

Essays in Philosophy

A Biannual Journal

Vol. 7 No. 2, June 2006

Book Review

Sweet Dreams: Philosophical Obstacles to a Science of Consciousness, by Daniel C. Dennett.
Cambridge, MA and London: The MIT Press, 2005. 199 + xiii pages (including references and index).
ISBN: 0-262-04225-8. \$28.00 (cloth).

This book, half of whose chapters had their genesis as the 2001 Jean Nicod Lectures, is vintage Dennett – engaging, erudite, clever, entertaining, and provocative. To some readers it will also appear frustratingly evasive. Dennett has now published three books whose titles or subtitles include the word “consciousness”, and although details and emphases have shifted, there has been a strong continuity in his fundamental message that what is frequently taken to be the obvious and central data of consciousness is illusory. Needless to say, this view is not universally accepted, and despite Dennett’s overall good humor, his exasperation with his adversaries occasionally shows through. For example:

The fact that philosophers would so much as *entertain* such an interpretation of such a casual exercise of the imagination fills me with astonishment. I had no idea philosophers still put so much faith in the authority of their homegrown intuitions. [pp. 107-108]

Nonetheless, recognizing that philosophers do respond as they do, *Sweet Dreams* takes on the charge of trying to explain to them why in many instances they shouldn’t.

Philosophers of all stripes agree that the topic of consciousness presents a supremely difficult challenge; it lies at the very core of the mind/body problem. There are many ways to respond to the challenge, but in recent years the following options have conspicuously emerged:

- (1) Embrace dualism.
- (2) Embrace physicalism, at least tentatively, with the proviso that we are not (currently) in a position to understand how physicalism could be true.
- (3) Embrace physicalism with the proviso that not all physical facts can be described in third-person language or reduced to third-person facts. (In contrast to (2), this does not mean that physicalism is unintelligible.)
- (4) Attempt to reduce consciousness (e.g., to behavior, brain processes, or functional states).
- (5) Deny the existence of conscious phenomena.

Dennett, as I read him, opts for a combination of (4) and (5), although he would certainly not endorse (5) as a component of his position. On the contrary, he vigorously resists any such characterization (“I don’t maintain, of course, that human consciousness doesn’t exist . . .” [p. 71]). In order to assess this, a key question must be addressed: What, if anything, is left unaccounted for after option (4) has been carried out? An examination of Dennett’s argument will repeatedly lead back to this issue.

A scientific study of the mind, Dennett insists, must adhere scrupulously to a third-person, objective method of investigation (“heterophenomenology”, introduced in Chapter 2). This method can be used by any investigators (human, Martian, or whatever) who are (a) capable of data collection, and (b) equipped to adopt the “Intentional stance” – i.e., capable of attributing beliefs, judgments, and other Intentional states to things whose behavior they seek to explain. Dennett describes his method as

the *neutral* path leading from objective physical science and its insistence on the third-person point of view, to a method of phenomenological description that **can (in principle) do justice to the most private and ineffable subjective experiences**, while never abandoning the methodological principles of science. [p. 36, bold emphasis added]

Applied to the study of human consciousness, heterophenomenologists begin by (i) cataloguing a person’s utterances and marks (e.g., writings and drawings); (ii) interpreting those behaviors via the Intentional stance as verbal judgments which express (iii) the individual’s beliefs about his conscious experiences. This yields a description of the subject’s “heterophenomenological world”. At this point, however, a chorus of resistance may object that those three steps would deliver exactly the same heterophenomenological worlds for both a conscious person and his zombie twin. The complaint, in other words, is that *conscious experience itself* has been left out of the picture. But this is the very move that Dennett wants to disallow:

People undoubtedly do believe that they have mental images, pains, perceptual experiences, and all the rest, and *these facts* – the facts about what people believe, and report when they express their beliefs – are phenomena any scientific theory of the mind must account for. [pp. 38-39]

