

# Daniel Dennett on the Nature of Consciousness

SUSAN SCHNEIDER

One of the most influential philosophical voices in the consciousness studies community is that of Daniel Dennett. Outside of consciousness studies, Dennett is well-known for his work on numerous topics, such as intentionality, artificial intelligence, free will, evolutionary theory, and the basis of religious experience (Dennett 1984, 1987, 1995c, 2005). In 1991, just as researchers and philosophers were beginning to turn more attention to the nature of consciousness, Dennett authored his *Consciousness Explained*. *Consciousness Explained* aimed to develop both a theory of consciousness and a powerful critique of the then mainstream view of the nature of consciousness, which Dennett called “The Cartesian Theater View.” In this brief discussion, I largely focus on Dennett’s influential critique of the Cartesian Theater View, as well as his positive view on the nature of consciousness, called the “Multiple Drafts Model.” In keeping with the themes of this section, I also discuss Dennett’s views on the hard problem of consciousness (Chalmers, chapter 17). As those familiar with Dennett’s views know, his work on consciousness is extensive. The reader is thus encouraged to turn to the suggested readings for further detail.

## Dennett’s Critique of the Cartesian Theater Model

Suppose that you are sitting in a café studying, right before a big exam or talk. All in one moment, you may taste the espresso you sip, feel a pang of anxiety, consider an idea, and hear the scream of the espresso machine. This is your current stream of consciousness. Conscious streams seem to be very much bound up with who you are. It is not that *this* particular moment is essential to you – although you may feel that certain ones are very important. It is rather that throughout your waking life, you seem to be the subject of a unified stream of experience that presents you as the subject, viewing the stream.

Let us focus on three features of the stream: it may seem to you, put metaphorically, that there is a sort of “screen” or “stage” in which experiences present themselves to you – to your “mind’s eye.” That is, there appears to be a central *place* where experiences are “screened” before you. Dennett calls this place “the Cartesian Theater.” Further, it seems that mental states being screened in the theater are in consciousness and that mental states outside of the theater are not in consciousness. Second, in this central place there seems to be a singular point in *time* which, given a particular sensory input, consciousness of the

input happens. For instance, there seems to be one moment in which the scream of the espresso machine begins, pulling you out of your concentration. Finally, there appears to be a *self* – that is, someone who is inside the theater, watching the show. Dennett calls this trifold view “Cartesian Materialism”:

... the view you arrive at when you discard Descartes’ dualism but fail to discard the imagery of a central (but material) Theater where “it all comes together . . .” Cartesian Materialism is the view that there is a crucial finish line or boundary somewhere in the brain, marking a place where the order of arrival equals the order of “presentation” in experience because *what happens there* is what you are conscious of. (Dennett 1991, p. 107, original emphasis)

Now, what if you are told that Cartesian Materialism is false? This is the negative or destructive ambition of *Consciousness Explained* – there is a very real sense in which our own first-person experience of consciousness leads us to Cartesian Materialism. Yet Dennett argues that given certain philosophical considerations, together with certain work in the psychology and neuroscience of consciousness, our sense of being in a Cartesian Theater is illusory.

Dennett’s critique of Cartesian Materialism can be understood against the backdrop of his own, positive view of consciousness, which he calls the “Multiple Drafts Model.” According to the Multiple Drafts Model the brain has many parallel information-processing streams (Dennett 1991, p. 111). At any point in time there are various narrative fragments, or “drafts,” which are at different stages of editing. According to Dennett, these drafts are not sent to a single place in the brain for viewing. But some or all of them may come together in the event that they need to determine a behavior for the organism. There is nothing like a Cartesian Theater, or Central Processing Unit (CPU) in the brain, in which all, or even most, commands are executed. Nor is there a viewer of such events, as they flow through the CPU. Furthermore, according to Dennett, asking “which events are conscious?” is to conceive of a Cartesian Theater in which one or more drafts comes before an audience. There is really no audience which has the experiences.

Of course, introspectively, we do have a sense of having sequences of events flowing through consciousness. Dennett does not deny this. But this sense is not due to there being a central place or time in the brain where consciousness comes together, or relatedly, to there being a self as viewer of the events, inside a Cartesian Theater. Instead, the self is a “center of narrative gravity” – a kind of program that has a persistent narrative, and in particular, “a web of words and deeds . . . The web protects it, just like the snail’s shell, and provides a livelihood, just like the spider’s web” (1991, p. 416). The sense in which there is a sequence of events in consciousness occurs when the stream is probed, for example, by asking a question. Consider the earlier example of studying in the café. Dennett would say that your consciousness of the scream of the espresso machine occurred when you probed the stream of multiple drafts at a certain point. This probe fixes the content of consciousness. On Dennett’s view, there are no facts about the stream of consciousness aside from particular probes (Dennett 1991, p. 113).

