

# Why Mary Doesn't Know All The Physical Facts

Daniel Sippel

---

## Abstract

The famous Mary's-room thought experiment by Frank Jackson features Mary, who knows all the physical facts but does not know what it is like to see red. I will argue, however, that it is impossible to know all the physical facts without knowing about qualia if we grant that qualia exist and are causally efficacious with respect to physical facts. As a result, physicalism cannot be refuted by Jackson's famous Mary's-room thought experiment, because it is flawed. To show this, I will firstly contextualise the question at stake and explicate the various notions and theories that are relevant to assess the question. Secondly, I will present Jackson's thought experiment which features Mary, the brilliant scientist. After that, I will reject Jackson's claim that qualia are causally inefficacious with respect to physical facts by using an example. With Goldman's Theory of Knowing, I will demonstrate that since Mary is ignorant of causally efficacious qualia, she cannot construe an appropriate causal chain linking her belief of some physical facts to the fact itself. Therefore, she does not not some physical facts.

---

## Introduction

In this essay, I argue that it is not true that someone could know all the physical facts without knowing what it is like to see red. As a result, physicalism cannot

*Daniel Sippel*

be refuted with this thought experiment. To show this I will firstly contextualise the question at stake and explicate the various notions and theories that are relevant to assess the question. Secondly, I will present Jackson's thought experiment which features Mary, the brilliant scientist. After that, I will reject Jackson's view that qualia are causally inefficacious with respect to physical facts by using an example. Then, I will argue that his thought experiment is paradoxical if we do not allow for his causally inefficacious qualia, but assume the existence of qualia that have causal power in respect to physical facts. I conclude that knowing only about all physical facts without knowing about qualia is impossible. Throughout the essay, whenever the formulation 'all physical facts' is used, I mean 'everything in completed physics, chemistry, and neurophysiology, and all there is to know about the causal and relational facts consequent upon all this [...]'<sup>1</sup>. I deliberately follow Jackson's definition here because of my dealing with his argument in which he uses this very definition.

## **I. Contextualisation**

### **a. Physicalism, Qualia and Epiphenomenalism**

Jackson prominently presented the thought experiment, which features a person who knows all the physical facts without knowing what it is like to see red. In his seminal paper 'Epiphenomenal Qualia',<sup>2</sup> he develops his famous Knowledge Argument. Jackson aims to refute Physicalism with his argument. Physicalism is the view that 'all (correct) information is physical information'<sup>3</sup> or, in other words, that everything is 'necessitated by the physical'.<sup>4</sup> It follows, according to this view, that if you imagine a possible world *w* which is an exact physical duplicate of the actual world *y*, then *w* and *y* are exactly the same in

---

<sup>1</sup> Jackson, 'What Mary Didn't Know'.

<sup>2</sup> Jackson 'Epiphenomenal Qualia'.

<sup>3</sup> Jackson 'Epiphenomenal Qualia' 127.

<sup>4</sup> Stoljar, 'Physicalism' 1.

## *Why Mary Doesn't Know All The Physical Facts*

every aspect: Biological, social, economic and psychological aspects of both worlds are the same, because the physical properties of both worlds are the same. Jackson is opposed to this view and argues that there is something additional to the mere physical things in the world. With his argument, he intends to show that there are non-physical entities, which contradicts physicalism. These entities are called qualia. Qualia are subjectively-felt or 'introspective accessible'<sup>5</sup> aspects of our bodily and perceptual experience, which Jackson does not include in the set of physical entities. When you walk on sand or when you burn yourself for the first time, more than just a release of ions into the synaptic cleft and the cytoplasm happens; more than the merely biological, 'physical' facts are triggered by the actions. *We experience, we feel* the sand under our feet, for example. Jackson therefore asserts that the physical events have a property, something that Jackson calls a 'quale'. Consequently, both physical events and qualia — as a consequence and property of physical events — happen when we walk on sand, for instance. Consider the following example: When you accidentally burn your hand because you touched a hotplate, c-fibres might fire and neurotransmitter are released in your brain. At the same time, as a consequence of the physical processes, you experience the pain. This experience is an entirely subjectively-felt 'quale'. This property of a physical event is caused by the physical events happening in one's body. However, the quale according to Jackson has no causal power of its own. It is merely a side-effect of the physical events. This position is called epiphenomenalism: It is the view that qualia are epiphenomena, i.e. caused by physical events, but without causal powers.

