

The Journal of Mind and Behavior

Summer 2005, Volume 26, Number 3, pp207-214

Sweet Dreams: Philosophical Obstacles to a Science of Consciousness. Daniel C. Dennett. Cambridge, Massachusetts: MIT Press, 2005, 216 pages, \$28.00 hardcover.

Reviewed by Leslie Marsh, Centre for Research in Cognitive Science, University of Sussex

And since the possession of qualities assumes a certain pleasure in their reality, we can see how a man who cannot summon up a sense of reality even in relation to himself may suddenly, one day, come to see himself as a man without qualities.

Robert Musil, *The Man Without Qualities*

The question of how a physical system gives rise to the phenomenal or experiential (olfactory, visual, somatosensitive, gestatory and auditory), is considered the most intractable of scientific and philosophical puzzles. Though this question has dominated the philosophy of mind over the last quarter century, it articulates a version of the age-old mind-body problem. The most famous response, Cartesian

My thanks to Chris Onof (Imperial College and Birkbeck College, University of London) and Geoffrey Thomas (Birkbeck College, University of London), for their pointed comments. Requests for reprints should be sent to Leslie Marsh, Centre for Research in Cognitive Science, University of Sussex, Falmer, Brighton BN1 9QH, U.K. Email: l.marsh@sussex.ac.uk

dualism, is on Daniel Dennett's view still a corrosively residual and redundant feature of popular (and academic) thinking on these matters. Fifteen years on from his anti-Cartesian theory of consciousness (*Consciousness Explained*, 1991), Dennett's frustration with this tradition is still palpable. This frustration is primarily aimed at philosophers. The "Sweet Dreams" of Dennett's title are the rationalist thought experiments of wishful thinking philosophers who, neglectful or unaware of empirical evidence, generate premature conclusions "of unexamined presuppositions and circularly defined elaborations" (p. 79). The nature of such presuppositions renders these thought experiments no more than "intuition pumps," ostensibly succeeding in stale-mating or in some cases check-mating any moves in the direction of a unified science of consciousness. The extent to which Dennett believes these "pumps" have skewed theorising about consciousness is captured in his remark: "I had no idea philosophers still put so much faith in the authority of their homegrown intuitions. It is almost as if one thought one could prove that the Copernican theory was false by noting that it 'seems just obvious' that the Earth doesn't move and the Sun does." (p. 108).

This small book affords Dennett an opportunity to clarify, refine, amplify and amend his theory as well as to assess the exponential growth in the field of consciousness studies in the interim fifteen years. In doing so, Dennett inevitably ratchets up a notch or two the already highly charged philosophical skirmishes that he partly initiated. His discussion brings a broader issue into focus – the relationship

between philosophy and the natural sciences: "I have spent more time than I would like explaining to various scientists that their controversies and the philosophers' controversies are not translations of each other as they had thought but false friends, mutually irrelevant to each other" (p. 21; see also pp. 134, 156). Since theorizing about consciousness has been the preserve for philosophers for centuries, how is it to respond and engage the sciences' new-found interest in consciousness? Until the 1980s, reference to consciousness by the brain sciences, was for the most part, a taboo.

There are two inextricably linked broad themes that cut across the eight chapters that comprise this book, the first negative, the second positive. The former offers a critique of the thought experiments that inform the aforementioned philosophers' "defeatist" conclusions. The latter, methodological in intent, offers a scientific third-person approach to the study of consciousness, what Dennett terms "heterophenomenology." The cross-weave deals with the issues of intentionality and personhood.

At the outset it is vital to keep firmly in mind Dennett's strategy. Dennett approaches consciousness tangentially: he is not trying to explain consciousness itself so much as he attempts to make sense of the judgments we make about conscious experience. With appreciation of this, one can avoid the common-place caricatural view that Dennett is denying that consciousness exists (p. 71). All Dennett is claiming is that consciousness is not what people often think it is – consciousness analogous to

magic or sleight of hand entices the theorist into postulating inelegant metaphysical puzzles. Furthermore, there is something comforting in upholding the mysteriousness (pp. 62, 75) of the “trick” and that any attempt to shine light in on the magic black box will somehow diminish us and turn us into mere things (p. 57).