Dennett explains his Multiple Drafts Model through the example of the Phi Phenomenon, and in particular, through the color phi phenomenon. Before reading further, it is best to view the phi phenomenon for yourself by searching for “color phi” on the internet or visiting the following website: <http://www.yorku.ca/eye/colorphi.htm>. In the colored phi illusion, two differently colored lights, with an angular separation of a few degrees at the eye, are flashed

one after the other. Two interesting things happen. First, the first light appears to move across to the position of the second light. And second, the light appears to change color as it moves. For instance, in the webpage cited above, which featured a green light followed by a red one, the green light seems to turn red as it appears to move across to where the red light is.

As Dennett notes, this is quite odd. For one thing, how could the first light seem to change color *before* the second light is observed? Dennett entertains two options, both of which he discards. First, he considers the possibility that the observer makes one conclusion, and then changes her memory when she sees the second light. Dennett calls this option “Orwellian,” after George Orwell’s *Nineteen Eighty Four*, where history was constantly revised by the Ministry of Truth (Dennett 1991, p. 116). In this scenario, shortly after the second spot goes into consciousness, the brain makes up a narrative about the intervening events, complete with the color change midway through. This new event sequence is encoded into memory, and the original event sequence is not (Dennett 1991, p. 121).

He then suggests a second alternative. According to this scenario, the events are held up in the brain’s “editing room” (if you will), before they go into consciousness. More specifically, the first spot arrives in preconsciousness, and then, when the second spot arrives there, some intermediate material is created, and then, the entire, modified sequence is projected in the theater of consciousness. So the sequence which arrives at consciousness has already been edited with the illusory intermediate material (Dennett 1991, p. 120). Dennett calls this second option “Stalinesque,” after Stalin’s show trials, in which bogus testimonies were staged, and the final verdict was decided in advance (Dennett 1991, p. 117).

Dennett then asks: What reason would we have for choosing one interpretation over the other? He contends that there is no way, even in principle, to select one interpretation over the other, for there is no way to demarcate the place or time in the brain in which material goes into consciousness (Dennett 1991, pp. 126–32). He further claims that since we cannot tell which is the correct interpretation, there really is no difference between the two interpretations; we are left with a “difference that makes no difference” (Dennett 1991, p. 132). He then concludes that the (putative) fact that there is no way of distinguishing between the two interpretations lends plausibility to the Multiple Drafts Model. For according to the model, there is no concrete place or time in which material is, or is not, in consciousness.

There has been much debate over both the plausibility of the above line of reasoning and concerning what Dennett’s precise argument is. (See the extensive peer review of Dennett and Kinsbourne 1992 in *Behavioral and Brain Sciences*; Korb 1993; Robinson 1994; Seager 1999). Unfortunately, Dennett’s discussion involved heavy use of metaphor, so the underlying argument was unclear. In any case, many critics have resisted Dennett’s verificationist suggestion that if there is no way to tell between the interpretations, there is no fact of the matter (Lycan 1992; Van Gulick 1992; Korb 1993; Robinson 1994; McGinn 1995; Seager 1999). Another major source of concern has been whether there is really no difference, even in principle, between the two interpretations. Block has suggested that Dennett’s rationale for this hinges on the rejection of phenomenal consciousness (Block 1992). Indeed many have interpreted Dennett as being an eliminativist about phenomenal consciousness (Block 1992; Van Gulick 1992; Seager 1999), a position which Dennett himself has disavowed in a response to critics (Dennett & Kinsbourne 1995b, but see below). In his response to critics, he explains that the reason that the two interpretations cannot be distinguished is *not* because, in general, there is no such thing as phenomenal consciousness, but because such an extremely small timescale is involved.

