

The knowledge argument¹

This handout follows the handout on 'Property dualism'. You should read that handout first.

Property dualism is the view that, although there is just one kind of substance, physical substance, at least some mental properties are not physical properties (as type identity theory claims) nor functional properties (as functionalism claims), nor are they behavioural dispositions (as philosophical behaviourism claims). While mental properties are possessed by physical substances, they are a fundamentally different kind of property from physical properties.

Physicalism is the view that everything that exists - every substance, every property that substances have, every event that occurs - is either physical or completely depends upon something that is physical. 'Physical' means something that comes under the laws and investigations of physics, and whose essential properties are identified and described by physics. According to physicalism, the physical properties of the world determine all the properties of the world, not just causally but metaphysically. Property dualism rejects this. It claims that some mental properties are properties that are not determined by physical properties in the way that physicalism claims. They are something new, something in addition to physical properties.

THE KNOWLEDGE ARGUMENT

In 'Epiphenomenal qualia', Frank Jackson defends property dualism on the basis of his 'knowledge argument'. He describes the following scenario. Suppose there is a neuroscientist, Mary, who has lived all her life in a room in which everything is black and white. She has never seen any colour other than black, white and shades of grey. However, she has specialised in the science of vision, and through textbooks and black-and-white TV, she has come to know every physical fact there is to know about colour vision - everything about the properties of light, everything about the eye, everything about the nerves and the brain related to vision. So, Mary knows all the physical information there is to know about what happens when we see a ripe tomato. She is then let out of the black-and-white room, and comes to see something red for the first time. Does she learn something new?

Jackson claims that 'it seems just obvious' that she will. She will learn about what it is like to see the colour red. And so she learns something new about our visual experience of the world. However, we said that she knew all the physical facts while she was in the room. So not all the facts are physical facts. It is possible to know all about the physical properties of the brain involved in having an

¹ This handout is based on material from Lacewing, M. (2017) *Philosophy for A Level: Metaphysics of God and Metaphysics of Mind* (London: Routledge), Ch. 3, pp. 291-99

experience and yet not know about the qualia.

- P1. Mary knows all the physical facts about seeing colours before being released from her black-and-white room.
- P2. On being released, she learns new facts about seeing colours.
- C1. Therefore, not all facts are physical facts, e.g. some facts about colours are not.
- C2. Therefore, phenomenal properties are non-physical and physicalism is false.

By 'all the physical facts', Jackson means not only what we already know about physics and neurophysiology. Mary knows all the physical facts as discovered by a completed physics and neuroscience. Furthermore, she has worked out all the causal and functional facts that are entailed by these facts. Because physicalism claims that the world is entirely physical (if we include causal and functional properties), it must claim that to have complete physical knowledge is to have complete knowledge. But no amount of physical information can enable Mary to know what it is like to see a ripe tomato.

RESPONSES TO THE KNOWLEDGE ARGUMENT

Physicalist responses to Jackson's argument point out that there is more than one meaning of 'to know', more than one kind of knowledge. We can and should accept that Mary gains new knowledge when she sees red for the first time. But this doesn't mean that she gains knowledge of some new fact. We will look at three different responses offering alternative accounts of just what Mary learns.

Mary does not gain new propositional knowledge, but does gain ability knowledge

The first response argues that instead of gaining knowledge of a fact, described by a proposition (e.g. 'that red looks like this'), Mary gains know-how - the knowledge involved in certain abilities. For instance, to see red for the first time is to gain the ability to know how to imagine or recognise red. So Jackson hasn't shown that there are any facts that are not physical facts.

We can challenge this objection as follows. Suppose that seeing red gives us these new abilities. Are such abilities all that is involved in knowing what it is like to see red? Suppose Mary wonders whether what it is like for others to see red is the same as what it is like for her. She isn't wondering about her abilities to imagine and recognise red. She is wondering about the truth of a proposition. So when Mary first learns what it is like to see red, she does gain knowledge of a new fact.

Is the objection even right to think that knowing what it is like to see red involves knowing how to imagine red? Suppose there is someone who (for whatever reason) has no ability to imagine seeing red. Now suppose this person looks attentively at something red. While they look at red, they know what it is like to see red. And yet they cannot imagine seeing red. This shows that the ability to imagine is not necessary for knowing what it is like to see red. Now suppose someone else has the most amazing ability to imagine seeing colours. They are told that there is a shade of red, e.g. burgundy, that is between plum red and tomato red. They are now able to imagine burgundy, but as long as they don't actually imagine burgundy, they still don't know what it is like to see burgundy. This shows that the ability to

imagine a colour is not sufficient to know what it is like to see it. (We can make similar arguments for recognising colours.)

If the ability to imagine seeing red is neither necessary nor sufficient for knowing what it is like to see red, then when Mary comes to know what it is like to see red, she learns more than simply knowing how to imagine seeing red. The response fails to show that Mary does not learn a new fact. It fails to show that the knowledge argument is mistaken.

Mary does not gain new propositional knowledge, but does gain acquaintance knowledge

A second response to Jackson's argument argues that Mary gains a different kind of knowledge again, not propositional knowledge (knowing that), but not ability knowledge (knowing how) either. Instead, she gains 'acquaintance knowledge' - knowledge given by direct awareness of something in experience, e.g. a person, a place, or one's own thoughts and feelings. To see red is a direct apprehension of red, as contrasted with descriptions of seeing red. How does the objection work?