In order to do justice to Dennett’s position we need to step back and briefly look at an ongoing concern he has had for close to forty years of writing – intentionality or the theory of content (pp. 26, 37-38, 43, 46, 55, 59). Dennett’s account of intentionality gives us one possible way to remove the distinction between the mental and the physical. For Dennett, intentionality *presupposes* a theory of consciousness, an emphasis that is at odds with prevailing philosophical orthodoxy. When philosophers say the mind exhibits intentionality they are referring to the fact that mental states can be about something. If one were to try and explain human behaviour by appealing to the laws of physics, the computational work involved in making such predictions would be impracticable. This is what Dennett terms the “physical strategy.” Perhaps a more appropriate explanatory strategy would be to invoke what Dennett terms the “design stance.” So, for example, we do not need to appeal to the laws of physics to understand the behavior of a computer. All we need know is that it has been designed to perform certain functions. This level of description still does not capture the complexity of human behavior. Dennett suggests that the only way to approach such complexity is to adopt higher levels of predictive strategy, what he terms the “intentional strategy.” Dennett’s theory of content and

consciousness require that the brain have certain capacities and structures to produce an array of behavior which we find in ourselves and others. Dennett's "intentional stance" is thus a functional, teleological view of content in the biological economy of an organism, amenable to analysis – a third-person perspective. To make a determination of what the mental content of a person is, is to make an interpretation of the person's external behaviour, *as if* they have rationality; rationality that is of course saturated with hopes, fears, desires, beliefs and intentions. The "as if" is important here. Dennett claims that our talk about minds does not refer to anything concrete, but to useful fictions, in much the same way as economists refer to entities like "the average taxpayer." Beliefs and desires are nothing more than a useful predictive tool: one is not then entitled to infer that we know what it is inside the system that is causing the theory to work. I think it is fair to say that Dennett's concern with intentionality has morphed into heterophenomenology – more on this later.

Dennett's Multiple Drafts Model (MDM) of consciousness as set out in his *Consciousness Explained* conceives of conscious experience as a result of micro-judgements made by various parts of our brain undergoing perpetual editorial revision. In *Sweet Dreams* Dennett not so much amends the MDM of perpetual editorial revision as he supplements the idea by invoking a new metaphor – the "fame in the brain" or "political power" or clout metaphor for consciousness (pp. 136-138, 142). The idea is that a theory of consciousness would need to explain how some relatively few contents become elevated to a position of political power,

explain why they hog time in the "limelight." Bernard Baars's (1997) not dissimilar Global Workspace Theory (GWT) postulates the idea that consciousness resembles the bright spot on a stage of fleeting memory and that this consciousness is the primary agent of global access (distributing information) to the rest of the darkened theatre. There is a two-way flow between conscious and unconscious brain activities.

Dennett makes the charge that many philosophers (*and* scientists), whatever their materialist credentials, harbor a residual Cartesianism (p. 101). This supposedly trades on the Cartesian myth that there is a central theatre, the idea that there is some central place in the brain where something like an "I" or the Self attends to and witnesses consciousness. The positing of some central authority or homunculus gives rise to a well-known infinite regress which was first pointed out by Gilbert Ryle in *The Concept of Mind* (1949). Ryle pointed out that an observing self must necessarily contain another observing self, and so on *ad infinitum*. To my knowledge, few philosophers or scientists posit a homunculus in this way – Baars's GWT doesn't. Of course, as Michael Wheeler (2005, p. 66) rightly observes, humuncular explanation is simply written into the rule book of classical cognitive science.

The consensus view is that there is no one place where consciousness happens; mental activity in the brain is accomplished as a result of parallel processes of elaboration and interpretation. Though the scientific literature is replete with references to the "executive regions" of the brain (at the time of writing entering these words into Entrez PubMed returns 421 references) these positions do not imply an

infinite regress. Francis Crick and Christof Koch (2003, p. 120) talk about parts of the prefrontal cortex having executive regions of interpretation, decision-making and voluntary control. They suggest that "it would be surprising if this overwhelming illusion did not reflect in some way the general organisation of the brain."