### **b. Jackson's "Mary's room" thought experiment**

In order to make his claim more substantive, Jackson created the thought experiment that the question at stake refers to. His experiment features Mary, 'a

---

<sup>5</sup> Tye, 'Qualia' 1.

*Daniel Sippel*

brilliant scientist',<sup>6</sup> who is locked in a black-and-white room. She has never been outside, and her education relies completely on black-and-white media, such as books without colours and lectures in a black-and-white television.<sup>7</sup> Thus, she gets to know all the 'physical information there is to obtain'.<sup>8</sup> As a consequence, in order to go back to the initial examples I gave above, she knows in every detail what goes on in her body when she would burn herself, or when she would walk on sand. All the neurological processes are known to her. *Nota bene*, Jackson points out, if physicalism was true, she would know everything what she could possibly know. Mary has all the physical information, and — according to physicalism — physical processes 'underlie all events and processes'.<sup>9</sup> So if physicalism was true, Mary would obtain 'full coverage' of the world,<sup>10</sup> because she knows all the physical facts. However, one day, she leaves the black-and-white room and sees the colour red for the first time. Naturally, she gets to know some new information, namely the information of how red looks like. Before, she did not know how red looks like, so in fact there was information which she did not have, although she had all the physical information. Hence, Jackson points out, physicalism is false: There is more to the world than mere physical facts, and this 'more' are qualia. Therefore, according to Jackson, it is possible to have all the physical facts without knowing what it is like to see red. He concludes that this refutes physicalism, because the thought experiment proves that qualia must exist, and these qualia are non-physical.

---

<sup>6</sup> Jackson 'Epiphenomenal Qualia' 130.

<sup>7</sup> cf. Jackson, 'What Mary Didn't Know' 291.

<sup>8</sup> Jackson 'Epiphenomenal Qualia' 130.

<sup>9</sup> Kim, 'Consciousness and the Mind-Body Problem' 301.

<sup>10</sup> This term is due to W.V. Quine and is quoted in Kim, 'Consciousness and the Mind-Body Problem'.

## **II. Refuting Jackson**

I will now uncover an assumed, though not compelling reason why Jackson asserts that qualia exist and are causally inefficacious with respect to the physical world. On the basis of the weak explanatory power of his claim, I will deconstruct his thought experiment in a second step. I will argue that if we grant that qualia actually exist, his view that qualia are causally inefficacious with respect to physical facts is plainly wrong. Jackson claims that non-physical qualia exist, and that they are 'causally impotent' with regards to physical facts.<sup>11</sup> That is, there is no causal chain leading from a quale back to any physical event in the world, including any physical event in your body. Note that he holds as well that qualia may have a causal effect on other qualia. I am not going to attack this latter view. Rather, I argue that the first assumption — that qualia are causally inefficacious with respect to physical facts — is not cogent. Indeed, in 'Epiphenomenal Qualia', Jackson merely asserts that 'you have to hold'<sup>12</sup> that qualia make no difference to anything physical. Instead of presenting a strong argument for his claim, it seems that Jackson tries to combine 'the world view of science' with the existence of qualia, without interfering too much with the scientific view that only accepts physical facts as facts with explanatory power. He refutes physicalism by declaring a 'low probability' of its truth: The view that exclusively physical facts, which humans 'in principle' can know about, are 'relevant in some way [...] to the survival of homo sapiens' is 'very low'.<sup>13</sup> In order to reconcile the scientific view and the purported low probability of its correctness, he allows for causal explanatory power of physical facts and, at the same time, holds that there are qualia which do not have causal power with respect to physical facts. However, I will argue that this argumentative balancing act fails to be compelling.

---

<sup>11</sup>Jackson 'Epiphenomenal Qualia' 133.

<sup>12</sup>Jackson 'Epiphenomenal Qualia' 133.

<sup>13</sup>All quotes from Jackson 'Epiphenomenal Qualia' 135.

