Physicalism, the knowledge argument and the colour of ripe red tomatoes[1]

Sigfrid Lundberg

sigfrid@sigfrid-lundberg.se
https://sigfrid-lundberg.se

ABSTRACT

A discussion of the knowledge argument based on papers by Jackson (1986) and Pereboom (1994). It seems that the former treatment of the problem survives attacks from the latter.

Physicalism is a moving target: The physics which is the result of a reduction varies with what physical theory is popular among philosophers at each moment in time; from billiard balls, to physico-chemical processes, molecular genetics and possibly, in the future, quantum gravity.

Introduction

Biology is, writes Smart (1959: 142), *"to physics as radio-engineering is to electromagnetism) except the occurrence of sensations seems to me to be frankly unbelievable."* The effect such a statement have on biologists and biologically inclined philosophers could be easily foreseen. I think it is safe to assume that Smart was wilfully provocative. Smart's (op cit.) and others' works led to a *"wave of reductionist euphoria"* or that is ho w Nagel (1974: 435) puts it in his seminal paper *What is it like to be a bat*. Nagel discusses at length that chiropterology (the scientific study of bats) has to investigate what it is like how to see the world using echolocation.

Nagel was later joined by Jackson who wrote a series of papers, including *What Mary Didn't know* (Jackson 1986). The argument furthered by the two is usually referred to as *the knowledge argument* (Ravenscroft 2005: 171). Formalising Jackson's (1986: 291) argument:

(A) Physicalism is the idea that the world is not only largely physical but **completely** physical. On this point Jackson (1986: 291, see his footnote 1) argues that if physicalism is true, and if you know everything expressible in a physical language, then you know **absolutely** everything.

(B) Suppose that our world W is **not completely** physical. Then there is a possible world W' which is.

(C) Since the world W' is completely physical, the difference between that world and our world W must be facts that cannot be expressed in a physical language (Jackson 1986: 291).

[1] Written as a deliverable for the course FTEA13:5 – Philosophy of Mind, which is a part of the introduction into theoretical philosophy at Lund university https://www.fil.lu.se/kurs/FTEA13/. The source code of this text can be found on https://github.com/siglun/term-paper-spring-2024.

Now, Smart (op cit.) regarded sensations as unbelievable but Nagel (op cit.) saw them as essential (at least for understanding bats: echolocation is a kind of sensation). Jackson's and Nagel's works provoked a response by Pereboom (1994). I first present Jackson's view, then I turn to Pereboom and compare the two. Finally I compare my conclusions with those by Ravenscroft (2005: 172)

What Mary Didn't Know

Mary's black-and-white room is a thought experiment proposed by Jackson. He has discussed it in a number of essays. I will concentrate on Jackson (1986), even though it is not his first treatment of the problem. His brief note was written as a reply to Churchland (1985), who in his turn had attacked those earlier papers. I will not further discuss Churchland.

Mary is an extremely gifted neuroscientist. Since birth she has been confined to a room furnished completely in black-and-white. She has spent her whole life there, studying physiology, cognition, physics, colour, optics etc, using black-and-white books, journals and TV set. If physicalism is true she knows absolutely everything there is to know that can be formulated in a physical language. Which is everything, simpliciter. If there is more to know, then physicalism is false. Or so claims Jackson (1986).

Eventually Mary is released from her confinement, and the questions are: What did she learn when she met the real world? In particular, what did she say when she first saw a ripe red tomato?

Following Jackson (1986), this how Mary's black-and-white room narrative can be boiled down to an argument:

(1) From (A) above: "Mary (before her release) knows everything physical there is to know about other people." (Jackson 1986: 293).

(2) From (C) above: "Mary (before her release) does not know everything there is to know about other people (because she learns something about them on her release)." (Jackson 1986: 293).

Hence

(3) "There are truths about other people (and herself) which escape the physicalist story." (Jackson 1986: 293).

So what did Mary learn? Jackson (1986: 293) claims that she learned something about other people. What she "*did not know until her release is not a physical fact about about their experiences. But it was a fact about them. That is the trouble for physicalism.*" Jackson does not tell us what she said. Only what **she did not say:** *"ho, hum"* (Jackson 1986: 291).

Pereboom's analysis

Pereboom (1994) discusses both Nagel's (1974) and Jackson's (1986) treatments of the knowledge argument (I'll concentrate on his discussion of the latter, though). More specifically, Pereboom (1994) starts with Churchland (1985). Pereboom (1994: 316) is convinced that his "*strategy can be enhanced to resolve the controversy in favor of the materialist.*"

Before attacking Jackson's (1986) argument, he presents how Mary's Room could be formalised. Quoting verbatim from Pereboom (1994: 318):

"

(1)′ Mary (while in the room) knows everything physical there is to know about human beings.

(2)′ Mary (while in the room) does not know everything there is to know about the nature of human sensations and their properties.

