

## ENGLISH TRANSLATION OF\*

**Bernal, R. J. et al.** (2013). “Précis de ‘E-Physicalism - A Physicalist Theory of Phenomenal Consciousness’ ”. *Ideas y Valores*. Bogotá: Universidad Nacional de Colombia. With contributions by Jérôme Dokic, Pierre Jacob, Max Kistler, Juan Diego Morales, David Papineau and Jaime Ramos. Translations to Spanish by R. J. Bernal.\*\*

---

### *PRÉCIS OF* E-PHYSICALISM - A PHYSICALIST THEORY OF PHENOMENAL CONSCIOUSNESS<sup>1</sup>

#### **Introduction and summary**

The so-called “mind/body problem” has occupied a central place throughout the history of philosophy. On the one hand, human beings are biological organisms. Their bodies are material entities, thereby subject to the laws of nature. But on the other hand, human beings have a mind: they are “rational”, they have feelings and emotions, and they have subjective perspectives on the world. They have a mental life that seems to evade the rigidity of the physical world. The mind/body problem concerns the relation between minds and bodies. *Prima facie*, these seem to have different metaphysical natures. But then, how it is possible for the mind to interact with the body? And if the mind is something physical, or the physical is something mental, why they seem to be so different?

Most modern philosophers were “dualists”: they considered that the mind and the body belong to different metaphysical categories. Some were “monists”—they claimed that mind and body belong to the same category—but usually they took the body to be some kind of mental entity, and not the other way around—they were “idealists”. Unfortunately, from a contemporary perspective, these philosophers were not able to support their convictions through a compelling solution of the mind/body problem.

It was especially during the 20<sup>th</sup> century that “physicalist” monism (or “materialism”), i.e., the idea that everything that exists has a physical nature, acquired many adherents. Physicalists are confident that the existence of the mind and its activity is, somehow, a natural phenomenon, and explore this possibility to its last consequences. Certainly, interesting and enlightening proposals have been put forward about how to “naturalise” mental phenomena, i.e., about how to account for them in a physicalist framework. Moreover, scientific research has provided useful empirical data and relevant theories in the areas of brain sciences and psychology. However, physicalism has received some compelling criticisms. It still faces the challenge of providing a persuasive solution for the mind/body problem.

“The problem of consciousness” is one of the aspects of the mind/body problem. It concerns the question of the nature of subjective experience and its relation with objective phenomena. Human beings happen to be such that there is something it is like to be one of them

\* Translator : R. J. Bernal.

\*\* The commentaries by J. Dokic, P. Jacob and M. Kistler were written in French, and the commentary by D. Papineau was written in English.

<sup>1</sup> This book is a revised version of the PhD thesis of Reinaldo Bernal. Its supervisor was Max Kistler, and it was defended the 5 of December of 2011 at the Institut Jean-Nicod (Paris). The jury was composed by: Ned Block (New York University), Jérôme Dokic (Institut Jean-Nicod), Pierre Jacob (Institut Jean-Nicod), Max Kistler (Paris 1 – IHPST), David Papineau (King's College London) and Jaime Ramos (Universidad Nacional de Colombia).

(Nagel 1974). Experiences like tasting wine, listening to music, looking to a painting, feeling cold, and feeling anxious, have a distinctive “what-it-is-like-ness” or “phenomenal character”. In this sense, a subject that is experiencing is said to be in a “phenomenally conscious” mental state. The problem of consciousness is about this type of states.

Until the seventies, most of the work in the analytic philosophy of mind was focused on “the problem of intentionality”: the fact that thoughts and words are about something else (Brentano 1874). Intentionality was taken as the distinctive characteristic of the mental, and thus as the core of the mind/body problem. The main questions were to determine what mental states are, what they represent, how they come about to represent something and, in general, how intentionality is possible. Subjective experience and, in particular, its phenomenal character, was not a primary topic. Certainly, a distinction between “conscious” and “unconscious” (or subpersonal) mental states was in place. Some mental states of a subject—the unconscious ones—were considered to be inaccessible, in one sense or another, to the subject himself. But there was no clear distinction between phenomenal consciousness and other notions of consciousness. And, more important, a comprehensive account of the contrast between conscious and unconscious mental states was not considered to be crucial for an understanding of the nature of the mind.

Behaviourism, in particular, dismissed the question of consciousness. Within this view, which until the late fifties provided the main theoretical framework in psychology, it was common to consider a discussion about subjective experience as close to nonsense. The realm of “the subjective” was taken, if not as fiction, as a pseudo-scientific category. Science was only concerned with what is directly “observable”. Everything “mental” had to be reduced in terms of behaviour.

During the sixties, with the advent of computationalism, the idea that mental states can be reduced in terms of behaviour was abandoned. In some sense or to some extent, mental states were considered to be internal states. But even though there was, consequently, a place for subjective experience in a theory of the mind, consciousness continued to be a secondary topic. The central question was to determine how mental representations are codified and processed in cognitive systems. Certainly, computationalism had to account for the distinction between conscious and unconscious mental states. But, firstly (and not surprisingly), this distinction was considered to be mainly functional, i.e., a question of access among different mental states or modules. The property of a mental state being phenomenally conscious was not clearly distinguished from its functional properties or role. Secondly, if a mental state had a phenomenal content, this characteristic was considered to be irrelevant for the functional role it could play.

But principally during the last three decades the interest in subjective experience and phenomenal consciousness increased. The fact that some mental states are phenomenally conscious is now taken as primordial for the understanding of the mind. The problem of intentionality continues to be central, and there is no general agreement about how to naturalise it. But it seems that the “hard problem” (Chalmers 1996) *par excellence* is to account for phenomenal consciousness. Indeed, some philosophers claim that consciousness is required for intentionality (e.g., Searle 2002), and others that it plays an essential role for the fixation of the reference of perceptual states (e.g., Campbell 2002). Much philosophical work, with contributions by many of the most prominent philosophers of mind, is been done nowadays on the question of the nature of phenomenal consciousness and the relation between subjective experience and objective reality.

The book *E-physicalism - A Physicalist Theory of Phenomenal Consciousness* advances a theory in the metaphysics of phenomenal consciousness. It is grounded on the convictions that subjective conscious experience—in the sense of Nagel (1974)—is a real phenomenon, and that some variant of physicalism ought to be true.

In *Chapter 1*, firstly, I elaborate the notion of phenomenal consciousness following Block’s (2007) distinction between access consciousness and phenomenal consciousness. Secondly, I argue for realism about consciousness by contrast with eliminativism. It is not possible to prove that consciousness is a real phenomenon, but neither can eliminativists prove that it is not. For the realist, consciousness is given as a brute fact. Thirdly, I argue that given

the mind-body problem, and despite our dualist intuitions, a physicalist monism is the most reasonable metaphysics. Accordingly, I claim that there is a property X, which is a physical property or a supervenient (on the physical) property, such that for an entity S to be conscious is for S to instantiate X. Finally, I criticise panpsychism and conclude that consciousness is a property of some complex physical entities.

*Chapter 2* concerns Strong AI and computational (or “machine”) functionalism about consciousness. Both take consciousness to be a supervenient property and thus are compatible with physicalism. But I argue, firstly, that the behaviour of an entity S supervenes on a base that includes not only S but also physical systems other than S, and secondly, that a function realised by some hardware H is not an intrinsic property of H. By contrast, consciousness has an “internal character”: it is an intrinsic property of the conscious entity. Therefore, I conclude that consciousness is neither a behavioural nor a functional property and thus I reject both Strong AI and functionalist views.

