Abstract

In his recent (2005) book "Sweet Dreams: philosophical obstacles to a science of consciousness," Dennett renews his attack on a philosophical notion of qualia, the success of which attack is required if his brand of Functionalism is to survive. He also articulates once again what he takes to be essential to his notion of consciousness. I shall argue that his new, central argument against the philosophical concept of qualia fails. In passing I point out a difficulty that David Rosenthal's "higher-order thought" theory of consciousness also faces in accounting for qualia. I then contrast Dennett's newest account of consciousness with interestingly different conceptions by contemporary neuro-scientists, and I suggest that philosophers should take the recent suggestions by neuro-scientists more seriously as a subject for philosophical investigation.

For a long time philosophers have tried to understand the notions of the self, the mind, the body, consciousness, and self-consciousness. In the post World War II period, under the influence of Gilbert Ryle, Ludwig Wittgenstein, and Sir Peter Strawson, then of Alan Turing and Hilary Putnam, and since the mid-1970's, as philosophy of mind became an increasingly popular

---

1 I am indebted to David Rosenthal, Jason Linder, and John Francis Walter for suggestions and comments. I am indebted to Prof. Dr. John Nerbonne, Dr. Laurie Stowe, to Prof. Dr. Jack Hoeksema, and to Rector Magnificus Dr. Frans Zwarts for their support and hospitality during my May 2005 research in the University of Groningen, The Netherlands.
professional subject, with the participation of contemporary philosophers like Thomas Nagel, Daniel Dennett, Ned Block, David Rosenthal, William Lycan, Colin McGinn, Jerry Fodor et al., philosophers in the United States and Great Britain have attempted to construct an account of the conscious mind that is anti-Cartesian and, roughly speaking, “naturalistic.” By “anti-Cartesian” I mean a view that rejected the Cartesian claim that the mind and the body were different “substances,” that, in Gilbert Ryle’s famous phrase, the mind was the “ghost in the machine.” On Ryle’s view of the late 1940’s, mental properties were no longer purely “inside” the mind, since he viewed thought, sensation, and emotions as necessarily manifested in actual doings of the person -- and viewed character traits like “being brave” not merely as a mental item embedded in the mind of the Canadian soldier but as a *disposition to act* in the Normandy hedgerow or on the streets of Groningen in a certain way *when* the occasion demanded. On this view Wittgenstein, Ryle, and Heidegger all agree. By “naturalistic” I mean a view that is compatible with a scientific world-view, one articulated by fundamental theories of physics and evolutionary biology. In the 1950’s in the United States, this meant a view that was compatible with Behaviorism in the psychology of Watson and B.F. Skinner, and then, with the Cognitive Revolution of the late 1950’s, a view that was compatible with a theory of mind that modeled mental states as computational states of a Turing Machine, a view first articulated by Hilary Putnam in 1960-67. The inputs to the Mind as Computer were perceptual states: information about the external world and about one’s own body; the transition rules of the Machine Table modeled causal relationships among beliefs, desires, and intentions to act that were the inner states, and the outputs were bodily states and events: movements that were essential to actions in the world. This “software” view of the mind identified mental states with Turing Machine states, causal connections with state-transitions, perceptual experience with inputs and behavior with outputs.

A somewhat looser description of the mind, derived from the original Turing Machine model, which identified mental states like belief as those “black box” states that mediated perceptual inputs with behavioral outputs through internal connections with other “black box” states, like desires, preferences, and intentions, a mediation that was “causal” if the relationships were to be “naturalistic,” became known as Functionalism. Daniel Dennett, in his April 2005 book *Sweet Dreams: philosophical obstacles to a science of consciousness* (Cambridge: MIT Press), summarizes the view in the English idiom “Handsome is as handsome does.” Functionalism remains the dominant view of the mind among philosophers in the English-speaking world, though there are internal debates about the adequacy of Functionalism to explain sensations, “non-conceptual” mental contents, and animal “sentience” or awareness, what Ned Block (1995) calls ‘phenomenal consciousness’ as contrasted with “access consciousness.” This is the famous problem for Functionalism of “qualia,” a term of art for philosophers, the exact theoretical sense and reference of which remains uncertain. Qualia are traditionally described as those features of our experience that we know by direct acquaintance with them: the taste of pineapple, the redness of a rose. Traditional views claim that they are essentially and solely first-person, ego-centric properties, subjective features that each of us is immediately aware of in his or her own experience. Since Functionalism inherits from

---

3 David Rosenthal’s (1986, 1991, 2002) view of consciousness as a second-order property, i.e. if I consciously believe that *p*, then I believe that I believe that *p*, suggests to him that what we know when we “know directly” the taste of pineapple only justifies the claim that it *seems* to us that we know directly the taste of pineapple. A commonsense description of my experience in which the word ‘seems’ arises is this: If I am awake and functioning
Behaviorism an attempt to be “scientific,” and to connect inner states with outer behavior or facts about the world, it tries to characterize such a mental state as belief by its causal consequences in observable behavior, and by its being the effect of other beliefs and perceptions, as much as by one’s introspective awareness of its content and the meta-belief that one has the belief in question.

Daniel Dennett’s version of Functionalism tries to be even more anti-Cartesian and

