

# THE MIND-BODY PROBLEM: AN OVERVIEW OF PROPOSED SOLUTIONS

## Abstract

The Philosophy of Mind consists of problems concerning aspects and properties of the human mind. The most important of these problems is that of the relation between mind and body, or, more generally, between mental and physical phenomena. Usually referred to as the *mind-body problem*, this has been one of the fundamental problems in Philosophy since Descartes (1596-1650) and his critics introduced it four centuries ago. The mental seems, at first glance, completely different from the physical. Physical properties are public, i.e. equally observable by everyone, but mental properties are not. It can be deduced that someone feels pain by his behaviour, but only that person can feel it directly. Conscious mental events are private in the sense that the subject has privileged access to them that no one has for the physical. Conscious experiences, such as the smell of jasmine, are completely different from the configurations and movements, however complex, of particles, atoms and molecules, or the physical changes of cells and tissues. Despite this, conscious phenomena do not seem to arise out of nothing, but from physical-biological processes in the body, especially from neural processes in the brain. But how can physical-biological systems have states such as thoughts, fears and hopes?

## 1 Introduction

Naturalism, an increasingly widespread school in Philosophy, claims that everything that exists can be explained in physical and natural terms. At the same time, since Descartes, many perceive human consciousness as the most self-evident reality. Many authors have expressed that maintaining both claims simultaneously - naturalism and the existence of consciousness - is not easy. We are puzzled that the mental has such different properties from the physical because these differences defy the naturalistic stance. Thus, for example, McGinn writes: “Somehow, we feel, the water of the physical brain is turned into the wine of consciousness, but we draw a total blank on the nature of this conversion” [34]. We therefore assume that there is an unresolved tension, conflict, or problem.

## 2 Origin of the problem

René Descartes in the 17th century was able to reduce all that exists (except God himself) to two types of substance: the *res extensa* and the *res cogitans*. At that time this was a formidable philosophical achievement because for the first time the diversity of the known world could be simplified to this duality (second meditation of his *Metaphysical Meditations*). The *res extensa* is composed of the bodies, animate or not,

that occupy extension in space, while the *res cogitans* can be identified with the soul, mind or consciousness of thinking subjects. To arrive at this simplification, Descartes relied on the reduction to mechanistic explanations that at that time were beginning to describe much of the physiology and function of living beings. However, everything related to the human soul appeared to Descartes as irreducible to such explanations and constitutive of an independent and differentiated type of substance. One way to summarise the history of the mind-body problem is to say that in these four centuries, we have been unable to complete the reductive program. Simplifying the two substances to a single substance is, from a philosophical point of view, undeniably attractive.

Soon after the publication of Descartes' ideas, criticisms appeared concerning the way in which mind and body interact. This issue has been called *the problem of causal interaction*. Gassendi (1592-1655) was the first to point this out in 1641 and can be said to be the true creator of the mind-body problem, since Descartes had proposed a substantial dualism which he had not, in principle, seen as problematic. The main objection to the Cartesian proposal is that if the two substances are really distinct and independent, it is difficult to explain the interaction between them. However, the body and the mind seem to interact in such a way that, for example, blows received by the body are perceived with a subjective sensation of pain and, conversely, our will is translated into movements of the body.

### 3 Definition of the problem

Although I have defined the mind-body problem as the issue of the relation between the mental and the physical, the definition in the literature is multiple and not without some complexity. Sometimes the term 'problem of consciousness' is used as an equivalent, but depending on the author, the mind-body problem and the problem of consciousness may not be exactly the same. In fact, some authors speak of problems, in the plural, of consciousness. If we add to this Chalmers' hard problem [7] as a reformulation that might not coincide with any of the previous ones, we can see that the definition of the problem is not simple.

In recent decades, some have renamed the problem the 'mind-brain problem', as the capacity to produce thought is more specifically attributed to the human brain. When the focus is on the 'problem of consciousness', the brain and the body are seen as unproblematic physical realities, whereas the emergence of consciousness is. In 1996 David Chalmers renamed the problem the 'hard problem of consciousness'. [7]. The hard problem is the problem of experience, subjective experience, and the difficulty of explaining it in terms of physical events. For Chalmers, the ambiguity of the term 'consciousness' is often exploited by both philosophers and scientists writing on the subject. Thus, they declare that consciousness is tractable after outlining their own theory of consciousness, but for Chalmers the real hard problem is never really addressed. In any case, there are authors who consider the hard problem as a reformulation of the classical mind-body problem. The new 'hard problem' would be no more than an improved version of an old problem that appeared with Descartes and his critics in 1641, and the hard problem is hard simply because it is the mind-body problem. [54].

