AN INTRODUCTION TO DUALISM: THE MENTAL TO THE FORE?

Abstract: the contemporary debate around consciousness presents us (and often leads us to embrace) a specific current of thought: physicalism, which states that everything in our reality is physical\(^1\). In this paper I want to introduce the main points of the opposite view, dualism, according to which there are two different realms of reality: the mental and the physical one. In the introduction I give the main idea and sketch the general intuition behind dualism. In Section 1 I present substance dualism and its more refined version, which is interactionism. In sub-section 1.1 I describe property dualism, which is the mainstream form of dualism adopted nowadays. In sub-sections 1.1.2 and 1.1.3 are proposed two variations of property dualism: epiphenomenalism and panpsychism. Finally, in Section 2 and its two sub-sections (2.1 and 2.2.) I recall Frank Jackson’s famous 1982 thought experiment and two classical objections to it, which are the Ability Hypothesis and the phenomenal concept strategy one. I conclude with a short reflection if we were to take under consideration the acceptance of dualism as an explanation of consciousness and its nature.

Introduction

Dualism is roughly a metaphysical thesis which claims that reality can be divided into two realms. First of all, there’s the physical, concrete reality described by physics and the empirical sciences like chemistry, neurophysiology, and others. In this category can be put stuff like chairs, mountains, atoms, brains. On the other hand, there’s a completely different aspect with as many different properties, that is to say the mental. In this case we can refer to the mental component using mental states, such as being happy, believing that the moon is red or thinking that 2+2=4. But there’s a further part which is philosophy of mind main concern: consciousness. Today we don’t have a full description of consciousness and there’s no general consensus on

---

1. Physicalism, according to a survey made by David Chalmers in 2014, is the philosophers of mind favourite option to describe consciousness in the contemporary debate. The percentage of philosophers (on a scale of 931) who embraces physicalism is 64, more than half of the candidates.
whether or not we should just accept a materialist one, even though it is currently incomplete. But on the other hand, how do we explain consciousness if it can’t be done using a physical-dependent terminology?

In trying to overcome these difficulties, dualists sustain that consciousness (in its most general form) has to be considered not as a physical part of reality, but as if it belonged to the mental realm. And why is that? A possible response could be because of the difference between the inner world (the first person point of view, composed of mental states, feelings, sensations and consciousness) and the reality which is external to it and which insists on an objective and third person description of everything that surrounds us. Just consider analyzing parts of the brain via neurocognitive instruments such as PET, EEG, or describing a natural phenomenon by observation, or postulating a fundamental theory of reality in combination with quantum mechanics).

There are many ways to prepare a dualist recipe for consciousness and in Section 1 (which is split into four parts: Section 1 and its sub-sections 1.1, 1.1.2 and 1.1.3) I consider the three most influential and spread variants of dualism nowadays, starting from the classical and well-known position of substance dualism. Unfortunately, it is not as popular as the other two, as we will have occasion to see in the next part of the paper. Still, there’s a minority of philosophers who think that the intuition behind substance dualism is overall correct and who affirm to be substance dualists to all effects (a brilliant example of a substance dualist is Professor Richard Swinburne).

1. Substance dualism and interactionism
Dualism is usually said to be born with Descartes and its Meditations, even though in philosophy there have been various attempts to put down a dualist-like explanation of the mental. Just take a look at Plato’s formulation of the Theory of Forms and Thomas Aquinas works in the XIII century. Whereas the first one thought that our world couldn’t just amount to the physical reality known by the five senses, the second one brought dualism to develop into a minor position which is known in the contemporary debate as “hylomorphism”, which roughly states that we are a conjunction of body and soul, which is a sort of function of the physical body.

Let’s turn back to the XVII century. What Descartes and its famous thought experiment (a more recent variant of which can be found in Kripke 1980) landed to in the Meditations is a form of “substance dualism”. Its main claim is that reality has to be distinguished in two separate parts, the mental and the physical. Not only that, but these two substances (here comes the name of the position) are totally different from each other: physical bodies, e.g., have physical properties such as weight, temperature, pressure, are composed of atoms and are governed by laws of nature. Mental aspects (thinking, feeling and being conscious) on the other side have to be ascribed as something more. Minds have no extension, no physical measure and are accessible only by the first-person point of view, by a subject of experience. No empirical analysis will ever reveal that you are conscious, neither by looking closely at each of your neurons nor by using logical inferences (e.g. I am conscious, and I have the same physical structure as you. So you must be conscious too).

