

Lecture 1: The Knowledge Argument

This lecture we consider the case for a physicalist view. We look at how physicalists have responded to the arguments for property dualism set out in the previous chapter.

Let us then look at some physicalist responses to the arguments in support of property dualism, beginning with the knowledge argument.

I. Honing in on the Phenomenon of Consciousness

Intuitive Motivation

Suppose you have just had a dental procedure under general anesthetic and are coming round. You are aware of a dazzling light above you and of a muffled voice echoing in your ears. There is sickness in your stomach and a sharp metallic taste in your mouth. You feel a moment of panic as you struggle to work out what happened. Moving your head, you recognize the dentist's face and realize that he is speaking your name and asking if you want a glass of water. You remember where you are, sit up shakily and take the glass. (Frankish, Consciousness)

What happened as you regained consciousness?

- Various body processes resumed: sense organs started functioning again,..
- Brain processing started up.

But in addition you started to have conscious experiences – experiences with a certain feel to them.

Different term to refer to this aspect of experiences: qualitative feel phenomenal feel, phenomenology, subjective character, 'what-it's-like', 'qualia', etc.

When we talk about consciousness, we are not interested in the nature of perceptions, sensations and thoughts as such but rather on what is special about those perceptions, sensations, and thoughts that have a feel to them.

We need to distinguish consciousness from some related phenomena.

Creature Consciousness: When we say that we are conscious, we mean that we are awake, as opposed to being asleep or knocked out. This is creature consciousness.

State Consciousness: When we say that an experience is conscious, we mean that they are of the sort that have a phenomenal character to them. This is state consciousness.

According to Ned Block, we also need to distinguish two kinds of state consciousness:

Phenomenal Consciousness: A mental state is phenomenally conscious if it has a phenomenal character.

Access-consciousness: A mental state is access-conscious if the information it carries is directly available to other mental processes, including reasoning, behavioral control and speech.

II. The Problem of Consciousness

The Easy and the Hard Problem

Recall functionalism: according to functionalism mental states and processes can be defined functionally, in terms of the causal role they play in the operation of the mind.

Functionalism can explain many types of consciousness. According to Chalmers, it can explain:

- the ability to discriminate, categorize, and react to environmental stimuli
- the integration of information by a cognitive system
- the reportability of mental states
- the ability of a system to access its own internal states
- the focus of attention
- deliberate control of behavior

Part II, Consciousness

- the difference between wakefulness and sleep

But can it explain phenomenal consciousness? Here's a reason to be skeptical: functionalism characterizes mental states by what they do, rather than by how they feel. And it seems that a brain state could play the functional role of an experience without having any phenomenal character to it.

Why are the easy problems easy, and why is the hard problem hard? The easy problems are easy precisely because they concern the explanation of cognitive *abilities* and *functions*. To explain a cognitive function, we need only specify a mechanism that can perform the function. The methods of cognitive science are well-suited for this sort of explanation, and so are well-suited to the easy problems of consciousness. By contrast, the hard problem is hard precisely because it is not a problem about the performance of functions. The problem persists even when the performance of all the relevant functions is explained. [...] How do we explain the performance of a function? By specifying a mechanism that performs the function. Here, neurophysiological and cognitive modeling are perfect for the task. If we want a detailed low-level explanation, we can specify the neural mechanism that is responsible for the function. If we want a more abstract explanation, we can specify a mechanism in computational terms. (Chalmers, 1995)

III. The Explanatory Gap Argument

This argument exploits the distinction between the "easy" and "hard" problem of consciousness. Versions of this argument have been presented by Levine, McGinn, and Nagel. Here's Chalmers's version:

Premise 1: Physical accounts explain at most structure and function.

Premise 2: Explaining structure and function does not suffice to explain consciousness.

Conclusion: No physical account can explain consciousness.

If we add the premise that what cannot be physically explained is not itself physical, then physicalism about consciousness is false.

