CONSCIOUSNESS AND ITS PLACE IN THE COSMOS

Christopher Diak

MAY 18, 2017
PHILOSOPHY OF MIND
Middlebury College
Introduction

David Chalmers wanted to be a materialist. So did I, until I seriously reflected upon the arguments Chalmers presents in *Consciousness and its Place in Nature*. In that article, Chalmers gives an organized account of three arguments against reductionist materialism, explores the philosophical problems which arise from various epistemic positions in relation to materialism, and posits three nonreductionist, nonmaterialist solutions. The philosophical attractions of materialism are myriad, and from the inside - speaking as a former materialist - the framework is so powerful and the evidence so compelling and intuitive, that any other position seems ridiculous. As Chalmers reminds us, “this is not to say that highly counterintuitive claims are always false, but they need to be supported by extremely strong arguments.”

In this paper I will explore the strengths of the three primary arguments Chalmers gives against materialism: the explanatory argument, the conceivability argument, and the knowledge argument. I will consider David Papineau’s phenomenal concepts strategy as a reply to the knowledge argument and evaluate Chalmers’ rebuttal against the phenomenal concepts strategy. I will offer my own critique of Papineau’s strategy in tandem with Chalmers’ reply. Finally, I will discuss Chalmers’ monism, the strengths and weaknesses thereof and posit a fundamental law of consciousness derived from his interpretation of consciousness as the intrinsic property of matter.

The Hard Problem of Consciousness

Perhaps the strangest part about being human is that there is something it *is like* to be human. Human beings are equipped with a system of self-consciousness: metacognitive awareness that enables us to step back from our experience and conceptualize our consciousness as an intentional object. I’ve written on the reification of consciousness in this way before, so I won’t focus on that problem in this paper. Instead, I’ll explore the epistemological problem presented by being able to think about consciousness, commonly called the “explanatory gap”. Chalmers defines the hard problem of consciousness as “the problem of experience.”
Why is it that the physical processes of the brain and body are accompanied by a subjective, phenomenal character? “A solution to the hard problem would involve an account of the relation between physical processes and consciousness, explaining on the basis of natural principles how and why it is that physical processes are associated with states of experience,” (Chalmers, 2002). While he does not posit a solution in Consciousness and its Place in Nature, he does suggest a framework, Type-F Monism, which may be able to accommodate such a solution.

In contrast with the “hard problem of consciousness” Chalmers characterizes the problems of cognitive function and neural structure as the “easy” problems (much to the chagrin of natural science PhDs who would characterize their work as anything but “easy”) because they are in principle solvable by a reductive physicalist account of the world. As an example, during the summer after my freshman year I worked in a biochemistry laboratory at Harvard Medical School, where I investigated the structure and function of a protein called Cdc48. Whatever the fine structure of Cdc48 turns out to be, it is in principle knowable because a set of propositions describing physical facts about the world (i.e. the relative positions of all the atoms in the protein and a simple set of rules for how those atoms interact) should perfectly predict the behavior of the protein. The hard problem of consciousness is different: suppose there is something it is like to be that protein. Could the physical account of the spatiotemporal relations of all the various atoms and the list of rules for their movement demonstrate the phenomenal character of that protein’s experience? Is the phenomenal experience of a protein conceivable? If a protein could be conscious, would the fact of its consciousness be a further fact than all the physical facts? The answers to these three questions are the foci of the explanatory, conceivability, and knowledge arguments against materialism, respectively.

Arguments Against Materialism

The first, and in my view most successful argument against materialism is the explanatory argument. The argument posits a fundamental epistemological gap between what can be described in physical terms and the nature of consciousness. From this epistemological gap, an ontological gap may be presumed (on dualist or monist accounts) or denied (on physicalist accounts).
The Explanatory Argument

(1E) Physical accounts explain at most structure and function.

(2E) Explaining structure and function does not suffice to explain consciousness; so

(3E) No physical account can explain consciousness.