When speaking about problems, in the plural, of consciousness, these problems can be classified into three families [52]: i) the *descriptive* questions as What is consciousness? What are its principal features? And by what means can they be best discovered, described and modeled?; ii) the *explanatory* questions as How does consciousness of the relevant sort come to exist? Is it a primitive aspect of reality, and if not, how does (or

could) consciousness in the relevant respect arise from or be caused by nonconscious entities or processes?; and iii) the *functional* questions as Why does consciousness of the relevant sort exist? Does it have a function, and if so, what it is it? Does it act causally, and if so with sorts of effects? Does it make a difference to the operation of systems in which it is present?, and if so, why and how?

Finally, when we face the philosophical problem, mind or consciousness can be characterised either by focusing on its representational aspects, as Rosenthal [44] or Dennett [18] do, or on its experiential or phenomenal aspects, as Chalmers [7] or Nagel [37] do. The first characterisation is usually considered relatively more manageable in terms of cognitive explanations. Here I will therefore refer specifically to the second characterisation in order to directly address the perplexity it entails.

## 4 Proposed solutions

The following is a sample of the main solutions proposed and the main objections they raise. As an open problem, new solutions continue to be proposed today, and it is impossible to list them all. Nor is there space here to consider the replies and counter-replies to the objections raised to each solution. Moreover, I will focus on the proposed solutions to the philosophical problem, leaving aside the scientific theories of consciousness that have proliferated in recent years.

Solutions to the mind-body problem can be classified according to whether they maintain the existence of the mental and the physical separately, or whether they try to reduce all that exists to a single type of entity. The former are called dualisms and the latter monisms. Dualists assert that both the mental and the physical are real, and that neither can be reduced to the other. Therefore, mental phenomena would be, at least in some respect, non-physical. Monism, on the other hand, does not accept fundamental divisions. Some authors have focused on the nature of the problem itself, instead of looking for solutions. Thus, the so-called Mysterians [34, 37] attack physicalist positions and adopt an epistemic approach, arguing that the mind-body problem is currently unsolvable, and will perhaps always remain unsolvable for human beings. There are also proposed solutions that are not easily classifiable under one of the monism-dualism labels. Such is the case with panpsychism, emergentism, and functionalism.

### 4.1 Dualist solutions

Dualism is roughly the thesis that not everything is fundamentally physical, and things that are not fundamentally physical are fundamentally mental [9]. Dualist proposals place at least some aspects of consciousness outside the realm of the physical, but the specific forms of dualism differ in what those aspects are. They can be divided into substance dualisms and property dualisms.

#### 4.1.1 Substance dualism

Substance dualism involves the existence of non-physical minds or selves as entities. One of the problems posed by substance dualism is that of causal interaction, i.e. how the two different substances posited by the theory, the mental and the physical, can have an impact on each other. Thus, substance dualism is subdivided into three forms according to the directions in which the causal interaction between the mental and the physical takes place: interactionism, epiphenomenalism and parallelism.

**Interactionism** According to interactionism, the mental and the physical, although distinct substances, can interact in some way. For example, Descartes speculated on the idea that this interaction could take place through the pineal gland (*L'Homme*, 1664). More recently, Popper and Eccles have considered the self-conscious mind as an independent entity that reads active centres in the brain modules of the linking areas of the dominant hemisphere. The self-conscious mind would select according to its interests and integrate its selection to give the unity of conscious experience. It would also act back on the neural centres. Thus, the self-conscious mind would exercise a superior interpretative and controlling role over neural events by virtue of a bidirectional interaction [39].

The main objection to interactionism is that physical science has shown us that the physical world seems to be self-sufficient in explaining natural phenomena. It is the so-called *principle of physical causal closure* which, in particular, also seems to be applicable to a physical system such as the brain. Moreover, it is difficult to imagine how two different and ontologically independent substances could ever interact with each other. [50].

**Epiphenomenalism** According to epiphenomenalism, physical events are causal with respect to mental events. But the reverse is not true: the mental has no causal power over the physical. Epiphenomenalism tries to respect the causal closure of the physical world and proposes that the mental is an epiphenomenon, i.e. a non-reducible secondary phenomenon that accompanies the physical without influencing it [43].

There are two main objections to epiphenomenalism. The first is that epiphenomenalism seems to be incompatible with being aware that we are conscious, since for us to know that we have consciousness would have to produce some change in our brain. The second problem is that of the emergence of consciousness in biological evolution. If consciousness is epiphenomenal, then it has no effect on an organism's adaptive capacity and should not have been selected [2].

**Parallelism** Parallelism is the school of thought according to which the mental and physical realms function synchronously without the need for either to interact causally with the other. If epiphenomenalism respects the causal closure of the physical, parallelism preserves both the closure of the physical and the closure of the mental [42]. An example of a parallelist proposal is Malebranche's occasionalism, according to which the soul and the body do not act directly on each other, but it is God who produces a sensation in the soul when the body experiences it, and who gives the body a movement when the soul desires it [36]. Another example of parallelism is the pre-established harmony proposed by Leibniz according to which God arranged things from the beginning of creation so that both substances behave as if they were interacting, without the need for God's intervention for particular events [15].

The main objection to parallelism is that the theory otherwise requires belief in a deity who intervenes in physical and mental events or programmes them in advance. In fact Leibniz himself accuses Malebranche of *deus ex machina*, not realising that the same could be said of his solution [14].

#### 4.1.2 Property dualism

According to property dualism, there is only material substance, but it can instantiate two essentially different kinds of properties: physical properties and mental properties.

One of the advantages of this type of dualism is that by not positing an immaterial substance it avoids religious connotations. Another is that it seems to avoid the problem of mental causation; there is no interaction between two different kinds of things. Moreover, mental properties are accepted as real and distinct from physical properties. [53]. In this school of thought, it is generally accepted that mental characteristics are supervenient on physical characteristics. This means that no two events can be the same in the physical aspect but differ in the mental aspect, or, equivalently, that an organism cannot be altered in some mental aspect without being altered in some physical aspect [16]. On the contrary, there can be differences in the physical without changes in the mental, which allows for the multiple realisability of mental properties. The dualism of properties that accepts the supervenience of the mental in the physical is often called *non-reductive physicalism*, since phenomenal properties supervene on physical properties but cannot be reduced to them [27, 53]. In general, the dependence of the mental on the physical expressed by supervenience can be of many different types. As an extreme case, the mental could be identical to the physical [19].

The main objection to property dualism is that if we respect the closure of the physical domain, and causal exclusion, i.e., that no event can have more than one sufficient cause, mental properties have no causal efficacy. Therefore, the conclusion is that phenomenal properties that are irreducibly mental are also merely epiphenomenal, i.e. they have no causal effect on physical events [27].

## 4.2 Monisms 193

Dualisms become monisms when we reduce one of the two substances to the other. Thus, physicalism and idealism are the main forms of monism, although there are other options such as dual-aspect theories, neutral monism, and anomalous monism. An extreme case of physicalism is eliminativism.

### 4.2.1 Physicalism 198

Physicalism (or materialism) is, broadly speaking, the thesis that everything is fundamentally physical [8]. Physicalists claim that, despite appearances, mental states are only physical states. Physicalism offers a simple and unified view of the world, but seems to have difficulties in offering a satisfactory explanation of consciousness. These difficulties are expressed in the following well-known arguments against physicalism:

#### Objections to physicalism 204

**A) Nagel: External objective third-person and internal first-person subjective points of view** For Thomas Nagel [37], an organism has conscious mental states if and only if there is something that is like to be that organism. The physicalist view would require at least an idea of consciousness as the subjective character of experience. But the subjective character is not captured by reductive analyses of the mental. If physicalism is to be defended, phenomenological features must have a physical explanation. But when we examine their subjective character, it seems that such a result is impossible. The reason is that all subjective phenomena are essentially related to a single point of view, and it seems inevitable that an objective physical theory would abandon any particular point of view. To illustrate the

connection between subjectivity and point of view, and the divergence between  
subjective and objective conceptions, Nagel proposes the example of the bat. Its  
experiences have a specific subjective character that is beyond our ability to conceive.  
Even the subjective character of the experience of a person born deaf and blind and  
mine are mutually inaccessible. And that affects the mind-body problem because it  
does not seem possible that the character of subjective experiences can be revealed from  
the physical functioning of the organism as an objective fact that can be observed and  
understood externally. In physical science, the aim is to know the thing by eliminating  
the subjectivities of the scientist's particular point of view. But what would be left of  
what it was like to be a bat if the bat's point of view were eliminated? If the subjective  
character of experience is comprehensible only from one point of view, then any shift  
towards greater objectivity removes us from the real nature of the phenomenon [37].

