

# The False Problem of Consciousness

Alex Gomez-Marin<sup>1,\*</sup>, Juan Arnau<sup>2</sup>

<sup>1</sup>*Behavior of Organisms Laboratory, Instituto de Neurociencias CSIC-UMH, Alicante, Spain*

<sup>2</sup>*Facultad de Filosofía y Letras, Universidad de Granada, Granada, Spain*

\* Correspondence: [agomezmarin@gmail.com](mailto:agomezmarin@gmail.com)

## ABSTRACT

Philosophers have usually dealt with the problem of consciousness but, in the last decades, neurobiologists have undertaken the daunting task to address it scientifically. In particular, to answer *how* the brain *produces* consciousness. Here we question *whether* it actually does so, seeking to articulate the precise relation between neural activity and subjective experience. There is no doubt that they are intimately related. However, we argue that the thesis of parallelism (that consciousness tells no more than what is going on in the brain, but only in a different language), rather than enunciating an empirical fact, betrays a philosophical commitment. In addition, such equivalence between mental and cerebral states can be shown to lead to self-contradictions (the brain produces the world with itself in it; the brain, as an object in conscious experience, gives rise to conscious experience). Our approach endorses an integration of philosophical and scientific efforts where the scientist seeks to obtain empirically expedient answers to questions that the philosopher has previously formulated in a conceptually precise manner. Namely, that the questions we ask warrant the answers we provide. We posit that the success of an interdisciplinary quest for consciousness hinges on two more challenges: to accept negative proofs as ways forward and, especially, to avoid presenting metaphysical positions as scientific results. Differentiating between solidarity (as brain lesions demonstrate) and equivalence (as no data does) offers an alternative point of departure for understanding consciousness that does not, from the outset, outlay a false problem. The “easy problem” is hard-enough for the “hard problem” to be ill-posed.

**Keywords:** consciousness, neuroscience, philosophy of mind, brain-mind relation.

*“That there is a close connection between a state of consciousness and the brain we do not dispute. But there is also a close connection between a coat and the nail on which it hangs, for, if the nail is pulled out, the coat falls to the ground. Shall we say, then, that the shape of the nail gives us the shape of the coat, or in any way corresponds to it?” (Henri Bergson)*

## 1. INTRODUCTION

Understanding consciousness had traditionally been an exclusive task of philosophy. However, in the last decade of the 20th century, for a growing number of academics consciousness ceased to be a dirty word and became “largely a scientific problem” (Crick, 1994).

In a landmark paper, Crick and Koch proposed a research program based on empirical neurobiological work to provide a solid and arguably more fruitful understanding of the neural mechanisms of consciousness (Crick and Koch, 1990). In particular, the cornerstone of their proposal was to lay down the foundations and the momentum for “a research program whose supreme aim is to discover the neuronal correlates of consciousness” (Koch, 2004, p.xv). Such quest for consciousness postponed causal accounts and left aside theoretical speculation. It was a pragmatic starting point, where consciousness really meant visual perception. In the last decades, such a quest has been reinvigorated by the advances, opportunities and promises that neuroscience offers. For a review of the most recent findings regarding such attempts at the anatomical neural localization see (Koch et al., 2016).

Moreover, in the nineties, Chalmers coined the famous “hard-problem” of consciousness, realizing that “there is something further that needs explaining here” (Chalmers, 1996, p.xiii). One thing is to explain behavior in terms of neural activity, another to try to account for the experience of red or the pain in one’s tooth. The articulation of such an observation was important to fuel further interest in consciousness research, but it is both obvious and certainly not new. Many philosophers had been thinking hard about it, amongst whom we find Nagel’s famous essay “What Is It Like to Be a Bat?” (Nagel, 1974).

Shifting the focus from brain regions to neural processes, Tononi and Edelman proposed to use neural integration and concepts from complexity theory, in combination with neurological data (Tononi and Edelman, 1998). More recently, Tononi has gained notoriety due to his Integrated Information Theory of consciousness (Tononi, 2004; Tononi, 2008; Oizumi et al., 2014), which is one of the few (if not the only one) tangible mathematical theory of consciousness we have today. Besides, Tononi’s framework has the virtue of starting from consciousness itself, namely, based on phenomenology, it is experience which informs about the axioms and postulates from which one would then derive requirements for the physical world to sustain it. Distinguishing between different kinds of information-processing in the brain (global broadcasting, self-monitoring) has also been explored as a way to address whether machines can be conscious (Dehaene et al., 2017).

Other notorious researchers have proposed a middle way between the hard and the easy problem: “the real problem” of consciousness (Seth, 2016). Such research program does not seek to answer why we are conscious, but instead to account for the properties that makes us so. This goal is harder than explaining behavior but easier than diving into the conundrum of neural experience. As one more iteration of the usual dismissal of the tenets of vitalism, the logic reads as follows: as with life, so with consciousness. It is just a matter of time and resources until we can explain the *élan mental* way. In addition, Seth’s approach is inscribed in the nowadays popular school of thought that considers that reality is constructed by the brain. We seem to hallucinate every conscious percept, including our brains.

There is also a faction of intellectuals that champion a more extreme position for which the null-hypothesis for a theory of consciousness must be that it is an illusion

(Dennett, 2016; Frankish, 2016). It is not only that we build reality. It does not really exist. The violence that such a proposition makes to our immediate experience seems to be compensated by the erudition of its proponents. With a sleight of hand proper of the theater illusionist, the armchair illusionist defends that experience is not real, contorting arguments around the fatal caveat that illusion is itself a mode of consciousness.