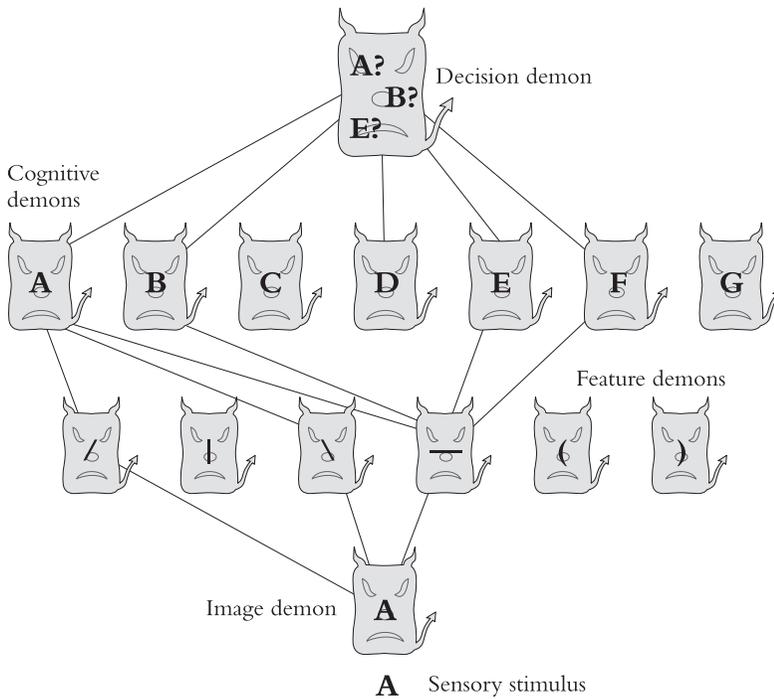
Conscious experiences are real events occurring in the real time and space of the brain, and hence they are clockable and locatable within the appropriate limits of precision for real phenomena of their type . . . Certain sorts of questions one might think it appropriate to ask about them, however, have no answers because these questions presuppose inappropriate . . . temporal . . . boundaries that are more fine-grained than the phenomenon admits. (Dennett & Kinsbourne 1995a, p. 235)

Here, the critic would probably object that in this case it is unclear why there would not be a fact of the matter about which interpretation is correct. For according to one version, even at such a small timescale, there would be conscious experience; the conscious events would simply not be remembered. In the other scenario, the conscious experience would not have occurred at all. Indeed, even if the subject herself could not report a difference because, for instance, she could not remember the experience, it seems there would be, at the very least, an in principle way to tell the difference (Korb 1993; Seager 1999). For if one sequence is held up, before entering consciousness, and the other is simply recalled differently, there would be underlying brain states which differ; otherwise, differences in mental processing would fail to supervene on physical states. No physicalist, including Dennett, would be prepared to accept this. In light of this, there should be, at least in principle, a measurable difference between the Orwellian and Stalinesque interpretations, and furthermore, such a difference may even fall in the realm of future, higher resolution, brain imaging techniques. It is only the claim that phenomenal consciousness itself does not exist, at least apart from probes, that would justify the strong conclusion that there is no difference between the two interpretations (Block 1992).

Leaving the phi illusion, let us now ask about the plausibility of the Multiple Drafts Model itself. It has been more than a decade since the Multiple Drafts Model was developed, and there are features of the model which have clearly withstood the test of time. It is widely accepted that processing in the brain is massively parallel and that there is no centrally located homunculus that views all experiences passing before it. However, it is worth mentioning that the idea of massive parallelism was certainly not original to Dennett, and even back in 1991 very few scientists believed that consciousness all came together at one place in the brain. But to fully judge the plausibility of the model, we might ask for the details of the model, because at this point in our discussion at least, we have not really laid out a model of consciousness, but an interesting contrast.

According to Dennett, consciousness is a sort of “virtual machine,” a sort of “evolved (and evolving) computer program that shapes the activities of the brain” (Dennett 1991, p. 431). But to have a model of consciousness, there needs to be an answer to the question: What sort of program is the machine running? Dennett has expressed strong sympathy with the *Pandemonium* model of Oliver Selfridge (1959), which was essentially an antecedent to connectionism. *Pandemonium* is a pattern recognition system that consists in four layers (see Figure 24.1). As the diagram illustrates, there are numerous units, called “demons.” Each of the members in the lower layers “shout” to be heard by the demons in the layer above. The second layer consists of simple feature detector demons. The “cognitive demons” in the third layer are sensitive to specific weighted features. The final layer consists in a decision-making demon that “hears” the shrieking of the layer immediately below, and decides what pattern was presented to the system (Selfridge 1959).

As Dennett surely knew, *Pandemonium* is far too simple to be a model of consciousness. But what fascinated Dennett was the parallel nature of *Pandemonium*, in which there is no central executive. Furthermore, in the case of *Pandemonium*, as with computational expla-



**Figure 24.1** Pandemonium

nation more generally, a cognitive or perceptual capacity is decomposed into extremely simple and unintelligent components. Indeed, explanation in cognitive science generally proceeds by the method of functional decomposition, a method which, put simply, explains a cognitive capacity by decomposing it into constituent parts, and specifying the causal relationships between the parts, as well as decomposing each part into further constituents, and so on (Cummins 1975). As with many in consciousness studies, Dennett is clearly opposed to functional decompositions that appeal to homuncular theories of the mind, whereby what is meant by “homuncular theories” are theories purporting to explain cognitive capacities by generating a decomposition that ultimately boils down to an internal agent, or homunculus, that has the cognitive capacity which was supposed to be explained by the decomposition in the first place. In the case of Cartesian Materialism, this homunculus is the conscious agent in the theater (Dennett & Kinsbourne 1995a, p. 85). The consciousness of the homunculus would itself need to be explained, so this sort of explanation is circular.