normally, and I am “tasting-pineapple”, i.e. experiencing a pineapple taste (normally the result of having a piece of pineapple in my mouth), then (potentially) I (consciously) believe that I am tasting-pineapple. Whatever the states involved in recognizing what the taste in my mouth is, there is no causal necessity to my forming a conscious belief, but I could form one. I experience the pineapple taste; someone asks me, ‘Are you sure that it’s the taste of pineapple?’, and I reply ‘Yes, I’m sure’. If I am asked, ‘Could it be the taste of a mutant guava?’, I reply ‘I don’t know what a mutant guava tastes like, but this thing in my mouth certainly has a pineapple-taste to me’. If asked ‘Does it seem to you that you are tasting-pineapple?’, I would understand the questioner to be conversationally implicating that what I have in my mouth is not a piece of pineapple but something that has been artificially flavored to taste like pineapple, and his question is how close to the taste of pineapple is the current flavor. If my answer is ‘So close that I cannot detect a difference between this and how I remember pineapple tasting’, then I assert felicitously ‘It seems to me that I am tasting pineapple’. Now let us consider a commonsense description in which the word ‘know’ is used for tastes: If I am aware and functioning normally, and I am tasting-pineapple, then if someone asks me, ‘Do you know whether you are tasting pineapple?’, the questioner conversationally implicates that it might not be pineapple that is causing the taste, or he implicates that my tasting and/or recognizing abilities are in doubt. If, as Tom Lehrer once put it, “my taste-buds were shot off in the war,” there might be reason to doubt my tasting abilities. But I could at least say felicitously ‘It seems to me that I am tasting pineapple’. Would I say, ‘It seems to me that I know directly that I am tasting pineapple’? No; that is philosophers’ jargon. I do not even know what that English sentence means. Would I say ‘It seems to me that I know directly the taste of pineapple’? If ‘know directly’ in that last question just means ‘am aware of’, the sentence ‘It seems to me that I am aware of the taste of pineapple’ is semantically interpretable, but it is semantically redundant. It is grammatically distinct from ‘I am aware of the taste of pineapple’, but in my speech it is a complicated paraphrase of it. So Rosenthal’s suggestion that there is an evidentiary difference between ‘It seems to me that I know directly the taste of pineapple’ and ‘I know directly the taste of pineapple’ cannot be true of the ordinary, ‘aware’ meanings of these sentences; if there is such a difference, it is one that is posited by an epistemological or psychological theory that adopts these special forms of sentence to express it. The natural guess to make about this jargon is that for Rosenthal awareness of a taste is either apparent or real, as in the remark ‘You are not really aware of that taste; you just think you are’, where ‘taste’ refers to an intentional object. I do not think I understand the difference between real and apparent awareness when one is awake, so not dreaming, and functioning normally. Dennett, in Consciousness Explained (1991: 96), says, “You are not authoritative about what is happening in you but only about what seems to be happening in you, and we are giving you total dictatorial authority over the account of how it seems to you...” By this Dennett merely means that as objective, third-person data he accepts first-person reports from a subject of experience. In this Dennett (2005: 147) does not differ from Benjamin Libet (2004); see the discussion of Libet later in this paper. Rosenthal wants to save Dennett from himself by interpreting Dennett as plausibly attacking “direct knowledge” of the pineapple taste and by suggesting that “seeming direct knowledge” of the pineapple taste would not be open to Dennett’s objections. For the reasons I have given about the ‘seems’ locution, I cannot see how to make Rosenthal’s suggestion work. Rosenthal also holds a view in which it would be true to say ‘You are really tasting-pineapple; you just are not aware that you are’. Awareness is a conscious state, so on Rosenthal’s view of conscious states as second-order states, we could paraphrase the sentence by ‘You are really tasting-pineapple; you just do not believe (are not aware) that you are aware that you are’. The latter analyses sentence might be true! The former analyses sentence? On the face of it, it seems false. Conclusion: Rosenthal’s (2002) theory of conscious states as second-order states is incorrect for qualia, or I have given the wrong, second-order paraphrase. See D. Rosenthal, “Two Concepts of Consciousness,” Philosophical Studies 94: 329-59 (1986), rev. version in D. Rosenthal, “Two Concepts of Consciousness,” in D. Rosenthal (ed.), The Nature of Mind (New York: Oxford Univ. Press, 1991), pp. 462-77, and David Rosenthal, “Explaining Consciousness,” in D. Chalmers (ed.), Philosophy of Mind: classical and contemporary readings (New York: Oxford Univ. Press, 2002), pp. 406-21.
scientific than the usual Functionalism. Rather than trying to find some account of the first-person, subjective “qualia” within a third-person, objective scientific perspective, Dennett (2005: 177) holds that if qualia are defined to be intrinsic properties of experiences defined independently of dispositional properties, of properties like ‘being brave’, ‘being kind’, or ‘being intelligent’ -- the latter being characterized in part by their causes and effects, just as ‘soluble’ as a property of salt is so characterized by dissolving if put in water or ‘fragile’ as a property of a delicate crystal vase is so characterized by breaking if dropped on a hard floor from a modest height in the earth’s local gravitational field -- then, on Dennett’s view, there are no phenomenal properties of experience, no qualia.

Dennett’s view might be summarized as follows: Only dispositional properties, and their normal effects (Dennett 2005: 142), are mental properties. Among the dispositions involved in “being intelligent,” for example, are saying intelligent things if one is in the appropriate contexts of utterance. One cannot be an intelligent creature without being an intelligent utterer of sentences, or at least engaging in intelligent reflection on the contents of one’s experience. To do that requires memory. And the ability to engage in reflection, recollection, remembering, or the reliving of experience does not arise innately; we get it from being social, language-using animals. Non-humans brains are not capable of this kind of memory, and they do not fill skulls of creatures that have social lives that produce reflection, and so are not, on Dennett’s (2005: 169) view, conscious.

Notice that Dennett’s notion of consciousness is essentially a notion of self-consciousness. Phenomenal awareness, or sentience with a first-person point of view, is not only something that we humans do not have, if it is something independent of functionally describable, dispositional mental properties, it is something that nothing has. Qualia, in the technical philosophical sense of being essentially subjective, first-person features of experience, are out – verboten.