This assertion rests squarely on the methodological point that beliefs are accessible to third-party investigators, whereas experiences are not. But Dennett does not conclude that there’s something real that scientific investigation simply can’t discover or describe – something that I have and a zombie lacks – for he thinks that there is no difference at all between me and my zombie twin. The two of us have the same beliefs, and searching for anything further produces only chimeras:

Moving from the third-person to the first-person point of view is just asking for trouble; you get no data not already available to all the rest of us from the third-person point of view, and you risk sending yourself off on wild goose chases trying to pin down conscious experiences that you only think you’re having. [pp. 49-40]

The most natural reading of this claim, it seems to me, is that consciousness is nothing over and above *beliefs about* consciousness (beliefs, not experiences, are available to third-person viewpoints, and they contain all the existing “data”). Here the strongly behavioristic orientation of

Dennett's position emerges, yet it masks the radical nature of his theory because it permits him to sound like everyone else in making attributions of consciousness. However, given the substitution of beliefs for experiences (or rather the reduction of the latter to the former), it is difficult to resist the conclusion that Dennett's treatment of conscious experience ultimately rests on its denial, and accordingly does not at all "do justice to the most private and ineffable subjective experiences". In fact, there's even less here than meets the eye, since for Dennett, *beliefs* about consciousness do not really exist in an organism either; rather, they emerge from an application of the Intentional stance – a heuristic, instrumental device that is useful for predicting an organism's behavior but not descriptive of its internal makeup.

Attempts to counter heterophenomenology's deliverances by appeals to first-person subjective experience are old hat to Dennett, and a primary aim of *Sweet Dreams* is to show what is wrong with them. Consciousness, Dennett contends, is less than it seems to be (Chapter 3 likens the "illusions" of consciousness to the illusions of stage magic tricks). His basic strategy is to argue that philosophers' efforts to identify something in consciousness other than *judgments* or *beliefs* about conscious experience are riddled with insoluble difficulties. A cornerstone of the argument is the notorious issue of qualia, which constitute the subject matter of Chapters 4 and 5. Here is Dennett's introduction:

[T]here are, however, some events that occur in my brain that I do know about, as soon as they occur: my subjective experiences themselves. And these subjective experiences, tradition tells us, have "intrinsic qualities" – qualia, in the jargon of philosophers – that I not only do have access to, but that are inaccessible to *objective* investigation. This idea has persisted for centuries, in spite of its incoherence [p. 78]

Although this passage seems to countenance subjective experiences while rejecting some of their alleged properties, it must be remembered that what heterophenomenology really countenances is not experiences themselves, but subjects' *beliefs* about experiences. In any case, in reply to the age-old view that experiences possess intrinsic subjective features or qualia, Dennett complains that the term "qualia" has never been adequately defined, and that philosophers really don't know what they are talking about when they use it. Although qualia are commonly referred to as the phenomenal characteristics of experience, this doesn't help unless a definition of "phenomenal" can be provided.

But why are dictionary definitions mandatory? Can't an adequate notion of qualia be captured ostensibly, e.g. by demonstrative references to such properties as the subjective yellow one experiences when seeing a lemon or having an after-image? Many philosophers have thought so, but Dennett disagrees. One of his major counterarguments involves problems that he thinks arise from *changes* in (putative) qualia. For example he discusses [pp. 82-91] an experiment in which subjects are shown two rapidly alternating photos of a kitchen, the only difference being that one of the pictures depicts a cabinet door that is brown rather than white. Interestingly, subjects typically require considerable time to notice the difference between the otherwise identical pictures. Once they succeed, the question posed to them by Dennett is whether their qualia changed prior to their noticing the difference (observe that this is actually a question about subjects' *memories* of qualia). There are three possible answers – (A) "Yes"; (B) "No"; (C) "I don't know" – all of which are intended to be embarrassing to qualophiles. (A) is problematic because it undermines subjects' presumptive authority about their own qualia. (C) is dismissed because it relegates qualia to "the curious position of being beyond the horizon of both third-person objective science and first-person

subjective experience” [p. 86]. And answer (B)

threatens to trivialize qualia as just logically *constituted* by your judgments or noticings, an abandonment of the other canonical requirement for qualia: that they be “intrinsic properties. [p. 86]