IV. The Knowledge Argument

Here is Jackson's thought experiment:

Mary is a brilliant scientist who is, for whatever reason, forced to investigate the world from a black and white room via a black and white television monitor. She specializes in the neurophysiology of vision and acquires, let us suppose, all the physical information there is to obtain about what goes on when we see ripe tomatoes, or the sky, and use terms like 'red', 'blue', and so on. [...] What will happen when Mary is released from her black and white room or is given a colour television monitor? Will she learn anything or not? It seems just obvious that she will learn something about the world and our visual experience of it. But then it is inescapable that her previous knowledge was incomplete. But she had all the physical information. Ergo there is more to have than that, and Physicalism is false.

Here's how we can set out the argument:

Premise 1: If physicalism is true, then on leaving her room Mary will not learn any new facts about colour vision (since, by hypothesis, she already knows all the physical facts about it).

Premise 2: On leaving her room, Mary learns new facts about colour vision (namely what it's like to see various colours).

Conclusion: Physicalism is not true.

A couple of clarifications:

- It's important that Mary learns all the physical facts about colour vision – not just the currently known facts. Mary lives in the future! This makes the case harder to imagine.
- Mary need not lack the ability to imagine what colour vision is like. What's important is that she lacks *knowledge* of what it's like.
- What's at issue isn't whether she will learn new things upon her release. Rather what's at issue is whether she will learn something about her specialist subject, colour vision:

the knowledge Mary lacked which is of particular point for the knowledge argument against physicalism is knowledge about the experiences of others, not about her own. When she is let out, she has new experiences, color experiences she has never had before. It is not, therefore, an objection to physicalism that she learns something on being let out. Before she was let out, she could not have known facts about her experience of red, for there were no such facts to know. That physicalists and nonphysicalists alike can agree on. [...] The trouble for physicalism is that, after Mary sees her first ripe tomato, she will realize how impoverished her conception of the mental life of others has been all along. She will realize that there was, all the time she was carrying out her laborious investigations into the neurophysiologies of others and into the functional roles of their internal states, something about these people she was quite unaware of. All along their experiences (or many of them, those got from the tomatoes, the sky,...) had a feature conspicuous to them but until now hidden from her (in fact, not in logic). But she knew all the physical facts about them all along; hence, what she did not know until her release is not a physical fact about their experiences. But it is a fact about them. That is the trouble for physicalism.

V. The Conceivability Argument

This is the second anti-physicalist argument we'll look at. Descartes made a version of it a while ago in support of substance dualism. It made a come-back under Kripke. We'll look at a version by Chalmers:

Premise 1: It is conceivable that there are zombies.

Premise 2: If it is conceivable that there be zombies, it is metaphysically possible that there are zombies.

Premise 3: If it is metaphysically possible that there be zombies, then consciousness is nonphysical.

Conclusion: Consciousness is nonphysical.

A couple of clarifications:

- Zombies here are not of the flesh-eating variety. Here's what Chalmers has in mind:

Consider my zombie twin. This creature is molecule for molecule identical to me, and identical in all the low-level properties postulated by a completed physics, but he lacks conscious experience entirely. [...] To fix ideas, we can imagine that right now I am gazing out the window, experiencing some nice green sensations from seeing the trees outside, having pleasant taste experiences through munching on a chocolate bar, and feeling a dull aching sensation in my right shoulder. What is going on in my zombie twin? He is physically identical to me, and we may as well suppose that he is embedded in an identical environment. He will certainly be identical to me functionally: he will be processing the same sort of information, reacting in a similar way to inputs, with his internal configurations being modified appropriately and with indistinguishable behavior resulting. He will be psychologically identical to me. [...] He will even be 'conscious' in the functional senses [...] – he will be awake, able to report the contents of his internal states, able to focus attention in various places, and so on. It is just that none of this functioning will be accompanied by any real conscious experience. There will be no phenomenal feel. There is nothing it is like to be a zombie. (pp. 94-98)

- What Chalmers needs is not that zombies actually exist or that they could exist in the actual world. All he needs is for them to be conceivable. To say that a situation is conceivable is to say that we can form a coherent and detailed conception of it, free from contradictions.
- Metaphysical possibility is not limited by the prevailing laws of nature. Premise 2 is an instance of the general principle that if something is conceivable, then it is at least metaphysically possible. This doesn't seem implausible.
- The last premise is the least controversial one. If it is possible for basic physical properties and phenomenal properties to come apart in this way, then those properties must actually be distinct – even if, as a matter of fact, they always go together. And to say that phenomenal properties are distinct from basic physical ones is to say that consciousness is non-physical.