Needless to say, this is not a popular view. As Dennett (2005: 91) reports his conversation with the American philosopher Wilfrid Sellars, “But Dan, qualia are what make life worth living.” Almost anyone would think that Dennett is crazy to deny the existence of how things feel, taste, or smell to someone, but, of course, that is not what he is denying. He is denying that these feels, tastes, or smells are essentially subjective, solely first-person, non-dispositional properties of one’s experience. That philosophical concept of qualia, he thinks, is incoherent. Why should he be a skeptic about this? I mean, apart from his philosophical bias in favor of his teacher Gilbert Ryle’s attack on the Cartesian Soul or on Hume’s Theater of the Mind in which the Self sits in the dark watching the mental happenings on the Stage or Screen, why should Dennett doubt the existence of phenomenal consciousness, that, in Thomas Nagel’s phrase, there is something it is like for me to taste pineapple? Even respectable cognitive scientists like Damasio, in his 1999 book A Feeling for What Happens (New York: Harcourt Brace), or Dahaene and Naccache (eds.) in The Cognitive Neuroscience of Consciousness (2001: MIT Press), seem to accept a notion of phenomenal consciousness, a brute, non-self-consciousness, as in passages like this quoted from Dahaene and Naccache (by Dennett 2005: 151):

---

4 A reductionist like Quine (1960) thinks that dispositional properties are just “promissory notes” to be redeemed by a future neuroscience, which would identify them with occurrent physical properties of the organism. I am leaving Dennett neutral on this point.
The flux of neuronal workspace states associated with a perceptual experience is vastly beyond accurate verbal description or long-term memory storage. Furthermore, although the major organization of this repertoire is shared by all members of the species, its details result from a developmental process of epigenesis and are therefore specific to each individual. Thus the contents of perceptual awareness are complex, dynamic, multi-faceted neural states that cannot be memorized or transmitted to others in their entirety. These biological properties seem potentially capable of substantiating philosophers’ intuitions about the “qualia” of conscious experience, although considerable neuroscientific research will be needed before they are thoroughly understood.

Dennett thinks he can spare all of you the trouble of doing the research, which he believes, would be like asking you to do research on witches and werewolves. If the question is whether in their entirety one can communicate the contents of perceptual states, and one calls the incommunicable contents ‘qualia’, then he admits that there are qualia. But he asserts – without argument or evidence – that these bits of content are just the familiar dispositional mental properties that he is willing to accept and that his Functionalism can explain. And if these bits of content are neither explained as causes of mental states (e.g. activation in a cortical region) nor explained as effects of mental states (e.g. reactions to particular colors, verbal reports, effects on memory), then they are indeed intrinsic, non-dispositional mental properties of experiences, but, Dennett (2005: 152) believes, there is no reason to think that there are such properties.

He (2005: 154) believes that neuroscientists are only willing to agree with philosophers who deny Dennett’s Functionalism and who believe that there are “qualia,” because the neuroscientists have come to believe that the kind of Turing Machine Functionalism that was once called by John Searle ‘strong AI’ cannot be right. And it cannot be right because Strong AI denies that neurochemistry matters. Strong AI believed that silicon chips would do as well as neurotransmitters across the synaptic gap.

In this University in June of 1995 I gave a lecture “What is it like to be a Chinese Room?”, sponsored by the BCN, in which I argued, in support of John Searle’s criticisms of Strong AI, that neurochemistry mattered. Now, ten years later, Daniel Dennett finally agrees. I also argued against Daniel Dennett’s version of Functionalism, which I called ‘Computational Cartesianism’. For Dennett, in his 1991 book Consciousness Explained (Boston: Little Brown), all that mattered were the computational properties of mental states. Your mind was just software. In 1995 I also pointed out, following the suggestions of Hilary Putnam (1988: 121-25), that such a view could not distinguish among computationally equivalent programs. If the brain realizes one program, there is an equivalent program that it realizes, in fact an infinite number of distinct but equivalent programs. Dennett’s view implied that if you had one mind, you had many. Now, in April, 2005, in Sweet Dreams (2005: 155), it turns out that neurotransmitters matter, but they matter only because they have computational properties,

---

5 David Rosenthal suggests making a distinction between mattering because it produces mental properties that functional relations cannot describe or explain and mattering because it produces mental properties of a special functional character. It is the former that is at issue here.

properties that causally affect the computational properties of neurons. So your neurotransmitters matter only because they are software. Crossing the synaptic gap turns out to be a programming problem. Dennett has managed to turn neurochemistry into computer science.

Dennett is an opponent of qualia; he is a Rylean behaviorist (Dennett 2005: 37) reborn as a Putnamian functionalist, and a student of Quine’s – Quine once said à propos B.F. Skinner that he agreed with Skinner’s view: “Mind, Schmind!” So Dennett ended up “Quining Qualia” (in Marcel and Bisiach, eds., Consciousness in Contemporary Science. Cambridge: CUP, 1988). He is not only a good Quinean, he is a Calvinist Quinean. So he (2005: 29) says things like this. “The third-person methods of the natural sciences suffice to investigate consciousness [by which he means self-consciousness – JDA] as completely as any phenomenon in nature can be investigated, without significant residue.” It is fair to ask whether Dennett (2005: 116), who complains that philosophers never offer non-question-begging arguments against him, has an argument against qualia.

He does. Dennett (2005: 82-7) appeals to a 1997 paper by R.A. Rensink, J.K. O’Regan, and J.J. Clark “To See or Not to See: the need for attention to perceive changes in scenes,” Psychological Science 8 (5): 368-73. One shows to a subject two photographs, which differ. For example one pair of photos shows a kitchen; what changes from one to the other is the color of a cabinet door, which changes from white to brown. The photos are shown for 250 milliseconds each, separated by showing a blank screen (a “mask”) for 290 milliseconds. The photos are shown continuously until the subject notices the difference between them and presses a button. Typically a subject looks at the alternating photos for twenty or thirty seconds before noticing the difference. Here is what Dennett asks an audience of philosophers who view the experiment:

Now before you noticed the panel changing color, were your color qualia for that region changing? We know that the cones in your retinas in the regions where the light from the panel fell were responding differently every quarter of a second, and we can be sure that these differences in transducer output were creating differences farther up the pathway of color vision in your cortex. But were your qualia changing back and forth – white/brown/white/brown – in time with the color changes on the screen? Since one of the defining properties of qualia is their subjectivity, their “first-person accessibility,” presumably nobody knows – or could know – the answer to this question better than you. So what is your answer? Were your qualia changing or not?

