue which is inaccessible (or not yet accessible) to which opponents can cling to, like shipwrecked sailors on a skerry which is gradually being engulfed by the rising tide, claiming that this constitute the essential core of the phenomenon which is 'consciousness'.

One may think of consciousness as being divine, or completely materialistically based, or a combination of the two, ultimately stemming from Descartes, and referred to as the mind(soul?)/body duality. As is usual with compromises, they are often more reprehensive to both extremes than the opposites thereof, and the guiding principle in Dennett's book is to liberate the mind from the straight-jacket of Cartesian thinking. The mind/body duality is an intellectual embarrassment, because ultimately it posits the possibility that physical causes may be non-physical, thus almost everyone seriously thinking about consciousness avoids it like the plague. On the other hand it honestly reflects on how we think of mind and matter. After all even purely materialistic theories depend ultimately on the mind, those theories are not simply given to us, but convenient and extremely successful stratagems of thinking of the external world born out of the mind. It is noteworthy that such an old-fashioned duality is proposed by hard-nosed realists such as Popper, and to some extent also by Penrose, in their positing of three interacting worlds⁶ For all intents and purposes we do in practice honor this split. No knowledge of the brain and its working is (at least so far) helpful to us when it comes to change and improve our thinking per se, in fact to think by taking the detour via brain-research, is about as helpful, as to explain the world through the dynamics of its constituent particles. Admittedly, sticking to this split between mind and brain, will forever conserve a fundamental mystery as to

me that the reading of Dennett's book is a waste of time, and that he owes his reputation to his unabashed flattery of scientists, something partly confirmed by the praise lavished on him by Dawkins.

⁶ The mystery of interaction is most strikingly exhibited by Penrose in his 'the Road to Reality'. Mathematics is but one of many of our intellectual fruits, and the physically applicable mathematics is but a small part of all mathematics, yet enough to generate a world rich enough, to make of human intellectual constructs just a tiny part of it.

the workings of the universe, at least the human one of most concern to us. And this very anxiousness as to its preservation is made the butt of Dennetts scorn. How indeed would an explanation of consciousness debase the mystery? In fact whenever a scientific explanation has replaced mere superstition and received belief, there has been an increase in the amount of fascination of the phenomenon under consideration. Thus, Dennett, confidently assures us, a scientific treatment of consciousness will rather enhance its interest than detract from it. Yet would there not be some grounds for such apprehensions? Maybe a materialistic understanding of consciousness may change it irrevocably, maybe even destroy it for ever? Those might seem rather farfetched, not to say ridiculous fears, yet they point to the very anathema of all logical thinking. Circularity and Contradiction. Thus the refuge provided by the Cartesian mind/body split, at least cuts some holes in the loops, in which contradictions can be drowned and replaced by intractable mysteries, in principle no different from those that underlie religious faith(s). Penrose in his 'the Emperors New Mind' actually makes the idea of self-contradiction crucial in his rebutting of an algorithmic explanation of the mind, through invoking Gdels theorem. He also disposes of the oldest materialistic view of the world stemming from the 18th century, as just a collection of particles with given positions and velocities (necessarily of indefinite accuracy) subjected to the deterministic unfolding governed by the laws of Newton⁷, by referring to the more modern Quantum Mechanical paradigm (incidentally another creation of the mind (governed by Newtons laws?)). Gdels theorem has been made to do much heavy philosophical work, far more than it was ever intended to do, and thus one should be leary of taking it too literally, or to use it as an axe to truncate all further discussions. But Gdels theorem is just another instance of the diagonal principle, the magic rope trick involving almost contradictory nonsense, backing off at the last moment⁸. It can also be taken as a manifestation of free will, a classical problem of a materialistic view of the mind. I personally believe that too strong an explanation of consciousness will inevitably embroil us in contradictions⁹.