VI. Connections between the three Arguments

The three arguments are closely related. They all start by establishing an epistemic gap between the physical and the phenomenal domains. In particular, each denies a different form of epistemic entailment from physical facts to phenomenal facts:

- i. explaining phenomenal facts in terms of physical facts
- ii. deducing (coming to know) phenomenal facts from physical facts

Part II, Consciousness

- iii. conceiving of phenomenal facts after reflectively conceiving the physical facts.

They then move to infer an ontological gap:

- i. from failure to physical explanation to nonphysicality
- ii. from failure of deducibility to difference in facts
- iii. from conceivability to metaphysical possibility.

VII. Assessing Property Dualism

Suppose we take the arguments at face value. If phenomenal properties are fundamentally distinct from physical ones, then it follows that consciousness cannot be reductively explained in the standard ways. The challenge for the property dualists is to go beyond mystery-mongering and explain how their view of consciousness fits in with our naturalistic outlook.

The central challenge here is causal efficacy. If consciousness is non-physical, does it have a causal role to play? Intuitively, it does play such a role:

- Jack made an emergency appointment with the dentist because his toothache was unbearably painful.
- Jill ate a second piece of cake because it tasted so delicious.
- Bob left the performance early because he couldn't bear the screechy voice of the lead actor.

But there is strong evidence that all events at the basic physical level are causally closed: we need never to appeal to anything outside the basic physical realm in order to explain events within it. This is compatible with high-level properties having causal powers insofar as they inherit their causal powers from the more basic properties in which they are realized.

The property dualist faces an inconsistent triad:

- I. Property dualism: Consciousness is non-physical.
- II. Efficacy of consciousness: Consciousness has effects in the physical world.
- III. Closure of the physical: Only physical phenomenal have effects in the physical world.

IX. What are the options?

Give up III: this is *Interactionism*.

The suggestion is that new causal powers arise in the brains of conscious creatures – powers which go beyond those of the brain's basic physical component and exert a 'downward' influence on the behavior of those components. If this is true, then the behavior of conscious creatures cannot be explained in the same way as that of inanimate things, since there will be new causal powers at work in the former. This is unattractive in light scientific evidence.

Give up II: this is *Epiphenomenalism*.

On this view, consciousness is merely a by-product of brain activity. It's important that on this view it's not that experiences do not have causal influence – only that certain properties of them do not. There may be no knock-down arguments against epiphenomenalism but it's quite counterintuitive. Here's Chalmers:

To see the problem in a particularly vivid way, think of my zombie twin in the universe next door. He talks about consciousness all the time – in fact, he seems obsessed by it. [...] He frequently gets into arguments with zombie materialists, arguing that their position cannot do justice to the realities of conscious experience. And yet he has no conscious experience at all! In his universe, the materialists are right and he is wrong. [...] To strengthen the sense of paradox, note that my zombie twin is himself engaging in reasoning just like this. He has been known to lament the fate of his zombie twin, who spends all his time worrying about consciousness despite the fact that he has none. (pp. 177-181)

A more subtle move: *Panprotopsychism*

We saw that one way to reconcile the efficacy of consciousness with the causal closure of the physical world is to maintain that conscious states are realized in basic physical ones. But, Chalmers, suggests, we could also take the opposite view and hold that basic physical states are realized in conscious ones – that phenomenal properties, or rudimentary versions of them, are found at the fundamental level of physical reality, in the basic

Part II, Consciousness

physical particles themselves. If so, then consciousness will be closely integrated with the physical world and the causal processes that occur there.

According to Chalmers, science must expand to incorporate a view of consciousness as a non-physical phenomenon. He calls this view 'naturalistic dualism'.

To make this more precise, we'll need some distinctions between different types of supervenience:

- Natural supervenience: To say that A properties supervene naturally on B properties is to say that in the real world it is a law of nature that B properties determine A properties.
- Metaphysical supervenience: To say that A properties supervene metaphysically on B properties is to say that a thing's B properties determine its A properties in all possible worlds, including those in which the laws of nature are different.
- Logical supervenience: to say that A properties supervene logically on B properties is to say that it is not even conceivable that a thing's A properties could vary without change to its B properties.

The suggestion is as follows:

- Physics characterizes basic physical particles in terms of their relations and dispositions – the way they interact with other particles and their tendencies to produce certain effects. It does not say anything about their intrinsic properties. Yet they must have some intrinsic properties. The suggestion is that these properties are in fact phenomenal ones. Crudely put, subatomic particles are little sparks of proto-consciousness. The physical and phenomenal worlds we know emerge from these particles – the former from their activity and the latter from their intrinsic properties.
- This is compatible with causal closure. But it also gives consciousness a causal role, since it treats phenomenal properties as essential features of the entities mentioned in the causal explanations given by physics.
- If the intrinsic properties are phenomenal ones, then we have to suppose that it is like something to be an electron. Chalmers isn't suggesting that electrons have rich phenomenal lives or that they engage in conscious thought, just that they have a tiny spark of consciousness.

Here's Chalmers' statement of his view:

The arguments do not lead us to a dualism such as that of Descartes, with a separate realm of mental substance that exerts its own influence on physical processes. The best evidence of contemporary science tells us that the physical world is more or less causally closed: for every physical event, there is a physical sufficient cause. If so, there is no room for a mental "ghost in the machine" to do any extra causal work. [...] The dualism implied here is instead a kind of property dualism: conscious experience involves properties of an individual that are not entailed by the physical properties of that individual, although they may depend lawfully on those properties. Consciousness is a feature of the world over and above the physical features of the world. [...] It remains plausible, however, that consciousness arises from a physical basis, even though it is not entailed by that basis. The position we are left with is that consciousness arises from a physical substrate in virtue of certain contingent laws of nature, which are not themselves implied by physical laws. [...]

Although it is a variety of dualism, there is nothing antiscientific or supernatural about this view. The best way to think about it is as follows. Physics postulates a number of fundamental features of the world: space-time, mass-energy, charge, spin and so on. It also posits a number of fundamental laws in virtue of which these fundamental features are related. Fundamental features cannot be explained in terms of more basic features, and fundamental laws cannot be explained in terms of more basic laws; they must simply be taken as primitive. Once the fundamental laws and the distribution of the fundamental features are set in place, however, almost everything about the world follows. That is why a fundamental theory in physics is sometimes known as a "theory of everything". But the fact that consciousness does not supervene on the physical features shows us that this physical theory is not *quite* a theory of everything. To bring consciousness within the scope of a fundamental theory, we need to introduce new fundamental properties and laws. [...]

There are two ways this might go. Perhaps we might take experience itself as a fundamental feature of the world, alongside space-time, spin, charge, and the like. That is, certain phenomenal properties will have to be taken as basic properties. Alternatively, perhaps there is some other class of novel

Part II, Consciousness

fundamental properties from which phenomenal properties are derived. Previous arguments have shown that these cannot be physical properties, but perhaps they are nonphysical properties of a new variety, on which phenomenal properties are logically supervenient. [...]

Where we have new fundamental properties, we also have new fundamental laws. Here the fundamental laws will be psychophysical laws, specifying how phenomenal (or protophenomenal) properties depend on physical properties. These laws will not interfere with physical laws; physical laws already form a closed system. Instead, they will be supervenience laws, telling us how experience arises from physical processes. (pp. 124-128)

VIII. The Causal Role of Consciousness

X. Readings

Chalmers, D. (1996) *The Conscious Mind*. OUP.

Jackson, F. (1982) Epiphenomenal Qualia. *Phil Quarterly*.