**B) Jackson: Epistemologic argument** The limits of the objective point of view  
are also emphasised by Frank Jackson [22] who illustrates his epistemological argument  
or *knowledge argument* through the fictional character Mary. She is a scientist who  
investigates the world from a black and white room through a black and white television  
monitor. She specialises in the neurophysiology of vision and has all the physical  
information that can be obtained about what happens when we see colours. What will  
happen when Mary leaves her black and white room or is given a colour television  
monitor? It seems obvious that she will learn something about the world and our visual  
experience of it. But then it is inescapable that her prior knowledge was incomplete. He  
had all the physical information, but it seems that this was not enough, so physicalism  
is false. For Jackson, the strength of this argument lies in the fact that you can have all  
the physical information without having all the information you need to have [22].

**C) Levin: The explanatory gap** Joseph Levine provides an epistemological  
argument to prove that any attempt to find psychophysical laws leaves an explanatory  
gap. This is because there is no way to determine exactly which statements about such  
laws are true. Consider, for example, the case of heat and the motion of molecules.  
Everything that needs to be explained about heat is explained as the motion of  
molecules. So, it is logical to conclude that heat and the motion of molecules are the  
same thing. On the contrary, there is nothing that we can determine about the physical  
substrate that explains why a conscious experience has the qualitative character that it  
has. Or, to put it another way, understanding its physical or functional properties does  
not explain or make intelligible what that particular experience is. It is therefore  
conceivable that there is a physical substrate without the usually associated experience,  
and vice versa [30].

**D) McGinn: Cognitive closure** We have seen how Levine's explanatory gap  
asserts that there is a practical limit to our current explanatory capacities. Colin  
McGinn goes further by claiming that, given our human cognitive limits, we will never  
be able to bridge the gap [34]. For this, he introduces the idea of cognitive closure: A  
type of mind  $M$  is cognitively closed with respect to a property  $P$  (or theory  $T$ ) if and  
only if the concept formation procedures available to  $M$  cannot be extended to an  
understanding of  $P$  (or an understanding of  $T$ ). Human beings would be cognitively  
closed to a natural explanation of consciousness, since we would always be puzzled as to  
how any property we discover instantiated in the brain could give rise to consciousness.  
For McGinn there is no philosophical mind-body problem because there is possibly some  
property of the brain that naturally explains consciousness. But let us be cognitively  
closed to it. The philosophical problem about consciousness arises from the feeling that

we have to accept that nature contains miracles. But the sense of miracle comes from us, not from the world. In reality, there would be nothing mysterious about how the brain generates consciousness [34]. One possible reason why humans are conceptually unable to grasp the nature of the psychophysical link is the intrinsically spatial nature of both our human perceptual concepts and the scientific concepts we derive from them. The mental, by contrast, appears to us as intrinsically non-localised in physical space [33].

**E) Chalmers' zombies** David Chalmers argues that consciousness escapes any reductive explanation in physical terms. For this, he turns to phenomenal zombies, who are physically and functionally identical to human beings, but who lack experience. There is no phenomenal sensation for them. They have no conscious experience: everything is dark inside them. Although this is probably empirically impossible, Chalmers argues that he is describing a coherent situation. And if no internal contradiction can be revealed, then the zombie world is logically possible. His argument goes like this: according to physicalism, everything in our world is physical. Therefore, a world in which all physical facts are the same as those in our real world must contain everything that exists in our real world. In particular, conscious experience must exist in such a possible world. But we can conceive of a zombie world and imply that such a world is possible. Therefore, physicalism is false [12].

**F) Inverted spectrum** A conceivability argument against physicalism can be found without the need to establish the logical possibility of a zombie world. It is enough to establish the logical possibility of a world physically identical to ours in which the facts about conscious experience are different from those of our world. One could imagine, for example, that where I have a red experience, my physical twin in another possible world has a blue experience, and vice versa. The mere fact that a subjective experience in our world is different in a physically identical world would refute physicalism. To achieve such a reversal in the real world, we would have to rewire the neural processes in the right way. But as a logical possibility, it would be consistent for subjective experiences to be reversed as long as the physical structure remained the same. Nothing in known neuroscience suggests that one type of visual information processing should be accompanied by green experiences instead of blue ones [12].

