

The New York Review of Books

The Mystery of Consciousness Continues

The New York Review of Books | June 9, 2011

[Review by:] John R. Searle

Self Comes to Mind: Constructing the Conscious Brain

by Antonio Damasio

Pantheon, 367 pp., \$28.95



How do neurobiological processes in the brain cause consciousness? I think this is the most important question in the biological sciences today. Two related questions: Where exactly is consciousness realized in the brain and how does it function causally in our behavior? Antonio Damasio is one of the leading workers in the field of consciousness research, and after having written a number of books on related problems, in *Self Comes to Mind* he addresses the problem of consciousness directly. He does not claim to have solved it but he believes that he has made advances and pointed in the right direction for a solution.

Musée d'Orsay, Paris/RMN-Grand Palais/Hervé Lewandowski

Odilon Redon: *Closed Eyes*, 1890; from the exhibition 'Odilon Redon: The Prince of Dreams, 1840-1916,' at the Grand Palais, Paris, March 23-June 20, 2011, and the Musée Fabre, Montpellier, July 7-October 16, 2011

What exactly is consciousness? There are a number of senses of the word in ordinary speech, but there is one that is most important for philosophy and science: consciousness consists of qualitative, subjective states of feeling or sentience or awareness. These typically begin when we awake from a dreamless sleep and they go on until we fall asleep again or otherwise become unconscious. Dreams are a form of consciousness. Consciousness, in short, is a matter of the qualitative experiences that we have. To understand qualitiveness, think of the difference between drinking beer, listening to music, and thinking about your income tax. Each experience has a distinct quality.

Because of this qualitative character all conscious states are essentially subjective in the sense that they exist only as experienced by a subject—human or animal. The problem of consciousness can now be stated somewhat more precisely: How does the brain produce qualitative subjectivity? How does it get us over the hump from the objective third-person character of neuron firings to the subjective first-person feelings we have when we are conscious?

There is so much confusion surrounding the notions of objectivity and subjectivity that I need to say a word to clarify them. In one sense, the objective/subjective distinction is about claims to knowledge. I call this the epistemic sense. A claim is said to be objective if its truth or falsity can be settled as a matter of fact independently of anybody's attitudes, feelings, or evaluations; it is subjective if it cannot. For example, the claim that Van Gogh died in France is epistemically objective. But the claim that Van Gogh was a better painter than Gauguin is, as they say, a matter of subjective opinion. It is epistemically subjective.

In another sense, the objective/subjective distinction is about modes of existence. I call this the ontological sense. An entity has an objective ontology if its existence does not depend on being experienced by a human or animal subject; otherwise it is subjective. For example, mountains, molecules, and tectonic plates are ontologically objective. Their existence does not depend on being experienced by anybody. But pains, tickles, and itches only exist when experienced by a human or animal subject. They are ontologically subjective.

I emphasize these two senses of the distinction because a common mistake is to suppose that because science is objective and consciousness is subjective, there cannot be a science of consciousness. Science is indeed epistemically objective, because scientific claims are supposed to be verifiable independently of anybody's feelings and attitudes. But the ontological subjectivity of the domain of consciousness does not preclude an objective science of that domain. You can have an (epistemically) objective science of an (ontologically) subjective consciousness. Much confusion has been created by the failure to see this point.

So our question is: How does the brain create ontological subjectivity? We know consciousness happens and we know the brain does it. How does it work? How do we approach this problem scientifically? The standard way is to go through three steps. First, try to find the neurobiological correlate of consciousness. A lot of work has been done on this. There is now even a commonly used abbreviation, NCC, for the neuronal correlate of consciousness. Second, try to test if the correlations are in fact causal. Do the neurobiological states cause consciousness? Third, try to formulate a theory. Why do

these processes cause consciousness at all, and why do these specific processes cause these specific conscious states? In recent years there has been a sizable number of important research efforts devoted to solving these problems, and I have reviewed several of the relevant books in these pages.¹

One depressing feature of this entire research project is that it does not seem to be making much progress. Most efforts to identify the NCC have concentrated on the thalamocortical system, the area including the thalamus and the different layers of the cortex. But the slowness of progress makes one wonder if we are, perhaps, proceeding on the basis of wrong assumptions. Damasio's book makes a new start in at least two respects. First, he emphasizes other areas of the brain in the production of consciousness, especially the brain stem. Other theorists ignore the brain stem, presumably because it is an evolutionarily primitive part of the brain. They think consciousness is the result of activity in more advanced neuroanatomical features, such as the thalamocortical system. Second, his entire book is built around the theme that the self plays a crucial role in the creation of consciousness.

2.