**a. The Placebo-Pill Example**

With the following example, I aim to demonstrate that there can be situations in which we observe a causal chain leading from a quale to physical events in human bodies. Imagine Paul takes part in a study to test the effects on humans of a new medicine against migraine. Paul suffered from migraine his whole life and expects great results from the newly invented pill whose effectiveness and mode of action has been outlined to him in a presentation prior to the study. The scientists point out that the pill prevents the patient of having migraine for at least a month. Excited and nervous, Paul then in the laboratory takes the pill. He does not know that he is part of the control group of the study and that the scientists actually gave him a placebo pill. With all he learned about the effects of the actual medicine, Paul now believes that the pill will free him from migraine for a month and takes the pill. As a consequence, he actually does not experience any migraine-symptoms for a month, although he only took a placebo pill. This effect of placebo pills is scientifically proven.<sup>14</sup> What happens in this example, in philosophical terms, is that Paul first learns with the presentation what effect his pill-taking would have. He gains knowledge about physical facts that happen to the human body after one takes the pill and believes that the pill will help him. Then he takes the pill. This is a classical quale, he learns how it feels like to take this pill — he experiences the pill on his tongue, when he swallows it etc.. Jackson would, accordingly, not include this gained knowledge in the set of knowledge about physical things. Finally, this experienced act of pill-taking causes a physical change in Paul's body, and actually releases him from any migraine-related pain. In conclusion it can be said that this example shows that Jackson's position is flawed: It is indeed possible that qualia are causally efficacious with respect to physical facts. Therefore, Jackson's position that qualia are causally inefficacious must be rejected.

---

<sup>14</sup>See Porro 'Open your mind to placebo conditioning' 2-3.

**b. Why Mary doesn't know all the physical facts**

Now that I established that qualia can be causally efficacious with respect to physical facts, where does it lead us? First and foremost, obviously, a part of the epiphenomenal position of Jackson's theory can be rejected. Qualia can be causally efficacious, not merely in relation to other mental states, as Jackson grants, but also with respect to the physical world. Building on that, I will establish that his whole thought-experiment collapses due to the causal efficaciousness of qualia. Jackson himself admits that if we grant causal role to qualia with respect to the physical world, it is hard not to sound 'like someone who believes in fairies'.<sup>15</sup> However, I claim that granting causal powers to qualia, which seems to be necessary considering the placebo-pill example, results in something worse than a fairy-like view. It will become evident that the thought experiment is, in fact, a paradox — in essence, it is not possible to know all the physical facts without knowing at least about some qualia, which Mary does not.

In order to prove this claim, I will use Alvin Goldman's 'Theory of Knowing'.<sup>16</sup> This is necessary in order to demonstrate clearly what it is that Mary did not know and why her missing knowledge about qualia causes her to be ignorant of certain physical facts. Essentially, given that we grant the possibility of causally effective qualia, the causal conflation of both scientific facts and qualia leaves Mary being ignorant of some scientific facts. Both are conflated to such a degree that the notion of pure knowledge about all physical facts, which does not include qualia, is impossible: it is full of holes, just like a Swiss cheese. This assumption however makes it necessary to clarify what it means to know something: how far is Mary's knowledge limited and 'full of holes'? According to Goldman's Theory of Knowing,<sup>17</sup>

---

<sup>15</sup>Jackson 'Epiphenomenal Qualia' 128.

<sup>16</sup>Goldman 'A Causal Theory of Knowing'.

<sup>17</sup>Goldman 'A Causal Theory of Knowing' 369.

*Daniel Sippel*

*‘S knows that p if and only if the fact p is causally connected in an ‘appropriate’ way with S’s believing p.’*

This ‘appropriate way’ is essentially, without going too much into detail, a ‘continuous causal chain’<sup>18</sup> that links fact p to the belief p by knowledge-producing causal processes like perception or memory.

It will be shown that Mary cannot know about all the physical facts without knowing about qualia. In the analysis, we use Goldman’s Theory of Knowing. To simplify the analysis, only the steps in the example relevant for Mary’s neuroscience are pointed out. (p) is the physical fact that the pill is inserted into the organism. This causes (e): The experience, or quale, of taking the pill, which then leads to the reshaping of Paul’s condition — the relief of his pain (r). Notice that it is necessary to incorporate (e) into the causal chain. The quale (e) ultimately caused Paul’s relief of pain, not the mere physical fact that the pill was inserted into the organism (p). Indeed, since the pill is a placebo, an insertion of the pill without Paul feeling and experiencing the pill on his tongue etc. would be without consequences. Consider the following hypothetical scenario to demonstrate the necessity of Paul’s quale for his recovery. Imagine the case where the doctors decide to narcotize Paul. Consequentially, Paul would be without conscience. Then the doctors insert the pill into his organism, for example with the help of a tube. The pill of course would not have any effect, because Paul would not have felt or experienced the taking of the pill. Without the quale, his physical condition would not change. Thus, the causal chain of the example can be formalized like this:

$$(p) \rightarrow (e) \rightarrow (r) \tag{1}$$

In contrast, the causal chain that Mary — being a brilliant scientist who knows physical facts exclusively — is able to develop however is not ‘appro-