In the midst of such an emerging zoo of exotic scientific and philosophical approaches to consciousness, Chalmers has recently introduced what he calls “the meta-problem” of consciousness: “the problem of explaining why we think that there is a problem of consciousness” (Chalmers, 2018). At first sight, this seems a curious way to avoid the hard task that Chalmers himself baptized in the nineties. We can leave aside the original preoccupation to account for our experience (say, of red, coffee, or pain), and concentrate on the preoccupation itself. Thus, studying our behavior in studying consciousness as a research topic, we avoid the hard problem. It is hard to have a great idea every decade. Now, the meta-problem has the advantage to welcome any set of contradictory attempts to explain consciousness. After all, it is the very drive to explain that Chalmers seeks to explain.

Last year, several prominent authors wrote a kind of manifesto discussing challenges and opportunities for consciousness as a maturing science (Michel et al., 2019). Rather than a conceptual or experimental contribution, the paper is a political statement that insists that scientific studies of consciousness require and deserve proper funding allocation.

It is impossible to recapitulate here all the many takes on the problem of consciousness, making justice to their history, complexity and subtleties. Nor is this the goal of the present manuscript. Before we move to our proposal (which is to question the unquestioned question that most orthodox neuroscientists seek to answer about consciousness, namely, *how the brain produces it*), let us now briefly discuss some heterodox approaches to the problem.

### **Other approaches to the problem of consciousness**

Varela’s reaction to Chalmers formulation of the “hard problem” led to a landmark methodological proposal (Varela, 1996). Negatively, the Chilean biologist, philosopher and neuroscientist argued that there is no “theoretical fix” or “extra ingredient” to bridge lived experience and the objects of science. Positively, he proposed the phenomenological method to explore the “mutual constraints” between them.

Within a similar stance, Noë and Thompson argued that both philosophers and neuroscientists should try to understand the relationship between brains and consciousness differently from the “matching-content doctrine”, where it is assumed that the content of visual experience matches the content of a neural representation (Noë and Thompson, 2004). In a remarkably deep analysis, Fink has deconstructed and identified some related problems with claims about necessity and sufficiency of neural correlates of consciousness (Fink, 2016). A similar analysis has been carried out regarding the conceptual flaws of counterfactual reasoning for the “easy” problem (Gomez-Marin, 2017).

Thompson’s approach to consciousness can be seen as a first order elaboration of Varela’s (which in turn is a reenactment of Husserl’s phenomenology in the context of neuroscience). Consciousness is certainly contingent upon what we call matter. But one must be vigilant with a subtle asymmetry in the use of “contingent upon”, “always used in the same biased way, namely with consciousness as a grammatical subject, and the brain as a grammatical object” (Bibol, 2015). The critique is that the implication of the data of neuroscience seems to acquire a higher ontological status than the immediate data of consciousness. And so, even when practicing the so-called neuro-phenomenological reduction, one must beware not to fall into physicalism.

Accepting that the solution of the hard problem requires its neuro-phenomenological dissolution has consequences for science in terms of recognizing its major blindspot, which is the construction of objective knowledge by the knower (Bitbol, 2002). In a recent eloquent essay it has been argued that the fundamental tenet of physicalism, that there is nothing but physical reality, “is either false or empty” (Frank et al., 2019). In a more radical position, it has been denounced that the “irrational nature of materialism” becomes patent when one denies experience in order to affirm abstractions (Spira, 2017, p.1). Bitbol has shown how most attempts to claim that conscious experience is not ontologically primary fail, while articulating half a dozen arguments against the claim that matter produces consciousness (Bitbol, 2008). Even more, one wonders whether consciousness has really an *origin* or, put differently, whether “il est impossible de s’avancer intact à travers elle, au-delà d’ella, ver une réponse plausible” (Bitbol, 2014, p.9). Qualia ceases to be a hard problem under the framework of Fundamental Awareness, in which awareness is “the underlying reality, not reducible to anything else” (Theise and Kafatos, 2016).

Within a “process ontology” framework, other authors have pointed out the ill-posed nature of the hard problem by noting it is indeed an “impossible problem”, namely, that “calling it a “problem” is in itself problematic” (Prentner, 2018). Other non-orthodox authors adopt a more belligerent position against materialism. The problem of consciousness is an artifact of “unexamined assumptions and fallacious logical bridges” inherent to physicalism (Kastrup, 2017). Trying to explain mind in terms of matter entails problems, costs and flaws (Kastrup, 2018).

Hardly anybody likes to be considered a dualist today. Virtually everyone is a materialist (at least, at work). Pushing materialism’s self-contradictions to the limit, Strawson has argued that materialism entails panpsychism (Strawson, 2006). Recently, Goff has been carrying the banner of panpsychism (Goff, 2017) amidst tides of reactionary materialists. Then, also at the other extreme of materialism, we have the hard-core idealist position. The proposal is to revert the direction of the quest from the material basis of mind to the mental basis of matter is interesting (Kastrup, 2019).

James did not claim that experience is an illusion but that in substantivizing consciousness (or when opposing matter to spirit, or body to soul) “there appears no universal element of which all things are made”; such entities are fictitious, “while thought in the concrete are fully real. But thoughts in the concrete are made of the same stuff as things are” (James, 1904). Not only scientists but also philosophers seem to have forgotten about James’ radical empiricism in favor of “pure experience”. Sometimes empiricism is not empirical enough.

## 2. SCIENCE AND PHILOSOPHY

The success (actual and potential) of current neurobiological attempts to dissect mechanisms and correlates of consciousness can give the impression —both to the general public but also to the research community— that fMRI scanners have rendered armchairs obsolete, namely, to “propose that a biological account of consciousness does not require metaphysical proposals” (Edelman et al., 2011). And yet, the philosophical bias is probably the only that science cannot avoid (Anderson et al., 2019).