What about the interaction between these two “worlds” or substances? Isn’t that a problematic aspect? The point was already raised by Princess Elizabeth of Bohemia (1615-1680), when in one of her letters to Descartes she asked him how a non-physical mind could interact with a physical body. How can it be that my desire of having a glass of water (which is a non-physical mental state, for a dualist) can help me interact with
the complex neurophysiological structure (my brain) that sends electric impulse down to my nerves, which then permit my hand to grab the glass? The objection was so powerful that Descartes, after turning this around a lot and being very vague, had to speculate that a particular organ of the epithalamus, the pineal gland, was the thing which allowed the mind to interact with the brain. To him, that was the location of the soul and the very organ which allowed the interaction between the two substances.

However, a more complete and elaborated response can be found in the evolution of substance dualism, which is a position known as “interactionism” or “interactionist dualism”. The main source for this position can be found in a work by Sir John Eccles and Carl Popper in 1977, even though in recent years it has received more attention. The main claim of interactionism is that the mind is a non-physical substance which can interact with the physical body, a fact that alone necessarily entails the falsity of the principle of causal closure (hereafter “PCC”). PCC states that the world is a physical system which is causally closed: every physical event has got to be described with a correspondent physical cause. Interactionism main objective is to break the wall of PCC and claim that science shouldn’t beg the question against it, accepting that it could turn out to be false. Even though it is a pretty risky claim and a step which many philosophers aren’t ready to make, Eccles and Popper find its denial not that much of a deal.

To summarize the proposal made by the two philosophers, I’ll skip the part concerning the metaphysics of reality conceptualized in their model (I am referring to the Three Worlds Theory), to focus instead on the interaction between mind and body, which is also the most important part of their proposal. According to Eccles and Popper, the notorious interaction which caused so many problems to substance dualism can be found occurring in a specific part of the brain: the liaison brain. It comprehends the major part of the left hemisphere (which in humans is well known for being the dominant one between the two), a part of the prefrontal lob and two important areas addicted to the production of language, that is to say the Broca and

the Wernicke’s areas. The liaison brain (abbreviated “LB”) interacts with a specific component, the self-conscious mind (SCM), which is roughly what Eccles defines the “I”, our central pole of identity, ourselves. When SCM and LB interact with each other, there’s an exchange of information. The mind isn’t in contact with the neurons of the brain, but with the so-called “modules”, which are small packs of about 10,000 neurons each and resulting from an activity of intertwining of inputs. The modules can be respectively “open” or “closed”, depending on the activity of the interaction between the two components just mentioned.

In sum, the mind can be considered as the center of consciousness and personal identity and her influence on the brain would be granted by spatiotemporal fields of influence, which are activated by the cerebral cortex (i.e. LB). This process would happen at a microscopic level and it has to be understood not as a sum of single actions, but as spatiotemporal modification of the neural activity. As we can see, denying PCC doesn’t seem like a crazy task, if we can give a coherent account of the interaction between the soul (Eccles believes in its existence and describes the Self essentially as his/her soul) and the physical body. Of course, Eccles and Popper’s ideas aren’t fully accepted in the contemporary debate, especially if we take a close look at dualism estimation nowadays, but their description is an important attempt not to allow our prejudice towards a non-physical mind to overcome us.

1.1 Property dualism

Substance dualism and its interesting interactionist variation haven’t been the only candidates for an explanation of the mind using non-physical items of knowledge. As a matter of fact, the closest position to this approach, but with some peculiar refinements, can be found in property dualism. If you accept materialism but aren’t ready to accept that reality can be fully grasped by physics and science, then property
dualism is the position well-suited for you. In the current debate it is actually the most influential and the most approved way to think of dualism. A few philosophers under this guise can be David Chalmers, Martine Nida-Rumelin and Philipp Goff, even though their positions differs in some respects of how they think of consciousness.