Dennett is obviously not out to explain in detail how consciousness arises, that would be a tall order indeed¹⁰, but to suggest along what lines an explanation is to be found, and in particular deprive the reader of all baggage that prevents a clear view. His starting point is to not only to reject the Cartesian duality, but also its implications, that of a central authority with a specific location in the brain. The author builds up his case slowly, obviously in order to create suspense, the idea being that the reader after some hard work and thinking (but luckily no math! Dennett assures the fainthearted) being instructed by parables and thought-experiments, will be ready for the final revolutionary insights. At least in my view the conclusion the author presents in the end is anti-climactic,

⁷ Formulated most succinctly by Laplace, although one senses, with a certain ironic detachment.

⁸ As most mathematical principles, only directly and explicitly known by its various particular manifestations, while the principle as such, resides in some Platonic heaven

⁹ The mechanical 18th century picture of the world deprives us all of free will. Is it clear that in a deterministic world we would as minds be predetermined to have a deterministic conviction, or even a suspicion?

¹⁰ He is critical of the only attempt so far to provide a solution drawing on all the four paradigms involved, namely that of Edelman.

and thus I feel entitled to give the whole story away at the onset. Dennett simply presents the solution as being provided by strong AI, so popular in the eighties, but, I suppose, much less so in the nineties¹¹ Our conscious minds are to a large extent, maybe even to its full extent, an illusion. What is going on is simply the implementation of a very powerful program. This is a supposition most people take instinctive exception to, after all how could my sense of myself, so tangible, as well as my sense of humor and the subtlety of my emotions, be the result simply of some computerprogram, some simple-minded mechanical algorithm? Dennetts rejoinder is that one cannot simply reject a possibility because one lacks the imagination to conceive of it. A program of only a few lines could be quite subtle, one of a few pages, obviously even more so, but if one has a very long and complicated program, phase-shifts in complexity may occur¹² Thus in effect Dennett proposes complicated programs as solutions, too complicated to be transparent. There are of course precedents, namely the DNA codes that in effect can be seen as programs regulating the very business of life¹³. To reject an algorithmic foundation of the way the mind works, puts the critic on the same level as someone who out of sentimentality rejects the prevalent DNA paradigm. Dennett clearly does not claim that the genetic code has been decoded and that we can make predictions, only that it provides the paradigm for all serious inquiry into the working of living bodies. Similarly one assumes that a study of the way the brain may be programmed is a reasonable, maybe the only reasonable way, to illuminate the various phenomena connected with consciousness.

The arguments of Dennett are often lengthy and repetitive, and it is not always clear what he really wants to express. His advocacy of the Multiple Draft model is one example, inspired by the fact that consciousness seems to revert temporal order of sensations, especially on small time scales¹⁴, either by some Orwellian method of false memories or some Stalinistic methods of fake ones, the distinction being impossible to make. He makes an effort to point out the trite fact that representation of space in the mind is not reflected by some corresponding spatial representation in the brain, then making the analogous suggestion that neither does temporal representation correspond to a corresponding temporal ordering in the brain, which if you think of it does not make sense, at least not on a global scale. The obvious solution that the unconscious brain acts as a short-time buffer, is rejected on grounds not entirely clear, on the strength of his multiple model, in which there is no authoritative script, but many competing ones. To me he never succesfully ties the loose ends together.

The most subjective aspect of consciousness is the presence of 'qualia', namely how sensations, thoughts, beliefs, pains 'occur to you'. You may know everything about color,

¹¹ This produces the somewhat uncharitable suspicion that had the author written his book ten years later, the conclusion would have been different.

¹² the digits of small numbers convey very little information, those of very large can be used to encode entire novels.

¹³ The development of organisms and their subsequent conduct cannot be reduced to the DNA as all biologists know, there are many other factors that enter; yet in principle the argument remains.