So, trying to further explain the Multiple Drafts Model, it appears that, in addition to appealing to massive parallelism, the model involves a kind of computational functionalism without a homunculus. While this sort of view has been regarded by many as a plausible approach to explaining cognitive capacities, an appeal to non-homuncular computational functionalism does not really make Dennett’s view distinctive, for computational functionalism is common throughout cognitive science. Nor would a mere appeal to functionalism do the needed theoretical work of serving to justify Dennett’s view that there is no way to differentiate between Orwellian or Stalinesque revisions; for the decompositions of the mental processes underlying Orwellian and Stalinesque accounts of the color phi

phenomenon would likely differ (see below). Furthermore, we do not yet have a model of consciousness, for although there is an appeal to the method of functional decomposition, not even the most basic functional decomposition of consciousness has been offered.

But perhaps the following details would yield the needed model, and separate Dennett's model from a generic appeal to functional decomposition. While Dennett shied away from proposing a particular theory of consciousness in *Consciousness Explained*, he expressed sympathy with the Global Workspace (GW) theory of consciousness, and the closely related Global Neuronal Workspace theory of consciousness, and he has recently re-emphasized his alliance with this position (Dennett 1991, 2001). To keep things simple, I will refer to both views as "the GW theory" as they are closely aligned and agree in the dimensions discussed herein. According to the GW theory, the role of consciousness is to facilitate information exchange among multiple parallel specialized unconscious processes in the brain. Consciousness is a state of global activation in a "workspace" in which information in consciousness is "broadcast" back to the rest of the system. At any given moment, there are multiple parallel processes going on in the brain which receive the broadcast. Access to the global workspace is granted by an attentional mechanism and the material in the workspace is then under the "spotlight" of attention. When in the global workspace the material is processed in a serial manner, but this is the result of the contributions of parallel processes which compete for access to the workspace. Introspectively, this seems intuitive, as our conscious, deliberative, thoughts appear to be serial (Baars chapter 18; Baars 1997; Dehaene & Changeux 2004; Dehaene & Naccache 2001; Shanahan & Baars 2005).

At least at first, there are commonalities between the GW theory and the Multiple Drafts Theory. The appeal to massive parallelism is in keeping with the Multiple Drafts Model. And one might find GW theory somewhat reminiscent of Pandemonium: as Dennett describes it, information is sent to the workspace when "demons" competing for access to the workspace "shout" sufficiently loudly to be granted access (Dennett 1991, p. 191). In a recent *Cognition* paper Dennett discusses the GW view, stressing its affinity with the Multiple Drafts Model:

... the specialist demons' accessibility to *each other* (and not to some imagined higher Executive or central Ego) ... could in principle explain the dramatic increases in cognitive competence that we associate with consciousness: ... This idea was also central to what I called the Multiple Drafts Model (Dennett 1991), which was offered as an alternative to the traditional, and still popular, Cartesian Theater model, which supposes there is a place in the brain to which all the unconscious modules send their results for ultimate conscious appreciation by the Audience. The Multiple Drafts Model did not provide, however, a sufficiently vivid and imagination-friendly antidote to the Cartesian imagery we have all grown up with, so more recently I have proposed what I consider to be a more useful guiding metaphor: "fame in the brain" or "cerebral celebrity." (Dennett 2001)

Interestingly, in a different place in the *Cognition* paper Dennett goes as far as likening the Global Neuronal Workspace model to "fame in the brain" (Dennett 2001). This does seem to suggest that Dennett thinks of the GW theory as filling in the details of his model, or at least bearing important similarities to it.

