

© 1999 Polger & Flanagan

Explaining the Evolution of Consciousness: The Other Hard Problem

Thomas Polger
Department of Philosophy
Email: twp2@acpub.duke.edu

Owen Flanagan
Departments of Philosophy, Psychology, and Neurobiology
Email: ojf@acpub.duke.edu

Box 90743, Duke University, Durham, NC, 27708, USA.

Explaining the Evolution of Consciousness: The Other Hard Problem

Abstract

Recently some philosophers interested in consciousness have begun to turn their attention to the question of what evolutionary advantages, if any, being conscious might confer on an organism. The issue has been pressed in recent discussions involving David Chalmers, Todd Moody, Owen Flanagan and Thomas Polger, Daniel Dennett, and others. The purpose of this essay is to consider some of the problems that face anyone who wants to give an evolutionary explanation of consciousness. We begin by framing the problem in the context of some current debates. Then we'll take a look at what a philosopher of biology, Robert Brandon, has to say about evolution and about adaptationist explanation. This will not only help us see why available accounts of the evolution of consciousness are lacking, but will show us why it is so hard to give a credible story.

Explaining the Evolution of Consciousness: The Other Hard Problem

The Other Hard Problem

David Chalmers, in his essay “Facing Up to the Hard Problem of Consciousness,” has helped to focus the current discussion regarding the nature of consciousness on the so-called “hard problems” of consciousness—those having to do with the phenomenal aspects of consciousness, the “what it’s like” quality common to all sorts of conscious experience.¹ Chalmers’ “hard problem” is the problem of why and how conscious experience arises. We believe that all questions that take qualitative experience seriously are hard problems. Among those, one of the most difficult is the question of why consciousness has come to be.

The purpose of this essay is to consider some of the problems that face anyone who wants to give an evolutionary explanation of consciousness. We begin by framing the problem in the context of the current debates. Then we’ll take a look at what philosophers of biology have to say about evolution, and about adaptationist explanation. This will not only help us see where available explanations of the evolution of consciousness are lacking, but will show us why it is so hard to give a credible story, and provide a framework for further discussion. This framework, we hasten to add, is not of our own invention; we are simply applying what philosophers of biology tell us about evolution—but more about this later.

Consciousness and Necessity

The thesis of *conscious inessentialism* presses the question of why consciousness has come to be. Conscious inessentialism, as framed by Owen Flanagan, is the thesis “that for any intelligent activity *i*, performed in any cognitive domain *d*, even if *we* do *i* with

¹ Chalmers, D. (1995) “Facing up to the hard problem of consciousness.” *Journal of Consciousness Studies*, 2, 1.

conscious accompaniments, *i* can in principle be done without these conscious accompaniments.”² This possibility has shuffled its way into philosophical discussion in the person of the philosophers’ zombie, that poor sap who is behaviorally indistinguishable from we conscious beings despite lacking any qualitative states. There is no “what it is like” to be a zombie.

The possibility of zombies is just the thesis of conscious inessentialism embodied. The possibility of not just isolated zombies, but of whole zombie worlds, presses the question of the origins of consciousness; for it seems to demonstrate that consciousness does not have to be.³

Some philosophers have thought that the possibility of zombies, of conscious inessentialism, undermines materialist philosophies of mind. But this is not so, and it is important to see why. Conscious inessentialism is a very weak claim. It is a claim about the mere possibility of some creature that can behave as we conscious beings do, but without consciousness. One way this might be true, of course, is if consciousness is an epiphenomena. But that is not the only way. It may be the case that consciousness is causally efficacious, but that the functions that it performs can be accomplished—at least in principle—by non-conscious mechanisms. So conscious inessentialism is compatible with a thorough-going naturalism about the mechanisms and subvenient basis of consciousness, and with a variety of claims about the causal efficacy of consciousness for us. According to this view, consciousness is a mechanism by which some important cognitive functions are performed in human beings. But the fact that we perform these functions consciously is contingent.

² Flanagan, O. (1992) *Consciousness Reconsidered*. (Cambridge, MA: MIT Press), 5. Note that this formulation makes a claim about ‘activity’ not ‘acts’ as defined by action theory. ‘Activities’ in the sense relevant to conscious inessentialism do not involve conscious intentions to act. See Flanagan pp. 129-30.