If (1E) and (2E) can be shown to hold, (3E) follows naturally from the premises. (1E) is relatively uncontroversial. (2E) is the critical premise in the explanatory argument. The crux of the issue is whether it is still a conceivable problem, if the structure and function is known completely, for the question of the phenomenal experience of a physical system to be intelligibly posed. The most hard-nosed materialists, Chalmers’ type-A materialists, maintain that “there is no epistemic gap between physical and phenomenal truths; or at least, any apparent epistemic gap is easily closed,” (Chalmers, 2002). The type-A materialist would thus deny premise (2E) outright on empirical grounds. Chalmers claims type-A materialists can only argue by analogy and that such appeals to empiricism beg the question. Two common analogies are 1) “in other areas of science, we accept that explaining the various functions explains the phenomena, so we should accept the same here,” and 2) “in the past, thinkers held that there was an analogous epistemic gap for other phenomena, but that these turned out to be physically explained,” (Chalmers, 2002). On closer scrutiny however, these analogies turn out to be specious. The first class of analogies begs the question by presupposing consciousness is explicable in physical accounts (a circularity) and the second class of analogies fails because structure and/or function (such as the spontaneous movement of living beings as opposed to nonliving objects) has always been the subject of an epistemic gap (where a nonphysical solution was proposed, such as the élan vital) whereas the hard problem posits a further explananda beyond structure and function. Thus, Chalmers says “the type-A materialist needs to address the apparent further explanandum in the case of consciousness head on: either flatly denying it, or giving substantial arguments to dissolve it.” Here, eliminativism and functionalism come to the fore, and I’m inclined to agree with Chalmers that these solutions “deny the manifest.” With regard to functionalism, my critique has always been that positing a
function requires understanding what a property *does*, but it is not clear at all what consciousness *does*. It may be argued that consciousness allows organisms to more efficiently operate in their environment, but if there is causal closure of the physical, and we can conceive of a world wherein the physical states of an organism are not accompanied by a phenomenal character, then an outside observer would not be able to describe a *functional* difference between an organism with a consciousness and an organism without.

The conceivability of physical duplicates of conscious beings lacking phenomenal experience, or *philosophical zombies*, presents a problem for materialists because conceivability is often thought to entail metaphysical possibility (though not necessarily physical possibility), and the meaning of a concept in one world must in some sense be transposable through all metaphysically possible worlds. To give a (relatively) concrete example, the property “gravity” could conceivably fail to exist in some metaphysically possible world, but the *nature of gravity*, the referent by which facts about the existence of gravity could be said to be true or false, would not change between worlds. If “gravity” means something different in world A than in world B, we aren’t making a meaningful metaphysical comparison to say “gravity exists in world A” and “gravity doesn’t exist in world B.” Therefore, if “consciousness” can fail to exist in some world A which is physically identical with the real world, then what we mean by “consciousness” must be nonphysical. Formulated more precisely, the conceivability argument runs as follows:

**The Conceivability Argument**

(1C) It is conceivable that there be zombies

(2C) If it is conceivable that there be zombies, it is metaphysically possible that there be zombies.

(3C) If it is metaphysically possible that there be zombies, then consciousness is nonphysical.

(4C) Consciousness is nonphysical.

Responding to the conceivability argument, type-B materialists (as Chalmers calls them) accept the epistemic gap between phenomenal and physical truths, but deny that this epistemological problem presents an ontological problem. They may argue that conceivability does not entail metaphysical possibility (along fairly technical lines which lie outside the scope of this paper. See (Chalmers 2002: *Does Conceivability Entail Possibility?*)). Or they may argue “phenomenal states can be identified with certain physical or functional
states,” empirically (Chalmers, 2002). To argue that this is not the case, Frank Jackson (1982) devised the famous example of the color-vision scientist, Mary. Mary knows all the physical facts about human color vision, how the brain processes visual information, the structure of electromagnetic radiation, etc. but she has had a strange life: she’s acquired all of this knowledge in a black and white room, on black and white computers, and has herself never seen any color. Jackson argues, and Chalmers agrees, that after Mary leaves the black and white room and experiences color for the first time, she knows something about vision she did not know before. Namely, she knows what it is like to see color. Formulated more precisely, the argument runs as follows:

**The Knowledge Argument**

1. There are truths about consciousness that are not deducible from physical truths.
2. If there are truths about consciousness that are not deducible from physical truths, then materialism is false.
3. Materialism is false.