So perhaps now we are equipped to return to the question of the plausibility of the Multiple Drafts Model. Many philosophers and scientists find the GW theory to be a promising informational theory of consciousness, although philosophers sympathetic to problems involving qualia may urge that while it might be a promising scientific theory of the infor-

mation processing involved in consciousness, by itself, it does not answer the hard problem of consciousness (Chalmers 1995). But let us set aside the hard problem for the moment, and pose the question: Would an alliance between Dennett's Multiple Drafts Model and the GW view finally provide the promised model of consciousness? Unfortunately, while the Global Workspace theory might provide the beginnings of an information-processing model of consciousness, there are significant points of tension between it and the Multiple Drafts Model. For one thing, the GW theory has been categorized as a kind of theater model (Blackmore 2004, p. 72). This may seem surprising, for according to the GW theory, consciousness is a highly distributed activity in the cortex, so there is no single spatiotemporal location in the brain where consciousness comes together. But Baars himself makes heavy use of theater metaphors, describing conscious events as happening in "the theatre of consciousness" and in the "screen of consciousness" (Baars 1997, p. 31). Are these metaphors merely misleading descriptions on Baars's part? It appears not, for according to the GW view, there is a definite sense in which certain mental states are *in* consciousness, while others are not: states are conscious when they are in the global workspace (Baars 1997 and chapter 18). This point of contrast is sharpened by reconsidering the color phi case. If one asks the question, "Which account of the color phi illusion is correct, the Orwellian or Stalinesque account?", the Global Workspace theory offers an (at least in principle) route to an answer. The question becomes: Did the subsystem that processed the first light broadcast the state into the GW, before the second light was processed by the subsystem, or was the broadcast held up, until the second light was processed?

There are two further sources of tension as well, the first concerning what the contents of consciousness are, the second concerning the role of a central system. First, the GW view does not seem to require a probe for a state to be broadcast into the workspace; what is conscious is not determined by what is probed. So the contents of consciousness will differ according to each theory. Second, although mental states are not processed in anything like a CPU, the global workspace has been likened to a central system, that is, a nonmodular, "horizontal," system in which material from different sense modalities comes together, and in which deliberation and planning occurs. (Here, it is important to note that a central system is not identical to a CPU. Again, a central system is a subsystem of the brain that integrates material from different modalities; a CPU, on the other hand, is a command center that executes every, or nearly every, command in a computational system.) As Stanislas Dehaene and Jean Pierre Changeux explain:

The model emphasizes the role of distributed neurons with long-distance connections, particularly dense in prefrontal, cingulate, and parietal regions, which are capable of interconnecting multiple specialized processors and can broadcast signals at the brain scale in a spontaneous and sudden manner. The concept of a "global neuronal workspace" . . . builds upon Fodor's distinction between the vertical "modular faculties" and a distinct "isotropic central and horizontal system" capable of sharing information across modules. (Dehaene & Changeux 2004)

Dehaene and Changeux's claim that the GW view builds upon Fodor's notion of a central system is particularly noteworthy, for it emphasizes that the GW theory has an element of centralization that Dennett has disavowed in the context of his Multiple Drafts Model. Indeed, the appeal to a central system by advocates of the GW theory is not limited to the work of Dehaene and Changeux. For instance, in a 2005 *Cognition* paper, Shanahan and Baars have offered a solution to one version of the frame problem (which arises for systems that involve a central processor) that appeals to the GW theory, presenting details

concerning the cognitive architecture of the GW that capture a clear sense in which the workspace operates as a Fodorian “horizontal,” or central system (Shanahan & Baars 2005).

But is this sort of centrality really compatible with a view like Dennett’s, which vehemently disavowed “central Headquarters” or “Central Meaners”? Indeed, Fodor’s central system is the subject of the following attack in *Consciousness Explained*:

By giving this central facility so much to do, and so much nonmodular power with which to do it, Fodor turns his modules into very implausible agents, agents whose existence only makes sense in the company of a Boss agent of ominous authority . . . Since one of Fodor’s main points in describing modules has been to contrast their finite, comprehensible, mindless mechanicity with the unlimited and inexplicable powers of the nonmodular center, theorists who would otherwise be receptive to at least most of his characterization of modules have tended to dismiss his modules as fantasies of a crypto-Cartesian. (1991, p. 261)

In light of the GW appeal to centrality, Dennett’s remarks underscore a clear source of tension between GW and the Multiple Drafts Model. It seems fair to say that Dennett’s Multiple Drafts Model faces the following dilemma: either, as noted, it lacks sufficient theoretical detail to be a genuine model of consciousness, or it borrows from the GW theory to yield the needed detail. However, given the points of tension, Dennett cannot incorporate GW detail into his theory.

### **Some Further Features of Dennett’s Views on Consciousness**

As noted in the introduction, Dennett’s contributions to consciousness studies are quite extensive. In addition to offering a critique of Cartesian Materialism and the Multiple Drafts Model, Dennett has attacked certain thought experiments, problems, and arguments that are commonly thought to lend support to the idea that consciousness is an irreducible feature of the world, going beyond the physical realm that science investigates. The hard problem of consciousness is the problem of why, in addition to the information processing that the brain engages in, there must be a feeling of what it’s like associated with the neural processing. For how can conscious experience emerge from the gray matter of the brain? (Chalmers 1995 and chapter 17). It is fair to say that the hard problem has been regarded as one of the most central philosophical puzzles about the nature of consciousness. Dennett has been a well-known critic of the hard problem: “Chalmers’ (1995) attempt to sort the ‘easy’ problems of consciousness from the ‘really hard’ problem is not, I think, a useful contribution to research, but a major misdirector of attention, an illusion-generator” (Dennett 1996).