In the first page preface of his first book on consciousness, Koch asks a question that neurobiology has sought expediently since: “How can a physical system, such as the brain, experience anything?” (Koch, 2004, p.xv). Koch goes on and writes: “The fundamental question at the heart of the mind-body problem is, *what is the relation between the conscious mind and the electro-chemical interactions in the body that give rise to it?*” (Koch, 2004, p.1-2; his italics). If one pays careful attention, the statement seems to contradict itself (as we will expand on the

following section). On the one hand what needs to be explained is the precise relation between brain and mind, whereas the sentence ends affirming that the relation is so that the former gives rise to the later. Actually, the main thrust of our manuscript is to agree with the first part of Koch's statement while discussing not only whether the second part is true or not, but also whether it is claim that really springs from neuroscientific data or from covert philosophical commitments.

A bit later we read that "[t]he working hypothesis of this book is that consciousness *emerges* from neuronal features of the brain" (Koch, 2004, p.10, our italics) However, as Koch's own words betray, what initially was a hypothesis has become a scientific assumption currently taken for a brute fact. "These provisional working hypotheses might well need to be revised or even rejected later on" (Koch, 2004, p.11). Well, it is clear that, to a great extent, they haven't been revised nor rejected, nor are they provisional nor working nor hypotheses anymore.

This is where a philosophical approach can play a crucial role amongst neuroscientists, and specially amongst those dedicated to consciousness. As we will argue below, the answers we get from brain scans are truly illuminating, only if the question is well-posed in the first place. By reading virtually any piece of neuroscientific research on consciousness today one realizes that Koch's hypothesis is not deemed a hypothesis anymore, but an established (and, to some extent, irrefutable) scientific fact: the brain *produces* consciousness, the only question is *how*. Isn't it astonishing that Crick's "astonishing hypothesis" (that "a person's mental activities are entirely due to the behavior of nerve cells, glial cells, and the atoms, ions, and molecules that make them up and influence them") has ceased to be astonishing?

Let us now quote Chalmers, as the quintessential philosopher of consciousness: "It still seems utterly mysterious that the causation of behavior should be *accompanied* by a subjective inner life. We have good reason to believe that consciousness *arises* from physical systems such as brains, but we have little idea how it arises, or why it exists at all." (Chalmers, 1996, p.xi, our italics). Again, we see the hypothesis immediately presented as a thesis.

Note another aspect here. Without engaging here in what would require an entire piece, it is safe to suspect that such "good reason to believe" is based, first, on correlational accounts (the famous "neural correlates of consciousness") and, second, on interventional accounts that seek to establish "necessity and sufficiency". Consciousness (or visual awareness in the monkey, as Koch himself concentrated for quite some time) goes off when certain brain areas are turned off, and it may go on again when other brain areas are turned on. The same is true for certain circuits regarding the image on a TV screen. Or, using an older metaphor: "That there is a close connection between a state of consciousness and the brain we do not dispute. But there is also a close connection between a coat and the nail on which it hangs, for, if the nail is pulled out, the coat falls to the ground. Shall we say, then, that the shape of the nail gives us the shape of the coat, or in any way corresponds to it?" (Bergson, 1896).

The reinvigorated quest for consciousness may have lead neuroscientists with an excess of confidence about the secondary importance of philosophy to their laboratory adventures. After all, they may say (not without a pinch of sarcasm), "after several thousand years of speculation, it would be very gratifying to find an answer to it". It is certainly the time for scientists to try to find the answer to the problem of consciousness. But, a kind of philosophical approach is still needed to (at least) help neuroscientists articulate their questions to avoid circularity or false problems. To that end (and before we present our "proof of impossibility" regarding the equivalence between brain and thought, neurons and consciousness, spikes and experience), let us very briefly revisit the origins of the gap between science and philosophy, and some hints of a possibly new conjunction.

## The bifurcation of science and philosophy

Historians of science generally locate the problematic between science and philosophy in the 17th century, with the confrontation between natural philosophers (the old name for physicists) and meta-physicists.

Descartes starts to detach himself from the concerns of medieval scholastics about the origin and nature of the world. Questions cease to be cosmogonic and start to become more personal. What am I doing here? How is it possible to know something? While philosophy tends to subjectivism (Descartes, Locke, Berkeley, Hume y Kant), the medieval and scholastic objectivism moves to physics and astronomy (only Spinoza and Leibniz keep the old habits). This is how physics, retreating from the subjective sphere, abandons “spirit”. Descartes expresses what was already in the air in his time: the essence of matter is spatial extension; that of the soul, its cogitation. From then on modern thought would see the world as set of things unfolded in space (*res extensa*). In turn, sounds, visions, smells and flavours would find shelter in a thinking substance (*res cogitans*). This uncanny situation erects the division between the domain and competences of science and philosophy.

Some thinkers such as Whitehead would object to this division, trying to avoid the bifurcation of nature (Whitehead, 1920), but in general from that moment the competences of each discipline would remain separate and the natural world would be deemed as a complex of entangled bodies. The atom has a mode with respect to the molecule, the molecule has it with respect to the cell, and the cell with respect to the tissue, tissue to organism, organisms to the planet, and so on. Everything belongs to a superior system, where nothing exists by itself and where things reflect themselves one another. The organism is then defined as the reciprocity and intimacy of the relations between the part and the whole. Under this new view, the question for neuroscience is to find out whether the molecules of living bodies have properties that are not found in the molecules of inorganic environments.