Dennett (2005: 85) offers his audience three answers: (a) Yes, (b) No, (c) I do not know. He then proceeds to argue that on each of the three answers, it is implied that qualia fail to have a property that, as qualia are normally understood, they should have. So he concludes that the notion of qualia is defective. In particular, on the ‘yes’ answer, that your qualia were changing but you did not notice, Dennett claims that qualia cannot be the contents or features of mental states about which the subject is “authoritative,” i.e. self-aware, about his mental states. The
‘yes’ answer would cast doubt on the subject’s “first-person access” to his mental states. It would then leave open the possibility of a third-person, scientific account of access, which Dennett (2005: 88-9) is happy to consider. Dennett (2005: 86) claims that the ‘no’ answer implies that qualia are the contents or features of mental states about which the subject is “incorrigible,” i.e. his beliefs about his mental states must be true. But if one who answers ‘no’ takes qualia to be the contents or features of mental states about which one is both “authoritative” (self-aware; see Dennett 2005: 89) and “incorrigible” (Dennett 2005: 85-6), Dennett (2005: 86) argues that this claim, an instance of which is that in the experiment your qualia shift if and only if you think they do, “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 [of experience].” By ‘intrinsic’ Dennett (2005: 78) means a property that one who does not have the experience “does not have access to” and “that is inaccessible to objective investigation.” (Note that for Dennett intrinsic-ness is an epistemic feature, one that describes how one knows about the property.) Finally, on the third option (c), the ‘I don’t know’ option, the option on which you do not know whether your qualia were shifting before you noticed the change, Dennett claims that “you put qualia in the curious position of being beyond the horizon of both third-person objective science and first-person subjective experience.” Thus, Dennett believes, on each of the three options the notion of qualia implies claims inconsistent with the intuitive properties qualia are supposed by philosophers to have – first-person authority in the ‘yes’ case, intrinsic-ness in the ‘no’ case, and first-person authority and intrinsic-ness in the ‘I don’t know’ case. Dennett (2005: 86-7) observes that he has:

found, in fact, that people confronted with these three choices don’t agree; all three answers find supporters who are, moreover, typically surprised to find that the other two answers have any takers at all. This informal finding supports my long-standing claim (Dennett 1988) that philosophers actually don’t know what they are talking about when they talk about their qualia. ...The philosophers’ concept of qualia is a mess. Philosophers don’t even agree on how to apply it in dramatic cases like this. This should be at least mildly embarrassing to our field, since so many scientists have recently been persuaded by philosophers that they should take qualia seriously – only to

---

7 Dennett (2005: 90) actually says that the ‘yes’ case as he describes it casts doubt on incorrigibility rather than first-person authority, but this was just a failure of consistent usage on his part.
8 I have cleaned up Dennett’s (2005: 85-6) argument. He did not notice that on his own argument the ‘no’ answer implies Cartesian Incorrigibility, i.e. if I believe a sentence about my mental state, the sentence is true, rather than first-person privileged access, what Dennett tends, not always consistently, to call ‘authority’, i.e. if a sentence about my mental state is true, then I believe the sentence -- (Fa ← B[Fa]) -- when he wrote, “since you noticed no shift, your qualia didn’t shift,” as contrasted with ‘since you didn’t notice any shift, your qualia didn’t shift’, which would imply first-person authority: ¬B[Fa] ← ¬[Fa]. We may express Cartesian Incorrigibility as follows: if ‘Fa’ is a predication of F-ness about my, i.e. i’s, mental entity α and ‘B,’ expresses “I believe,” then (B[Fa] ← Fa).
9 Terminology is not always uniform among philosophers. David Rosenthal prefers to call Cartesian Incorrigibility ‘infallibility’ and to adopt a weaker notion of incorrigibility, which allows incorrigible beliefs to be false though no-one is in a position to correct them. I use the notions of Cartesian Incorrigibility and privileged access, as explained here, because they are the notions that are employed in Dennett’s arguments, whatever his terminology. Throughout the text I shall refer to Cartesian Incorrigibility by the term ‘incorrigibility’. 
I have to say that were Dennett’s argument a good argument, philosophers and scientists who believe in the first-person authority and incorrigibility of “subjective experience”, “phenomenal consciousness,” and the like as coherent notions ought to become very worried. There is a short objection to Dennett’s argument against the coherence of the traditional notion of qualia, viz. holding up one hand containing a red tomato and exclaiming ‘I seem to be seeing a red tomato, and I am sure of it’, and holding up the other hand containing a red tomato and exclaiming ‘I seem to be seeing a red tomato, and I am if I think I am’. But that is not going to convince Dennett, hyper-sensitive as he already is to alleged refutations that “beg the question” against his views. There are four questions. First, how successful is Dennett’s argument? Second, does the traditional, philosophical notion of qualia necessarily possess the properties Dennett mentions in his argument. If not, what features are essential to the philosophical notion? Would such a notion survive Dennett’s objections?

If one is a materialist like Dennett or Quine, so that the basic stuff of the universe is matter as current physics understands it, should one’s commitments to materialism lead one to deny privileged access to some of one’s mental states, in a sense of ‘mental’ understood in an ordinary, commonsense way? In Quine and Ullian’s little book Web of Belief Quine explicitly permits a kind of privileged access, or first-person authority, for some of one’s mental states. What Quine rejects is the incorrigibility of such states. Quine has a modified, modest notion of privileged access: he believes that the subject of experience has privileged access to his own experience in the sense of more reliable knowledge of it than an external observer. What the traditional philosopher also claims is that it is difficult to believe that you could suddenly develop an excruciating pain in your chest without having any knowledge or awareness of its appearance. In the case of Dennett’s experiment, the traditional philosopher grants the facts: it can take time to notice even a significant change in what one sees. But he or she does not think that the experiment shows that he should grant, in general, that major changes in how things look, feel, taste, or smell to you can occur without your being aware of it, even if it takes thirty seconds to notice it. And he agrees with Quine that you are more reliably aware of it, but not necessarily more quickly aware of it, than an external observer. I shall come back to the issue of time in sensory awareness later in this essay, when I discuss the disagreement between Dennett and Benjamin Libet.