If there is no central control, then how is it that it seems as though I am a singular conscious agent? The supposed loss of the continuous and unitary Self is disturbing to many since on Dennett's account the work done by the imagined homunculi (pp. 137, 161) in the Cartesian Theatre is distributed amongst various lesser agencies in the brain, *none* of which is conscious (pp. 69-71). Dennett invokes Leibniz's factory metaphor: "a good theory of consciousness should make a conscious mind look like an abandoned factory, full of humming machinery and nobody home to supervise it, or enjoy it, or witness it." The "I" is neither something outside the physical world or something in addition to the team of busy, unconscious robots whose activities compose you; it should be included in the accounting by heterophenomenology (p. 75).

This loss of Self has strong resonance with Buddhist doctrine of *annatta*, no-self or no soul. This is the idea that the positing of a soul, a permanent and stable entity, is incoherent because all beings are subject to continuous change, death and decay. Beings are in a constant state of flux or "becoming." The human personality for *annatta* is an aggregate of several individual components. Buddhist-like conclusions are also to be found in the work of moral philosopher Derek Parfit as set

out in *Reasons and Persons* (1984). Parfit distinguishes two views about the nature of persons, one the Non-Reductionist, the other the Reductionist. The former is the Cartesian Ego whereby a person is distinct from her brain, body and experiences. The latter, Parfit's position, is that the existence of a person just consists in the existence of her brain, body, her thoughts, deeds and innumerable other physical and mental events – *personal identity is not a separate further fact*. My reference to the Buddhist conception of no-self is merely to make the point that contrary to some critics, Dennett's views are hardly novel or idiosyncratic.

To assuage the worry of the loss of Self perhaps one should take the view that personal identity, on Dennettian terms, is best understood not just as a kind of rational or intellectual presence, but as a conglomeration of ongoing goals, projects and commitments. One recognises oneself in part by keeping track of this flow of projects and commitments; others recognise me not only by my physicality but also by some distinctive nexus of projects and activities. Andy Clark (1987), a cognitive scientist, makes a similar proposal: all that should matter is that the conscious self has a broad sense of what the entire situated and embodied agent can and can not do. Our bundle of "taken-for-granted" skills, knowledge, and abilities structures and informs our sense of who we are and what we know, a sentiment echoed in Dennett. Dennett (pp. 138-139) cites with approval, Susan Hurley's view that "where we act and where we perceive is not funnelled through a bottleneck, physical or metaphysical, in spite of the utility of such notions as 'point of view.'" But there is a

sense in which there is a Self, a bio-evolutionary Self that “is the subject of interoceptive signals that alert an organism to its own homeostatic state – to its autonomic, neuroendocrinological, and hedonic condition” (Flanagan, 1992, p. 49).

To say one has an experience that is conscious (in the phenomenal sense) is to say that one is in a state of its seeming to one some way: it is to say that there is something it's like for one (an “I”) to have it. Because of the immediacy of these senses, consciousness has been taken to consist in the *in corrigible* monitoring of one's own states of mind – a residual first-person Cartesianism. For many philosophers a first-person methodology is the only way of explaining the seemingly ineffable subjective quality of our experience, or in philosophical parlance, *qualia* (quale singular): the definition of this term of art being highly contested (pp. 78-79, 86-88). How can one be wrong about one's own qualia? (p. 89). It is this first-person phenomenology that has informed philosophical discussion of consciousness, which in Dennett's view, can only be plausibly discussed within a third-person framework. Heterophenomenology as a methodology simply insists on anchoring subjective nuances to something – that can be detected and confirmed in replicable experiments (p. 149). Dennett's appeal to heterophenomenology is to commend its *neutrality* (p. 42). It seems to me that contrary to anti-naturalist sensitivities, Dennett's heterophenomenology, in its commending of *intersubjectivity*, is a universalizing impulse with positive socio-ethical implications – understanding subjectivity in others, means that one can enter into empathy with others!