---

<sup>18</sup>Goldman ‘A Causal Theory of Knowing’ 361.

appropriate' in order to claim that she knows the physical fact (r). The physical fact that the pill was inserted into Paul's organism (p) causes Mary to infer (dotted arrow) that Paul is, as a consequence, relieved of his pain. How else could she explain his recovery, given that she is ignorant of qualia? This belief is represented below as  $B_M(r)$ . Goldman's Theory of Knowing requires that there is a causal connection between (r) and Mary's belief of (r). However, since the fact (r) is not connected to her belief of (r), she does not know (r). Rather, Mary knows that (p) by definition, and that causes her to infer that (p) is causally connected to the observable physical changes in Paul's body. She is ignorant of the intermediate step, the non-physical quale:

$$(p) \rightarrow (B_M(r)) \quad (2)$$

The 'Mary's room' thought experiment therefore is a paradox — it leads to a conclusion that is self-contradictory. On the one hand, *by definition*, Mary knows all the physical facts that can be known. On the other hand, as I have established above, using Goldman's Causal Theory of Knowing, Mary cannot know a physical fact that is caused by a quale. In other words, there is no knowledge available to Mary which is inferred via a non-physical intermediary cause, a quale, in the causal chain. This conclusion of course requires the rejection of Jackson's claim that qualia are causally inefficacious. Therefore, the Mary's-Room thought experiment is paradoxical.

## **Conclusion**

To conclude, I refuted Jackson's claim that physicalism can be rebutted by the Mary's-Room thought experiment in two steps. Firstly, contrary to Jackson's view, I have shown that qualia can be causally efficacious with respect to physical facts with the Placebo-Example. In a second step, I have demonstrated that if we grant a causal role to qualia, it is logically impossible to know all the physical facts. Since Goldman's Theory of Knowing suggests that there

*Daniel Sippel*

must be a causal chain connecting the actual fact with the belief of the fact, I demonstrated that without knowing about qualia, pure knowledge of all scientific facts is impossible to obtain. Throughout the essay, I have accepted the existence of qualia, and thereby rejected physicalism, as a way of showing that Jackson's thought experiment is flawed. Yet his flawed thought experiment that is supposed to refute physicalism cannot do so effectively.

## References

- [1] Goldman, A.I., 'A Causal Theory of Knowing', *The Journal of Philosophy* 1967. Vol. 64, No. 12: 357-372.
- [2] Jackson, F., 'Epiphenomenal Qualia', *The Philosophical Quarterly* 1982. No. 32, No. 127: 127-136.
- [3] Jackson, F., 'What Mary Didn't Know', *The Journal of Philosophy* 1986. Vol . 83, No. 5.
- [4] Kim, J., 'Consciousness and the Mind-Body Problem', *Philosophy of Mind* 2010, (3rd Edition). Boulder: Westview Press.
- [5] Porro, C.A., 'Open your mind to placebo conditioning', *Pain* 2009. Vol. 145, Issues 1-2.
- [6] Stoljar, D., 'Physicalism', in *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), 2009, <http://plato.stanford.edu/archives/fall2009/entries/physicalism/> (accessed on 20.2.2014).
- [7] Tye, M., 'Qualia', in *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), 2013, <http://plato.stanford.edu/archives/fall2013/entries/qualia/> (accessed on 20.2.2014).

**Daniel Sippel** is a BSc Politics and Philosophy student at the London School of Economics and Political Science (2012-2016). In 2015, he spends a year abroad at Sciences Po in Paris. His areas of interest are Epistemology, Political Theory and the intercept point of both. He hopes that he might one day be able to contribute to solve the issues with deliberative democracy. You can contact him at [d.sippel@lse.ac.uk].