¹⁴ The experiment referred to is the blinking of two lights close to each other in space and time, but of different color, creating the illusion of one point moving changing color in the process, which presupposes some kind of advanced knowledge

all its physical properties, yet you would have no idea how the sensation of a particular color appears to you, if you have never any direct contact with it. In fact a celebrated thought-experiment poses exactly this puzzle, about a smart young woman being brought up in a completely monochromistic environment, yet learning everything objective there is to know about color, having no presentiment on how color would present themselves to her¹⁵. Dennett argues that such a woman would not be tricked by say a blue banana, provided that she has really learnt everything there is objectively to know about color. Qualia being confined to a particular consciousness cannot be compared across interpersonal boundaries as qualia. As every reflective child realises, the world may look very different to different people. Naively it may at first entertain the idea that what is blue to it, may be red to another, until it realises that this does not make sense, as no one else may experience the qualia of someone else¹⁶. Frege in his writings, even mathematical ones, is very clear about this point, namely that the subjective world of each person is a closed box, in which outsiders have no access. Communication is always carried on through the notion of isomorphisms, so familiar to all mathematicians¹⁷. The notion of 'qualia' is in fact an epiphenomenon, i.e. it has no discernible physical effects. Taking an instrumental point of view, it simply cannot exist¹⁸, any claims to the contrary is simple pure sentimentalism on the parts of 'qualephiles', a philosophical sub-species doomed to extinction due to the inevitable contradictions they necessarily embroil themselves in. By getting rid of the notion of 'qualia', you also get rid of the notion of zombies. There are no zombies, if they taste and feel like conscious beings (in fact if they pass the most taxing Turing tests) they are for all intents and purposes conscious beings. By excorcising the one non-objective feature of consciousness it makes the subject scientific. This is of course good science and physics, but hardly satisfying meta-physics. We do sense that other people are conscious, otherwise the world would be a very lonely place. This conviction has the ring of religious

¹⁵ This example is e.g. repeatedly elaborated upon by the novelist Lodge in his 'Thinks...'

¹⁶ As a less obvious example of a subjective quale, consider the way we distinguish between up and down. How do you really do it? The image projected on the retina is upside down, how come we do not walk around seeing the world upside down? Some elementary reflection reveals the stupidity of such a supposition. People who are equipped with glasses turning their views upside down, are at first quite disoriented. Soon, however, they adjust fully. How do they adjust? By simply turning the image upside down and thus reverting to their original state? This is the naive solution. A more sophisticated one is that the notion of the incomparability of qualia not only applies between different people, but also temporally in the same individual.

¹⁷ One may say that the question of equality of two objects only arise if they are a priori placed in the same set. Frege remarkably seems to make no explicit reference to isomorphisms in his writings.

¹⁸ Dennett apparently accepts a weak form of qualia, namely that induced by a general disposition. As an example he takes a pious churchgoer of 17th century Leipzig. Such an individual would experience the music of Bach rather differently from us. In order for us moderns to get an idea of how the music 'felt' to him, we would need not only to teach ourselves about the particular environment he was living in, but also to divest ourselves of all modern knowledge, especially that of music, postdating Bach. Another weak sense of qualia, not brought up by Dennett, is connected with the way we read texts in different languages. The ostensible contents may be the same, yet the words, although having similar meanings, would bring about different associations.

faith, without which we would find life intolerable. By restricting the inquiry to what is phenomenological we side-step (perhaps wisely) the main issue, and thus the problem becomes, meta-physically at least, trivial.

In addition to this Dennett speculates that consciousness is a recent biological invention¹⁹, basically contemporary with the emergence of language, without which consciousness as we know and treasure it would hardly be possible. Consciousness is simply the result of Dawkian memes that have found the environment of brains congenial to their reproduction. This is hardly an uncontroversial issue. We often tend to intuit some degree of consciousness, especially as a capacity for suffering, in some of our favourite fellow mammals; without perceiving any capacity for language. Also the date (10'000 BC) at which Dennett puts down the emergence of language seems to me far too recent²⁰. To Dennett, if I follow him right, selves are but illusions, similar to fictions, brought about by emerging narratives (multiple drafts again?) to which single authors are attributed, not unlike the abstractions provided by centers of gravities²¹. In particular, if I understand him right, language is not a genetic implication, our ancestors being biologically indistinct from ourselves, going back at least 100'000 years, but serendipitously fitted to the structure of the evolved brains.