Chalmers argues that a type-B materialist who accepts the epistemic gap between the physical and phenomenal, but considers the physical identical with the phenomenal, can only accomplish this identity by positing *epistemically primitive* identities, such as those which arise in the fundamental laws of physics. A non-epistemically primitive identity is one which could be empirically discovered, such as the identity of water with $H_2O$, but clearly an empirically discoverable identity does not actually present an insoluble “hard” epistemological problem, but simply a technological or “easy” epistemological problem. This is more akin to type-C materialism, which posits that the explanatory gap is closable “in the limit” as more information is gained about the world, than to true type-B materialism which wants to accept a *true* explanatory gap, but that may simply be because type-B materialism doesn’t work.

**The Phenomenal Concepts Strategy**

In *Thinking about Consciousness*, David Papineau suggests a type-B materialist framework for responding to the explanatory gap and the knowledge argument, called the *phenomenal concepts strategy*. The strategy, which Papineau also calls conceptual dualism, divides the kinds of concepts which exist into two
basic types, material and phenomenal, which both refer to material properties. According to Papineau, “material concepts are those which pick out conscious properties as items in the third-personal, causal world,” such as when we take our consciousness as an intentional object and say “I am.” The “I” is an object which doesn't need to be conceptualized phenomenologically. “When we use phenomenal concepts, we think of mental properties, not as items in the material world, but in terms of what they are like,” (Papineau, 2002). Papineau argues that these two concepts differ at the level of sense, rather than reference, and thus that a priori conclusions about the co-referentiality of a phenomenal and material concept are impossible. Empirical evidence is necessary. Responding to worries that Mary, the color-vision scientist, would not be able to imagine a color before having seen it herself (and thus confirming the truth of the conclusion of the knowledge argument) Papineau suggests a materialist explanation: “Suppose that imaginative re-creation depends on the ability to reactivate some of the same parts of the brain as are activated by the original experience itself. Then it would scarcely be surprising that we can only do this with respect to types of experience we have had previously. We can't form replicas, so to speak, if external stimulation hasn’t fixed a mold in our brains. Less metaphorically, we can only re activate the parts of the brain required for the imaginative re-creation of some type of experience, if some actual experience of that type has previously activated those parts,” (Papineau, 2002). Of course, the precise details of the empirical explanation are always subject to revision, but the basic idea of the phenomenal concepts strategy should be clear.

On Chalmers’ characterization, “proponents put forward a thesis C attributing certain psychological features—call these the key features—to human beings. They argue (1) that C is true: humans actually have the key features; (2) that C explains our epistemic situation with regard to consciousness: C explains why we are confronted with the relevant distinctive epistemic gaps; and (3) that C itself can be explained in physical terms: one can (at least in principle) give a materialistically acceptable explanation of how it is that humans have the key features,” (Chalmers, 2007). Given the story just offered by Papineau, it's easy to see how well Chalmers’ characterization fits Papineau’s picture:

(C) Humans imagine phenomenal concepts.

(1) (C) is true.
(2) There are no phenomenal properties; only phenomenal concepts which refer to material properties. Our explanatory gap with regard to consciousness stems from a lack of *a priori* coupling of phenomenal and material concepts.

(3) C can be explained in physical terms; if one knew how to activate the precise neural network involved in an experience, one could (at least in principle) imagine a phenomenal concept without prior experience.