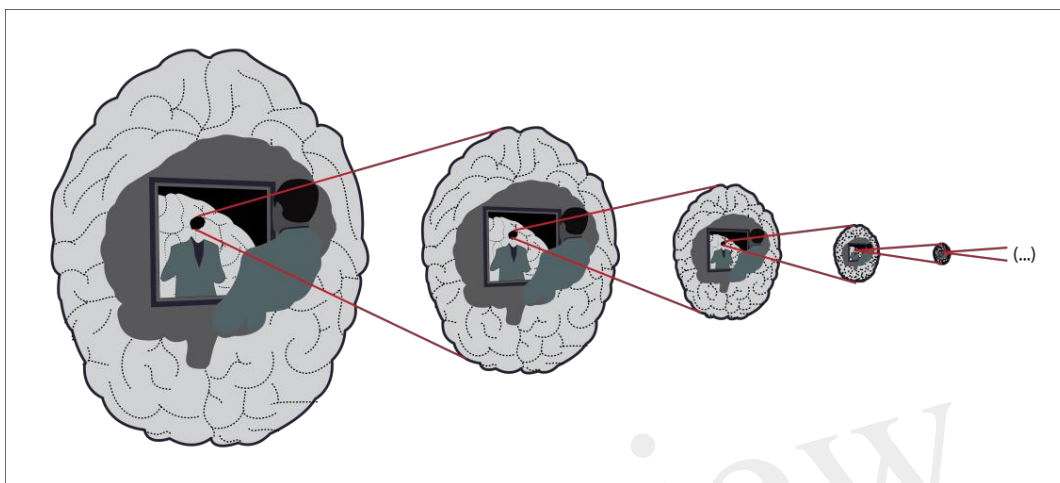
The subject-object dichotomy (philosophical version of the linguistic subject-predicate) can be disorienting. It presupposes a conception of differentiated subjects with privative predicates. It also suggests that there is a being underlying the objects. It would be better to refer, as Whitehead proposes, to an object-ego in the middle of objects. Such object-ego is a here-now of consciousness, aware of its essence as experience, internally related with the world things and ideas.

However, physics finds itself compelled to conceive the things of the world apart from values. Its justification is a methodological one. But the antithesis between things and values can be deceiving. Substances are the starting point of materialism. As such, they have independent existence (matter and spirit). Matter suffers modifications due to its external relations based on impact and movement; spirit due to its objects of contemplation. Science deals with the former, philosophy with the latter.

Since neuroscience has admitted consciousness as an object of research, things have changed. To this false division, Whitehead’s organic philosophy proposes the analysis of the process of events as an entangled society (Whitehead, 1929). The unity of the real isn’t anymore the subject, or the “I” as self or soul, but the event, the lived experience.

Modern physics presupposes a field of electromagnetic activity that fills space and time, and whose laws are the conditions observed by activity, namely, events. What is instructive about this scientific conception is that one can ignore what a thing is in itself in the Kantian sense and study beings as a function of its extrinsic reality. The observer is taken into account, but only in the sense of seeing red or blue. The intrinsic nature of red does not matter, but the diversity of experiences of colour. The acknowledgement of the futility to conceive such “ultimate beings” as individuals instead of events or experiences is productive. Here we can find points of intersection between science and philosophy.

We can now narrow the focus of the present manuscript. The problem of consciousness is triple: there is a *why*, there is a *how*, and there is a *whether*. Before those two, we find it necessary to concentrate on the latter.



**Figure 1. The paradoxical stance of the mainstream neuroscience of consciousness.** The paradox of parallelism (consciousness is nothing but brain activity): the brain generates the world including the scientists inhabiting it and also the brain that their study, which in turn creates the world again and also itself, and so on. As in Escher's lithograph the Print Gallery, this view of the relationship between brain and consciousness leads the Droste effect; an ontological and epistemological picture in which the brain and the world recursively appear within themselves.

### 3. BRAIN-CONSCIOUSNESS

As we have seen, the current dominant position in the science (and, arguably, also in analytical philosophy) of consciousness is that the brain produces consciousness. Neuroscientists should thus be busy figuring out how neural activity gives rise to subjective experience. Such a position can be cast in the following terms: "consciousness tells no more than what is going on in the brain; it only tells it in a different language". Let us call this the thesis of parallelism between cerebral and mental states.

Parallelism can actually be regarded as one of many materialistic attempts to surpass dualism by mutilating it. Amongst its variants we find: eliminativism (consciousness is a folk way of talking about what goes on in the brain), epiphenomenalism (consciousness is a by-product of the brain), illusionism (conscious activity is just an illusion, whereas brain activity is real), and cognitivism (consciousness, like cognition, is information processing, and the brain is, of course, a computer). Since virtually all academics reject dualism, materialism seems to be the only sensible alternative, despite embracing "simply located bits of inertly enduring stuff" (Gomez-Marín and Arnau, 2020). However, as an alternative to physicalism, panpsychism is now experiencing a revival (although consciousness is then deemed a fundamental property of matter, but a *property* nevertheless).

In what follows, we argue that parallelism is literally doomed to remain lost in translation, and crowded with paradoxes (**Figure 1**). To make our argument precise (see later, and **Figure 2**), in the remainder of this manuscript we follow Henri Bergson's take on this problem, exposing the fundamental self-contradictions of such a position. By pointing out a cystic theoretical dead-end, we then hope not only to enable alternative ways forward, but we also wish to spur a more conscious neuroscience of consciousness (**Figure 3**).

## The parallelism between brain and thought

In a lecture given in 1904, at the Geneva Philosophy Congress, and published under the title *Le parallélisme psychophysique* (Bergson, 1904), Bergson argued that the brain does not produce nor contain representations, that brain function is limited to selecting images and memories useful for the action of the body. Bergson went so far as to claim that memory is not stored in the brain (a position we subscribe but cannot address here). This view is in sharp contrast with the pervasive idea that the brain ought to produce them. Therefore, what we say may sound preposterous (specially before one waits to consider the arguments). We beg the reader to hold on.

For Bergson the thesis of parallelism does not find its origins in anatomical or physiological studies. It is a legacy of the metaphysics and Descartes. Kepler and Galileo had prepared the terrain, reducing the astronomical problems to mechanical problems. From this arose the aspiration (recovering in passing old Platonic-Pythagorean doctrines) to represent the totality of the material universe subjected to mathematical laws; a solution that pleased physicists and placed them at the very centre of the axis of knowledge. Descartes gave the final blow: living bodies are clockwork mechanisms, and nothing of what we will do escapes a Universe mathematically determined beforehand.

According to the strong version of parallelism, the cerebral and psychological state correspond point by point, without it being necessary to attribute to one the creation the other. In its weak (but still problematic) version, consciousness ensues from brain activity, and so, in a revival of Laplace's demon, knowing the precise location and momentum of every atom in the brain would give us consciousness. In other words, were we to see the full dance of brain activity, we could know how it feels like to be.

As we have formulated it above, parallelism between mental and brain states affirms that "consciousness tells no more than what is going on in the brain; it only tells it in a different language". This belief was a metaphysical hypothesis adopted by physiologists and physicians of the eighteenth century, who did not have other models to choose from. However, the problem is less the inability to articulate a viable alternative but the fact that it suffers from internal contradictions. Parallelism is not empirically adequate nor logically consistent. We contend that, despite the tremendous advances in neuroscience over the last century, Bergson's logic remains valid. Let us see how.

To advance the core of the contention, let us say now that the internal contradiction of the psycho-physiological parallelism lies in the surreptitious passage from one "system of notation" to another that, by construction, precludes the former, without noticing or recognizing such displacement.

By these two "systems of notation" we mean, in a general sense, "realism" (R) and "idealism" (I). Because we are well aware of the many different possible and subtly complex definitions of such terms, let us use R and I as flags referring to two ways of considering the ultimate nature of the world: either made of ideas (I) or made of stuff (R). Of course, deeper definitions could be provided for such opposed tendencies, but this shall not be necessary for our purposes here.

If we adopt the R-notation, we treat the objects of the world as *things*, external, "out there", independently or beyond the mind that observes. In the I-position we would treat them as *ideas*, as there is nothing out there but all is mind or in mind. For the user of the I-notation it is absurd to speak of a property of matter that is not the object of representation. There is nothing virtual or possible, all that exists is current representation. For the user of the R-notation, the question is very different. Matter exists independently of the representation, and so something inaccessible under the represented must be postulated (paradoxically, realism favours the occult). Whereas the idealist is more prudent and limited



to that which can be represented, the realist ambitions what is outside of representation since, behind the perceived, there are powers and hidden virtualities.

What is of relevance for our purposes here is that such views are incompatible with respect to each other. If you subscribe to one, you can't embrace the other, and vice-versa. It is not legitimate to apply the two approaches to the same object at the same time. Note that we *do not* call the realist or the idealist world-view into question here. We follow the logical consequences of affirming the thesis of parallelism from any of such positions.

As we intend to show now, whether one opts for the R-notation or, alternatively, for the I-notation, the thesis of parallelism only seems sustainable if both notations are used at the same time. But, by definition, this is not possible. And so, the switch is illegitimate, and so the thesis implies a fundamental self-contradiction.

First, let us note that these two notations suggest two points of view: the term *brain* suggests the word *matter*, while the term *thought* suggests the word *representation*. This causes grave confusions and deep misunderstandings.

From the I-position, external objects are considered representations and the brain is one of them. But when the brain is analysed in the laboratory, it is considered a thing, so we move to the realistic framework. What was representation becomes a thing. Notably, such surreptitious shift from the I-notation to the R-notation is favoured by the facts. When we remember the objects after having perceived them, when they are no longer before us, their image becomes visible in the form of a memory. It seems necessary that some part of my body can evoke that image, if not create it (or at least arise it). How could it be done if a certain memory did not correspond to a certain brain state?

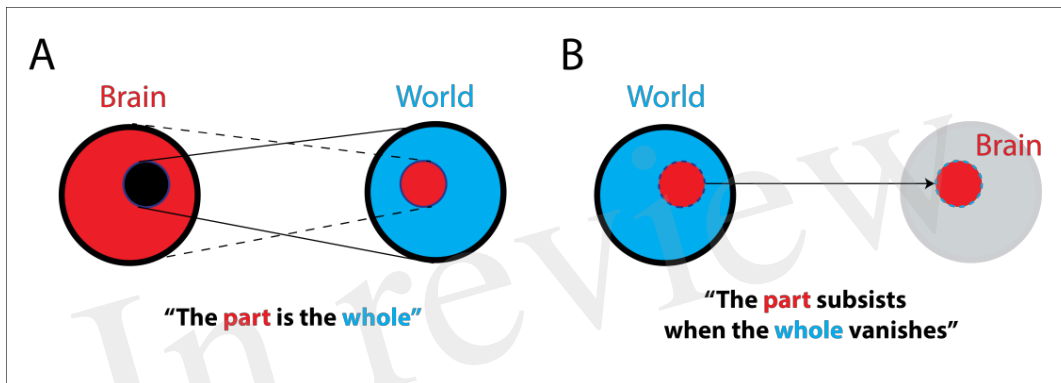
From the R-position, on the other hand, objects are considered to fill the field of vision, including the brain and sensory centres, the displacements of molecules and atoms caused by the interaction with external objects. But the essence of realism is to assume behind our representations a cause that differs from them. Some even affirm that brain states create representations, which are only an *epiphenomenon*; others assume that brain movements cause the appearance of conscious perceptions.

And so in considering the brain (or the neuronal interactions themselves) as an isolated system, an idealist can legitimately declare the object isolated because, from this notation, the object is not distinguished from the representation. But realism consists precisely in rejecting that pretension, in considering arbitrary or relative the lines of separation that our representation traces in things (since it involves reciprocal influences "below" them). He begins by saying that the brain undergoes modifications of external objects so that representations arise in it. Then *tabula rasa* is made of those objects external to the brain and the ability to trace the representation of objects is attributed to it. But when the objects are removed, the brain state that takes their reality from them is also removed. It is only conserved because it passes from the idealist framework (which allows a representation to be isolated) to the realist (who does not admit it). With the external objects present to the brain, representation occurs. Admittedly, this representation is not only a function of the brain state, but of both. The brain state and the presence of the objects constitute an indivisible block. But the thesis of parallelism consists in isolating the brain states and in supposing that they could create by themselves the representation of the objects.