³ As suggested by Todd Moody (1994) in his article, “Conversations with zombies,” *Journal of Consciousness Studies*, 1, 2, pp. 196-200.

There is a confusion that arises in discussions of consciousness and necessity that we want to get clear about. Suppose there is some organism that performs function f by going into physical state p . Suppose further that p is a conscious state. That is, whatever the relationship between conscious states and physical states turns out to be (identity, supervenience, etc.), p has that relationship to conscious state c . Now, we can ask two sorts of questions about p . One kind of question is: Is it logically, metaphysically, or nomically necessary that any system in state p is thereby in conscious state c ? The other sort of question is: Is it logically, metaphysically, or nomically necessary that f be accomplished consciously? The answer to questions of this second sort is: *no*. Conscious inessentialism is a view about this second sort of necessity. To whatever degree one thinks that f does not have to be accomplished consciously, consciousness will be contingent. As such, the presence of consciousness in us requires explanation.

If consciousness is not logically or nomically necessary for the performance of the cognitive functions that it does in fact perform in human beings, what can be said about why consciousness has come to be in us? We are conscious organisms. If it did not have to be so, then there is an interesting question: Why are we (and presumably, to a greater or lesser extent, many other creatures) conscious? There is theory that is well suited to explaining why it is that certain contingent features of organisms come to be: the theory of evolution by natural selection.

It is a consequence of taking seriously the idea that consciousness is a natural phenomenon that we must treat it like any other naturally occurring trait of a living organism. That means that consciousness is a contingent product of evolution, itself a contingent process in our world.⁴ Even if there are certain configurations of matter (or matter*, depending on your favorite theory) that are necessarily conscious states (*qua* the

⁴ Beatty, J. (1993) "The Evolutionary contingency thesis," in *Concepts, Theories, and Rationality in the Biological Sciences*, Wolters & Lennox (eds.) (Pittsburgh: University of Pittsburgh Press), 43-81. Note that this claim—that consciousness is a product of evolution—is distinct from the claim that consciousness is an adaptation. See below.

first sort of necessity distinguished above), *those* configurations did not have to be realized.

Consciousness and Necessity

In their article “Zombies and the Function of Consciousness,” Owen Flanagan and Thomas Polger suggested some reasons that certain sorts of purported evolutionary explanation of consciousness can be undercut.⁵ Some philosophers were not happy. Daniel Dennett, for one, insisted that credible accounts *have* been given, and was kind enough to put his own account forward again for our consideration. To demonstrate that Flanagan and Polger’s claims about evolutionary explanations of consciousness are wrong-headed, Dennett revived his functionally complex zombies: zimboes.⁶ Zimboes are zombies that have higher-order, self-reflective states.⁷

Dennett tells a story about how zimboes could evolve from zombies, based on the adaptive advantage of their capacity for learning and plasticity. Dennett thinks that Flanagan and Polger reject learning and plasticity as possible advantages of consciousness on the basis of the possibility that a non-conscious system, like a zimbo, could have learning and plasticity. But this is not the point. It may be that learning and plasticity will turn out to be among the advantages of some varieties of consciousness. Nevertheless, the story “Consciousness is necessary for learning and plasticity, which are selectively advantageous” is not a credible evolutionary explanation of consciousness.

This is important: The answer to the question, “Why did consciousness come to be?” cannot be, “Because consciousness is necessary for learning and plasticity?” This is *not* because consciousness doesn’t give us learning and plasticity—maybe it does. It is because answers of the form, “Because consciousness is necessary for *x*” are undermined

⁵ Flanagan, O. & T. Polger. (1995) *Journal of Consciousness Studies*, 2, 4.

⁶ Dennett, D. (1995) “The Unimagined preposterousness of zombies,” *Journal of Consciousness Studies*, 2, 4, 322-26.

⁷ Dennett, D. (1991) *Consciousness Explained*. (Boston: Little, Brown and Company), 310-1.

by conscious inessentialism. This simple point is absolutely crucial. What conscious inessentialism (i.e., the possibility of zombies) demonstrates is that the answer to the question, “Why did consciousness come to be?” cannot be of the form, “Because consciousness is necessary for x .” This is not because consciousness is not for x , it is because consciousness is not *necessary* for x .

Note that this is an observation about the *form* of recent proposals and the *form* of evolutionary explanations. It is not just that explanation of the form “consciousness is necessary for x ” is wrong—although we think it is. More importantly, “ T is necessary for x ” is not the form of a proper evolutionary explanation of any trait. If T is necessary for x then one does not need to appeal to evolutionary theory to explain the presence of x in T . Evolutionary explanations account for contingent features of organisms.

This exposes the fundamental flaw in every proposal for an evolutionary explanation of consciousness that we have seen. They are all explanations in terms of one or another cognitive function that consciousness is alleged to be necessary for. How the defenders of consciousness got into this position is an interesting question about the recent discourse in the philosophy of mind. The important thing for our purpose is to realize that the winning move is to abandon the notion that consciousness is necessary and look for an explanation of it as a contingent feature of our world.

Now comes the hard part. So far we’ve sketched the case against treating consciousness as a necessary feature of the world, never mind of human beings. And along the way we’ve been less than subtle about what sort of explanation we favor. But making the case for consciousness as a product of evolution, and moreover as an adaptation, is very difficult. Robert Brandon begins his book, *Adaptation and Environment*, by noting that, “The existence of adaptations, the fit between organisms and their environments, is one of the most striking features of the biological world.”⁸

⁸ Princeton: Princeton University Press, 1990, 3.

Explanation in terms of evolution by natural selection is not the only way to explain how adaptations come to be, but, he argues, it is “the only general and scientifically legitimate theory of adaptation.”⁹ Let’s take a look at a possible account of the evolution of consciousness.

The Theory of Evolution by Natural Selection

The following is a possible account of the evolution of consciousness: In a finite population of interbreeding organisms, random mutation caused a portion of the population to have some sort of conscious states (i.e., for those states, there is something that it is like for the organism to be in that state.) In each case, the new phenotypic trait (speaking generally, *consciousness*) was heritable. Sadly, a nearby volcano erupted. By chance, the eruption killed all and only the non-conscious organisms. The conscious organisms, however, survived and reproduced successfully, passing on the consciousness traits. Consciousness evolved.

Evolution occurred because there was phenotypic variation, heritability, and differential reproduction. This is an evolutionary explanation. But is it the sort of explanation that we are seeking? No. Although evolution of consciousness has occurred, it was not evolution by natural selection but rather by *random drift*. Only by chance did the conscious organisms out-reproduce their non-conscious counterparts; it was not *because* they were conscious that they survived. Evolution by natural selection requires a further element: there must be a cause other than chance for the differential reproduction that leads to evolution.¹⁰ We need to specify what the adaptive advantage of the feature in question was for a particular organism in a particular selective environment.

Dennett claims that it is a mistake to ask what the adaptive advantage of consciousness might be:

⁹ *Ibid.*, vii.

¹⁰ Brandon, 6-9.

The question of adaptive advantage, however, is ill-posed in the first place. If consciousness is (as I argue) not a single wonderful separable thing ('experiential sensitivity') but a huge complex of many different information capacities that individually arise for a wide variety of reasons, there is no reason to suppose that 'it' is something that stands in needs of its own separable status as fitness enhancing. It is not a separate organ or a separate medium or a separate talent.

To see the fallacy, consider the parallel question about what the adaptive advantage of *health* is. Consider 'health inessentialism': for any activity *b*, performed in any domain *d*, even if *we* need to be healthy to engage in it (e.g., pole vaulting, swimming the English Channel, climbing Mount Everest), it could in principle be engaged in by something that wasn't healthy at all. So what is health *for*? Such a mystery!¹¹

Dennett's "health inessentialism" parody misses its mark. If consciousness were a "single wonderful separable thing" then there would be a single evolutionary, perhaps even adaptationist, explanation for its existence. But we do not think that is the case with consciousness. Conscious states are various; we group them together as an interesting phenomenon on the basis that they share the Nagel property: there is something "that it is like" to be in a conscious state.

The fact that some feature has multiple realizations, that its name covers a "huge complex" of capacities, does not render foolish the question of its adaptive advantage. It just means that there has to be an answer for each instantiation—and very possibly not the same answer. We agree that various sorts of consciousness, as Dennett says, "individually arise for a wide variety of reasons." When we ask about the adaptive advantage of consciousness, we are asking for some of those reasons to be specified.