Suppose that what it is like to see red is a physical property of the visual experience, which itself is a physical process. In other words, the phenomenal property of what it is like to see red is some property of the brain (type identity). Mary can then know all about this physical property, about what it is, when it occurs, and so on, before she leaves the room. However, she is not acquainted with the property - she doesn't have direct knowledge of it because her brain has never itself had this property. When she sees red, this property occurs in her brain and she becomes acquainted with it. She gains new knowledge, but she hasn't learned any new fact. She already knew all about this property before she left the room. (Compare: a friend describes someone you have never met. When you first meet the person and become acquainted with them, you think of them in a new way. But the person you meet was someone you already knew about.)

In *Brainwise*, Patricia Churchland puts the two responses together. Knowing the neuroscience won't help you experience or identify phenomenal properties in consciousness. For this, the theory needs to be true of your brain, i.e. your brain needs to undergo the processes that the theory describes as constituting colour experience. This fact doesn't mean that there is something that the theory misses out. When Mary's brain actually undergoes the processes that she knows all about, then she will be acquainted with colour and gain abilities of recognition etc. But that is all the colour experience is. Nothing in addition to the physical processes is needed or occurs.

There are two possible responses to this objection. First, we can argue that acquaintance knowledge involves propositional knowledge. What it is to be acquainted with red is to know that seeing red is like this (having the experience). Becoming acquainted with red involves learning some new fact. So Mary does learn a new, and therefore non-physical, fact when she becomes acquainted with red. So what it is like to experience red can't simply be a physical property of the brain.

Second, we can argue that the objection misunderstands the argument. The

knowledge argument isn't about Mary's experience. The argument is that Mary didn't know everything about other people's experiences before she left the room, even though she knew everything physical about their experiences. Mary doesn't know what it is like for anyone to experience red. This is a fact about experiences that Mary doesn't know. When Mary leaves the room, she realises how impoverished her conception of people's colour experiences has been. So there are facts about other people's experiences of seeing red that Mary learns.

- P1. Mary (before her release) knows everything physical there is to know about other people when they see colour.
- P2. Mary (before her release) does not know everything there is to know about other people when they see colour (because she learns something about them on her release).
- C1. Therefore, there are truths about other people (and herself) when they see colour which escape the physicalist story.
- C2. Therefore, phenomenal properties are non-physical and physicalism is false.

Mary gains new propositional knowledge, but this is knowledge of physical facts that she already knew in a different way

A third response to Jackson's argument distinguishes between two ways we might talk about 'facts' on the basis of the distinction between concepts and properties.

Suppose I know that there is water in the glass. Is that the same as knowing that there is H₂O in the glass? No - because someone may know one of these truths without knowing the other. Someone can have the concept WATER without having the concept H₂O. Or again, someone may have both concepts, but not know that water and H₂O are the same thing. So we can say that to know that there is water in the glass and to know that there is H₂O in the glass is to know two different facts. In this sense of 'fact', we count facts in terms of concepts.

However, in another sense of 'fact', the fact that there is water in the glass just is the fact that there is H₂O in the glass, because water and H₂O are identical - one thing. Both of these claims are made true by just one state of affairs in the world. In this sense of 'fact', we count facts in terms of how the world is, not how we think about it. Another way of expressing this is to say that the property of being water and the property of being H₂O are one and the same property. We will use 'fact' in this sense from now on.

We can now apply this to the knowledge argument. Before leaving the room, Mary has a concept of red in physical terms - wavelengths of light, neurons firing, and so on. Call this the 'physical' or again a 'theoretical' concept of red, RED_{TH}. Or again, using Chalmers' distinction between psychological and phenomenal concepts, Mary knows what it is to see red in the psychological sense of 'seeing red'. We can contrast this with a 'phenomenal' concept of red, RED_{PH}. A phenomenal concept of something is the concept by which you recognise something when you experience or perceive it. So we gain the phenomenal concept RED_{PH} by seeing red. Before she leaves the room, Mary doesn't know what it is to see red in the phenomenal sense.

When Mary comes out the room and sees red, she acquires the phenomenal concept RED_{PH} for the first time. She is now able to think about red in a new way,

in terms of what it is like to see red. She couldn't know what it is like to see red before because she didn't have the phenomenal concept. But, we can claim, the phenomenal concept RED_{PH} is a concept of the same thing that her theoretical concept RED_{TH} is a concept of - they are two different concepts of a physical property of the brain (like WATER and H_2O are two concepts of the same physical substance). Mary gains new propositional knowledge about seeing red in one sense (because she gains a new concept) but her new knowledge is about a property that she already knew about under a different concept. The theoretical concept RED_{TH} and the phenomenal concept RED_{PH} are two concepts that refer to the same property.

Let us accept that the knowledge argument shows that there are different ways of thinking about physical things, some of which depend on experiencing, rather than describing. To know what it is like to see red, you need to have the phenomenal concept RED_{PH} , and this you can only gain from experience. So Mary gains knowledge of a new fact, in the sense of fact that relates to concepts.

However, physicalism and property dualism are claims about what exists. They are claims about properties, not about concepts. The knowledge argument does not show that Mary gains knowledge of a new property. It doesn't show that Mary learns about something in the world that she didn't know about before. And so it doesn't show that what it is like to see red cannot be a physical property. So the argument fails to show that there are any non-physical properties. So it fails to show that physicalism is false.