However, all scientific work always involves certain idealism, because it requires isolating phenomena to study them, abstraction being one of the prerequisites of experimental activity. In this way the realist is forced to hypostatize the representations that appear to her and to call them things. These *things* are receptacles of hidden virtualities, which allow to consider the intra-cerebral movements (already converted into things and not into representations) as if potentially, or virtually, they had the whole representation. Thanks to

it, one can affirm the psycho-physiological parallelism. She forgets, however, that she had placed the receptacle outside the representation and not in it, outside space and not in space, that is, she forgets that an undivided reality has been supposed, articulated in a different way than the representation. And by making each part of the representation correspond to a part of reality, she articulates the real as representation. Reality is unfolded in space and realism is abandoned to enter into idealism. The relationship of the brain with the rest of the representation is that of the part with the whole.

For the realist the brain cannot be an independent entity but in the hypothesis of parallelism it considers it as such. Therefore, a tacit shift from idealism to realism and from realism to idealism occurs. We move astride the two systems, apparently reconciling two irreconcilable systems. The whole argument is summarized in **Figure 2**.



**Figure 2. Brain-consciousness parallelism leads to self-contradictions.** Parallelism is the thesis that consciousness tells no more than what is going on in the brain. **(A)** Let us first consider the so-called “I-notation” system, in which reality is made of ideas beyond matter (*all is in the mind*). Then, according to the thesis of parallelism, the brain (in red) would produce the world (in blue), with itself in it (in red). One is forced to claim that “the part is the whole”. **(B)** Let us consider the “R-notation” instead, in which reality is made of matter beyond ideas (*there is something “out there” that is not mind*). Then, according to the thesis of parallelism, the brain (in red), which is inextricable from the world (in blue), would produce the world (in light blue) when isolated from it. One is forced to claim that “the part subsists when the remainder of the whole vanishes”. In order to avoid such logical contradictions, the proponents of parallelism in the I-notation (in A) would suddenly switch to the R-notation. The proponents of parallelism in the R-notation (in B) would do the same switching to the I-notation. But each position is by definition incompatible with the other. The contradiction remains. Therefore, from both notation systems, the thesis of parallelism must be abandoned.

### Solidarity, not equivalence, between brain and consciousness

If we accept the above logical reasoning, we may feel puzzled and protest with an irrevocable empirical fact: the brain is “necessary” for consciousness (as some lesions seem to demonstrate). Apart from the pin and coat image we quoted above, Bergson gives another example: let us say that the presence or absence of a nut makes a machine work or stop working. Can it be inferred from this that each part of the nut corresponds to a part of the machine, and that the machine has its equivalent in the nut?

All that experiments really show is a certain solidarity between the life of consciousness and the life of the body (neuronal activity, but also bodily functions such as breathing, etc). All we are saying here is that it is a logical fallacy to state that such mutual

dependence implies the *equivalence* between both (or, as it is currently presented, that consciousness can be reduced to brain activity).

For Bergson the brain does not draw all the details of consciousness. Rather (and as unacceptably dualistic as this may sound) the brain is the place where the insertion of the mind into matter occurs. Another simile may serve to clarify this point.

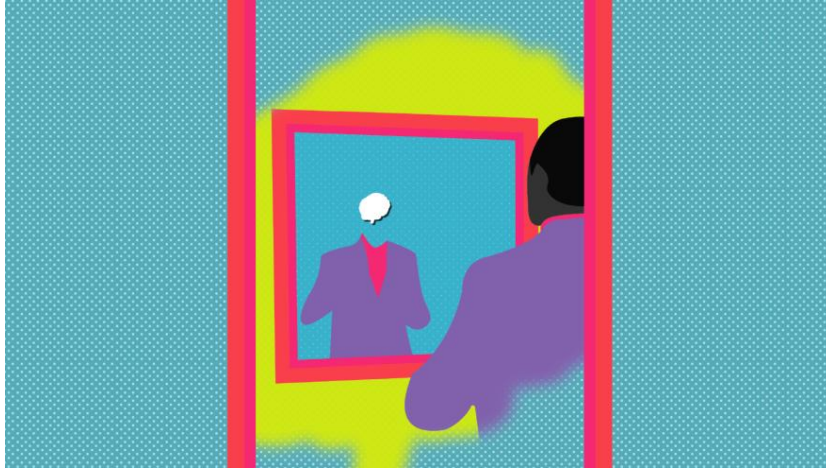
Thought is oriented towards action. Actions (real or virtual) are simplified projections of thought in space. Motor activity is what is manifested in brain activity, which in turn is overflowed by mental activity in much the same way as the movements of the conductor's baton fail to exhaust a symphony. Such correspondence obfuscates that the brain is not, strictly speaking, the organ of thought, feeling or conscience (not even a store of memories) but only—and this is not a minor point—“the place of attack” of consciousness in matter. This is, by the way, consistent with the effects of drugs, intoxications and other altered states of consciousness, as well as with lesions. In these phenomena what is affected is less the mind, but its mechanism of insertion in the body. This has consequences not only for consciousness studies but for neuroscience in general.

Apart from exposing the self-contradiction of those who claim that consciousness equals brain activity, is it possible to address the so-called “hard problem” of consciousness from a stance that does not fall into similar flaws or formulate it as a “false problem”?

In his 1896 work, *Matter and Memory*, Bergson not only points to a dead-end (with a different argument than the one presented here), he also proposes a way forward (Bergson, 1986), which we cannot cover here. Suffice to say that so false would it be to reduce matter to *representations* (as in Berkeley), as to make it a *thing* that produces them (as in Descartes). Matter is, for the French philosopher, a “set of images”. But what he means by *image* has a greater existence than what idealism calls representation and less than what realism calls thing. The solution of the problem of consciousness then starts to look more like a dissolution, a reformulation.