Neither Quine nor the traditional philosopher would have been prepared to make an exception to their views for a specially designed task in which the observer is currently monitoring you in an fMRI experiment. The question for Quine is whether an external observer

9 Jason Linder has pointed out to me that in the 8 May 2005 issue of The New York Times Magazine, in an article by Jim Holt, “Of Two Minds,” Holt reports that “In the current issue of Nature Neuroscience, however, Frank Tong, a cognitive neuroscientist at Vanderbilt University, and Yukiyasu Kamitani, a researcher in Japan, announced that they had discovered a way of tweaking the brain-scanning technique to get a richer picture of the brain's activity. Now it is possible to infer what tiny groups of neurons are up to, not just larger areas of the brain. The implications are a little astonishing. Using the scanner, Tong could tell which of two visual patterns his subjects were focusing on -- in effect, reading their minds. In an experiment carried out by another research team, the scanner detected visual information in the brains of subjects even though, owing to a trick of the experiment, they themselves were not
would be more reliable in detecting the existence and character of the subject’s experiences than the subject himself. It is not for Quine a pressing question whether an external observer could also detect the subject’s experiences or that oneself could by external means detect one’s own internal states. Quine grants that an external observer could do that. And he simply eliminates as irrelevant a case in which the external observer but not the subject himself had awareness of something. For the traditional philosopher it might be a question whether it is only the subject of the experience that can know the existence and character of some of the subject’s experiences. But that is not the issue for Quine.

So neither Quine nor the traditional philosopher answers ‘yes’ to Dennett’s question ‘Were your qualia changing?’ Both accept some form of privileged access for the subject of experience. In this Dennett differs from them both. He (Dennett 2005: 102) claims that “first-person accessibility or subjectivity of qualia is in trouble … since, as change blindness demonstrates so vividly, one’s first-person access to one’s color qualia counts for nothing if it cannot be relied on to secure authority for simple judgments of the sort that elude people in these circumstances.” (In the case of Dennett’s audience of philosophers, all three answers were offered; a consensus eluded them.)

Would Quine and the traditional philosopher answer ‘no’ to Dennett’s question ‘Were your qualia changing?’ about his experiment instead? If answering ‘no’ committed one to incorrigibility, Quine would not answer ‘no’. The traditional philosopher would answer ‘no’, since he is committed to the incorrigibility of his judgments of how things look, feel, etc. If he noticed no shift in his experience, how things looked to him did not shift. Dennett assumes that the ‘no’ answer involves both a commitment to first-person authority and to incorrigibility – the necessary correctness – of the subject’s beliefs about his experience. It is assumed that your qualia shift if and only if you think that they do: if you have privileged access to your experience, then if your qualia shift, then you think that they do; if you make incorrigible judgments about your experience, then if you think that your qualia shift, then you are right – they do. On those assumptions Dennett says that the notion of qualia is trivialized – deprived of the features that traditionally make it “intrinsic.”

Dennett thinks it follows that your qualia, how things look and feel, are now “logically constituted” by your judgments or noticings, and so not “intrinsic” features of your experience. He thinks this because there is now a necessary equivalence (if and only if-ness) of the description of your experience and the description of your belief that you are having the experience. So, Dennett infers, your qualia are “constituted by” your beliefs about them. From this, Dennett believes, it follows that qualia cannot be “intrinsic” to your experience. As I noted earlier, by ‘intrinsic’ Dennett (2005: 78) means a property that one who does not have the experience “does not have access to” and “that is inaccessible to objective investigation.” (Note that for Dennett intrinsic-ness is an epistemic feature, one that describes how one knows about the property.) But your beliefs about your experience are not inaccessible to someone else or inaccessible to objective investigation. On these assumptions of first-person authority and incorrigibility, Dennett believes that qualia lose the very property of intrinsic-ness that make...
This Quinean argument relies on an epistemic, best-evidence, form of privileged access in the second premise of the argument for mental-state sentences $p$. The logical form of the epistemic argument is: $\neg B^* p, \Diamond B \rightarrow B^* p$; so, $\Box B^* p$ (where $\Diamond$ expresses evidential possibility, $B$ expresses belief, $B^*$ expresses first-person belief, $\Box \rightarrow$)

This is an interesting argument, but it is a bad argument. First, the central inference from the truth of a biconditional (if and only if) sentence to the claim that what is described on one side of the biconditional is “constituted by” what is described on the other side, is just a fallacy. Such a claim of reduction could just as well be symmetrical, so that we constituted beliefs by qualia rather than qualia by beliefs. Moreover, no such reduction is justified by an “inference to the best explanation” from this biconditional; reduction isn’t the best explanation! One might as well say, on the grounds of the soundness and completeness of a formal system for first-order logic, viz. in system L, $\vdash A$ if and only if $\vdash A$, that logical truth was “constituted by,” or reducible to, provability from the empty set of premises. And that is an obvious absurdity; only a 1920’s, pre-Gödel, Viennese Logical Positivist could believe it. Second, the argument that Dennett (2005: 86) offers addresses only incorrigibility: “since you noticed no shift, your qualia didn’t shift”. The addition of the converse, the principle of first-person authority or privileged access, is gratuitous. The denial of first-person authority was a consequence of the ‘yes’ answer, that one’s qualia were changing without one’s knowledge of it. Dennett is thinking that the corresponding ‘no’ answer commits one to first-person authority. But that is to assume that the (near) converse “if ‘no’ then authority holds” follows from the conditional “if ‘yes’ then authority does not hold”. That is just a fallacy. (In any case the actual converse of the conditional is “if the answer is not a ‘yes’, then authority holds”, also a fallacious inference from the ‘yes’ answer case.) The ‘no’ answer just says that the qualia were not changing in the experiment. The reason for the ‘no’ answer is a claim for the incorrigibility of your perceptual judgment. As a consequence, Dennett has made no case for the reduction of qualia to beliefs, and so no case that a ‘no’ answer commits one to the failure of qualia to be “intrinsic” to experience.