According to Chalmers, “Papineau sees phenomenal concepts as *quotational concepts*, which represent their referent as *That state* — , where the blank space is filled by an embedded phenomenal state in a way loosely analogous to the way that a word might be embedded between quotation marks. Papineau suggests that even if the embedded state is a neural state, this quotational structure will still give rise to the familiar epistemic gaps,” (Chalmers, 2007). Chalmers reply to the phenomenal concepts strategy is to show that it poses an insoluble dilemma, and that “no account can simultaneously satisfy (2) and (3). Taking P to be the total physical truth, and C to be a psychological feature of human beings, “the key question will be: is P&~C conceivable? That is, can we conceive of beings physically identical to us (in physically identical environments, if necessary) that do not have the psychological features attributed by thesis C?” (Chalmers, 2007). The argument against conceptual dualism is a variation on the conceivability argument discussed above:

**Chalmers’ Conceivability Reply to Conceptual Dualism**

(1) If P&~C is conceivable, then C is not physically explicable.

(2) If P&~C is not conceivable, then C cannot explain our epistemic situation.

(3) Either C is not physically explicable, or C cannot explain our epistemic situation.

Thus, epistemically non-primitive identities of the type suggested by the phenomenal concepts strategy, those which are open to empirical verification, have not brought us any closer to an argument for materialism. It is for this reason that Chalmers posits, as noted above, that the type-B materialist must posit epistemically primitive identities between phenomenal and physical truths. But the very notion of epistemically primitive identities between phenomenal and physical truths belies the proper ontological status that phenomenal truths ought to have. “In effect, the type-B materialist recognizes a principle that has the
epistemic status of a fundamental law, but gives it the ontological status of an identity. An opponent will hold that this move is more akin to theft than to honest toil: elsewhere, identifications are grounded in explanations, and primitive principles are acknowledged as fundamental laws,” (Chalmers, 2002).

**Chalmer's Solution: Type-F Monism**

Perhaps the fundamental is the correct realm for a theory of consciousness. In *Consciousness and its Place in Nature*, substance dualism and epiphenomenalism are explored in addition to the three materialist philosophies so far discussed, and they are found wanting. Substance dualism, while a commonsense religious viewpoint for billions of people, faces ontological and logical problems such as the coupling problem, interaction problem, Ockham’s Razor (why should there be two types of fundamental substance rather than one?) and the problem of pluralism with respect to religion (if souls exist, why should this fact be known to some but not all people?). Epiphenomenalism is a dark specter in the field of consciousness studies, which threatens a disturbing roller coaster of a universe wherein phenomenal subjects are locked into bodies, entirely lacking mental agency, for a crazy conscious ride through the cosmos. On purely aesthetic grounds, we should search for any other tenable position before being relegated to epiphenomenalism. Chalmers' monism could suggest a brighter picture.

Premised on the work of Bertrand Russell’s *The Analysis of Matter*, type-F monism “is the view that consciousness is constituted by the intrinsic properties of fundamental physical entities: that is, by the categorical bases of fundamental physical dispositions,” (Chalmers, 2002). Fundamental physical dispositions, such as mass or electric charge, are concepts which allow researchers to describe the behavior of objects; they are not ontological claims about the nature of reality. “Mass” is the dispositional property of an object to resist acceleration – a description, not a referent. The nature of the referent, the intrinsic property which makes particles display the disposition of mass (or whatever else) is thus an open metaphysical problem. “At the same time, there is another metaphysical problem: how can phenomenal properties be integrated with the physical world? Phenomenal properties seem to be intrinsic properties that are hard to fit in with the structural/dynamic character of physical theory; and arguably, they are the only intrinsic properties that we have direct knowledge of,” (Chalmers, 2002). Importantly, this view doesn’t alter the *structure or function* of the
physical world, because all that is being supposed about that structure is that it has intrinsic phenomenal properties.