When Bergson raises the mind-body relationship in terms of memory and matter, his point of departure is to state that what we call the universe is a set of images. Among those images there are perceptions for the outside and affections for the inside (a stomach ache, for example). But the brain and the nervous system are in turn images, and it would not make sense to say that an image can produce other images. The brain is part of the material world, not the material world part of the brain.

Thus, from the outset we must discard the representation of the world as a brain production. The brain is not so much a source, but a channel; it does not produce representations, it conduits movement. The people of the movie you watched last night were not in your living-room, nor in the cable that actually brought the TV signal to your house. We then face an even thornier problem: the possibility to conceive “stuff” that does not occupy space, but only time. *How* does the brain produce consciousness? *Where* is memory? Both questions are entangled. And both are most certainly ill-posed. In sum, making the brain the condition of *images* would be a contradiction.



**Figure 3. Towards a conscious neuroscience of consciousness.** The orthodox neuroscientific approach to the study of consciousness is paradoxical in so far as scientists, examining consciousness with their own consciousness, are infatuated with neurons and claim that brains must be the source of everything, declaring consciousness epiphenomenal.

#### 4. OUTLOOK

A great deal of neuroscientists studying consciousness have found their truth before their research actually begins: the brain produces consciousness, the only question is how. Despite the vast empirical evidence that seems to point to parallelism, the modern neuroscientist is, perhaps unknowingly, a “metaphysical trickster”: a philosophical hypothesis is presented as a scientific brute fact.

But that is not all. Consciousness is not brain activity. As we have argued, it can’t be. In embracing the thesis of parallelism, one is either forced to sustain that “the part is the whole”, or that the “the part subsists when the remainder of the whole vanishes”. Both are self-contradictory propositions.

Despite glimpses of alternative ways of articulating the brain-consciousness relation that do not lead to ill-posed questions, we believe that articulating the conceptual dead-end of parallelism is already a productive effort. We have no doubt that consciousness is real. And so is its special relation to the brain. And yet, differentiating between solidarity (as lesions demonstrate) and equivalence (as no data does) offers an alternative point of departure for an understanding of consciousness that does not, from the outset, outlay a false problem. In order to be in position to solve the problem of consciousness, one must first understand it.

Putting forth “proofs of impossibility” has proven powerful in mathematics (i.e. Gödel’s theorem) and physics (i.e. Bell’s theorem). But in biology and neuroscience, “repulsive approaches” have bad press. After all, empirical expediency can lead to conceive scientific progress in a narrow sense. However, to show that something is impossible in principle is as important, if not more, than to believe that it is possible in fact. As we hope to have illustrated in this manuscript, progress is can also be made when one knows where not to look at. Dead ends are ways forward.

Furthermore, we believe that integrating scientific and philosophical research styles places the researcher of consciousness at a sweet spot where one can empirically seek for answers to conceptually precise questions. Clarity should not be at the expense of precision.

Given the main thrust of the current neuroscience of consciousness, our manuscript is a “minority report”, which challenges the very premises onto which the neural bases of consciousness are sought. However, the reason for our critique coincides with our fervent support of neuroscience in its efforts, given its privileged position in the elucidation of the brain-mind relation. We believe that allowing for critical views and voices that are at odds with some of the dominant fundamental tenets of consciousness studies is a sign of intellectual courage and academic health. Few things are more pernicious for science and philosophy than echo-chambers.

If the hard problem is the problem of explaining how consciousness is derived from matter, the solution to the problem does not seem hard but impossible. Matter is, by definition, that which is devoid of mentality. By which intellectual contorsion would one affirm that experience can be obtained by a particular arrangement of bits devoid of experience? Many still think that the aquosity of water is a logical side-product of H<sub>2</sub>O. Ironically, the starting point of materialism is that of substances whose existence is independent. Its destiny, then, when confronted with mind and matter, cannot be other than to negate one of them. Of course, as one could have guessed, materialism will end up denying mind. And yet, in order to overcome dualism, a more sophisticated strategy is required than rushing to negate one side of it. One cannot get rid of the north pole of a magnet by splitting it into two.

Bergson rejects materialistic and idealistic monism. Both are excessive and precipitate. If one assumes that the world is made of things, then establishing the precise role of the brain in consciousness becomes an arduous task haunted by contradictions. The same happens if one assumes it is made of ideas. If, however, and along the lines of the radical empiricists, one sees the world as made of images (as Bergson does) or prehensions (as Whitehead does), then a new field of theoretical possibilities opens up, along with a different research program.

In sum, we think that the problem of consciousness is hard because it is false. It is elusive because it is ill-posed. Figuring out the nature of the brain-mind connection remains as an exquisitely wonderful problem. According to Bergson, this is "the greatest problem that humanity can propose". The French philosopher has the elegance to give up any kind of final solution. The approach is perfectible and this is the maximum one can aspire to.

Although one could name several alternatives (i.e. dual aspect monism, pan-experientialism, enactive approaches, process metaphysics), we insist that it is not necessary to promise a solution now. This contrasts with the belief that neuroscience already understands the nature of consciousness in principle, only to fill in the neural details in practice (an excess of humility after so much pride). The undeniable technological boldness of neuroscientists may be necessary but certainly not sufficient in the times to come. Conceptual audacity is needed more than ever to understand consciousness. To build an ever taller ladder is not the best way to get to the moon. The future neuroscience of consciousness will require a conscious neuroscience. We should question everything, including the very questions we ask.

\*

**Author Contributions.** AGM and JA wrote the manuscript. AGM made the figures.

**Funding.** This work was supported by the Spanish Ministry of Science (SEV-2013-0317 start-up funds to AGM, and BFU-2015-74241-JIN grant project to AGM).