The third answer to Dennett’s question, ‘Were the qualia changing?’, is ‘I do not know’: neither ‘yes’ nor ‘no’. ‘Yes’ would entail the denial of a strong privileged access thesis, and ‘no’ would be justified by an incorrigibility thesis. It would not follow, contrary to what Dennett (2005: 86) suggests, that the ‘I do not know’ answer would “put qualia in the curious position of being beyond the horizon of both third-person objective science and first-person subjective experience.” Quine, as much a materialist and as naturalistic as Dennett, would deny the strong privileged access thesis. So a ‘yes’ answer is consistent with Quine’s view of privileged access; that does not mean that he would say ‘yes’. For other reasons, I suggested above that Quine would not say ‘yes’. And though Quine would never justify a ‘no’ answer by holding an incorrigibility thesis, he might well have been tempted by the slightly weaker justification for a ‘no’ answer that since you did not notice any shift in qualia, and as the subject of the experience you are in the best position to notice, you are justified in asserting that your qualia did not shift.\footnote{This Quinean argument relies on an epistemic, best-evidence, form of privileged access in the second premise of the argument for mental-state sentences $p$. The logical form of the epistemic argument is: $\neg B^* p, \Diamond B \rightarrow B^* p$; so, $\Box B^* p$ (where $\Diamond$ expresses evidential possibility, $B$ expresses belief, $B^*$ expresses first-person belief, $\Box \rightarrow$)} Such a justification would be similar to a methodology of which he
expresses the subjunctive conditional, and ‘\#’ expresses evidential necessity). Of course, the argument is not deductively valid, but the epistemology does not require it to be. The argument tacitly assumes a maximal consistency (or decidability) condition for justified belief or assertion: for any \(p\), \(\#B^*p \lor \#B^*[\neg p]\).

11 David Rosenthal (2002) is not happy with even this qualified version of Cartesian Incorrigibility. One of his counterexamples is a case of anaesthetized nerves in a tooth in which the dental patient reports pain. Rosenthal thinks that patients mistakenly call sensations of vibrations in the tooth, combined with their fear and anxiety from undergoing dental surgery, ‘pain’. Rosenthal goes so far as to say that the patients “confabulate” pain. In William Hirstein’s *Brain Fiction: self-deception and the riddle of confabulation* (Cambridge: MIT Press, 2005), after describing clinical data on Capgras’s syndrome, Cotard’s syndrome, anosognosia, Korsakoff’s syndrome, Anterior Communicating Artery aneurysm, et cet., Hirstein (2005: 186) suggests an epistemic account of confabulation as best accounting for the clinical data. On his account S confabulates that \(p\) if and only if (1) S claims that \(p\), (2) S believes that \(p\), (3) S’s thought that \(p\) is ill-grounded, (4) S does not know that the thought that \(p\) is ill-grounded, (5) S should know that the thought that \(p\) is ill-grounded, (6) S is confident that \(p\). Clinical confabulations require more conditions, depending on the type of syndrome. A distinctive feature of Hirstein’s analysis is that ill-groundedness rather than falsity is the second condition and that clinical memory-defects are not necessary. The dental patient’s claim ‘I am in pain’ is a confabulation if and only if the six conditions are satisfied. Conditions (1), (2), and (6) are satisfied. Conditions (4) and (5) depend on (3). Is the patient’s thought ill-grounded? Aren’t fear and anxiety normally associated with pain? The dentist is drilling into your tooth, and you experience “sensations of vibration,” which are not pain according to Rosenthal. Why not? Where is it written that such sensations cannot be denoted by the ordinary meaning of the word ‘pain’, especially if they are strong enough, and especially when they are accompanied by anxiety and fear. I believe that the patient’s thought ‘I am in pain’ is well-grounded. Hence (3) is false; (4) is vacuously true; and (5) is false. The patient is not confabulating pain. Unlike Rosenthal (2002), I do not think that the ordinary word ‘pain’ is being misused in these cases. So I am willing to assert a somewhat restricted form of Cartesian Incorrigibility. See David Rosenthal, “Explaining Consciousness,” in D. Chalmers (ed.), *Philosophy of Mind: classical and contemporary readings* (New York: Oxford Univ. Press, 2002), pp. 406-21.
Qualia are known only to the subject? Yes; they could fail to be private in that sense, for Quine and nearly all neuroscientists, though not for the traditional philosopher. For Quine and like-minded philosophers, as well as for Benjamin Libet (in *Mind Time: the temporal factor in consciousness*, Cambridge: Harvard Univ. Press, 2004) and other neuroscientists, “knowledge of qualia” does not require that, in Bill Clinton’s expression, I feel your pain. They have rejected the incorrigibility of mental-state reports in the first-person. Like Bertrand Russell in the first decade of the 20th century, they distinguish between knowledge by acquaintance and knowledge by description; in the case of first-person, pain reports, even if they grant that a sincere expression of belief that you are in pain suffices for the truth that you are in pain, the incorrigibility feature of a first-person report does not entail that *only* you have knowledge that you are in pain, even if only you know by acquaintance that you are in pain because “only you are acquainted with your pain.” Such a restricted, Cartesian notion of privacy of mental-states does not make epistemological or semantic sense to Wittgenstein or to Quine. Even if they thought that your sincere first-person pain reports must be true, their knowledge of the existence and character of your qualia only requires that they understand your introspective report of your pain; in Russell’s terms, it only requires that they have knowledge by description of your internal state. The crucial claim is that there is such knowledge by description. Even if I do not know what it is like for you to feel your pain – that is your first-person subjectivity – it does not follow that I do not understand your sentence ‘I am in pain’ nor that I cannot know that you are in pain.

Dennett thinks that there is no subjectivity, no first-person authority or privileged access, that will survive the challenge of his “change blindness” experiment. But, as he so often says of others, he has not provided an argument to support it. In a nutshell, Dennett’s “change blindness” argument is a failed, reductio argument against the “intrinsic-ness” of qualia. His notion of intrinsic-ness is epistemic, defined in terms of solely subjective knowledge and privileged access. But at the end of his (2005: 177) book *Sweet Dreams* he offers the following, parting-shot argument:

> Phenomenal properties are, by definition, not dispositional but rather intrinsic and accessible only from the first-person point of view. …If we define qualia as intrinsic properties of experiences considered in isolation from all their causes and effects, logically independent of all dispositional properties, then they are logically guaranteed to elude all functional analysis – but it’s an empty victory, since there is no reason to believe such properties exist.