In *Chapter 3*, firstly, I argue that higher-order representation theories of consciousness (HOR) fall short as accounts of the existence of phenomenal consciousness. The occurrence or possibility of a higher-order mental state M' representing a mental state M is not sufficient to account for the fact that there is something it is like to be in M. Secondly, I discuss the unity of consciousness (Bayne 2010) and, primarily, “phenomenal unity”. I claim that any theory, and in particular higher-order thought (HOT) theories, must account for this unity; it stands for one of the essential characteristics of subjective experience. Finally, I discuss the “explanatory gap” (Levine 1983). I suggest that the gap appears, at least in part, when we take the subjectivity of consciousness as an ontological condition and not as an epistemological one. The exclusively subjective access there is to phenomenal contents can be explained by the very particular nature of the epistemological relation holding between a subject and his own mental states. Thus, the property of having phenomenal content can be objective despite the subjectivity of phenomenal experience.

*Chapter 4* is the core of the work. I argue that consciousness does not supervene on physical items,<sup>2</sup> but is a physical property of the conscious entity that emerges from its fundamental constituents. The emergence of properties is conceived as resulting with nomological necessity from the emergence base, and emergent properties are thought as not reducible to fundamental items and endowed with causal powers of their own. This thesis—the “e-physicalism” view—is in conflict with “microphysicalism”, i.e., with the idea that every property of a complex physical system supervenes on fundamental items. Therefore, I argue against microphysicalist metaphysics, and show the plausibility of the emergentist view I advance, through the elaboration of two examples—one in classical physics and one in quantum mechanics. My argument does not show that consciousness is an emergent property, but opens this possibility. The metaphysics of e-physicalism gives a plausible framework for a realist and physicalist view of consciousness that avoids a commitment to panpsychism.

In *Chapter 5*, firstly, I criticise the strategy of using the “conceivability” of a metaphysical world to drive metaphysical conclusions. To determine whether a “world” is metaphysically or physically possible is a nontrivial and uncertain matter. Secondly, I reject—on the base of e-physicalism—Chalmers’ (1996) “zombie argument”. I conclude that an exact physical replica of the actual world cannot be “a zombie world”, and throw doubts about its very metaphysical possibility. Thirdly, I show that Kim’s (2005) “supervenience argument” does not threaten the thesis that consciousness has “original causal powers”, i.e., causal powers that are not reducible to the ones of the fundamental constituents of its emergence base. The e-physicalism view avoids, in particular, the tension between vertical determination and horizontal causation.

*Chapter 6* concerns phenomenal character and *qualia*. Its purpose is not to advance a thoroughly elaborated account of phenomenology, but just to make explicit the commitments and consequences of e-physicalism for this difficult question, and to provide the grounds for a further development of the theory. I try to make plausible the idea that *qualia*, which I define as the ingredients of phenomenal character, are physical properties. First, I argue that phenomenal

<sup>2</sup> An “item” is an entity, property, fact, event, or law that governs the behaviour of entities.

character is different from representational content. It can have the function of representing, and in this case the representational content it conveys is nonconceptual. But it can also comply with nonrepresentational functions. Secondly, I suggest that consciousness has biological functions that result from natural selection, and I sketch a model of “phenomenal space”, i.e., of the structure of the phenomenal character of conscious experiences, in order to illustrate in what sense phenomenal properties could be physical properties. Thirdly, I address Jackson’s (1982) “knowledge argument”. I agree that the what-it-is-like-ness of having a given experience can only be known by having the experience, as the argument assumes. However, I argue that this does not prove physicalism to be false. Physicalism is compatible with the idea that not everything that can be known about natural phenomena can be captured in scientific theories. In particular, scientific theories cannot capture phenomenal contents since these are not propositional contents, but nonconceptual ones.

The objectives of this work do not include a historical synthesis of the discussion about consciousness, or a recapitulation of the totality of influential arguments that have been given in different directions. I discuss some views, many of them in the most general form, some of them more in detail, as they become relevant as I advance, step by step, in the discussion and elaboration of e-physicalism. I expect some of the arguments I present to be original to some extent and, even though I advance some controversial conclusions, I hope that the view put forward is at least coherent.

## COMENTARIES AND REPLIES

*Commentary by Jérôme Dokic*  
Institut Jean-Nicod – France

Bernal did a significant contribution to contemporary debates about consciousness. The book is very well structured and it is pleasant to read it, primarily because the author is concerned with guiding the reader through all the stages of an unavoidably complex argumentation. Bernal belongs to the group of philosophers who consider the problem of consciousness as central in the philosophy of mind and preceding the problem of intentionality. He takes seriously Thomas Nagel’s intuition that “there is something it is like” to be a conscious organism, and affirms a manifestly realist position about consciousness. Like many other contemporary philosophers, Bernal tries to accommodate realism about consciousness within physicalism, but he argues for a particularly original view. He considers that physicalism must be emergentist, that is, that it must acknowledge non-fundamental physical properties endowed with *sui generis* causal powers. The core of his work is precisely the rigorous articulation between the theory of consciousness and metaphysics.

Alongside his argumentation, Bernal exhibits a very impressive mastering of the contemporary literature about consciousness. All the important authors and the most discussed problems are mentioned and analysed in a frequently original manner. The quality of the arguments is excellent. The specific thesis defended by Bernal is really new and deserves to be widely diffused in the contemporary philosophic community. This thesis will undoubtedly give rise to numerous discussions.

According to “e-physicalism”, the relation between emergent properties and fundamental properties is nomologically and not metaphysically necessary. Therefore, if we follow the definition of supervenience in terms of metaphysical necessity, emergent properties do not supervene on fundamental properties. I would like to advance two questions that follow.

1. The author would certainly admit that the thesis saying that consciousness is a physical property is not a *direct* consequence of the thesis that consciousness is an emergent property, in the sense he attributes to this term. At this point, three positions can be considered. The strong option is to claim that emergentism is incompatible with physicalism. For instance, Tim Crane writes about this: “whatever emergentism is, it is not physicalism (quoted in Macdonald & Macdonald 23). A more intermediate position acknowledges that some additional premises are necessary in order to conclude that consciousness has a physical nature. From this

point of view, some emergent properties are physical, but it is possible for others not to be so. Finally, a third option tries to establish, probably based on an *a priori* argument, that all the emergent properties are physical. Bernal rejects without ambiguity the first option, but I wonder if he would accept the second or the third.

2. According to Bernal's definition of emergence, an emergent property is related only in a nomological manner to its base. I wonder thus if it would be metaphysically possible, in principle, for the entity to exist without this base. The question is, in other terms, whether an emergent property is *necessarily* emergent. Another option is for emergent properties to be necessarily linked to a base, but not to a physical base. If these considerations apply to consciousness, then, in one or other case, consciousness *could* exist without a base, or at least without a physical base, which would imply that it is a physical property in our world but not in every possible world.

### **Reply by the author**

J. Dokic states, rightly, that it does not follow from consciousness being a property which emerges from the physical (if it does) that it is itself a physical property. He presents three possibilities for the nature of properties that emerge from the physical: a) they are not physical, b) they are physical if they meet some conditions besides being emergent, c) they are physical (this would be established by *a priori* arguments). Dokic wonders if I would accept the second or the third option, given that I clearly reject the first one.

I consider that e-physicalism is committed to the third option. First, because it is a form of physicalism according to which everything that exists has a physical nature (is physical or supervenes on the physical). Secondly, because of the way in which the relation of emergence is defined: An item U emerges over a set of items {Pi} if U is *nomologically* entailed by {Pi}. In this definition, the nomological necessitation involves exclusively laws of nature. Accordingly, if the emergence base (the set {Pi}) is constituted by physical items, the result (the emergent item) must have a physical nature. In fact, in my opinion if a given property is related with physical properties *via* laws of nature, then it has a physical nature (I argue for this in *Chapter 1*).

The second issue proposed by Dokic is the question of whether emergent properties are necessarily emergent (in the metaphysical sense). If this is not the case, there is the possibility of a world where consciousness exists in the absence of a base it emerges from, or where its emergence base has not a physical nature.

Effectively, for e-physicalism the emergence of consciousness only happens, *in the way it happens in the actual world*, in those worlds that share with the actual one the laws of nature. In worlds governed by other laws, as Dokic notes, there is the possibility for consciousness to exist despite the absence of a base it emerges from, or to emerge from a base that has not a physical nature. However, in *Chapter 5* I argue that the "conceivability" of these worlds does not prove that in our world consciousness has not a physical nature. As Dokic notes, I consider that in our world consciousness has a physical nature and I do not pretend to generalise this position to other worlds.

**Commentary by Pierre Jacob**  
Institut Jean-Nicod – France

The author's general project is to defend a conception both physicalist and realist of phenomenal consciousness. I was greatly impressed by the aptitude of the author to develop and defend original and coherent theses about difficult and deep metaphysical matters.

Now, I have the following question. There seems to be a tension between two lines of thought of the author, one internalist and the other externalist. On the one hand, the author seems to subscribe to an internalist conception of phenomenal consciousness (*Chapter 2*). But, on the other hand, e-physicalism attributes a crucial role to nomological relations between physical properties of the brain of the agent of experience and properties instantiated in its surrounding. For example, the author writes:

Obviously, and importantly, experiencers are not isolated systems. Like every physical system they are nomologically related with other systems. Which sub-state the consciousness property occupies depends on the interaction between the experiencer and the environment. For instance, the what-it-is-like-ness of seeing a ripe tomato is nomologically related with physical properties of the tomato. (159)

It thus follows that the intrinsic properties of the brain of the agent of experience do not constitute a sufficient supervenience base for the phenomenal experience of the tomato's redness. But the author also writes: "if S is conscious this is an intrinsic (or "internal") property of S" (55). "The rejection of both Strong AI and computational functionalism is based on the remark that it is by definition that consciousness is an intrinsic (or "internal") property" (70). "[C]onsciousness has an "internal character": it is an internal (or "intrinsic") property of the conscious entity" (73). In short, in the second chapter (70-73), the author seems to endorse an internalist conception of consciousness that is incompatible with externalist conceptions (both of vehicle and of content). However, he also claims that it would be a mistake to believe that this internalism about consciousness is incompatible with externalism.

### **Reply by the author**

P. Jacob is perfectly right when he notes, regarding mental states, the existence of a tension between an internalist and an externalist position. On the one hand I claim, following Nagel (1974), that consciousness is an intrinsic property of the subject of experience. But, on the other hand, I am sympathetic with the externalist conception of mental states. I try to accommodate both views with a distinction (that I develop in *Chapter 6*) between a representation's vehicle and a representation's content. I claim that consciousness is a physical property of the vehicle of representation, and not a constituent of the representational content.<sup>3</sup> Thereby, the phenomenal character would be internal while the representational content would be (at least partially) external. Additionally, I do not claim that all the properties of a mental state are physical. Some properties may supervene on physical states that include brain ingredients but also other ingredients that are external to the subject.

To claim that the brain should not be conceived as a closed system—because brain states are nomologically related with physical states external to the subject—is to state what follows: brain states, and particularly the ones on which conscious mental states supervene, are determined not only by the physical properties of the brain but also by the external stimuli (which are physical interactions) on the brain. The state a brain occupies is determined by the complex net of external relations it is involved in. Without taking into account the existence of these relations, it is not possible to understand the brain's dynamics. In this sense, properties like consciousness are linked intimately with external properties.

***Commentary by Max Kistler***

Université Paris 1 – IHPST – France

This is an excellent work on metaphysics and philosophy of mind. Bernal has a clear objective: to develop and defend the view that consciousness is an intrinsic high-level property, characteristic of some complex physical individuals that are subjects of experience. He defends this view in a systematic and perfectly structured manner; each argument is presented with a detailed analysis of its logical structure.

Alongside his defence of what he calls "emergentist physicalism", Bernal shows that the crucial difference between it and microphysicalism lies in the answer to the question about whether microphysical entities (objects and properties) are metaphysically sufficient for the set

<sup>3</sup> I also hold that the phenomenal character of a mental state serves representational purposes in many cases, but not always. For instance, visual experience (arguably) has the function of representing a scene, but an experience like feeling anxiety (arguably) does not represent anything.

of facts, including biological and cognitive ones. According to microphysicalism, when the set of the fundamental physical facts is determined, the set of all the other facts is determined as well. Now, the role of the laws of nature in the metaphysical determination is frequently ignored, or not taken seriously enough. Bernal then argues, persuasively, that microphysicalism would not be convincing if no law was among the fundamental entities. Without the laws that determine the interactions between the entities at the fundamental level (in virtue of their properties), the mere existence of these entities and facts does not determine the higher-level facts. The mere existence of a set of H<sub>2</sub>O molecules, together with the set of facts about them, when taken separately or independently, does not determine any high-level fact about water. Liquidity, transparency, and all the remaining macroscopic properties of water are determined by the facts concerning the H<sub>2</sub>O molecules *in virtue of laws* that govern the interactions among these molecules. Once this has been established, Bernal can illustrate better in what consists the opposition between emergentist physicalism and microphysicalism: the question of finding out which laws are necessary to determine the facts at the highest levels. According to microphysicalism, the laws at the microphysical level are sufficient; following emergentist physicalism, further laws are necessary: emergence laws. Bernal develops two examples of non-fundamental laws that cannot be derived in turn from fundamental laws. The first shows that non-fundamental laws (irreducibly statistical) are necessary to determine thermodynamic properties like temperature. The second, concerns the laws that determine the global and non-local properties of some superposed (entangled) systems in quantum mechanics, like those introduced by Einstein, Podolski and Rosen in 1935.

Now, despite its virtues the thesis advanced by Bernal faces, like every theory, some difficulties. A first difficulty results from the fact that the proposed view can give the impression that the property of being conscious is placed at the same level that quite modest physical properties like having a temperature. This would deprive consciousness of the very particular status that leads to the “hard problem of consciousness” (to use Chalmers' expression). The task for the physicalist consists in avoiding both horns of the following dilemma: on the one hand, presenting consciousness as something so special that it has no place in the physical world; on the other hand, claiming that, given that it has its place in the physical world, it is thus an “illusory” property. The view proposed by Bernal seems to avoid the first horn but not the second. If molecules and gases also have emergent properties, how can we explain the particular metaphysical status of consciousness?

An important aspect of the metaphysics elaborated by Bernal is the thesis, shared with the usually called “British Emergentists” philosophers, that reality has a “layered” structure. Bernal does not define explicitly in which sense he understands “level of reality”, but explains this notion through the concept of “emergence”. An entity belongs to a high level, relative to a lower level, if it is necessary to appeal to a law of emergence in order to explain its existence and properties. Now, the notion of level of reality, therefore, suggests that two entities which interact belong to the same level. However, this seems to be incompatible with another traditional idea concerning levels: the idea that these form a (exclusive) partition of reality. The problem is that many entities seem to belong to different levels at the same time. Persons, while they are conscious, belong to a higher level than fundamental physical particles. Nevertheless, persons can interact directly with fundamental physical entities, as happens when a person perceives an individual photon. This seems to have the consequence that persons belong also to the level where photons—which are fundamental entities—belong.

### **Reply by the author**

The first critique advanced by M. Kistler is perfectly fair and points to the limits of the theory I advanced. The thesis that consciousness is an emergent property is not sufficient to explain its startling features and the truth of emergentism does not entail at all the reality of consciousness. Indeed, the examples I use to argue for emergentism are poorly related to consciousness: temperature and non-locality in quantum mechanics.

To be sure, if consciousness is in fact an emergent property, it is *a sui generis* one. To answer the question of why there is this property in a material world, I do not know what to say

apart from “this is just the way our world is”. We live in a world where there is a physical property which, when instantiated in some entity, makes this entity to be such that “there is something it is like” to be it. To ask why does consciousness exist is like asking why there are atoms with such and such properties. That is just the way things are, and things must be somehow.

Now, I think that succeeding in showing that there are emergent properties in nature is a significant movement towards the development of a physicalist and realist theory of consciousness that avoids panpsychism (which I do not take as plausible). This success opens the possibility for consciousness to be an emergent property as well: a physical property of complex systems not instantiated in the individual constituents of these systems.

The second critique of Kistler points to a significant difficulty that every theory saying that nature has levels faces. One reason for postulating levels is the following: there seem to be complex systems that interact with each other in virtue of causal powers that cannot be reduced in terms of the causal powers of their respective constituents. Along these lines, each level is defined as a set of elements causally closed under irreducible causal relations. Now, it is easy to find counterexamples where an entity belonging to some level causally interacts with an entity that should be considered to belong to another level.

Because of the mentioned difficulty, I avoided associating with e-physicalism a metaphysics that structures reality in levels. I have doubts about the possibility of establishing the existence of levels differentiated in a clear and precise way. Certainly, the emergence of a property places it at a “higher” level than its emergence base. But I take this difference in levels as something that has local significance. I am not committed to the existence of different global levels that are related by laws of emergence. Consequently, nothing prevents in e-physicalism the possibility of a conscious person interacting causally with a fundamental particle. Physical properties, whether fundamental or emergent, can interact with each other in virtue of the laws of nature. Now, in Chapter 5 I try to show why e-physicalism would escape from the influential argument by Kim against the existence of “vertical” causal relations.

**Commentary by Juan Diego Morales**  
Universidad Nacional de Colombia

*E-Physicalism* is an audacious book that, in general terms, shows its main thesis: phenomenal consciousness should be understood as a property *both physical and emergent*. It is a book that should be read by anybody interested in the mind-body problem, that is, in the question about the place that consciousness and mind have in our physical world, and about the relation between the different sciences from physics to sociology. In particular, it is a book that should be read by those who are working and researching in these topics. I say it is audacious because its main thesis, the idea that consciousness is both physical and emergent, has been considered as an oxymoron by a big part of the philosophical tradition of the last century: How could emergent minds and consciousness be different from the merely physical and at the same time be something physical? Beyond the conceptual clarifications necessary to dissolve this initial paradox, a big part of the philosophical work of the last decades—and I would say the most influential—insists in its incoherence (see Bennett 2008, Stoljar 2008 and Kim 2010).

In this short commentary, I will focus on the analysis of three fundamental concepts in Bernal’s thesis: supervenience, emergence and reduction. I will try to argue basically for two points: (a) that there is an ambiguity in the treatment of the referred concepts that can lead to a misinterpretation of the general argument, and (b) if we solve this ambiguity and follow the characterisations and definitions proposed by Bernal, we would see the strength of an argument that combines in a clear and methodical way premises both metaphysical and scientific, to successfully ground his physicalist thesis.

Let us begin with the bricks of the structure. Bernal follows the contemporary use of the term “physicalism”; it refers to “a *metaphysical* position that says that all the entities inhabiting the actual world, their properties, and all the facts and events involving them, have a physical nature” (19). Now, he says that “[a]n item T has a physical nature if it is a physical item, or it supervenes on a set of physical items” (20), where “an item” is “any kind of entity, property,

fact, event, or law ruling phenomena. Any element of any ontology is subsumed under the category of ‘items’” (20).

Now, the author says that “[i]ntuitively, we may consider the physical items to be the subject matter of physics” (20). This characterisation may lead to what has been dubbed “Hempel’s Dilemma” (1969). In the words of G. Hellman:

[...] either physicalist principles are based in actual physics, and thus it is the case that there are plenty reasons to believe that they are false; or else they are not, and thus in the best case it is difficult to interpret them, given that they are based on “physics” that does not exist (1985, 609)

Even though Bernal states that he will not “attempt to provide a positive characterisation of what it is for an entity or property to be physical [...] [and that he expects that] our intuitive conception of ‘the physical’ will do for present purposes” (21), I think it is fair and relatively important to align his position with the one that D. Papineau presents in several works. He believes that, beyond the characterisations that appeal to actual or future physics, there are different approximations that avoid the problem and are useful to present positively the questions under discussion. In particular, Papineau (2008) favours the idea of defining the “physical” as “anything that can be directly identified without using mental or biological terminology”, opening thus the possibility, as he says himself, for “some parts of this physical kingdom to be identified *also* with mental or biological terms”. This position can be extended coherently to the idea that “the ‘physical’ can be understood as equivalent to something ‘that exhibits a mathematically simple and precise behaviour’”.

Let us continue with the hearth of the proposal. Following D. Chalmers, Bernal characterises supervenience as follows: “An item U supervenes on a set of items Pi if U is entailed by {Pi} with *metaphysical* necessity and  $U \notin \{Pi\}$ ” (109). But several philosophers have claimed that there are important differences between the relations of implication and supervenience, since the former is neither sufficient nor necessary for the latter (*cf.* McLaughlin & Bennett 2010). That is why the very idea of supervenience was historically introduced by anti-reductionist philosophers. As J. Kim (1994) says, “the principal appeal of supervenience for physicalists has been to deliver dependence without reduction” (578). But if we have entailment among two sets of properties, we would thereby have reduction, since the former set is entailed by the latter. Nevertheless, we can accept Bernal’s characterisation and follow its consequences.

Bernal supports a realist view about conscious properties, i.e., the idea that these properties have an irreducible place in the world, which entails that they have causal powers that are novel and irreducible. In particular, it entails that consciousness cannot be derived from (or be reduced to) microphysical properties, thereby denying microphysicalism. Now, at this stage of his argument, I think there is an ambiguity that enables two different understandings of the author’s thesis.

1. On the one hand, Bernal claims that the difference between microphysicalism and his view, the *emergence physicalism (e-physicalism)*, concerns *the modal force that is involved*. He claims that the former is the idea that “[e]very HP [high-level property]<sup>4</sup> is entailed by fundamental items with *metaphysical* necessity”, while for the latter “[s]ome HPs [particularly consciousness] are entailed by fundamental items with *nomological* necessity” (131). We saw that Bernal characterizes supervenience in terms of *metaphysical necessity*, and defines emergence in terms of *nomological necessity* (128). Thus, his emergentist stance leads him to assume that consciousness *emerges* from fundamental items, because it is entailed by these, but does not *supervene* on these, because it is not entailed with metaphysical necessity.

Now, how can we understand the difference between both types of modal force? We could say that if two types of properties are related with *metaphysical* necessity, this relation holds in every possible world,<sup>5</sup> and if two types of properties are related only with *nomological*

<sup>4</sup> That is, a property that can be instantiated only by non-fundamental entities.

<sup>5</sup> Even though Bernal rejects the method of modal evaluation for possible worlds, I will use it for simplicity. I assume that this difference is harmless for my analysis.

necessity, this relation holds exclusively in worlds governed by the same laws of nature that the actual world. But there is here a problematic ambiguity that Bernal does not discuss. The microphysicalist would say that the laws of nature that can be involved in the nomological necessity relation are the ones governing microphysical entities—of course, since for him these are all the laws that exist. But then, notice that this is precisely the emergentist position of Bernal! Consciousness “is implied by fundamental items with *nomological* necessity”, where “item” stands for “any kind of entity, property, fact, event, or law that governs phenomena”. Therefore, it seems that Bernal’s proposal permits to get rid of consciousness once we have all the fundamental items, especially particles, their properties, and the laws governing them. This would be so, simply, because the former are entailed by the latter. To say that they are entailed with nomological necessity only means that this entailment is exclusively valid in the worlds governed by our microphysical laws or, in other words, that once we know the microphysical laws of our world, we know as well all the other facts, and in particular, the facts about consciousness. It is not necessary to add anything else.

2. On the other hand, this interpretation of Bernal’s thesis does not match what he really has in mind. I will take the following claim as the centre of his emergentist thesis: “there are emergent properties in the sense of *e-PHc* [emergence physicalism], i.e., HPs [high-level properties] that cannot be instantiated without the mediation of laws that are not fundamental (133), that is, laws of emergence that do not apply to fundamental entities *and cannot be derived from their microphysical laws*. But if the laws of emergence are not derivable from the microphysical laws, then we have in fact two different meanings for *nomological necessity*: a) one when the relation between two types of properties has *micro-nomological* necessity: it holds in every possible world governed by the same microphysical or fundamental laws that our world; and b) one when the relation between two types of properties holds in every possible world governed by *all the natural laws* that govern our world, whether microphysical or emergent (“special” in J. Fodor’s sense).

Bernal develops in detail two scientific examples concerning the physical and chemical level, to show that *the current scientific theories are committed* to the existence of macroscopic properties that are not derivable exclusively from microphysical properties and their nomological relations; they are derivable from these items only if we add macroscopic properties and the corresponding laws. The first example is about the putative reduction of the macroscopic property of a gas having a temperature, in terms of the microphysical properties (kinetic energy) that have the molecules that constitute gases. The second example, is an attempt of Bernal to show that quantum mechanics must suppose that some systems of particles have non-local and holistic properties that, accordingly, cannot be derived from (or reduced to) the intrinsic properties of the constituting particles. But notice that, if Bernal is right concerning emergentism, e.g., in the first case, it is not enough, in order to state that the macro-property of temperature is emergent, to claim that “the relation between the micro-level of the molecules of a gas and the macroscopic property of temperature is nomologically mediated” (133). We must add that this nomological relation goes beyond the microphysical level and includes macroscopic or emergent natural laws.

If the considerations above are right, then we should characterize emergence physicalism as follows: there are high-level properties (like consciousness) which are entailed with nomological necessity by fundamental items accompanied exclusively by laws and emergent properties that relate these fundamental items with the emergent ones. If I understand correctly, this is the idea that Bernal has about emergence laws (153): the ones that Broad (1923), one of the classical British Emergentists, dubbed “trans-ordinal laws”. In any case, I believe that the fundamental virtue of Bernal’s book is to show in a clear and detailed manner that nowadays science should accept an emergentist ontology and that, contrary to the widespread opinion, the advancements of microphysics and quantum mechanics do not imply reductionist metaphysics. It is clear for the author and for the theoreticians that discuss these problems that, because the question depends on the scientific developments and findings, it is an empirical point that can be falsified at any moment.

I would like to finish simply by introducing a question that has been fundamental for the non-reductionist theoreticians of the last decades: the question of the multiple realisation of

mental properties. H. Putnam and J. Fodor are well known for having developed, in the sixties and seventies, arguments against what was considered at that time the official view concerning the mind-body problem, i.e., the identity theory, which states that mental properties are in fact physical properties, and more specifically, neuronal properties. In principle, Bernal's theory is compatible with properties physicalism because consciousness is a physical property and, moreover, it can be argued that, given this, mental properties can be locally reduced to their realization bases by functionalizing them (this is the method developed by K. Lewis (1980) and J. Kim (2010), among others).<sup>6</sup>

### Reply by the author

First, J. D. Morales correctly notes that supervenience was introduced by anti-reductionist philosophers. He finds thus the definition I propose of supervenience problematic, because in this definition the supervenience base *entails* (with metaphysical necessity) what supervenes and then, according to Morales, the supervenient could be reduced in terms of the supervenience base.

I consider that reduction can be conceived in two ways, one epistemological and the other ontological. The kind of supervenience I propose certainly involves an ontological reduction of the non-physical in terms of the physical. That is why e-physicalism is indeed a form of physicalism; recall that it claims that everything has a "physical nature" (is physical or supervenes on the physical). But the conception of supervenience I propose does not imply at all an epistemological reductionism. I explicitly reject, in Chapters 1 and 4, the possibility of reducing the *categories* belonging to an ontology of the supervenient in terms of physical categories. In Davidson's anomalous monism every mental state token is identical to a brain state token, but mental categories cannot be reduced in terms of scientific categories like the ones to be found in the neurosciences. Similarly, I consider that every non-physical entity necessarily supervenes on a given set of physical items, but that this does *not* imply that it is possible to give necessary and sufficient conditions in the language of physics to capture every supervenient entity. Indeed, note that the definition I propose of supervenience does not relate categories or sets. It relates an individual item (the supervenient) with a set of individual items (the supervenience base): "An item U supervenes on a set of items Pi if U is entailed by {Pi} with *metaphysical* necessity, and  $U \notin \{Pi\}$ ".

Secondly, Morales notes the existence of an ambiguity concerning the distinction between metaphysical necessity and nomological necessity, which is crucial in the argument for emergence that denies microphysicalism. Surely, the definition of the emergence relation I propose involves a notion of nomological necessity that is put in contrast with one of metaphysical necessity: "An item U *emerges* on a set of items Pi if U is entailed by {Pi} with *nomological* necessity."

Probably, the ambiguity can be solved by focusing on a detail concerning the application of the definitions of supervenience and emergence for the characterization of microphysicalism and emergentism. I define microphysicalism as the thesis that every item that is not *fundamental*, including complex physical entities and their systemic properties, supervenes on a set of items *exclusively* constituted by fundamental items. The fundamental items are: fundamental entities (i.e. indivisible entities), fundamental properties (i.e., *intrinsic* properties that get instantiated in fundamental entities), fundamental laws (i.e., laws that *directly* relate fundamental entities in virtue of their intrinsic properties), and the facts or events necessary to fully determine boundary conditions. By contrast, e-physicalism claims that *some* items belonging to our reality do not supervene on sets of fundamental items, but *emerge* from them in virtue of some laws (of emergence) that are not fundamental. Now, if we add to the emergence base {Pi} of an item U the corresponding emergence laws, we obtain a new set

<sup>6</sup> Bernal develops arguments against a functional characterization of consciousness in terms of Strong AI and computationalism. There is another sense that is relevant here: the Ramsey-Lewis functionalisation method, which depends on the causal powers of the properties to be functionalized, which are powers that Bernal claims to be necessary (*cf.* Kim 2010).

{Pi}\* that does constitute a *supervenience* base for U. However, this new base {Pi}\* is not exclusively constituted by fundamental items and hence does not correspond to the supervenience base set involved in the definition of microphysicalism.

Now, as Morales correctly notes, given that the relation of emergence involves nomological necessity it would not hold in any possible world but only in the ones that share with our world its natural laws, *including the laws of emergence*. In fact, the argument I propose against Chalmers' "zombie argument" is based precisely on the possibility there is for two worlds to share everything related to fundamental items and yet differ regarding the laws of emergence. I develop this argument in *Chapter 5*.

Finally, Morales claims that to sacrifice the idea of multirealisability proposed by functionalism is to abandon a valuable intuition for our conception of the mind. In fact, many philosophers consider, and I share this opinion, that in principle it should be possible for some systems to instantiate minds despite significant physical differences. Moreover, it should be possible for them to occupy mental states of the same type and with the same phenomenal character in spite of their physical differences.

Surely, in *Chapter 2* I reject every form of functionalism about consciousness and, therefore, the corresponding multirealisability theses. However, I believe that e-physicalism does not sacrifice the valuable intuition underlying functionalism. There is nothing in my view that prevents a given emergent property of emerging from *different* emergence bases. Indeed, this is the case with temperature: gases that differ in their molecular constitution share a macroscopic property (temperature). That is why, despite rejecting Strong AI, I leave open the possibility of Weak AI.

**Commentary by David Papineau**  
King's College London

In my view this is an excellent book. It outlines an original and cogent view with clarity and in great detail. It also covers a great deal of ground and says intelligent things about a wide range of topics without getting bogged down in the kind of irrelevant technical detail that occupies much writing in this area.

At the centre of Bernal's argument is an intriguing suggestion about the metaphysics of the mind. Bernal defends a position he calls 'emergentist physicalism', according to which the conscious facts are not metaphysically determined by the fundamental physical (or 'microphysical') facts.

At first sight it might seem surprising that Bernal calls this a species of 'physicalism'. After all, determination by the microphysical facts is normally considered to be definitive of physicalism. But Bernal argues that anything that causally interacts with the physical realm should itself be counted as physical, and it is certainly part of his view that conscious facts causally interact with physical facts.

There are further complexities to Bernal's position. He does not in fact assume that anything that metaphysically supervenes on the physical (or even the microphysical) realm thereby qualifies as physical. Thus he holds that economic facts and institutions (e.g. money and banks) supervene on the physical (and so have a 'physical nature') but are not themselves physical.

How so? Surely money and banks can have physical effects (e.g. deforestation), and so why don't they count as physical by Bernal's basic criterion of physicality? But he explains that such merely supervenient facts have the wrong kind of causal powers to qualify as physical. The causal powers of money and banks derive from the causal powers of their physical realisers. To qualify as physical in virtue of causing physical effects, these effects must be produced by original causal powers (in the way that Bernal takes conscious states to produce their effects) and not by derived causal powers (of the kind possessed by money and banks).

So this then makes Bernal's position clear. Conscious facts do not supervene on microphysical facts. Rather the microphysical facts generate the conscious facts in virtue of synchronic contingent laws. Nevertheless, despite this metaphysical independence, the

conscious facts qualify as physical because of their original causal powers to produce novel physical effects.

Some might object to Bernal's position on the grounds that there is strong reason to believe in the microphysicalist thesis that everything is metaphysically determined by the microphysical facts. But Bernal argues convincingly that this microphysicalist thesis does not even hold within physics. He cites the example of temperature and quantum-mechanical non-locality. Temperature is perhaps not an entirely convincing case, but Bernal drives his point home with a lengthy and highly compelling demonstration that entangled quantum states do not supervene on their local microphysical components.

If I had to criticise Bernal's position, it would be that it is hostage to some strong empirical predictions. It is one thing to posit that there are emergent properties like temperature or quantum entanglements which are ubiquitous throughout nature. It is another to posit emergent properties with causal powers that are found specifically inside the brains of sentient organisms. This hypothesis carries the implication that there are physical effects to be found within such brains that can only be predicted on the basis of laws that apply inside brains. This is a striking prediction, but not an outlandish one. It is open to Bernal to respond that in principle there are indeed such novel effects to be found inside the brains of sentient beings, but that there is no practical possibility of empirically demonstrating their existence. In any case, it is scarcely a demerit in a philosophical position that it take a definite stance on important matters.

### **Reply by the author**

D. Papineau is perfectly right when he notes that the physicalist metaphysics I propose should be developed further concerning what I call "non-physical" items (which supervene on the physical). All the items that e.g. Searle (1995) calls constituents of "social reality" belong to this kind. I did say that once every physical item of our world is determined, every thing belonging to the realms of biology, psychology, and to the social real in general, is thereby determined as well in a unique and necessary manner. However, the question of how different types of items supervene (if they do) on bases exclusively composed by physical items should be analysed in greater detail. Indeed, some arguments against physicalism (e.g. Robinson in Stich & Warfield 2003) point precisely to the alleged impossibility of establishing these relations of supervenience in a coherent manner. They claim that it is not possible, simultaneously, to privilege in our world view the ontology of physics (or of ideal physics) and to accept the existence of entities belonging to categories that cannot be described in physical terms.

Papineau also points out that e-physicalism has an empirical prediction among its consequences: there is a set of physical items which, together with emergence laws, gives rise to the instantiation of the consciousness property. This property, in principle, should be accurately describable with the terminology of the natural sciences, and its instantiation should be able to be established or predicted following empirical criteria.

Indeed, e-physicalism opens, *in principle*, the possibility of a scientific theory of consciousness. If we knew what characteristics a system must comply with in order to instantiate consciousness, we would be able to determine, from the third-person point of view, if a given entity has or not consciousness and what is the character of a corresponding experience. However, the phenomenon of consciousness, given the subjectivity of experience, faces important epistemological obstacles (unsurpassable according to McGinn 1989): it is not clear if it possible to empirically confirm a theory that proposes empirical criteria for the attribution of consciousness. The thoughtful discussion—within the theoretical framework of e-physicalism—of these epistemological problems about consciousness is still pending.

***Commentary by Jaime Ramos Arenas***  
Universidad Nacional de Colombia

Reinaldo's book on phenomenal consciousness is solid and well structured. He did a fine job at extensively reviewing all the relevant specialised literature. He presents his ideas clearly, the structure of the arguments is sound, and he is able to build a plausible version of physicalism.

Although I am not persuaded by his proposal, and most of my comments below are critical of this project, this must be understood as a philosophical exercise which does not seek to diminish its academic or theoretical value.