If qualia can be identified with a physical property, this would secure a causal role for them without violating the causal closure of the physical world. We want qualia to be efficacious: if they don't, it becomes increasingly difficult to understand why we should have them at all. Dennett is perplexed by the faultline that divides the educated mind: on the one hand there are those to whom it is obvious that a theory that leaves out the Subject is disqualified as a theory of consciousness; on the other hand there are those to whom it is just as obvious that any theory that doesn't leave out the Subject is disqualified (p. 145). Dennett's heterophenomenological commendation is thus eminently moderate.

Olfaction, vision, somatosensation, gestation and audition are taken as the paradigmatic examples of phenomenally conscious states. For those who work in the empirical sciences, philosophers' talk of Zombies, "what is like to be a bat?," and scientists with omnipotent knowledge must seem at best whimsical, at worst indulgent. These thought experiments have generated a voluminous industry of parametric responses which, for reasons of space, I cannot touch upon.¹ The common denominator of these thought experiments is their supposed illustration that the phenomenology of experience cannot be (and may never be) accounted for by science: qualia are ineffable, intrinsic (i.e., knowable independently of knowing how the state relates to other states), private (i.e., directly knowable only by the person who has them), and directly introspectible (i.e., consciousness of them does not involve inference or interpretation).

¹ By far the best online resource can be found at <http://consc.net/online.html>

The Zombie thought experiment turns upon the idea that it is logically possible that there are beings that are objectively indistinguishable from people with consciousness, except that they do not have phenomenal consciousness. Inverted spectrum thought experiments say that what appears red to me appears green to someone else without this difference showing up in our cognition or behaviour in any way. We draw the same inferences from what we are seeing, and use the same words – where I say “red,” the other person says “red” (pp. 21, 31).

I share Dennett’s annoyance with the “Zombie hunch”: it amounts to the logical equivalent of saying “that zombies have streams of unconsciousness where the normals have streams of consciousness!” (p. 150; see also pp. 14-17, 22-23, 47-48, 91-92). The point of the Zombie thought experiment is to claim that heterophenomenology must be leaving something out. I reiterate. Dennett is not denying one must account for the first-person perspective. On the contrary, it only makes sense to do so within a framework of heterophenomenology which *includes* taking the reports of subject seriously (pp. 146-147). Dennett's critical point is that no property of any cognitive state or process in isolation could be a property of something seeming a certain way. Something coming to seem a certain way is a result of a complex cognitive process, not a property there to be “read off” individual cognitive states.

The omnipotent scientist thought experiment is a version of the “what is it like?” thought experiments. In a nutshell it poses the following question. What, if

anything, would be experientially different for someone on her release into a full color world that had heretofore lived her whole life in a black and white world, *even though* this person was in possession of a complete physical description of reality outside of her confined world? The puzzle is intended to “prove” a priori that even with the possession of a complete physical description of reality, Mary on her release would still experience something new, in this case the color red. Therefore qualia cannot be reduced to the level of physics and hence there cannot be a unified theory of consciousness. In other words, qualia have been declared outside of science altogether (p. 91). It’s worth noting that Frank Jackson (2004) who formulated the “Mary” experiment has, twenty five years on, retracted his original conclusion. He is now of the view that the sensory side of psychology is, in principle, deducible from the world’s physical nature. Not dissimilar to Dennett he thinks that the puzzle posed by the so-called “knowledge argument” is to explain *why* we have such a strong intuition that Mary learns something new, an intuition that outruns that which can be deduced from the physical account of how things are. Again, like Dennett, he takes the view that much of the contemporary philosophical debate is concerned with the clash of strongly held intuitions. In Dennett’s view, philosophers have spent a disproportionate amount of energy on what appears in retrospect a relatively trivial definitional issue: “nothing is going to count as knowing what it’s like to see red without also counting as an experience of red.” (p. 120). This tautology can be resolved by appreciating that *consciousness doesn’t happen in addition to the*

experience of being conscious. Jackson (2004, p. 421) now claims that “most contemporary philosophers, when given a choice between going with science and going with intuitions, go with science.” The *raison d’être* of this book is that Dennett doesn’t think this is the case.