However, if one finds the hard problem to be a rich and compelling problem, Dennett’s Multiple Drafts Model does not yield a satisfying answer. For, as discussed, what Dennett is ultimately defending is a sort of computational functionalism, together with the view that consciousness is a highly distributed activity in the brain. And one could still look to a penultimate functional decomposition of consciousness and ask: But why does experience need to exist, in addition to all this information processing? However, in his 1996 paper, Dennett offers three responses to those who find the hard problem compelling. First, in the context of a comparison between the hard problem and what he called “the hard question for vitalism,” he asks us to imagine a vitalist who says:

The easy problems of life include those of explaining the following phenomena: reproduction, development, growth, metabolism, self-repair, immunological self-defense, . . . These are not all that easy, of course, and it may take another century or so to work out the fine points, but they are easy compared to the really hard problem: life itself. We can imagine something that was capable of reproduction, development, growth, metabolism, self-repair and immunological self-defense, but that wasn't, you know, *alive*. (Dennett 1996)

Dennett's vitalist believes that being alive is something above and beyond the other phenomena. Clearly, the vitalist is mistaken. And according to Dennett, those who find the hard problem compelling are making a similar mistake. For their view is that the explanation of functions is not sufficient to explain experience. However, "it is precisely the 'remarkable functions associated with' consciousness that drive them to wonder about how consciousness could possibly reside in a brain" (Dennett 1996).

Dennett further contends that positing something beyond functions is a form of humuncularism. Francis Crick, at the close of his book on consciousness, wrote, "I have said almost nothing about qualia – the redness of red – except to brush it to one side and hope for the best" (1994, p. 256). In light of Crick's remark, Dennett asks us to imagine a neuroscientist, whom he calls "Crock," who substitutes "perception" for "qualia" in the above quotation, saying instead: "I have said almost nothing about perception – the actual analysis and comprehension of the visual input – except to brush it to one side and hope for the best" (Dennett 1996). Dennett points out that Crock is, of course, in error, creating a "hard" problem of perception (Dennett 1996). Analogously, claims Dennett, Crick makes a similar error, because he thinks he can make progress on the easy problems without making progress on the hard problem.

I make the parallel claim about the purported "subjective qualities" or "qualia" of experience: if you don't begin breaking them down into their (functional) components from the outset, and distributing them throughout your model, you create a monster – an imaginary dazzle in the eye of a Cartesian homunculus. (Dennett 1996)

Finally, Dennett charges that Chalmers's claim that consciousness is fundamental is not justified, as the decision to take physical properties as fundamental involves an appeal to independent evidence. However, in the case of consciousness, there is no independent motivation. "It is a belief in a fundamental phenomenon of 'experience'" (Dennett 1996). Dennett charges that this sort of reasoning is circular, comparing Chalmers's proposal to what he calls "cutism": the proposal that

since some things are just plain cute, and other things aren't cute at all – you can just see it, however hard it is to describe or explain – we had better postulate *cuteness* as a fundamental property of physics alongside mass, charge and space-time. (Dennett 1996)

Concerning Dennett's first point, Chalmers has denied that the analogy with vitalism holds. In the case of the problem of life, it is clear that the only thing that needs explaining is structure and function. There is no further property, beyond reproduction, metabolism, adaptation, and so on, that requires explanation. According to Chalmers this is not analogous to the case of consciousness, as what seems to need explanation is experience, and the general view is that experience seems to outrun the functions (Chalmers 1996). Chalmers then asks for a non-question begging argument for the conclusion that function exhausts

the nature of consciousness. In light of our discussion of Dennett's third problem above, it appears that both sides believe that their opponent's assertion concerning whether the functions of consciousness are exhaustive is question begging. Dennett, for instance, asks for independent evidence, in his third point above. Chalmers, on the other hand, believes that consciousness is a phenomenon that needs explaining in its own right (Chalmers 1996). "And if it turns out that it cannot be explained in terms of more basic entities, then it must be taken as irreducible" (Chalmers 1996).