But although this response to answer (B) is congenial to Dennett’s reduction of conscious experiences to judgments, it seems vastly overstated. Why should I not say that I failed to notice the qualia shift because for awhile my visual experiences caused by the two pictures were the same? The suggestion here is not that qualia are “constituted by” judgments, but rather that they provide the *basis for* those judgments, and this is not intended as a hypothesis, but rather as a description of my experience. And to Dennett’s question “Why would a zombie’s judgment be any less authoritative than yours?”, the answer is that it would *not* be any less authoritative to a heterophenomenologist. But that in a nutshell is just what is wrong with heterophenomenology. Later on I will argue that, contrary to what Dennett seems to suppose, abandoning his methodology does not entail abandoning science.

Any book that takes on qualia can be expected to confront Frank Jackson’s famous example of Mary the Color Scientist, and Chapter 5 of *Sweet Dreams* takes careful aim at it. Aficionados will remember Mary as a brilliant neuroscientist who knows everything about the neurophysiology of vision, and who also knows everything physical that takes place when people see blue skies, red tomatoes, etc. For unspecified reasons Mary has been confined all her life to a monochromatic environment, gaining her knowledge through a black-and-white television monitor (perhaps she is also permitted read books, albeit none containing colored plates). Jackson maintains that, despite her mastery of all “physical information”, Mary will learn something new when she is released from her prison and sees colored objects for the first time. He concludes that physicalism is false, since Mary’s new knowledge was not captured by the physical information that she possessed prior to gaining her freedom.

Dennett’s rejoinder to this begins by reprinting (from his earlier book *Consciousness Explained*) a retelling of the Mary story. In Dennett’s variant version, experimenters try to trick Mary by displaying a bright blue banana for her very first color experience. Mary immediately replies that she can see through the deception; she knows that bananas are yellow, not blue. This, Dennett continues, amazes her captors, but as Mary explains, due to her omniscience concerning all things physical,

I already knew exactly what *thoughts* I would have I realize it is *hard for you to imagine* that I could know so much about my reactive dispositions that the way blue affected me came as no surprise. [p. 105]

Although Dennett regards his version of the tale as incompatible with Jackson’s (“two little fantasies pulling in opposition directions”), there’s no reason to accept that claim, for Mary’s antecedent knowledge of how she would respond verbally (or through other behavior) to a colored object is arguably independent of her knowing what it would be like to experience that color. Dennett, however, explicitly denies any such distinction:

I was saying that Mary had figured out, using her vast knowledge of color science,

exactly what it would be like for her to see something red, something yellow, something blue in advance of those experiences. [p. 106; see also footnote #3 on that page]

Well, yes, she had figured it out *if there's nothing more to experiences than judgments and reactive dispositions*. But that's exactly the point at issue. In any case, Dennett clearly realizes that his skeptical audience will not yet be won over, and the remainder of the chapter abounds with further examples. Here I will consider just two more of them.

In the first, "Swamp Mary" is introduced. She is like the original Mary, but just prior to her release from the monochromatic room, a

bolt of lightning puts Swamp Mary's brain into the dispositional state, the competence state, that an experience of a red rose *would have put her brain into* had such an experience (hallucinatory or not) occurred. [p. 121]

Accordingly, Swamp Mary registers no surprise when she undergoes her first experience of red. Dennett has little use for this scenario per se, but he improves it by replacing the bolt of lightning with

good hard work: Mary puts all her scientific knowledge of color to use and *figures out* exactly what it is like to see red (and green, and blue) and hence is not the least bit surprised when she sees her first rose. [p. 122]