What is property dualism main point? The idea can be expressed like this: physicalism at its finest form, but with the exception of mental states. According to a property dualist, everything in our concrete reality is physical and can be successfully interpreted appealing to chemistry, physics and all the other empirical sciences. The main difference with respect to a materialist doctrine is that there are some extra-facts to account for, which are the notorious non-physical mental states (as we have already seen in the introduction, the ones such as believing that God exists, thinking about a pineapple or feeling pain in your left toe). Mental states are to be understood as non-physical properties of the brain, which can have or not physical effects on the external reality, depending on how you approach the matter.

Watch out: we are not talking about any kind of cartesian substance in this case, but of properties, that is to say qualities of physical substances like brains and bodies. Properties are attributes of the physical substance; they are not a different part of reality. Otherwise, we would slip again in the problems concerning interaction between a non-physical mind and physical reality. In this sense, the claim made by property dualists seems to be softer and less radical than a substance dualist one): the mind can be considered as a property of the brain (in some stronger version a function, if we embrace the functionalist proposal) or as a sophisticated physical system which can perform a lot of different actions. But to a property dualist that is not sufficient to account for qualia and mental states, which are a product of neural activity but with no correspondence to a specific brain part or function. In a way, it can be considered a stepbrother of substance dualism, but with a strong distinction regarding properties instead of substances.

5. Functionalism states that the mind is a mediator between the external inputs (e.g. the stimuli derived when observing a landscape) and the outputs (such as a behavioral response). A mental state is defined by its function and its causal role.

6. Qualia are subjective, intrinsic and private properties of experience. An example can be the smell of coffee, what is like to see a determinate color or the particular feeling of pain in your left hand. It seems that experience is necessarily followed by qualia for a subject of that same experience.
What are the most discussed and potential approaches that were given in the literature? There are currently many variations of property dualism and in the following sub-sections I would like to focus on a couple of them: epiphenomenalism (defended for example in Gertler 2019) and panpsychism, which finds it best defender and one of its current and most valid sustainers in Philipp Goff (2017 and 2019).

1.1.2 Epiphenomenalism

One of the main varieties of property dualism, which proved to be successful in the early 20’s, is epiphenomenalism. This position is very strange, as it maintains a form dualism which denies the most problematic aspect for it, which is mental causation. We have already seen that interactionism, at least the one expressed by Eccles and Popper, tried to solve the issues around PCC using specific set of neurons and areas of the brain. Epiphenomenalism instead claims that there’s no such problem: qualia and mental states, which are non-physical according to dualism, do not have any sort of causal effect on the physical. If we had to consider a concrete example, if I put my hand on the fire, the feeling of pain doesn’t cause my hand to retreat. The action is caused by neurophysiological states and can be fully described using a scientific-like jargon. There’s no mention of mental states and non-physical qualia whatsoever. This is exactly the crucial point: qualia and mental states are epiphenomena of the brain and they can’t intervene on the physical. The physical causes the non-physical, but we can’t say the same for the contrary. At first, this seems like an implausible and extremely counterintuitive opinion. Why should anyone adopt it?

A first reason to accept epiphenomenalism which comes to mind almost instantaneously is that it allows to keep the dualist intuition, while denying any problematic aspect involving causation. If qualia are non-physical and mental causation can be accounted in completely physical terms, property dualism is safe and
we can have our cake and eat it too. After all, it is reasonable to say that consciousness and mental states do not cause anything: our current physics doesn’t show any contradiction with non-physical properties being present in our universe and not doing anything. What’s more, consciousness can’t be empirically described using traditional methods, as we only have access to it by introspection (recall the introduction). The epiphenomenalist could insist on the fact that this radical difference between physical facts and mental states can be put in terms of causation: the physical causes the non-physical (causality can be considered as an intrinsic property of matter), whereas the non-physical can’t cause anything (because of the ontological gap between the physical and mental states in regard of their properties, qualia’s intrinsic properties can’t be the same as of a standard physical object or component like a rock).

Gertler (2019) also brings up another point in support of epiphenomenalism, with a special focus on the PCC. Here’s the passage, which consists of a defense against a powerful objection by materialist: are we really ready to accept that qualia are non-physical? Are we sure that my feeling pain on my hand doesn’t cause my hand to retreat from the fire? Gertler responds along these lines:

Papineau suggests that accepting would violate epiphenomenalism would violate simplicity considerations. (...) This argument relies on an inference, form the fact that science doesn’t recognize non-mental epiphenomena, to the conclusion that there aren’t any. But that inference is unjustified. After all, standard scientific methods detect entities by their effect and Occam’s razor surely does not prescribe to deny that that specific phenomenon occurs. (...) To use Sober’s terms, what applies in this case is not the “razor of denial” but instead the razor of science. (page 78)
There’s then a second line of objection: epiphenomenalism would be, according to it, against natural selection. The point was already raised by William James, as we can see in the following passage:

If pleasure and pains have no efficacy, one doesn’t see... why the most noxious act, such as burning, might not give trills of delight, and the most necessary ones, such as breathing, cause agony (1890).

The problem here is that if phenomenal proprieties (i.e. qualia) aren’t causally efficacious, then they will not affect the organism which undergoes that experience, which then results in a being who has way less chances of survival. Jackson, in his (1982), already tried to give an answer to this problem, claiming that qualia may be a by-product of something that it is selected for, like the heaviness of a polar bear’s coat, which doesn’t seem to be extremely helpful when poor boy has got to move to another point of his habitat.

How does Gertler save epiphenomenalism? Her idea is that the process involving qualia doesn’t consist in causation, but on association instead. Gertler brings an example of someone who may experience a new taste (let’s assume the person is enjoying a yogurt): when that person claims that it is unpleasant, his/her quotation isn’t given by the taste itself, but via the association made when elaborating the taste. It is not its phenomenology the responsible for him/her saying “This yogurt tastes really bad”. What that does is contributing to non-causally create an association between the stimulus and the response. The creature makes an association, which can be either negative or positive, between the stimulus and the hedonic value that she grants. Since our candidate registers the association directly and being acquainted with the relevant phenomenal properties, there’s no causation and qualia seem to be causally inefficacious.
1.1.3 Panpsychism

The second and final version of property dualism that I will consider in this section is panpsychism. The 
understand the meaning of the name, we have to divide in two parts: “psyche”, which states that we are 
talking about the mind, and “pan”, which in Greek means “everything”. Panpsychism is the thesis according 
to which the mind is everywhere in the universe. It is important to stress out that between 
epiphenomenalism and others positions I haven’t analyzed (such as Chalmers’ important naturalistic dualism) 
that panpsychism is compatible with physicalism, and whether or not is it a form dualism depends on the role 
of qualia. The general assumption made by the panpsychist is that consciousness can be found anywhere: in 
chairs, mountains, seas, socks, atoms, and pasta. Of course, many point out that it would be counterintuitive 
to say that a chair possesses consciousness and that it has got the relevant qualia, when confronted with the 
quale of a red experience, for example. That’s not what the panpsychist is saying, as the point is a bit more 
subtle: consciousness is a fundamental component of our universe and it is as essential as space, time, and 
spin charge. It is exercised by every component of matter as our laws of nature do on it.

In a way, we could say that the mental substrate represented by consciousness is at the base of matter and 
that it can be considered as an intrinsic property which is common to everything in the universe. Some 
panpsychists prefer to talk of micro-phenomenal properties, which together constitute our fully and unite 
experience of everyday life. Taken alone, these properties don’t constitute anything: they must be combined 
all at the same time in order to have what we call experience, accompanied by the relevant qualia.

This is one of the most important points of a variation which is known as “Russellian Monism”. The name 
comes from one of Bertrand Russell’s most famous books, *The analysis of Matter* (1927), in which he exposes 
is thesis. The core of this variation of panpsychism is that physics can only explain dispositional properties,
such as the way an electron behaves and its functions (therefore their *disposition* to act in a certain way), whereas it is completely silent on the intrinsic ones (like phenomenal properties). A stronger version of it even asserts (as Chalmers 1996 states in his *Conscious Mind*) that phenomenal properties are formed by even smaller properties which are called “proto-phenomenal properties”, presenting us with a position called panprotopsychism.

A powerful objection that is usually raised against panpsychism is the following: how can phenomenal properties combine together to give rise to a conscious experience of, for example, a green apple? And, most importantly, what are they made of? This is the so-called “combination problem”, by far one of the most important and impressive challenges for panpsychism. If physics can’t get a grip on them, we should expect a philosopher to analyze them and describe what their role is. This is a very strong objection which is brought by Nagasawa (2020, forthcoming) specifically against the panprotopsychism Chalmers is sympathetic to.

2. The Knowledge Argument

I have already analyzed Frank Jackson’s amazing knowledge argument in another work\(^7\). For reasons of space, I will briefly present the argument and a few objections to it, even though I will omit many. There are many other arguments to sustain dualism, but I find Jackson’s the most intuitive and powerful one. In his “Epiphenomenal Qualia” (1982), Jackson presents us probably the most famous neuroscientist in the history of thought experiments. Mary is a brilliant girl who has been kidnapped and kept her entire life in a black and white room. From there, she starts studying on black and white books and using a black and white monitor. Her research focuses on the human color vision: she comes to know every single fact about it, from all the variations of the spectrum of colors to how the light reaches our optic nerve.

---

7. See Pelucchi (2020) *The Knowledge Argument: acquaintance and experience towards a world of colors.* It hasn’t been published yet, but you can find on my page on Reserachgate.
There’s one very crucial feature of the experiment: Mary has never seen the color red. When she is finally released from her room, she experiences a ripe tomato (I prefer a rose, but Jackson originally used that image, so we’ll stick with it) and comes to know something new about the world (i.e. what is like to see red). If so, physicalism — the doctrine which asserts that everything out here is physical and only physical—must be false, otherwise Mary would have known from her room what is like to see a ripe tomato. Then, if that’s the case and the argument is sound, physicalism, is false. Qualia are non-physical items of reality.

Let’s put the argument in a more refined structure:

**Knowledge argument**

P1. Mary knows all the physical facts about the human color vision.

P2. Mary doesn’t know what is like to see red before her release.

P3. Mary comes to know something new after her release (i.e. what’s is like to see red).

Conclusion: physicalism is false.

The argument is valid, that is to say that the premises force the conclusion and are logically linked to each other. But is it correct? Are all its premises true? Physicalism of course must deny the conclusion of the knowledge argument: to do so, there have been many attempts in the past forty years.

In the next tree sub-sections I will focus on two different replies: the Ability Hypothesis, which find its defenders in Lewis and Nemirow, and the phenomenal concept strategy, defended for example by Michael Tye and John Perry, even though with slight variations.
2.1 The Ability Hypothesis

David Lewis and Laurence Nemirow are advocates of one of the very first objections to the knowledge argument. The main idea in their writings can be summarized as follows: Mary doesn’t come to know any new fact, as inside her room she knew everything there was to know on vision and colors. Her new knowledge once she is released consists of a set of abilities: after her release, she learns how to recognize, for example, red instead of blue. She also learns how to remember the experiences of red and how other people think of the color that they see when they experience roses, strawberries, and blood. Qualia are therefore physical, as Mary doesn’t know any new form of factual knowledge-that (e.g. that water is H20), but some new knowledge-how, that is to say an ability (like when we learn how to ride a bike).

Here are two segments, respectively from Lewis 1983 and 1990:

...knowing what it is like is the possession of abilities: abilities to recognize, abilities to imagine, abilities to predict one’s behavior by imaginative experiments (Lewis 1983, 131).

The Ability Hypothesis says that knowing what an experience is like just is the possession of these abilities to remember, imagine, and recognize. ... It isn’t knowing-that. It’s knowing-how (Lewis 1990, 516).

*Reply:* the usual reply to the Ability Hypothesis insists on the fact that an ability isn’t *per se* sufficient to understand Mary’s epistemic progress. Once she comes out of her room, she seems to learn something other than a series of abilities. Sure, what is like to see red implies that she is then able to remember the color red, recognize it, etcetera. But does it only consist of that? After all, we can imagine a scenario (see e.g. Alter
Mary experiences red, but when we ask her to make a comparison with a slightly different shade of the same color, she’s in trouble. How can she discriminate that with such a precision? It seems that what is like to see red isn’t exhausted by an ability, or at least that isn’t perfectly equal to that.

A second case which strengthens a possible and well-structured response states that after losing her memories for a short period of time, Mary wouldn’t be able to remember her first experience of a red rose. But she can understand whether she’s staring at a red or a blue rose. Once again, it seems that her new knowledge of colors isn’t fully explainable in terms of abilities like Lewis and Nemirow would say. If the arguments offered are sound, then Mary’s new knowledge of qualia can’t amount to a set of abilities. Thus, both Lewis and Nemirow fail to explain what the real meaning of the knowledge argument is.

2.2 The phenomenal concept strategy

A popular response, which in the last years has received way more attention than other types, has got to be the phenomenal concept strategy. The key idea behind this position is that Mary’s scenario doesn’t work properly: what she comes to know after her release isn’t a new fact on our world, but a new mode of presentation of something she already knew inside of her room. To use an analogy, it’s like you knew that the Name of the Rose was written in 1980 by Umberto Eco. If you had met Eco (who unfortunately died in 2016), you would have known him in first person. You would know, in a way, what is like to see the author in front of you, but there’s no new piece of information. You already knew that he was the author of the work you read before meeting him in that specific occasion. There’s no new knowledge-that, to use a quote from the precious section, and consequently no new facts of any nature.
Mary’s knowledge can be accounted as follows: what she missed in her room was a certain type of knowledge, that is to say, “indexical knowledge”. She didn’t know that red is like this, where “this” refers to her specific experience of the rose (or the tomato) outside of her room. There are no non-physical qualia or any new fact she wasn’t aware of: she simply learns that a physical experience has got a specific phenomenal character, which corresponds to what she experiences in first person after being released. She already knew that people usually have experience of red during their life when they stare, let’s say, at a strawberry, but she didn’t have access to that specific phenomenal character. Once outside, she experiences it and can infer that this is equal to what she studied during her confinement. This strategy can be found for example in Tye (2009), Perry (2001) and Lycan (1990 and 1996). How can a dualist counter this strategy?

Reply: one could insist on the fact that the phenomenal concept strategy can’t account for what is like to see red in Mary’s case. Sure, she learns that a specific experience has that particular phenomenal character. But an explanation of why she reacts that ways (the so famous “aha!”) and what her phenomenal character consists of is nowhere to be found. What’s more, if qualia were really physical, Mary would have been able to deduce what is like to have them in her room and we wouldn’t need a new mode of presentation.

Third, if there simply was a difference in the mode of presentation, those who embrace the explanation given by Tye and colleagues should explain why there’s a particular need of an experience in first person, something you don’t need in order to acquire a complete knowledge of physical facts in general. This line of objection is sustained in Alter (2013). It’s true that Mary has got to be acquainted with the relevant qualia in order to form the phenomenal concept, but why does it have to be in first person? And why should a concept fully explain a phenomenal fact like the smell of coffee and what is like to see the sky, when there’s a strong distinction between the mere concept (e.g. the word “blue of the sky”) and its phenomenal counterpart?
3. Conclusion

In this paper I have briefly presented many ways a dualist could account for the phenomenon of consciousness. I have analyzed two main groups, i.e. substance dualism and property dualism with its sub-categories, and I also presented one of the main arguments in support of dualism.

What is left to say? There would be way more to discuss about, not to talk about all the other arguments in support of the dualist cause, such as the conceivability argument and the inverted qualia cases. My purpose in this paper wasn’t however to fully describe dualism and its privileges compared to physicalism and all the similar positions, but rather present it in a coherent and clear way to anyone new to the debate.

If we are ready to agree on the failure of materialism, then we should also expect to embrace a general form of dualism. I think the mental can’t be accounted in physical terms yet, which is by far one of the biggest challenges for science and the philosophy of mind. A strong consequence can be immediately apparent: if there’s more to say thank what physics actually tells us, then we could speculate on afterlife, nature of consciousness and the possible existence of a non-physical soul. This would highly impact the way we perceive our life and the relationship with other people, with a different perspective on how we should spend our time and how we conceive death and the whole matter of God’s existence, for example.

I am not so confident that materialism is the definitive answer for consciousness and if you agree with me and the rest of this writing, then the only option left is to conclude that we are more than matter, that there’s a veil of mystery which can’t be dismissed with equations and physics which embraces ourselves and our lives.
References

➢ Descartes R., 1641, Meditations on First Philosophy.


➢ Soriani T., 2018, "La mente e il suo cervello": la proposta dualistico interazionista di J.C. Eccles e il mind-body problem, Universitas Studiorum.