To summarize Damasio's argument² is not an easy task because the book is densely argued and to me at least often unclear. Here is the basic framework: the brain creates an (unconscious) mind. The brain also creates the self. When the self encounters the mind, consciousness results. As with most important theories there is an underlying intuition that drives his theory, though he does not say it explicitly: whenever I have a conscious experience I always experience it as *mine*. I do not just have a sequence of unrelated neutral qualitative states that could belong to anybody, but I have them as part of a coherent unity that is constitutive of and experienced as *myself*. So if consciousness is somehow always related to the self, then it seems natural to think that maybe the key to understanding the neurobiology of consciousness is by way of the neurobiology of the self.

The book addresses two problems: (1) How does the brain construct a mind? and (2) How does the brain make the mind conscious? The brain creates a mind by creating images, which are unconscious momentary patterns on sheets of neurons called maps. The images may be either of parts of the body or of things outside the body, but in general, perception is the result of mapping. Damasio says, "The distinctive feature of brains such as the one we own is their uncanny ability to create maps." Brain maps are not static; they change from moment to moment. The mind is a consequence of the mapping activity of the brain. "Minds emerge," Damasio writes, "when the activity of small [neuronal] circuits is organized across large networks so as to compose momentary patterns. The patterns represent things and events located outside the brain." The term "map" applies to all of these patterns, and though they are mental, they are at this stage still totally unconscious, according to Damasio.

Body mapping is the key to the problem of consciousness, because by mapping the body the brain manages to create the critical component of the self. Having made a mind by making maps, the brain makes the mind conscious by creating a self, and when the self encounters the mind, consciousness results. This is the source of Damasio's title, *Self Comes to Mind*:

The decisive step in the making of consciousness is not the making of images and creating the basics of the mind. The decisive step is *making the images ours*, making them belong to their rightful owners... [Italics in original.]

Damasio's two crucial notions are consciousness and the self.

(1) *Consciousness*. In actual practice I think his idea of consciousness is essentially the one stated above. Its essence is qualitative subjectivity. But when Damasio defines it explicitly it comes out a bit differently: it is "*a state of mind in which there is knowledge of one's own existence and of the existence of surroundings*" (italics in original). I do not believe this definition is correct. My dog, Gilbert, is plainly conscious, but in what sense does he have *knowledge* of his own existence? He is certainly aware of his surroundings when he perceives anything. But it is hard to say that when he is dreaming he has knowledge of the existence of his surroundings. It is Damasio's right to define a word any way he likes, but I think in practice he uses "consciousness," as I do, to refer to ontologically subjective states such as pains, and does not use it just to describe epistemic states, such as my knowing that I am in Berkeley.

(2) *The Self*. The self is a much harder notion to define, and I do not find his definitions entirely clear. He says the self is decomposable into three components, the protoself, the core self, and the autobiographical self. Each of these can come in two forms, the "self-as-object" and the "self-as-knower." But the self-as-object can also operate as knower.

The protoself is constituted by special kinds of mental images of the body produced in body-mapping structures, below the level of the cerebral cortex. The protoself is "*an integrated collection of separate neural patterns that map, moment by moment, the most stable aspects of the organism's physical structure*" (italics in original). The first product of the protoself is "primordial feelings." Whenever you are awake there has to be some form of feeling. The second form of the self, the "core self," is about action. "The core self unfolds in a sequence of images that describe an object engaging the protoself and modifying that protoself, including its primordial feelings." These images are now conscious because they have encountered the self. Finally there is the autobiographical self, constituted in large part by memories of facts and events about the self and about its social setting. The protoself and the core self constitute a "material me." The autobiographical self constitutes a "social me." Our sense of person and identity is in the autobiographical self.

When we put this all together we get the following result: conscious minds begin when the self comes to mind, when brains add a process involving a person's sense of self to the mind mix. Specifically, the neurology of consciousness is organized around the brain structures involved in generating three features: wakefulness, mind, and self. Three major anatomical features are the brain stem, the thalamus, and the cerebral cortex. There is no direct alignment between, on the one hand, each of these anatomical features, and, on the other, each component of the mental trio of wakefulness, mind, and self. All three anatomical divisions contribute some aspect of wakefulness, mind, and self. To be fully conscious you have to have three features: (1) to be awake, (2) to have an operational mind, and (3) to have a sense of self as a protagonist of the experience.

Finally, after more than 250 pages, we get to the problem of qualia—a term often used for qualitative conscious states. Damasio tells us that there are two kinds of qualia:

Qualia I and Qualia II. Qualia I is about pain and pleasure, but the problem of Qualia II is why there should be any feelings at all. He thinks Qualia I is not a mystery but that the Qualia II problem is more difficult. About Qualia II we get the following stunning passage:

Qualia is part of the contents that come to be known as the self process, the self construction illuminating the mind construction. But somewhat paradoxically, Qualia II is also the grounding for the proto-self and thus sits astride mind and self, in a hybrid transition. The neural design that enables qualia provides the brain with *felt* perceptions, a sense of pure experience. After a protagonist is added to the process, the experience is claimed by its newly minted owner, the self.

It is a bit hard to see how this is consistent with the rest of the book. The self is introduced to explain consciousness, but if it is to explain consciousness we cannot assume that the self is already conscious. How did it get to be conscious? Yet he tells us that qualia are the "grounding" of the self. But qualia just are conscious states. So this account, if we are to take this passage seriously, would be circular. We would be assuming consciousness to explain consciousness.

This is only a brief summary of the main argument of the book and I am leaving out Damasio's rich and rewarding discussion of many other issues, such as the emotions, perception, memory, and mirror neurons.³

3.

I have a large number of criticisms to make of the book, but I will confine them to three topics: Damasio's account of the self, his conception of the mind, and his theory of consciousness. I think all are open to question.

The Self. The project is to give an account of consciousness by showing how the interaction between the mind and the self produces it. In order to do that one would have to give an account of the mind and the self that did not already presuppose that either was conscious and then show how their interaction produces consciousness. One would have to explain the mind as a set of ontologically objective biological processes, then do a similar explanation of the self, and then specify the mechanisms by which the structures of the self interact with the mind structures in order to produce qualitative subjectivity.

As far as I can tell Damasio does not succeed in doing this; indeed he does not even really try. He does try to give an account of the mind as a set of (unconscious) mapping activities of the brain, and this does not presuppose, or at least it does not obviously presuppose, that these activities are conscious. But when he gets to the self, it is hard to understand any of his three divisions of the self, protoself, core self, and autobiographical self, without supposing that they are already conscious.

The problem can be stated succinctly by presenting his account with the following dilemma: Is the self, as he describes it, unconscious or conscious? If it is unconscious then he has nothing to say about how its encounter with a mind results in consciousness. But if you look at the text closely it seems pretty clear that there is no way to understand the sort of self that he describes without supposing that it is already conscious. He frequently uses words like "primordial feeling" and "emotion" to describe

the self. It is hard to understand these in a way that does not imply consciousness. This account is therefore circular because we are assuming a conscious self in order to explain the conscious mind, but this uses consciousness to explain consciousness.

The Mind. My objection to Damasio's account of the self is that he has tacitly, if unconsciously, smuggled consciousness into his conception of it without explaining how it got there. My objection to his account of mind is that he does not see how much consciousness is essential to our understanding of the mind.

He says the brain creates the mind by making maps. On the standard understanding of the causal relations between brain and mind, that is not true. The brain creates the mind by making thoughts, feelings, perceptions, pains, memories, sensations, and all the rest of it, both conscious and unconscious. The creation of neurobiological patterns is an essential part of this process, but he gives no reason to suppose that the map, qua map, has any psychological reality at all. When he tells us that the mind consists largely of unconscious maps, one has to ask: What fact about these maps makes them mental? When we read words like "image," "perception," and "feeling" in his account of maps, we tend naturally to connect them with the conscious formation of images and the experience of perceptions and feelings. But that is not what he means when he talks about the mapping activity of the brain. The problem, to put it in a nutshell, is that he has given us no reason to suppose that these maps have any mental or psychological reality at all.

Consider an obvious example of a map that we are all familiar with, the retinal image inside the eyeball when we see anything. As far as I can see, this map/image has no psychological reality whatever. We do not see it or otherwise experience it. The retinal image is a step from the introduction of optical stimuli to conscious perceptions. But most of these steps, as far as I know, have no psychological reality whatever. They are not the mind, or part of the mind. In any case he gives us no reason to suppose they have any mental reality at all. They are just "momentary patterns." He tells us that they are "of" something and that they "represent" things and events outside the brain. But how do they represent? What fact makes them "of" anything?

At any given point in waking life there are two sorts of things going on in the brain, neurobiological processes and consciousness. Some of those neurobiological processes are unconscious mental states. What fact about them makes them *mental*? The question about how these processes become mental is not trivial. Damasio assumes that a pattern that occurs on the way to a mental state is already mental. There is no justification given for this assumption. He tells us that subjectivity is not required for mental states to exist. Yes, but what is required? What fact about them makes them mental?

Whenever we talk about mental states and events, conscious or unconscious, we have to be able to say what exactly the *content* of the state in question is. I see the tree in front of me, I think about the lunch I am going to have later, I feel a pain in my thigh. In each case, the mental reality, conscious or unconscious, has a specific content. The problem with Damasio's notion of the mind is that he does not specify any actual or possible contents. Maybe there is a psychological reality to the maps, but he gives us no reason to suppose there is. A more natural way to describe at least many of the maps would be to say that they are stages on the way to the construction of mental states.