Reinaldo undertakes a very difficult task; perhaps an impossible one: to defend a realist stand with respect to phenomenal consciousness, i.e., to defend that *qualia* are real, and at the same time to defend physicalism. To do that, he constructs what may be a new brand of physicalism that he calls “e-physicalism”. This is a theory that must be distinguished from microphysicalism (the thesis that all reality can be reduced to microphysical entities), given that it is a version of emergentism. Consciousness is conceived as a physical emergent property of some complex physical systems. Reinaldo skilfully uses his background in physics to defend that even in well known fields such as thermodynamics and quantum mechanics, we can find cases of emergent properties. Since I am not a specialist in these areas I can hardly evaluate those arguments, but they look persuasive to me.

Notwithstanding his dear attempt to reconcile physicalism with the reality of phenomenal consciousness, the thesis still appears quite problematic to me. In what sense is phenomenal consciousness physical? Reinaldo claims explicitly: “For e-physicalism consciousness is not a correlate or something caused by brain activity. It is a physical property instantiated in physical bodies while they are in some dynamical state” (159). What does it mean to say that a property is (or is not) physical? Perhaps one thing one may mean is that it is a property of physical entities. A property would be non-physical if it is a property of a non-physical entity. Another possible meaning for “physical property” would be a property definable in the language of physics. It seems that it is the first notion the one Reinaldo has in mind. He wants to avert the spectre of dualism, and since what can have conscious experiences are physical systems he infers in turn that conscious experiences must be physical events. The possibility of non-physical properties (of physical events) is unacceptable for him (one may wonder, though, if a real emergentist should be so scared by the same ghosts that torment physicalist reductionists). At any rate, it seems to me that his thesis that phenomenal experiences are identical to physical processes is undermined by his deviant notions of “experience” and “experiencer”. The experiencer is defined as: “Any minimal physical system that instantiates (a single token) of the c-consciousness property is an experiencer” (85). By “instantiating one token of the c-consciousness property” is understood having a conscious experience. Reinaldo explicitly denies that the experiencer is a person “or something of the sort” (84-85). Then he claims that some subpersonal system, such as some part of the nervous system, is the experiencer. But that is at least misleading, since “experience” is an intentional term. Of course, we need a functioning nervous system to have sensations, but the nervous system itself does not have sensations. What is the minimal physical system that sustains an experiencer? I would say the body of an animal complex enough to have sensations. But it is the animal, not its nervous systems that has sensations. The problem is one of individuation. An organ qua organ is always part of a body, and we cannot attribute to the organ what can be said only of the individual as a whole. Reinaldo says he subscribes to a holistic view of the mind, but his approach seems rather atomistic to me.

Reinaldo’s physicalism leads him to reject supervenience. Qualitative mental states should not be conceived as supervenient on the physical states of a system; they are identical to higher physical states of the system. Reinaldo correctly avoids a mistaken causal account between the neurological and the experiences. It is not the case that mental states supervene on or are caused by physical processes; they are identical to physical processes. At any rate, the question remains why *qualia*, if they are physical properties, are not describable in the language of physics. Why can’t we use physical concepts to capture them? Reinaldo’s response is that they are non-conceptual (they are too fine grained) and hence not accessible by the relatively thick concepts of science. I doubt, however, the truth of the claim that everyday language cannot capture the fine grained discriminations that we have in experience (see what McDowell says about this in *Mind and World*), or that it is not possible to capture them with a mathematical language we should develop with this purpose (consider the great success of the digitalisation of information in contemporary technology). In any case, notice that we do use a public language and a set of concepts to describe pains, smells, colours, and so on, although those concepts are

alien to physics. Those concepts, which are public (not part of some impossible private language), describe accurately phenomenal properties. They describe the way things appear to the members of a community, that is, the manner in which all of us we normally experience objects. Thus, I believe that Reinaldo's book has a problem that indeed is shared by most of the literature about experience, but here I cannot argue in detail for this opinion. The problem is the lack of a resolute acknowledgement of the fact that the subject of experience is always an individual situated in a historical and social context that brings a background of meaning and the very condition of possibility of experience itself (indeed, the content of experience changes as the subject evolves in a social context; thus, when we grow up, beer does not taste any more to us as a disgusting bitter but instead as an enjoyable bitter). However, to be sure, it is very common to conceive phenomenal experience, implicitly, as the experience of a pure, preconceptual, and private "object", as did Russell in the sense data theory that was persuasively criticised by W. Sellars.

To finish, I just want to introduce a difficult question that Reinaldo's book, and emergentism in general, does not address in deep. The putative emergence of causal processes and related events in complex systems is purely ontological or partly conceptual? In other words, when we say that during the evolution of a complex whole new events with new properties do emerge (for instance *qualia* in qualitative experience), are we using the concepts in the same sense that when we talk of the parts of the system, or are we talking about causality and emergent properties in a categorically different language? If we describe with the language of neurophysiology the causal processes that take place in the organism when it ingests alcohol and we describe these same processes in a psychological intentional language saying "to drink beer causes to x this experience", are we using "causality" in the same sense? I do not want to suggest that it is just a way of talking; we have indeed qualitative experiences and I share the decisive rejection of eliminativism by Reinaldo. But the question here is about the relation between ontological and conceptual issues. The question is if, from a physical point of view, it is necessary for something new to emerge when beer starts to taste good to me (a physical process  $\Delta$  is always replaced by a process  $\delta$ ), or if from a psychological point of view, which is the only one that can capture experience *qua* experience (for reasons that we would have to discuss), there was such and such alteration.

Realists about subjective experience believe that the case of *qualia* is different from the case of the reality of pieces of art or of laws: the reality of the former is independent of the description and conceptualization we can make of them, but this conception brings them back to "the myth of the given" that Sellars criticised.

In my opinion, Reinaldo could explore later with greater detail what is being tried to be said when talking here of the "emergence" of a new reality. But—of course!—nobody can do everything in a single book.

### **Reply by the author**

J. Ramos points out, first, a problem that I have very present: the need of a definition of what it means for an entity to be "physical". I admit that physicalism has an urgent need to solve this problem, but in the book I do not undertake the difficult and extensive task of advancing a proposal. I just hope that the intuitive notion we have of the "physical", which we use to refer to the object of study of physical science, is clear enough for my purposes. Now, I also present some examples of non-physical entities (which have a "physical nature" according to e-physicalism), in order to explain further—by contrast—in what sense something is "physical".

Secondly, Ramos criticizes the notion of "experiencer" defined as a physical system having, notably, the property of consciousness. He rightly points out that we attribute awareness to people, animals or, eventually, to entities that have a "mind" and are *subjects* of experience. Certainly, we do not attribute consciousness to a nervous system or a part of it. We attribute it to organisms that are complex enough to be possible objects of this attribution.

However, since I take consciousness to be an *intrinsic* property of certain physical systems, these systems would instantiate it regardless of whether or not we attribute it to them, and regardless of which criteria mediate the attribution. I think there are entities such that "there

is something it is like” to be one of them—for example persons, and that there also entities such that “there is nothing it is like” to be one of them—for example a stone (assuming panpsychism is false). Now, if physicalism is true, the entities of the first type are physical systems, or supervene on the physical, and it is in virtue of their physical properties (or of the physical properties in their base of supervenience) that they are conscious.