Of course this notion of intrinsic-ness is not the notion that he had appealed to in his “blindness change” argument; that was an epistemic notion. This notion of an intrinsic property, defined by its logical independence from all dispositional properties, is a metaphysical notion. Whether there could be properties of experience logically independent of *all* dispositional properties – whatever ‘logical’ or ‘independent’ is supposed to mean here—is not entailed by the traditional epistemological notion of subjective features of experience to which the subject has privileged access. This is not the knockdown argument that it pretends to be; it is a sleight of hand argument, creating a “strawman” opponent that the slightest breeze will blow over. On this
account qualia are epiphenomena, playing no causal role. They are not even Cartesian entities; at least Descartes had the “animal spirits” tickling ever so gently the pineal gland. These qualia are non-scientific entities, causally isolated from intentional mental states. No science could take them seriously, not even Descartes’s. Perhaps Dennett thinks that Thomas Nagel (“Conceiving the Impossible and the Mind-Body Problem,” *Philosophy* 73: 337-52) and other non-reductionists are just “village metaphysicians” who not only make unreasonable demands on explanatory theories but have 17th century metaphysical views that even Descartes would have thought crazy.

As I mentioned earlier, for Dennett, only dispositional properties and their normal effects are mental properties. And a theory of the mental need only explain dispositional properties and their effects. Among the dispositions involved in “being intelligent”, for example, are saying intelligent things if one is in the appropriate contexts of utterance. One cannot be an intelligent creature without being an intelligent utterer of sentences, or at least engaging in intelligent reflection on the contents of experience. To do that requires memory. And the ability to engage in reflection, recollection, remembering, or the reliving of experience does not arise innately; we get it from being social, language-using animals. Non-humans brains are not capable of this kind of memory, and they do not fill skulls of creatures that have social lives that produce reflection, and so are not, on Dennett’s (2005: 169) view, conscious. As Dennett (2005: 171-2) writes:

I have ventured (1) the empirical hypothesis that 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; and (2) the empirical hypothesis that this echoic capacity is due in large part to habits of self-stimulation that we pick up from human culture, that the Joycean machine in our brains is a virtual machine made of memes. These are independent claims. If the meme-hypothesis were roundly defeated by the discovery – the confirmation – of just such echoic systems at play in the brains of nonhuman animals, I would then agree, for that very reason, that the species having those echo-chambers were conscious in just about the way we are – because that’s what I say consciousness is.

And for the sake of clarity, we must note a further comment (Dennett 2005: 167):

…I am impressed – perhaps over-impressed – with the power of self-consciousness, or reflective or introspective consciousness, at the expense of just plain animal sentience or, echoing Block again, phenomenal consciousness, but when I talk of reflective power here, I am not talking about the highly intellectual (and arguably language-dependent) capacity for – shall we say – musing about our musings. I’m talking about the capacity of a dog, for instance, to be reminded of its owner …by an aroma that provokes an echo that provokes a re-identification…if we could be sure that the reminding aroma really does operate by triggering in the dog the sort of echoic, Proustian events that we report to each other. Does – can – the dog recollect the earlier encounter? Are dogs capable of episodic memory, or is there just summoned
up in the dog a “visceral” echo, of either joy or fear? Minimal recognition of this sort need not involve recollection in our own case, so it need not involve recollection in the case of other species. It need not bring in its ensemble the Proustian trappings and surroundings of the earlier encounter that normally ...decorate our own episodes of episodic memory. ...[O]ne of our greatest talents [is] episodic memory and “one-shot learning” that is not restricted to special cases [e.g. the Garcia Effect in rats].

It seems obvious to me that Dennett should never have raised the “red herring” of the capacities of dogs to recognize the aroma of their owners, particularly of the capacity of a dog to re-identify an odor. First, as I understand the neuroscience, recognition of a stimulus requires only that an immediately present stimulus match a representation of a similar stimulus in memory; this does not require conscious awareness, much less conscious awareness of a self that engages in the recognition of the stimulus. Second, while mammals have medial (anterior cingulate) and ventral (orbital) pre-frontal cortices, only primates have a dorso-lateral prefrontal cortex. So a major region supporting working memory in primates does not exist in dogs. As Joseph LeDoux (2002: 196) puts it, “The key element that distinguishes working memory from sensory consciousness is that the former allows for the simultaneous interrelation of temporarily stored information across domains and the flexible use of such information in decision-making, capacities that prefrontal circuits seem to make possible. By this analysis, something akin to human consciousness would be present in other animals with well-developed working memory systems (nonhuman primates) but not in other creatures.” So, not dogs.

Alan Baddeley in the 1970’s revised the traditional notion of short-term memory with the notion of “working memory,” the notion that I have appealed to above. Working memory is not just temporary storage; it is a workspace for executive functions, which change the focus of attention or direct specialized sensory systems to process certain stimuli and ignore others. The analogy is with a DOS operating system that does not multi-task. Working memory is supported by interconnections among the dorso-lateral, medial, and ventral pre-frontal cortices.

Long-term memory is of two types, declarative—that is, consciously verbalizable (Neal Cohen and Larry Squire in 1980) -- or, equivalently, explicit (Daniel Schacter and P. Graf in 1986) long-term memory and non-declarative or, equivalently, implicit long-term memory. Implicit long-term memory covers conditioning, skills, priming, and similar phenomena. Explicit long-term memory was characterized by Endel Tulving, in the 1970’s, as being of two

---

15 Dogs, like other mammals, have some of the neural support structures for working memory, just not enough of them. As I have noted, there is more debate about rats.
kinds: episodic memory and semantic memory. Episodic memories are about personal experiences and require conscious recollection of the time and place of some event that happened to you; the oddly named “semantic memory” is about facts that you come to know but not from personal experience. Tulving believed that episodic memories were characteristic of human beings but that semantic memories were within the capacities of other animals.\footnote{See J. LeDoux (2002: 108), op.cit}