Consciousness. We have to keep reminding ourselves that any type of qualitative subjectivity is a form of consciousness. The possession of such states is necessary and sufficient for being conscious. Consciousness comes in degrees, and these range all the way from fanatic intensity to just barely being awake. But all of these are degrees *within consciousness*. There is no such thing as a "hybrid" form of consciousness. But the present task in the neurobiology of consciousness is to explain exactly how brain processes create consciousness.

I think the fact that Damasio finds dreams and states such as jet lag "exceptions" to his account is a clue that something is wrong. About jet lag he says: "there is a mind but not quite yet a mind organized with all the properties of consciousness." He tells us that wakefulness is essential for consciousness. But in the ordinary sense it is not. I can have dreams when I am in REM sleep. He says this is a "partial exception" but there is nothing partial about it. It is a total exception.

He sees correctly that dreams and jet lag have to be distinguished from normal, healthy, goal-directed states of consciousness. But the question "What makes a state of consciousness healthy?" is quite different from "How does the brain create consciousness in the first place?" Our task is to explain consciousness, i.e., qualitative subjectivity. Damasio claims that wakefulness is a necessary condition for consciousness. But what these cases show is that it is not. His answer, that they are not normal, is quite correct but does not address the question: What are the necessary and sufficient conditions for the creation of qualitative subjectivity?

4.

Damasio and I agree that in addition to conscious mental states there are unconscious states. What fact about certain brain processes makes them mental? This is one aspect of the traditional "mind-body problem." What is its solution? Without going into detail, the best way is to consider an example. Even when I am in a sound, dreamless sleep it is true to say of me that I believe/remember that George Washington was the first president. What fact about me makes it the case then and there that I have that unconscious mental state? I think the answer is that my brain is in a condition such that I can produce that state in a conscious form and in conscious behavior. I can, for example, give a correct answer to the question "Who was the first president?"

The notion of the unconscious is the notion of a capacity of the brain to produce states and actions in a conscious form. Sometimes these capacities are blocked by brain damage, repression, loss of memory, etc., but the basic idea of an unconscious mental state, I think, is clear enough. I believe that Damasio uses the notion of the unconscious rather uncritically, and he does not see the importance of distinguishing between those neurobiological processes that are genuinely unconscious mental states, such as my unconscious memory when I am sound asleep that Washington was the first president; and those that are totally nonconscious and nonmental, such as the secretions of the neurotransmitters when I move my body.

I am sympathetic with the basic intuition that drives Damasio's investigation, namely that in any account of consciousness we need to explain how our conscious states are experienced, not just as a sequence of isolated qualitative subjective events, but as "my experiences." This is part of what we need to explain. He proposes that we should take

this characteristic of the self and treat it as the basis of consciousness. In the end that may be the right approach, but he does not give convincing reasons to suppose that it is. I believe a more plausible approach is to suppose that nonpathological forms of consciousness already come with a sense of the self. Our sense of self is a product of a certain sort of consciousness, not conversely. That is why we can lose that sense in certain pathological forms of consciousness. I have great admiration for this book and its author. I think it is an adventurous, courageous, and intelligent effort. I do not think he has made a convincing case that this is the right way to solve the problem of consciousness.

1. For my review of Francis Crick, *The Astonishing Hypothesis* (Touchstone/Simon and Schuster, 1995) and Roger Penrose, *Shadows of the Mind* (Oxford University Press, 1994), see *The New York Review* , [November 2, 1995](#) ; for my review of Gerald M. Edelman, *The Remembered Present* (Basic Books, 1989) and *Bright Air, Brilliant Fire* (Basic Books, 1992), Daniel C. Dennett, *Consciousness Explained* (Back Bay/Little, Brown, 1991), and Israel Rosenfield, *The Strange, Familiar and Forgotten* (Vintage, 1993), see *The New York Review* , [November 16, 1995](#) ; for my review of Christof Koch, *The Quest for Consciousness* (Roberts and Company, 2004), see *The New York Review* , [January 13, 2005](#) . Several of these are collected in John Searle, *The Mystery of Consciousness* (New York Review Books, 1997). ↵
2. I am grateful to Damasio for an e-mail exchange about the correct interpretation of his views. ↵
3. When a primate engages in some physical activity and is watched by another, some of the neurons in the watcher match or mirror some of the neuronal behavior of the actor. These matching neurons are called mirror neurons. ↵

The logo for the Institute for Cognitive and Systems Studies (ICSS) features the acronym 'ICSS' in a bold, sans-serif font, centered within a circular emblem. The emblem is partially obscured by a large, faint watermark of a human head profile in the background.

پژوهشکده علوم شناختی