There are two major problems that these intuition pumps generate. The first, and one with which I am not out of sympathy with is that “it is a mistake to inflate practical indescribability into something metaphysically more portentous” (p. 111; see also pp. 107, 115). It seems that at the root of these puzzles is a contentious philosophical distinction between intrinsic and extrinsic properties. Very briefly, the former are those which can be isolated from everything else going on in the brain (or elsewhere) and is not dependent on relations to other mental states. The latter, synonymous with “relational,” is a property which depends on (and consists of) a relation to something else. If one assumes that qualia take the form of intrinsic properties, then there is every reason to question one’s existence much like the naïve American abroad who thinks of dollars as “real” money, having intrinsic value, in contrast with euros, yen, etc. (p. 177).

The second problem is the question of just how “introspectively accessible” must an experience be to count as a quale? Which aspects of our experiences are the “phenomenal aspects” and which are not? Is our *enjoyment* of a good meal, the deliciousness, or the aesthetic appreciation of a work of art, itself a phenomenal aspect? (pp. 79, 89). How does one account for the young adult’s first taste of beer

which is met with repulsion. Later, this same person develops an appreciation and perhaps a connoisseurship for this taste. Did the taste remain the same or has it changed? These intuition pumps, on Dennett's view, illustrate that we do not have any clear idea what a change or lack of change in an experiential state would consist in. Either we would have desisted from future drinking of beer or if beer is not a taste we learn to enjoy, and we admit that our reactions or attitudes to experience are at all constitutive of their experiential character, then experiential character ceases to be intrinsic.

I don't think Dennett is suggesting that thought experiments have no role to play in philosophical reflection. He is after all himself a master of such contrivances. I think the moral that he intends us to take away is that they are open to misuse in the sense that they are, more often than not, masquerading as argument because assent to them seems so easy and compelling. One could argue that there wasn't a mind-body problem until Descartes conceptualised it as such, and this has since been a defining feature of our thinking about such matters.

So how does Dennett accommodate qualia? Recall Dennett's claim that we tend to employ useful fictions: there is no reality to the subjective quality of our experience over and above the fact that there *seems to be* that subjective quality. A word of caution. Though he does have his eliminativist moments, it must be remembered that Dennett is *not* an outright eliminativist about consciousness. Dennett, it seems, deploys Wittgenstein's argument against the notion of a private language

here along with a distinct flavor of verificationism. Take pain, surely the paradigm example of a private experience. Wittgenstein pointed out that for the word “pain” to have any meaning at all presupposes some sort of external verification, a set of criteria for its correct application, that must be accessible to others as well as to myself. And this is applicable to all other inner experiences. It doesn’t make sense to say that it *seems as if I’m in state of pain*. *There is no such phenomenon as really seeming – over and above the phenomenon of judging in one way or another that something is the case.*

Dennett is *not* denying that there is this experiential aspect to our lives. What he is saying is that there is nothing in the process of perception which is ultimately mysterious or outside the normal causal system. He proposes that instead of mystifying ourselves with first-person phenomenology we adopt a third-person perspective – heterophenomenology. In other words, instead of trying to talk about our ineffable inner experiences, we should talk about what *people report* as being their ineffable inner experiences.

Part of Dennett’s diagnosis as to why discussion of these matters is so skewed lies in the multidisciplinary nature of cognitive science (philosophy, psychology, computer science, neuroscience, linguistics, and anthropology). While there is fascinating and fertile work being undertaken under the aegis of cognitive science, the waters are muddied because disciplinary concerns don’t easily translate into another’s concerns without remainder (p. 156). The corollary being that “some

philosophers have misappropriated those same controversies within cognitive science to support their claim that the tide is turning against functionalism, in favour of qualia, in favour of the irreducibility of the 'first-person point of view' . . ." (p. 156)

Perhaps the fetishising of qualia is a dead-end enterprise: it is only a partial albeit prominent aspect of consciousness – emphasising the intense special effects rather than the diffuse narrative. Let science pursue the search for the neural correlates of consciousness. And if in twenty-five years science has nothing to show for it, then I'd say that all bets are off. This said, one needs to be wary of physicists who try to apply exotic physics to the brain. As Crick and Koch (p. 124) point out: "they seem to know very little, and even less about consciousness" than many philosophers.

The penultimate and final chapters in this book are somewhat superfluous – they merely restate, rephrase or summarise much of the preceding book. Perhaps the publisher deemed the 157 pages that comprise the first six chapters of the book too brief. If the reader feels cheated, it has to be on the grounds of the substantive argument, not the wordage. Unlike many of the new books jostling for reader attention that are either recently reheated Ph.D. theses, or old hands cobbling together too diverse a collection of papers, Dennett is always good value for money. Dennett is without doubt a supreme stylist: but his tongue-in-cheek humor and flippancy are not to academic philosophy's taste. Indeed, Dennett is the best advert

for philosophy: neither the logic chopper nor the turgid obscurant, he still thinks about the BIG questions with clarity.

And whether or not one subscribes to Dennett's position *tout court*, he offers a welcome corrective to philosophical immodesty and smugness. I'm not suggesting that scientific immodesty should go unchecked, an immodesty associated with scientism, if by scientism we mean a dilettantish engagement with science. Even if some critics call Dennett a "pop scientist," the charge of dilettantism cannot be applied to Dennett – whatever their disagreements Dennett commands respect by distinguished scientists as diverse as Richard Dawkins, Nicholas Humphrey, Marvin Minsky, Roger Penrose, Steven Pinker, and the late Francisco Varela.

Dennett, a best selling author, now very much a public figure, is subject to the lingering suspicion of public intellectuals in Anglophonic culture. Perhaps it is in this public realm that Dennett's most important contribution is being made. Dennett is a lightning rod for a whole cluster of conflated issues and internecine battles across the broader socio-cultural landscape. The reader, thus, needs to be alerted to the fact that Dennett engenders criticism not on purely philosophical considerations. Dennett is cast by diverse critics disparagingly as polemical, deflationary, iconoclastic, huffy and a belligerent gadfly. He is all these things – he is, by his own admission subversive and radical. But on any metric Dennett is hardly a pariah. I suspect that Dennett's belligerence is pronounced because of his willingness to speak up against some dubious intellectual currents, creationism now masquerading as "intelligent

design" and relativism to name but two. In many ways, Dennett has the iconoclastic tendencies of his Oxford supervisor Gilbert Ryle – if one may pardon the pun, Dennett continues to *rile* tired philosophical orthodoxies, fashionable postmodernisms and junk science. Long may this continue.

References

Baars, B. (1997). *In the Theatre of Consciousness: The Workspace of the Mind*. Oxford and New York: Oxford University Press.

Clark, A. (1987). *Being There: Putting Brain, Body, and the World Together Again*. Cambridge, Massachusetts: MIT Press.

Crick, F. and Koch, C. (2003). A framework for consciousness. *Nature Neuroscience*, v. 6, no. 2.

Flanagan, O. (1992). *Consciousness Reconsidered*. Cambridge, Massachusetts: MIT Press.

Jackson, F. (2004). Mind and Illusion. In P. Ludlow, Y. Nagasawa & D. Stoljar (Eds.), *There's Something About Mary: Essays on Phenomenal Consciousness and Frank Jackson's Knowledge Argument* (pp. 421-442). Cambridge, Massachusetts: MIT Press.

Parfit, D. (1984). *Reasons and Persons*. Oxford: Oxford University Press.

Ryle, G. (1990). *The Concept of Mind*. Harmondsworth: Penguin Books. (Originally published 1949)

Wheeler, M. (2005). *Reconstructing the Cognitive World: The Next Step*. Cambridge, Massachusetts: MIT Press.