This is a good time to note that although some forms of consciousness (conscious visual perception seems a good candidate) will likely turn out to have an adaptive advantage, others may not. Flanagan suggests, for example, that dreams are a biological epiphenomena.¹² Dreams, he argues, are a byproduct of pre-existing conscious systems

¹¹ Dennett, 1995, 324-25.

¹² Flanagan, O. (1995) "Deconstructing dreams: The Spandrels of sleep," *Journal of Philosophy*, 5-27. Reprinted with modifications in O. Flanagan, *Self Expressions: Mind, Morals, and the Meaning of Life*, (New York: Oxford University Press, 1996), 32-52.

being activated while the brain does the important things that brains do while we sleep. The point is that not only will there not be *one* adaptationist explanation for all sorts of consciousness, but there is good reason to think that there is *no* adaptive advantage at all to some sorts of consciousness.

Specifying what the adaptive advantage of consciousness might be is difficult. But it is child's play compared to finding the sort of evidence that would indicate that any such "how possibly" story could in fact be how some variety of consciousness *actually* had an adaptive advantage in a selective environment. This problem, the problem of justifying a "how possibly" explanation as a "how actually" explanation, requires solid empirical data.¹³ The problem is not peculiar to the study of consciousness—giving credible adaptationist explanations is notoriously difficult. That problem is compounded in the case of consciousness on account of a certain philosophical skepticism regarding consciousness that has widely infected the sciences of the mind. Some researchers feel compelled to conclude that the valuable data they have is not data about *consciousness* at all.

Ideal Adaptationist Explanation

By specifying the adaptive advantage of some variety of consciousness, we give an ecological account of its relative adaptedness. But even if we can discover the adaptive advantage of some variety of consciousness, that is only one piece of an adaptationist explanation. Robert Brandon formulates five elements for a minimal ideally complete adaptationist explanation, explanation in terms of evolution by natural selection:

- 1) Evidence that selection has occurred.
- 2) An ecological explanation of relative adaptedness.
- 3) Evidence that the traits in question are heritable.
- 4) Information about population structure.
- 5) Phylogenetic information about trait polarity.¹⁴

¹³ The notions of "how possibly" and "how actually" explanations are from Brandon, § 5.3.

¹⁴ Brandon, 165. These five elements are listed in the order that Brandon presents them, which is not to indicate any relative importance. In fact, he argues that the relative importance of the various elements will

What does all this mean? And what, in particular, does it mean for explaining the evolution of consciousness?

The second element of adaptationist explanation, the ecological explanation, is the one that we have devoted a great deal of this essay to discussing. Ecological explanations of relative adaptedness tell why some trait should be thought to increase the fitness of its bearer in a particular selective environment.

The first and third elements are straight forward. If one is going to give a story on which consciousness is favored by natural selection, one had best have some evidence that selection for consciousness has occurred. Some sorts of evidence that would fit the bill would be experimental or fossil evidence. It is difficult to imagine what the fossil evidence of consciousness might be. There are no ancient but well-preserved experiences of red, or of the taste of a fine wine. And experimental evidence is hard to get. Even if we knew the relevant variables and were willing to ignore some philosophical skepticisms, the sorts of experiments that would have to be conducted are generally thought to be unethical (even on undergraduates.) Similar ethical problems arise for getting evidence of heritability. Since consciousness is (we believe) fixed in our population, we can't observe selection for it.

The fourth element requires information about population structure that is relevant to patterns of gene flow and of selective homogeneity and heterogeneity. These can be relevant because some selective models directly appeal to population structure. The fifth element requires information about trait polarity; that is, what evolved from what. We need evidence that non-conscious creatures evolved into conscious creatures, not vice versa.¹⁵

If all of this seems rather complicated and a bit too much to grasp as we quickly run it by you, that's because it is. And that's the point. Understanding and taking seriously what it would require to give an evolutionary account of consciousness makes it very clear

vary on a case by case basis. For example, although trait polarity is sometimes in question, he has told me that he believes it is not an issue in the case of consciousness.

¹⁵ *Ibid.*, 165-174.

why it's so hard. Good evolutionary explanations are difficult to give for any trait. The problem of giving an explanation for consciousness inherits all those difficulties intrinsic to adaptationist explanation, and complicates them with all of the philosophical and scientific problems attendant to consciousness, and to gathering empirical evidence regarding consciousness. That is what makes giving an evolutionary explanation of consciousness a hard problem.