**Acknowledgments.** We thank Gil Costa for help in designing Figures 1 and 3.

**Conflict of Interest.** The authors declare no conflict of interests.

## REFERENCES

- Anderson F., Anjum R.L., Rocca E. (2019) Philosophical bias is the one bias that science cannot avoid. *eLife* 8:e44929.
- Bergson, H. (1896). *Matière et mémoire. Essai sur la relation du corps à l'esprit*. Presses Universitaires de France, Paris 2012.
- Bergson, H. (1904) Le parallogisme psycho-physiologique. *Revue de Métaphysique et de Morale* 12 (6):895-908.
- Bitbol, M. (2002) Science as if situation mattered. *Phenomenology and the Cognitive Sciences* 1:181-224.
- Bitbol, M. (2008) Is Consciousness primary? *NeuroQuantology* 6:1, 53-72.
- Bitbol, M. (2014) *La Conscience a-t-elle un origine?* Flamarion.
- Bitbol, M. (2015) *The Journal of Mind and Behavior* 36:1, 101-11.
- Chalmers, D.J. (1996) *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press.
- Chalmers D.J. (2018) The Meta-Problem of Consciousness. *Journal of Consciousness Studies*, 25(9–10): 6–61.
- Crick, F., Koch, C. (1990) Towards a neurobiological theory of consciousness. *Seminars in The Neurosciences* 2:263-275.
- Crick, F., Koch, C. (1998). *Consciousness and Neuroscience*. *Cerebral Cortex* 8:97-107.
- Dehaene S., Lau H., Kouider S. (2017) What is consciousness, and could machines have it? *Science* 358, 486-492.
- Dennett, D.C., (2016) Illusionism as the obvious default theory of consciousness. *Journal of Consciousness Studies*, 23 (11–12): 65-72.
- Edelman, G.M., Gally, J.A., Baars, B.L. (2011) Biology of consciousness. *Frontiers in Psychology* 2:4 1-7.
- Fink, S.B. (2016) A Deeper Look at the “Neural Correlate of Consciousness”. *Frontiers in Psychology* 7:1044 1-13.
- Frankish, K. (2016) Illusionism as a Theory of Consciousness. *Journal of Consciousness Studies* 23 (11-12):11-39.
- Frank, A., Gleiser M., and Thompson E. (2019) *The Blind Spot*. Aeon Magazine.
- Fuchs, T. (2011) The Brain—A Mediating Organ. *Journal of Consciousness Studies*, 18(7-8): 196-221.
- Goff, P. (2017) The Case for Panpsychism. *Philosophy Now* 121:6-8.
- Gomez-Marin, A. (2017) Causal circuit explanations of behavior: Are necessity and sufficiency necessary and sufficient? *Decoding Neural Circuit Structure and Function*, 283-306, Springer.
- Gomez-Marin, A., Arnau, J. (2020). When the part mirrors the whole: interactions beyond simple location. Submitted. Pre-print: <http://philsci-archive.pitt.edu/16258/>
- James, W. (1904) Does “consciousness” exist? *The Journal of Psychology* 1:18 477-491.
- Kastrup, B. (2017) The Quest to Solve Problems That Don't Exist: Thought Artifacts in Contemporary Ontology. *Studia Humana* 6:4 pp.45-51.

- Kastrup, B. (2018) Conflating Abstraction with Empirical Observation: The False Mind-Matter Dichotomy. *Constructivist Foundations* 13 (3):341-361.
- Kastrup, B. (2019) *The Idea of the World: A Multi-Disciplinary Argument for the Mental Nature of Reality*. iff Books, UK.
- Koch, C. (2004) *The Quest for Consciousness: A Neurobiological Approach*. Roberts & Company Publishers.
- Koch, C, Massimini, M., Boly, M., Tononi, G. (2016) Neural correlates of consciousness: progress and problems. *Nature Reviews Neuroscience* 17:307-321.
- Laplane, L., et al. (2019) Why science needs philosophy. *PNAS* 116 (10) 3948-3952.
- Michel, M. et al. (2019) Opportunities and challenges for a maturing science of consciousness. *Nat Hum Behav* 3(2):104-107.
- Nagel, T. (1974) What Is It Like to Be a Bat? *The Philosophical Review* 83:4 pp. 435-450.
- Noë, A., Thompson, E. (2004) Are There Neural Correlates of Consciousness? *Journal of Consciousness Studies*, 11:1 3-28.
- Oizumi, M., Albantakis, L., Tononi, G. (2014) From the Phenomenology to the Mechanisms of Consciousness: Integrated Information Theory 3.0. *PLoS Comp Biol* 10(5) e1003588.
- Prentner, R. (2018) Process Metaphysics of Consciousness. *Open Philosophy* 1:3-13.
- Seth, A. (2016) The real problem. *Aeon Magazine*.
- Sheldrake, R. (2013) Setting Science Free From Materialism. *Explore* 9(4):211-218.
- Spira, R. (2017) *The Nature of Consciousness: Essays on the Unity of Mind and Matter*. Sahaja Publications.
- Strawson, G. (2006) Realistic monism — why physicalism entails panpsychism. *Journal of Consciousness Studies* 13 (10-11):3-31.
- Theise, N.D., Kafatos, M.C. (2016) Fundamental awareness: A framework for integrating science, philosophy and metaphysics. *Communicative & Integrative Biology* 9:3 e1155010.
- Tononi, G., Edelman, G.M. (1998) Consciousness and Complexity. *Science* 282: 1846-1851.
- Tononi, G. (2004) An information integration theory of consciousness. *BMC neuroscience* 5 (1), 42.
- Tononi, G. (2008) Consciousness as integrated information: a provisional manifesto. *The Biological Bulletin* 215 (3), 216-242.
- Varela, F.J. (1996) *Neurophenomenology: A Methodological Remedy for the Hard Problem*. *Journal of Consciousness Studies*, 3:4 330-49.
- Whitehead, A.N. (1920) *The Concept of Nature*. Cambridge University Press.
- Whitehead, A.N. (1929) *Process and Reality*. New York: The Free Press, 1978.