Therefore

(3)′ There is information about human sensations that is not physical, and thus human sensations are not completely physical. ”

The differences between Jackson's and Pereboom's first premises could hardly be larger. Jackson is focusing on that Mary knows everything about **other human beings** that can be formulated in a **physical language,** see (1). Pereboom is much more vague, see his premise (1)′. For him, Mary just knows everything physical about **humans in general.** The difference between the consequents are different as well. Jackson's Mary realises, upon her release, that there were important things she did not know about other people's perceptions. See (3). Pereboom's consequent (3)′ just concludes that there is information about human sensations that are not completely physical.

Clearly, Jackson's views have evolved over time. It seems that Jackson (1986), successfully refutes Churchland (1985), or that is what he claimed. Jackson is using a new release of the knowledge argument. Pereboom's (1994) objective was to give support to Churchland, but however I read Pereboom I cannot see that he succeeds. Basically, he tries to defeat the previous version of the knowledge argument. He tries to do so using minor variations of an argument Jackson had already refuted several years earlier.

Pereboom (1994: 316) claims that Jackson implicitly assumes *"that if materialism is true, then complete physical knowledge will provide cognitive access to everything about a subject's mental states as they are apprehended in introspection"* (Pereboom 1994: 316). Pereboom (op cit.) claims it to be false. False or true, that argument seems irrelevant to me; when Mary is released she realises something about others, not about herself (3), a realisation which is impossible to formulate in a physical language and that does not require introspection.

Another core argument is that Mary through, her complete knowledge of everything physical related to human beings, as Pereboom (1994: 320) puts it (cf. (1)′ above), would be able to gain important knowledge about others perception through inter-subjective accessibility. Since the Mary's room narrative is about what happens after she has left the confinement, I cannot see that this is relevant.

Ravenscroft (2005: 172) reached a similar conclusion, albeit he does not say that Jackson is right. He points out that Jackson's (op cit.) describes Mary's lack of knowledge as a lack of understanding of what others knew about the visual cues of ripe red tomatoes. Jackson evaded the objections to his argument. He evaded not only Churchland (1985), but also Pereboom (1994). Basically, Pereboom refer to Jackson (1986) but does not really respond to it. Pereboom's paper is basically about earlier versions of the knowledge argument.

Conclusions

Smart (1963: 651) wrote that he does *"not hold materialism to be wedded to the billiard-ball physics of the nineteenth century."* I think that is good. However, a few years earlier he wrote that *"science is increasingly giving us a viewpoint whereby organisms are able to be seen as physico-chemical mechanisms"* (Smart 1959: 142). Depending on what you actually understand by physico-chemical mechanisms, this might be seen as an improvement from billiard-ball physics. From my point of view it seems only marginally so. He then continues: *"it seems that even the behavior of man himself will one day be explicable in mechanistic terms."* Who,

if anyone, doubts that? To me the question is not whether behaviours should be described mechanistically. The question is what mechanisms should be used when doing so. Brigandt & Love (2023) mention that in late 20th and early 21st century philosophy of biology, one important problem has been if "*classical genetics can be reduced to molecular genetics.*" Now we have billiard balls, physico-chemical mechanisms and molecular genetics. Three data points: With some inductive expansion we can conclude that we are on the right track. If the trend continues we might be relieved from seeing attempts to explain the theory of natural selection using quantum gravity.

Brigandt & Love (2023) also provide nice explanations of some of the theories of theory reduction, and from my point view they (and references in their article) save philosophy. A theory reduction occurs when we have two theoriesT_A andT_B. The theoryT_A is said said to reduce theoryT_B if the former entails the latter. This implies that the reducing theory should be capable of explaining the problems addressed in the reduced theory.

Think of physics asT_A and what is in Mary's head when she holds a ripe tomato in her hand as theoryT_B. She understands that all other people have eaten this wonderful fruit each early summer, while she ate black hamburgers with white bread and black-and-white ketchup. That experience cannot be experience cannot be described in a physical language according to Jackson (1986).

References

Brigandt, Ingo, & Love, Alan (2023). Reductionism in Biology. In: Edward N. Zalta, & Uri Nodelman (ed.) *The Stanford Encyclopedia of Philosophy*, Metaphysics Research Lab, Stanford University https://plato.stanford.edu/archives/sum2023/entries/reduction-biology/

Churchland, Paul M.Reduction, Qualia, and the Direct Introspection of Brain States. *The Journal of Philosophy*, 82(1), 8-28.

Jackson, Frank (1986). What Mary Didn't know. *The Journal of Philosophy*, 83(5), 291–295. https://doi.org/10.2307/2026143

Nagel, Thomas (1974). What Is It Like to Be a Bat? *The Philosophical Review*, 83(4), 435–450. https://doi.org/10.2307/2183914

Pereboom, Derk (1994). Bats, scientists, and the limitations of introspection. *Philosophy and phenomenological research*, 54(2), 315–329. https://doi.org/10.2307/2108491

Ravenscroft, Ian (2005). *Philosophy of mind: A beginner's Guide*, Oxford University Press

Smart, John J. C. (1959). Sensations and Brain Processes. *The Philosophical Review*, 68(2), 141–156. https://doi.org/10.2307/2182164

Smart, John J. C. (1963). Materialism. *The Journal of Philosophy*, 60(22), 651–662. https://doi.org/10.2307/2023512