Are we thus at a dialectical stalemate? The burden of argument does indeed seem to fall on Dennett, for he is denying the commonplace view that experience seems to outrun function. As Chalmers notes, "Such *prima facie* intuitions can be overturned, but to do so requires very solid and substantial argument. Otherwise, the problem is being 'resolved' simply by placing one's head in the sand" (Chalmers 1996). Dennett does attempt to make his case in the context of a discussion of the second argument considered above. Here, Dennett appeals to his own phenomenology, claiming that introspectively, only functions need explaining (Dennett 1996). But as Chalmers has observed, the various mental states that Dennett raises, for example, "feelings of foreboding," "fantasies," "delight and dismay," are not at all clearly functional issues (Chalmers 1996). Why would the functions associated with experience be all that needs to be explained? Further argumentation does seem to be required.

In addition to being a vocal critic of the hard problem, Dennett has offered related concerns with philosophical discussions about the possibility of "zombies." Consider a thought experiment in which there is a molecule for molecule duplicate of you. That is, imagine a creature which has your precise neural configuration, having all the same internal brain (and other bodily) states, and which evinces precisely the same behaviors as you do when put in the same situations. And suppose that this duplicate inhabits a world much like ours except, in this case, that the duplicate lacks consciousness. That is, the duplicate has the same type of neural states as you do, including those which are invoked in our best scientific theories of the nature of consciousness, yet the duplicate lacks inner experience, or what philosophers have called "qualia." Some find such a scenario to be conceivable and indeed possible, supporting the dualist position that a state of being conscious is not essentially a physical or neural state (Kirk 1974a, 1974b; Chalmers 1996). That is, it is possible, at least in some broad sense of possibility, that there are duplicates which lack qualia. Such have been called "zombies" by philosophers (where such are understood to be different from Haitian zombies).

Dennett has argued vehemently against this view. "It is an embarrassment to our discipline that what is widely regarded among philosophers as a major theoretical controversy should come down to whether or not zombies (philosophical zombies) are possible/conceivable" (1995d, p. 325). Furthermore, he claims that philosophers discussing this issue frequently assume that there is some physical or behavioral difference between zombies and conscious humans. Zombies are physical and behavioral duplicates (Dennett 1995d; Chalmers 1996). Doing so would indeed be erroneous. Dennett also contends that there has been no plausible argument that zombies differ from humans which employs the correct conception of what zombies are; those who think zombies are possible have not really imagined them correctly. "My conviction is that the philosophical tradition of zombies would die overnight if philosophers ceased to misimagine them, but of course I cannot prove it *a priori*" (1995d, p. 325).

A further feature of Dennett's extensive views on phenomenal consciousness deserves mention as well. Many have construed Dennett as being an eliminativist about phenom-

enal consciousness (Block 1992; Seager 1999; Velmans 2006). Dennett and his coauthor Kinsbourne have denied this charge, as the earlier quoted passage indicates. As it turns out, Dennett and his critics may have been talking past each other, for there is both a sense in which Dennett is an eliminitivist about phenomenal consciousness and a sense in which he is not. The expression “qualia” has commonly been used by philosophers to denote the qualitative features of experience; the feelings of “what it’s like” (Nagel 1974) associated with experience, such as the taste of chocolate, the experience of seeing bright red, or the sound of a clarinet. In our discussion of the phi illusion, it was noted that Dennett denied being an eliminitivist about phenomenal consciousness. However, Dennett has in fact argued for eliminitivism about qualia (Dennett 1993) where by “qualia” he has in mind a narrower construal of qualia than the more generic view sketched above. According to this more specific conception of qualia, qualia are the intrinsic, ineffable, private, features of mental states of which we are immediately or directly aware (Dennett 1993). Dennett has argued through the use of extensive thought experiments that there is nothing which satisfies this description; hence, he is an eliminitivist about qualia, where qualia are understood in this more specific sense (Dennett 1993). However, this view is in fact compatible with the reality of qualia, when construed in the more general sense (Tye 2003; Van Gulick, chapter 30).

*See also 17 The hard problem of consciousness; 18 The global workspace theory of consciousness; 20 Representationalism about consciousness; 29 Anti-materialist arguments and influential replies; 30 Functionalism and qualia.*

## Further Readings

- Dennett, D. C. (1993) Quining qualia. *Readings in Philosophy and Cognitive Science*, 381–414. Cambridge, MA: MIT Press.
- Dennett, D. C. (1991) *Consciousness Explained*. Boston, MA: Little, Brown and Company.
- Dennett, D. C. (2005) *Sweet Dreams: Philosophical Obstacles to a Science of Consciousness*. New York: MIT Press.
- Seager, William (1999) *Theories of Consciousness: An Introduction and Assessment*, chapters 4 and 5, 85–131. New York: Routledge.
- Velmans, Max (2006) Heterophenomenology versus critical phenomenology. *Phenomenology and the Cognitive Sciences* 5: 3/4 (in press).