Dennett commits himself to a characterization of aspects of self-consciousness that is constitutive of his view of consciousness: episodic memory and a general capacity for “one-shot” learning. As Tulving characterized it, episodic memory is conscious recollection of the time and place of a personal experience. Though associated with self-consciousness, it is not the core of the notion. The conscious retrieval of episodic memories requires activity of the pre-frontal cortex (PFC), and LeDoux reports that A.D. Wagner\footnote{Neuron 1999: 19-22} suggests that the PFC is involved in the encoding of episodic memory. So LeDoux (2002: 192) concludes that “in order to have an explicit, conscious memory, in addition to the involvement of the medial temporal lobe [the rhinal cortex and hippocampus], two conditions need to be satisfied: you have to be conscious of the information constituting it at the time of the original experience (that is, the experience had to be represented in working memory at the time it occurred), and during retrieval, you have to transfer the information from cortical storage circuits into working memory.” LeDoux’s “working memory model of consciousness” seems to me a better characterization of a materialist, naturalistic account of consciousness than what Dennett has offered in emphasizing episodic memory. LeDoux’s emphasis on the role of the pre-frontal cortex in “working memory” subsumes what Dennett (2005: 171) himself is committed to as part of his description of consciousness, “our capacity to relive or rekindle contentful events.”

<table>
<thead>
<tr>
<th>Sensory system A</th>
<th>Sensory system B</th>
<th>Sensory System C</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\uparrow)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Temporary storage</th>
<th>WORKING MEMORY</th>
<th>executive functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\uparrow)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

LeDoux’s (2002: 179) model of Working Memory

The temporary storage areas seem to be distributed in various ways, with various domain-specificities, in the prefrontal cortex. LeDoux (2002: 187) also observes that the \textit{executive functions} involve both the lateral, prefrontal cortex and the medial prefrontal cortex, specifically the anterior cingulated cortex. These regions have been described by Michael

17
Posner as the “frontal lobe attentional network.” LeDoux cautions that the interactions may involve the orbital cortex as well. What is interesting is that LeDoux (2002: 191) is willing to go out on a limb, saying that the “stuff we are conscious of is the stuff that working memory is working on.” It is characteristic of working memory that its executive functions control shifts of attention. The connectivity of the layers of the prefrontal cortex, enhanced by dopamine, allow stimuli to remain represented as long as one is attending to the event or task at hand.

LeDoux (2002: 192) admits that there is a less attention-laden form of consciousness, what he describes as a “passive awareness of events,” that might not depend on the pre-frontal cortex (see A. Damasio, The Feeling of What Happens, New York: Harcourt Brace, 1999.)

This is where the controversial work of Benjamin Libet on sensory awareness becomes relevant. In his recent book Mind Time (Cambridge: Harvard Univ. Press, 2004: 66-7) Libet is as explicit as LeDoux when he writes, “Awareness is not a function of a memory process. It is not the equivalent of a formed, declarative memory trace. Nor is the absence of a report of awareness due to a rapid forgetting of an early actual sensory experience.18 The proposal that remains most compatible with all the evidence is the hypothesis that awareness is the emergent result of appropriate neuronal activities when these persist for a minimum duration, of up to 0.5 seconds.” And Libet (2004: 62) explains the reason for part of his conclusion, saying, “…the
retention of awareness in people with the severely reduced ability to form new explicit memories indicates that the phenomenon of awareness is not a function of a memory process.”

The case Libet (2004: 61) has in mind is that of a patient with bilateral loss of hippocampal structures, a patient he describes as having “virtually no recallable awareness of any event or sensory image that has just happened (although long-term memories formed before the loss are recallable.) However, such an individual retains the ability to be aware, both of the immediate present and of himself…He is even aware [not in the same sense, I suspect; a bad pun by Libet – JDA] of his own inability to remember what just happened, and he complains about that…”


Much more problematic for the neuroscientist is Dennett’s “commonsense” remark that what characterizes self-consciousness is “one-shot learning.” Given the number of cognitive systems in the brain taking in all kinds of information, and given the complex interactions between those systems and the emotion systems by which learning is deeply affected, it is hard to imagine just what Dennett is talking about when he uses the phrase ‘one-shot learning’. Would one, as a neuroscientist, be better served by asking what “one-shot developmental plasticity” or “one-shot synaptic plasticity” might be? Not being a professional neuroscientist, I have no idea whether that makes any sense or not. As a philosopher, however, I am sure that the notion of “one-shot learning” has been too under-described by Dennett to be explanatorily useful.

Speaking as a philosopher reading with interest and, I hope, with some understanding what neuroscientists have been writing for the education of their non-specialist colleagues and for the public, I have been persuaded to focus less on episodic memory, learning, and culture (memes) – the foci of Dennett’s theorizing -- in my attempt to understand consciousness and more on separating the types or objects of consciousness or awareness, as when Damasio (1999) distinguishes between a “transient or core self” and an “autobiographical self,” and more on understanding the notion of the executive functions of non-verbal working memory, in particular the notion of selective attention, and the way in which a primate dorso-lateral prefrontal cortex has transformed the operation of the mammalian prefrontal cortex. The philosophical lesson I draw from this discussion of Dennett and his critics among neuroscientists is a new emphasis on the notion of selective attention as an executive function of working memory, both because it distinguishes primate consciousness and because it bears a family resemblance to notions of focused attention and of “vigilance” that characterize our biological cousins.

But there is one place where Descartes, Wittgenstein, Quine, Dennett and neuroscientists like Joseph LeDoux and Michael Gazzaniga share common ground: their
emphasis on human language.\textsuperscript{19} The existence of \textit{verbal} working memory necessarily makes the human primate a special animal. As LeDoux (2002: 197-8) says:

It is, in my opinion, the structuring of cognition around language that confers on the human brain its unique qualities. Other animals may be consciously aware, in some sense, of events going on in their world. They may have domain-specific consciousness, or in the case of nonhuman primates, domain-independent nonverbal consciousness, but lacking language and its cognitive manifestations, they are unlikely to be able to represent complex, abstract concepts (like “me” or “mine” or “ours”), to relate external events to these abstractions, and to use these representations to guide decision-making and control behavior. …[T]he emergence of the cognitive capacities underlying language changed the way the brain works, making it possible for human brains to think and experience events in ways that other brains cannot. The addition of language into the human brain involved a revolution rather than an evolution of function.

It thus becomes possible for us (courtesy of John F. Walter) to enjoy the following cartoon: