

## A Challenge to the thesis that conscious states are physical states

### **The explanatory gap**

We may know that certain conditions are necessary and sufficient for a particular phenomenon to obtain and yet not know why those conditions are necessary and sufficient for that phenomenon. To explain that phenomenon, we need an explanatory theory.

For instance, we may know that fuel, oxygen, and heat are each necessary and together sufficient for fire and yet not understand fire. To understand fire, we need an explanatory theory that describes why fire appears always and only when fuel, oxygen, and heat are present.

Now we might know that certain conditions are necessary and sufficient for a particular phenomenon while not knowing how those conditions produce and sustain that phenomenon while nonetheless having a pretty good idea what a true explanatory theory would look like.

For instance, the early chemists knew what was necessary and sufficient to produce water and had no theory that would explain how bringing about sufficient conditions for water produced water. But they did have good ideas what would count as an explanation here.

The explanatory gap with regard to consciousness is this: Even if we did know the necessary and sufficient physical conditions for the production and sustaining of consciousness, not only would we lack an explanatory theory of how those conditions produce and sustain consciousness, we would have no idea how those conditions produce and sustain consciousness.

**We simply have no idea what a theory that explained consciousness on the basis of its necessary and sufficient physical conditions would look like.**

How is this claim an objection to the thesis that conscious states are physical states?

### **The identity thesis**

The physicalistic identity thesis of mind and body is that each conscious state is (is identical to) some physical state. There are not two states, says the physicalist about mind, however intimately related the two states might be supposed to be. There is one state and only one state.

### **The physicalist commitment**

Each physical fact is either brute or dependent on other facts (and ultimately on brute facts). Facts dependent on other facts can be explained by citing those other facts. That an animal is conscious is too high-level a fact to be brute. Thus, it must be explicable through a theory that mentions only lower-level physical facts.

### **The argument from the explanatory gap that challenges the thesis that conscious states are physical states**

1. The fact that an animal with such-and-such a nervous system is conscious is not a brute fact.

Thus (from 1): 2. Certain facts about that nervous system, if we knew them, would explain why that animal is conscious.

3. All facts about that animal's nervous system are entirely physical facts.

Thus: 4 (from 2 and 3). If we knew those facts, we would understand how a nervous system can produce and sustain consciousness.

5. We have no idea how physical facts about an animal's nervous system, whatever those facts are, could explain how that animal is conscious. (The explanatory gap.)

Therefore (from 4 and 5): 6. Likely, either premise 1 or premise 3 in the argument is false.

In short: If minds are brains and mentality can be explained by physical processes only, then we would have at least some confidence that we could recognize a plausible theory of how brains produce mentality. But we can have no idea how a physical system could be the site of consciousness. Thus, likely, one of our assumptions is mistaken.

### **Which assumption is false?**

1) Maybe minds are not brains, even though minds are in intimate connection with brains (or with whatever physical states are correlated with consciousness). Perhaps the sufficient physical conditions for consciousness are not all the conditions for consciousness. Perhaps consciousness arises from physical conditions together with some non-physical condition.

If minds are not identical to brains, we would not be able to explain the presence of consciousness entirely by noting physical processes; we would have also to note the presence of some non-physical thing.

2) Maybe it is a brute fact that when sufficient physical conditions are present, consciousness obtains. Consciousness, then, is an emergent property of properly organized neural matter. We can have no

explanatory theory of its basis, no more than we can have of any basic brute fact.

### **How can we get metaphysical conclusions from an epistemic claim?**

That there is an explanatory gap is a claim about our understanding of the world. But the challenge it poses is to a metaphysical theory, namely, the physicalist identity thesis of mind and brain.

The gap can be understood as arising for one of three reasons.

1) We simply have not yet acquired the concepts necessary to understand how physical systems could be conscious ones. Investigators might eventually come to fashion the needed concepts. (Our ancestors might have had good knowledge of the necessary conditions for procreation and yet had no idea what an explanation of the phenomenon of procreation would look like.)

If this is why there is a gap, then its existence poses no challenge to physicalism.

2) Our human brain and mind evolved in a certain range of circumstances and because of its or their ability to help us live long and prosper. We cannot conceive of the theories that would explain how physical systems can become conscious simply because of our cognitive limitations.

If this is why there is a gap, then, again, its existence might pose no challenge to physicalism, for physicalism might be true. We just are condemned never to understand how it is true.

3) Either physicalism is false, and some sort of dualism is true, or consciousness is a brute emergent feature of a physical world. That

would explain the gap. And then the existence of the gap would falsify either physicalism itself or the idea that the emergence of consciousness can be explained.