Natural selection and evolution also provides an alternative path to study and explain the emergence of consciousness. Our brains have in fact been evolved from simpler prototypes²² Thus what we have in fact is an inductive setting, starting out with something extremely simple, and then in principle evolving with small steps to the complexity of our own brains. The initial step we are bound to handle, and then each subsequent. In practice we are never presented with such an almost continuous array. And besides there is always the phenomenon of phase-shifts, an innocuous addition having momentous and unpredictable consequences.

A book on the mind and consciousness can nevertheless be quite instructive and interesting, even if you would disagree with the main tenets of its author. What makes it interesting is the presentation of empirical data, either from cognitive psychology or neurobiology, which challenges your views on what it means being conscious. There are many such strange things exposed by pathological brains²³. One example is the notion of 'blindsight', when not all nerve-connections are severed, but not enough remaining to make the visual sensations conscious. Blind-sighted people can guess the nature of simple visual stimuli above what chance alone would predict, just as we, Dennett reminds us,

¹⁹ One may compare this to the theory of the psychologist Jaynes, as to the recent emergence of consciousness as a result of the breakdown of the bicameral mind

²⁰ His argument appears to be that with the introduction of language, agriculture and civilization suddenly came into being.

²¹ Dennett seems to treat centers of gravities as fictions, useful for computation, but by no independent meaning. Physicists and mathematicians tend to view such an attitude as somewhat peculiar

²² What are brains for? Anticipatory machines, and in one of the books most amusing jokes, he refers to the sea-squirt, which is equipped with a brain in its juvenile state, but once it has found a place to attach itself to, the brain becomes useless and is simply eaten up. Not unlike people obtaining tenure.

²³ The neuropsychologist Oliver Sacks has made himself a name by sympathetically presenting such pathological misfits to the view of the curious. And the subject is inescapably fascinating.

would have the ability to guess the make of the cars owned by people presented to us in the phone-directory with somewhat higher degree of accuracy than the truly blind guess, without having the slightest cue as to our strange presentiments. Such phenomena provide interesting feed to the philosophical mind. I am far more sceptical about the light computer science and artificial intelligence could throw on the real problems of human consciousness²⁴. But we do not have to resort to pathologies to whet our curiosity and challenge deeply held prejudices; even more interesting things are to be found just around the corner. Hearing and understanding speech has been well-studied, but not (at least at the time of the books writing) the production of speech. How do you formulate yourself? Often you find that you have a sense of what you want to say, and there is a problem of formulation. Ordinarily though when you are chatting away, there seems to be no time for thoughtful considerations, the words just come out of you, before you have time to consciously phrase them, and some of it may strike you as surprising and even original, and you wonder how you could have thought of it. In this context Dennett posits a pandemonium of voices, out of which swift selections are made. There are once again all those multiple drafts, and it is hard to predict what gets eventually sanctioned by being uttered. Other topics of engrossing interest are supplied by visual imagery. Dennett opposes the idea of brains filling up missing information. The 'invisibility' of the blind spot²⁵ is due partly by redundancy of vision (you have two eyes after all, and your eyes moves around a lot) but also to the fact that the brain does not 'expect' any information to come from it, and thus 'experiences' no absence of information, that otherwise might 'appear' as a hole. When it comes to the subconscious workings and organizations of the brain, hard enough as they may be, there no longer is felt this impenetrable barrier of a possibility of understanding. This is of course a topic rich enough to supply interest independantly upon the light it could possible shed on consciousness. The empirically minded might even hope that in due time enough empirical understanding might have been amassed as to allow the emergence of an understanding of consciousness, or at least the illusion of such a thing. But there is a very long way to go, and books like Dennetts really makes very little progress in elucidating the issue to critical readers.

December 11, 2005 **Ulf Persson:** *Prof.em, Chalmers U.of Tech., Göteborg Sweden ulfp@chalmers.se*

²⁴ A consequence of Dennetts attitude is that in principle down-loading would be feasible, that the workings of the brain can be theoretically mimed by a supposedly giant computer. Such claims are of course hard to falsify, but that does of course not mean that they are true, much less realistic

²⁵ A rather big entity subtending an angle of six degrees