Since some qualophiles cited by Dennett have framed Jackson's example in terms of Mary's being "surprised" by her first color experience, it's perfectly fair for Dennett to play that game as well. However, this strikes me as a distraction. What kind of "surprise" is being envisaged? Mary must have known that normal percipients (claim to) experience qualia, so the fact that she encountered a color quale – her very first – in a normal environment would not at all be unexpected. Furthermore, there's no reason to doubt that she would correctly identify a red rose, based on her prior knowledge of the reactive dispositions that would be induced in her upon seeing one. In short, one does not have to grant Dennett's reduction of experiences to judgments in order to accept his conclusion that in several important senses, hard-working Mary (or Swamp Mary) would fail to be surprised by what she encountered in her new visual environment. But this concession does not show that, while previously ensconced in their monochromatic rooms, either Mary or Swamp Mary knew what it is like to experience red. What they subsequently learned is that the experience is like *this* (where *this* is being used demonstratively to refer to my color experience as I look at a red tomato). Whether or not such knowledge is "surprising", the important point is that it cannot be captured in third-person, heterophenomenological terms. Dennett's standard retort is that this claim cannot be proven, although in one passage he evinces a slightly conciliatory attitude:

[A]re we really so sure that what it is like to see red or blue can't be conveyed to one who has never seen colors in a few million or billion words? What is it about the experience of red, or blue, that makes this task impossible? [p. 115]

The second question here makes plain that the first is not merely rhetorical, and I will come back to both of them shortly.

Dennett's final foray into the Mary thicket introduces RoboMary, a computer analogue of "her" human counterpart. In the first phase of this scenario, RoboMary is outfitted with monochromatic vision (e.g., black-and-white TV cameras), and allowed to roam around freely. However, drawing from her vast compendium of knowledge, RoboMary manages to reprogram herself to "colorize" her black-and-white input, thereby putting herself into exactly the same states that other such robots, equipped with color vision, go into straightforwardly. Thus, if RoboMary's monochromatic inputs were later replaced by color cameras, she would notice no change. In order to meet the objection that RoboMary has illicitly created self-induced color experiences (prohibited by Jackson's ground rules), Dennett replaces her with "Locked RoboMary" whose tamper-proof (!) filtering software prevents her from creating any self-induced color experiences. But this doesn't prevent the ingenious Locked RoboMary from using some of her considerable spare memory capacity to build an internal model of herself sans the filtering device; she can do this in part by studying her unlocked robotic counterparts. And after minutely studying the total response state of her model when it looks at a red tomato, Locked RoboMary programs herself to go directly into that state. Dennett writes:

But now she can know just what it is like to see a red tomato, because she has managed to put herself into just such a dispositional state – this is of course the hard-work analogue of the miraculous feat wrought by the Cosmic Accident of the lightning bolt in the case of Swamp Mary. [p. 128]

To be sure, Locked RoboMary provides a vivid illustration of Dennett's central thesis that there's nothing to phenomenal states beyond judgments or reactive dispositions. But this hardly settles the issue, since the thesis itself remains highly contentious.

As has already been noted, Dennett's reductionist position derives directly from his heterophenomenological method, an important tenet of which is that everything worthy of acceptance in a study of consciousness must be observable by third parties and expressible in third-person descriptions. He apparently thinks that an abandonment of that principle leads to some sort of dualism:

It is a mistake to inflate practical indescribability into something metaphysically more portentous, and I have been urging people to abandon this brute hunch, tempting though it may be. [p. 111]

Among the many philosophers who might be accused of having such a "brute hunch" there has been a division of opinion over the reason why first person experience cannot be captured in third-person descriptions. Some think it is a matter of principle; others regard it as a practical limitation that may in time be overcome. But it is not necessary to settle that disagreement in order to address Dennett's worry that the hunch leads to metaphysical excess. First, it should not seem surprising that each individual subject has access to his or her own experience in a way that no one else does, for each of us is an intact physical system with integrated physical components. Given the relationship between my senses, brain, and central nervous system, my experiences result from a particular set of integrated causal processes, just as everyone else's do. Moreover, since each person is a unique configuration of such causal processes, the experiences arising from (supervening upon) each system could well have aspects or qualities that are available *only* from the perspective of the organism consisting of that causal configuration. Consequently, those qualities might be