To be sure, in the case of a human organism for example, it is not the whole body what instantiates consciousness. One can lose an arm, the kidney, or the eyes, without losing consciousness. Obviously, if the visual system of a human organism is in a completely dysfunctional state, this organism will not have visual experiences. To that extent, the *character* of its experience will be affected by the visual impairment. But the organism will not cease having experiences, whatever their character. The visual system, therefore, does not belong to the part of the organism that, somehow, instantiates consciousness. In neurology there are different proposals about which human biological system (and in which state it must be) is necessary and sufficient for consciousness. It is that system what I call “the experiencer” in the absence of more creativity for the invention of a label. The available empirical research suggests that the experiencer is to be found in the central nervous system, but without being localized in any specific region, and that it essentially involves the thalamocortical system.

Finally, Ramos criticises the theory I outline (in Chapter 6) about *qualia*. It happens that, according to e-physicalism, *qualia* are physical properties (or supervene on physical properties). From e-physicalism it follows, notably, that *qualia* are such. This is, undoubtedly, hard to accept. Influential arguments like “the knowledge argument” by Jackson (1986), the “conceivability argument” by Kripke (1972), and the “explanatory gap” by Levine (1983), show precisely the deep difficulties that physicalism faces when applied to *qualia*.

Some philosophers proposed the idea that the mental can appear under different “modes of presentation”. Under one mode, the mental appears as something physical; we observe brains, neurons, etc., which can be described with the language of natural sciences. But under another mode, the mental appears as a set of states and processes with intentional contents and/or phenomenological characters. In particular, the phenomenological characters (or “*qualia*”) are captured by “phenomenal concepts”. But then the question arises about how are related the different modes of presentation of the mental, and if a description corresponding to one mode can be reduced to a description corresponding to another mode. Ramos, as far as I understand, is convinced that you cannot make this reduction, since each mode of presentation is associated with a language categorically different from the others. Personally I fully agree with this statement, even though I do so based on considerations that Ramos does not fully share.

Now, if I understand correctly, Ramos goes further. In principle, it can be said that to accept the existence of different modes of presentation of the mental is just to accept that the mind can be *known* from different perspectives. The mind would exhibit different “aspects”, because of some epistemological particularities about it. But, for Ramos, the existence of different modes of presentation has a deep implication for the *ontology* of the mental. To accept that the mental has different aspects is to accept that any description of it that focuses on a single one will be unavoidably incomplete: it will leave out what is captured and can only be captured through approaches directed to the other aspects. In fact, the different aspects complement each other, and then all of them must be considered to fully describe the mental. Now, from the idea that no particular description of the mental should be favoured in order to fully capture it, Ramos arrives to the conclusion that no particular ontology underlying some description should be favoured either. But this is precisely what physicalism does: it privileges descriptions in terms of physical entities and their properties and thus claims that, in its ultimate essence, the mental has a physical nature. In short, for Ramos psychology and neuroscience are alternative approaches to the mind, and none of them should be privileged from the epistemological or from the ontological point of view.

I find the position of Ramos attractive and I deeply respect it. However, I believe that the philosophy of mind should explain how and why the different aspects of the mind interact with each other. Even though I agree that a psychological description cannot be reduced in terms of a neurological description, I also insist in the following point: there are strong

correlations between psychological and neurological events. For example, aspirin (which is a chemical product) soothes a headache.

Now, to explain the correlation between psychological and neurological events, I argue that the psychological has a physical nature. I could have argued for the contrary, proposing some form of idealism, or I could have claimed that the metaphysical nature of the mental is neither material nor immaterial. The main reason why I privilege the physical is that I observe that everything that exists in the world has a material aspect, while only some entities have, in addition to the physical, a psychological aspect. Thus, physicalism provides, in my view, the more austere ontology.

## Acknowledgements

I am sincerely thankful to Alejandro Rosas, who had the initiative that led to the realization of this *précis*; to *Ideas y Valores*; and to those who contributed by providing their valuable comments and questions: Jérôme Dokic, Pierre Jacob, Max Kistler, Juan Diego Morales, David Papineau and Jaime Ramos.

## References

- Bayne, T. *The Unity of Consciousness*. Oxford: Oxford University Press, 2010.
- Bennett, K. "Exclusion Again". *Being Reduced: New Essays on Reduction, Explanation, and Causation*, Kallestrup, J. & Hohwy, J. (eds.). New York: Oxford University Press, 2008. 280-305.
- Block, N. *Consciousness, function, and representation*. Cambridge, Mass.: MIT Press, 2007.
- Broad, C. D. *Scientific Thought*. New York: Humanities Press, 1923.
- Brentano, F. *Psychology from an Empirical Standpoint*. London: Routledge and Kegan Paul, 1874.
- Campbell, J. *Reference and Consciousness*. Oxford: Oxford University Press, 2002.
- Chalmers, D. *The conscious mind*. Oxford: Oxford University Press, 1996.
- Hellman, G. "Determination and Logical Truth", *Journal of Philosophy* 82/11 (1985): 607-616.
- Hempel, C. "Reduction: Ontological and Linguistic Facets". *Philosophy, Science, and Method: Essays in Honor of Ernest Nagel*. New York: St. Martins Press, 1969. 179-199.
- Jackson, F. "What Mary Didn't Know", *Journal of Philosophy* 83/5 (1986): 291-295.
- Kim, J. "Supervenience". *A Companion to the Philosophy of Mind*, Guttenplan, S. (ed.). Cambridge, Mass.: Blackwell Publishers, 1994.
- *Physicalism, or Something Near Enough*. USA: Princeton University Press, 2005.
- *Essays in the Metaphysics of Mind*. New York: Oxford University Press, 2010.
- Kripke, S. *Naming and necessity*. Cambridge, Mass.: Harvard University Press, 1972.
- Levine, J. "Materialism and Qualia: The Explanatory Gap", *Pacific Philosophical Quarterly* 64 (1983): 354-361.
- Lewis, D. "Mad Pain and Martian Pain". *Readings in Philosophy of Psychology*. 2 Vols., Block, N. (ed.). Cambridge, Mass: Harvard University Press, 1980. 216-222.
- Macdonald, C. & Macdonald, G. (eds.). *Emergence in Mind*. New York: Oxford University Press, 2010.
- McGinn, C. "Can we solve the mind-body problem?", *Mind* 98/361 (1989): 349-366.
- McLaughlin, B. & Bennett, K. (2011). Supervenience. The Stanford Encyclopedia of Philosophy. <http://plato.stanford.edu/entries/supervenience/>
- Nagel, T. "What is it like to be a bat?" *Philosophical Review* 83/4 (1974): 435-456.
- Papineau, D. (2008) "Must a Physicalist be a Microphysicalist?" *Being Reduced: New Essays on Reduction, Explanation, and Causation*, Kallestrup, J. & Hohwy, J. (eds.). New York: Oxford University Press, 2008. 126-148.
- Searle, J. (2002). *Consciousness and language*. Cambridge: Cambridge University Press.
- Shoemaker, S. *Physical realization*. New York: Oxford University Press, 2007.

- Stich, S. & Warfield, T. (eds.). *The Blackwell Guide to Philosophy of Mind*. Oxford/Melbourne/Berlin: Blackwell Publishing, 2003.
- Stoljar, D. "Distinctions in Distinction". *Being Reduced: New Essays on Reduction, Explanation, and Causation*, Kallestrup, J. & Hohwy, J. (eds.). New York: Oxford University Press, 2008. 263-279.