## References

- Baars, B. J. (1997) *In the Theater of Consciousness: The Workspace of the Mind*. New York: Oxford University Press.
- Blackmore, Susan (2004) *Consciousness: An Introduction*. New York: Oxford University Press.
- Block, Ned (1992) Begging the question against phenomenal consciousness. *Behavioral and Brain Sciences* 15, 205–6.
- Chalmers, David (1995) Facing up to the problem of consciousness. *Journal of Consciousness Studies* 2, 3, 200–19.
- Chalmers, David (1996) Moving forward on the problem of consciousness. *Journal of Consciousness Studies* 4, 1, 3–46.
- Cummins, Robert (1975) Functional analysis. *Journal of Philosophy* 72, 741–65.

- Dehaene, Stanislas and Changeux, Jean-Pierre (2004) Neural mechanisms for access to consciousness. In Michael Gazzaniga et al., *The Cognitive Neurosciences III*, 3rd edn. 1145–59. Cambridge, MA: MIT Press.
- Dehaene, S. and Naccache, L. (2001) Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition* 2, 79.
- Dennett, D. C. (1984) *Elbow Room: Varieties of Free Will Worth Wanting*. Cambridge, MA: MIT Press.
- Dennett, D. C. (1987) *The Intentional Stance*. Cambridge, MA: MIT Press.
- Dennett, D. C. (1991) *Consciousness Explained*. Boston, MA: Little, Brown and Company.
- Dennett, D. C. (1993) Quining qualia. *Readings in Philosophy and Cognitive Science*, 381–414. Cambridge, MA: MIT Press.
- Dennett, D. C. and Kinsbourne, M. (1992) Peer review in *Behavioral and Brain Sciences* 15, 183–247.
- Dennett, D. C. and Kinsbourne, M. (1995a) Time and the observer: the where and when of consciousness in the brain. *Behavioral and Brain Sciences* 15, 2, 183–247.
- Dennett, D. C. and Kinsbourne, M. (1995b) Escape from the Cartesian theater. *Behavioral and Brain Sciences* 15, 234–46.
- Dennett, D. C. (1995c) *Darwin's Dangerous Idea: Evolution and the Meanings of Life*. New York: Simon and Schuster.
- Dennett, D. C. (1995d) The unimagined preposterousness of zombies. *Journal of Consciousness Studies* 2: 4, 322–6.
- Dennett, D. C. (1996) Facing backwards on the problem of consciousness. *Journal of Consciousness Studies* 3, 1, 4–6.
- Dennett, D. C. (2001) Are we explaining consciousness yet? *Cognition* 79, 221–37.
- Dennett, D. C. (2005) *Sweet Dreams: Philosophical Obstacles to a Science of Consciousness*. New York: MIT Press.
- Kirk, R. (1974a) Sentience and behaviour. *Mind* 83, 43–60.
- Kirk, R. (1974b) Zombies v. materialists. *Proceedings of the Aristotelian Society*, supplementary vol. 48, 135–52.
- Korb, Kevin (1993) Stage effects in the Cartesian theater: review of consciousness explained. *Psyche* 1, 4.
- Lycan, William (1992) UnCartesian materialism and Lockean introspection. *Behavioral and Brain Sciences* 15, 216–17.
- McGinn, Colin (1995) Consciousness evaded: comments on Dennett, *Philosophical Perspectives* 9: AI, *Connectionism and Philosophical Psychology* 241–49.
- Nagel, T. (1974) What is it like to be a bat? *Philosophical Review* 83, 435–50.
- Robinson, W. (1994) Orwell, Stalin, and determinate qualia. *Pacific Philosophical Quarterly* 75, 151–64.
- Seager, William (1999) *Theories of Consciousness: An Introduction and Assessment*. New York: Routledge.
- Selfridge, Oliver (1959) Pandemonium: a paradigm for learning. In *Symposium on the Mechanization of Thought Processes*. London: HM Stationery Office.
- Shanahan, Murray and Baars, Bernie (2005) Applying global workspace theory to the frame problem. *Cognition* 98, 157–76.
- Tye, Michael (2003) Qualia. In Edward N. Zalta (ed.), *Stanford Encyclopedia of Philosophy* (Summer 2003 edn.), <<http://plato.stanford.edu/archives/sum2003/entries/qualia/>>.
- Van Gulick, Robert (1992) Time for more alternatives. *Behavioral and Brain Sciences* 15, 228–9.
- Velmans, Max (2006) Heterophenomenology versus critical phenomenology. *Phenomenology and the Cognitive Sciences*. 5: 3/4 (in press).