There is one important, “principled” problem for Chalmerian monists to address before a panpsychism or protopanpsychism can be honestly advocated: the combination problem. The combination problem is the problem of how our highly structured, specific phenomenological experience could arise from the interplay of myriad microphysical systems. To understand how consciousness can be unified at a very high level of complexity and at the most fundamental scale, I suggest we consider a fundamental law of consciousness that would arise in a Chalmerian framework: conservation of consciousness.

**Conservation of Consciousness and a Thought Experiment**

1. Consciousness is the intrinsic property of all matter.
2. By the mass-energy equivalence, all matter is energy.
3. Consciousness is the intrinsic property of all energy.
4. Energy is conserved over time.
5. Consciousness is conserved over time.

The combination problem arises because the microphysical systems which comprise our bodies are not being considered in the proper context. “To answer it, it seems that we need a much better understanding of the compositional principles of phenomenology: that is, the principles by which phenomenal properties can be composed or constituted from underlying phenomenal properties, or protophenomenal properties,” (Chalmers, 2002). If, speculative as it may be, the conservation law of consciousness was shown to be true, then it can solve the combination problem as follows:

1. Consciousness is the intrinsic property of all energy.
2. All energy is causally continuous with the Big Bang.
3. All consciousness is causally continuous with the Big Bang (consciousness is conserved).
4. The combination problem asks: how can individual microsystems produce unified consciousness?
5. By (2) there are no individual microsystems. Every physical system is a concrete particular with a specific spatiotemporal relationship with all other energy.
There is a unified physical system of behavior in the universe, so, consciousness is unified.

This solution dissolves the combination problem by attacking its logic. The conclusion of the argument, however, may seem to border on solipsism or idealism, but as far as I can see it is the only principled solution to the combination problem, as any attempt to correlate qualia with the structure (however context specific) of a physical system will ultimately fall prey to the same epistemic problems we considered earlier in this paper. If this is the state of affairs, then how would it be possible for different conscious beings to exist?

My answer, however speculative, is that some sort of experiential relativity, akin to General Relativity, must exist which could translate between the phenomenological states of various positions and combinations of particles. The motivation for this experiential relativity is a thought experiment: suppose there is something it would be like to be the Big Bang. What would it feel like? The answer is that it would feel just like this. Here’s a vivid illustration of the idea:

It’s like you took a bottle of ink and you threw it at a wall. Smash! And all that ink spread. And in the middle, it’s dense, isn’t it? And as it gets out on the edge, the little droplets get finer and finer and make more complicated patterns, see? So in the same way, there was a big bang at the beginning of things and it spread. And you and I, sitting here in this room, as complicated human beings, are way, way out on the fringe of that bang. We are the complicated little patterns on the end of it. Very interesting. But so we define ourselves as being only that. If you think that you are only inside your skin, you define yourself as one very complicated little curlicue, way out on the edge of that explosion. Way out in space, and way out in time. Billions of years ago, you were a big bang, but now you’re a complicated human being. And then we cut ourselves off, like this, and don’t feel that we’re still the big bang. But you are. Depends how you define yourself. You are actually--if this is the way things started, if there was a big bang in the beginning--you’re not something that’s a result of the big bang, on the end of the process. You are still the process. You are the big bang, the original force of the universe, coming on as whoever you are. (Watts 1960).

That every being in the Universe would answer the question in the same way, “just like this,” seems to me the same kind of constant that Einstein obsessed over when he began formulating his theory of relativity: the constancy of the speed of light. Along with Type-F monism, experiential relativity is an under-investigated possibility of life. Considering Type-F monism, Chalmers comments, “no-one has yet developed any sort of detailed theory in this class, and it is not yet clear whether such a theory can be developed,” (Chalmers, 2002). That the hard problem really is hard means a drastically different sort of explanatory technology will need to be developed if serious scientific work is to be conducted on the problem. The nature of this framework will be the subject of my next series of philosophical investigations.
REFERENCE