indescribable in third-person language, but this fact would itself be so describable, and a third-person physicalist science of consciousness could offer an account of why this is the case (in fact, a primitive sketch has just been provided). Dennett laments that a large number of philosophers have believed there to be some sort of gap between the subjective and the objective, and he is entirely right to reject an adjudication of this issue via undefended intuitions and votes. The point to insist on here, however, is that a subjective/objective gap need not imply anything metaphysically extravagant; in fact, it may be congruent with physicalism, and it is not tantamount to a “first-person science” which Dennett rightly disparages. This conclusion is a fairly familiar alternative to Dennett’s, and it amounts to adopting what I called option (3) at the beginning of this review. One careful and persuasive case for it may be found in John Perry’s recent book *Knowledge Possibility, and Consciousness* (my review of which appeared in Vol. 4, No. 2 of this journal: [Review of Perry](#)).

In Chapter 6 (which originally appeared as an overview of neuroscientific papers on consciousness, published in *Cognition*), Dennett turns to a discussion of what goes on in the conscious brain. As a metaphorical introduction he proposes that consciousness is like “fame” or “clout”: Various neural structures compete for control of the body, and consciousness *consists in* the emergence of victorious “attention-grabbers” that produce some lasting effects. Underlying this functionalistic characterization of consciousness are solid empirical questions about how neural structures account for “the cognitive competence that we associate with consciousness” [p. 136], and Dennett provides brief but useful descriptions of a number of recent proposals. When all this is said and done, however, the old question resurfaces: Where in all that does one find first-person, phenomenal experiences or qualia? Dennett quotes with approval from a paper by S. Dehaene and L. Naccache who write that philosophers’ belief in qualia may stem from the fact that

the flux of neuronal workspace states associated with a perceptual experience is vastly beyond accurate verbal description or long-term memory storage . . . Thus the contents of perceptual awareness are complex, dynamic, multi-faceted neural states that cannot be memorized or transmitted to others in their entirety. [p. 151]

To this Dennett adds that philosophers may choose to call the incommunicable remainder “qualia”, but in that case qualia “are just more of the same, dispositional properties that have not yet been entered in the catalog” [p. 152]. And to qualophiles who will surely reject this reduction, he offers his familiar retort that no reason has been given to believe that any further irreducible phenomenal properties exist.

The book’s penultimate chapter elaborates on the idea that consciousness is analogous to “fame”, insofar as events in the brain don’t become conscious unless they acquire some staying power. Perhaps the clearest illustration of such lasting events is the case of memory, which Dennett considers most essential to consciousness:

[Our] capacity to relive or rekindle contentful events is the most important feature of consciousness – indeed, as close to a defining feature of consciousness as we will ever find...” [p. 171]

Although Dennett acknowledges the existence of Proustian episodic memory, it turns out once again that the only thing really being acknowledged is what succumbs to a functionalistic analysis. Phenomenal consciousness – the “what-it-is-like-to-be” aspect of experience – is dismissed. Simply

having an experience (e.g., a sharp pain) wouldn't prove anything, since experiences themselves (vs. beliefs about them) can't be communicated to anyone else, and consequently don't make up any part of a subject's heterophenomenological world.

The main philosophical battle line should be clear enough by now. Dennett aptly encapsulates it in his final chapter, which contains a short imaginary debate between a qualia enthusiast ("Phil") and a heterophenomenologist ("Sy"). The participants' names are an obvious allusion to the book's subtitle.

Sweet Dreams is an energetic and accessible presentation of an interesting point of view on a vitally important philosophical problem. It is also far more radical than it may appear at first glance, given Dennett's thesis that first-person subjective consciousness does not exist except in the form of judgment or belief. The methodical precepts of heterophenomenology impose very tight ontological constraints. This may be a good antidote to the excesses of philosophers who have argued from qualia to dualism or to the ultimate unintelligibility of physicalism. However, there are alternative and credible approaches that attempt to give full-blooded acknowledgment to first-person experiences – not just to *descriptions* thereof – while at the same time not running afoul of physicalism or sound scientific principles. Some such approach, I believe, is required if a theory of consciousness is ever to “do justice to the most private and ineffable subjective experiences”.

Charles E. M. Dunlop
The University of Michigan-Flint