

Consciousness

I. The Knowledge Argument (Jackson)

1. Mary learns all the physical facts in the black and white room. (Assumption)
2. If Mary learns all the facts in the black and white room, Mary doesn't learn something new when she leaves the black and white room. (Assumption)
3. Mary learns something a new fact when she leaves the black and white room. (Assumption)
4. Mary doesn't learn all the facts in the black and white room. (2,3)
5. There are some facts which aren't physical. (1,4)
6. Physicalism is false. (5)

Objections: The main way of objecting to this argument is by rejecting Premise 3. People try to distinguish different kinds of knowledge: "knowledge that" and "knowledge how." What kind of skill could Mary have learned to account for her new knowledge?

II. A Big Problem with Non-Physicalism: Causal Closure!

Jackson's way out: epiphenomenalism about qualia (not all mental states – beliefs and desires may be different). What's wrong with epiphenomenalism? Three possibilities:

First Consideration: It's obvious that the hurtfulness of pain is partly responsible for the subject seeking to avoid pain.

Jackson's Response: It seems obvious because of the correlation. But we all know that correlation doesn't imply causation. Like in many cases, this correlation might be best explained by a common cause which causes the hurtfulness of pain, and the subject seeking to avoid it. When there are good theoretical reasons to give up on a causal relation that we thought existed because of a correlation – we should do so.

Second Consideration: Evolution – the traits that evolve over time are those that are conducive to survival. Qualia evolved over time – but they can't help us survive if they're causally inert.

Jackson's Response: Not every trait that evolved is advantageous for survival. Some traits are byproducts of other traits that were advantageous.

Third consideration: How could we come to know about other people's qualia if they don't cause any observable behavior?

Jackson's Response: Common cause.

III. The Identity Theory

According to the identity theory the property being in pain is the same as the property being in some brain state B (example: C-fibers firing). Some seeming problems with this theory: couldn't we imagine discovering aliens, with brains that differ completely from our own, but that still feel pain? Or, couldn't

we imagine possible worlds where we are in pain even though we have completely different brains – or no brains at all?

Response: here's an argument which is clearly bad: water isn't identical to the main substance that we drink. Why? We can imagine a world in which water isn't the main substance we drink. So water isn't the main substance we drink.

This argument does seem structurally analogous – but is this a good response? Surely, sometimes learning that one thing could exist without the other *does* show that they're not identical. For example, suppose we came to think that the property of having a heart is the same as the property of having a kidney because every being we found with a heart also has a kidney. Couldn't someone reasonably say – but these can't be the *same* property – because we can imagine one existing without the other? Similarly, suppose a bunch of people lived in a world where whenever they were in pain they cried. The scientists came to think that being in pain is the same as having chemical x emerge from the eyes. Couldn't someone reasonably come along and say – wait – we can imagine there being pain without chemical x coming from the eyes – so pain can't be identical to chemical x coming from the eyes. What's going on here? Why are some arguments of these form good and some arguments of this form bad?

MIT OpenCourseWare
<http://ocw.mit.edu>

24.00 Problems in Philosophy
Fall 2010

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.