#### 4.2.2 Eliminativism

In philosophy of mind, it is often claimed that one of the distinguishing characteristics of the mental from the physical is that the mental is *intentional*. The misnomer intentionality referring to mental states means that these states are always *about something*. In many cases, that 'something' is a proposition, i.e. the meaning of a declarative sentence such as 'my arm is broken'. In these cases, the fundamental units of thought are called *propositional attitudes*. Thus, the content of a propositional attitude is a proposition that can be true or false from the perspective of the subject. And the subject in turn can have different attitudes towards that proposition such as belief, desire or fear. An example of a propositional attitude is that the subject fears that her/his arm is broken.

This view of other minds (and our own) composed of propositional attitudes as units is called *folk psychology* (FP). FP is embedded in our common sense and constitutes the shared body of wisdom that allows us to explain and predict other people's behaviour, desires, beliefs, fears, intentions, perceptions, etc. However, *eliminativism* argues that FP is fundamentally false, i.e., that common sense misleads us about psychological

phenomena and that we will need future neuroscience to truly understand them [13].  
Eliminativism is an extreme physicalism that asserts that the neuroscience of the future  
will eventually be integrated into physical science by eliminating references to mind or  
consciousness. For eliminativists, intentionality and propositional attitudes are at the  
core of FP and what makes the mental seem so different from physical phenomena.  
Therefore, propositional attitudes are the target of their criticism.

The eliminativists argue that FP is an empirical science analogous to the old  
Aristotelian physics, which expressed our common sense of the physical. For  
eliminativists, intentionality would not be a mystery but a structural feature of FP.  
Thus, FP and mathematical physics are sciences whose only difference is the abstract  
entities they handle: numbers in the case of physics and propositions in the case of  
FP [13].

But, from the eliminativist point of view, as an empirical theory FP is false since, for  
example, conceiving learning as the manipulation and storage of propositional attitudes  
we would be unable to explain pre-linguistic learning. Nor does it have explanatory  
power for phenomena such as the nature and dynamics of mental illness, creative  
imagination, differences in intelligence between individuals, the nature and functions of  
sleep, the construction of three-dimensional visual images from two-dimensional stimuli,  
perceptual illusions or memory. In addition, the attribution of propositional attitudes  
has lost strength over the course of human history as we have moved from a generalised  
*animistic* approach to nature to one restricted to higher animals. FP as a research  
programme would be stagnant. Finally, materialistic neuroscience fits better than FP  
into the framework of natural history and the physical sciences. FP cannot be part of  
this framework because its intentional categories are not reducible to it [13].

A major objection to eliminative materialism is that it is self-refuting. If there really are  
no propositional attitudes such as beliefs, then the eliminativists' belief that there are  
supposedly no beliefs would not exist [41]. One can also refute eliminative materialism  
by arguing that FP is highly successful in predicting human behaviour. Its success could  
be compared to that of the natural sciences, and improves on that of most recent  
psychological and neurobiological theories. Moreover, FP not only predicts but also  
justifies, evaluates, praises and rationalises [28]. Finally, all eliminativist reasoning is  
based on FP being an empirical theory subject to refutation, but there is an alternative  
view that FP is more a simulation our mind makes of what the other would do with the  
beliefs and desires we think they have, i.e. a putting ourselves in the other's situation  
rather than a complete theory of mind [20].

### 4.2.3 Idealism

Idealists say that the physical can be reduced to the mental, since the supposed physical  
world is empirical and therefore a social construct created from shared subjective  
experiences. This school of thought has its classical example in Berkeley (1685-1753) for  
whom the objects of human knowledge are: a) ideas impressed on the senses, b) mental  
ideas, and c) ideas formed by composing and/or dividing others. In today's language,  
Berkeley's ideas would be equivalent to the contents of consciousness or mental objects  
in the broad sense. By the sense of sight I have the ideas of light and of colours with  
their various degrees and variations. A colour, a taste, a smell and a figure observed  
together are considered as a distinct thing, signified by a name: apple, stone, tree,  
book.... [4]. For Berkeley, to exist is to be perceived, and although it is possible to  
conceive of something existing other than in a mind that perceives it - for example, we  
can imagine trees in a park and nobody that perceives them - it would just be framing  
ideas in your mind that you call trees and framing the idea that no one perceives them.



This just goes to show that you have the power to form ideas in your mind. When we strive to conceive of the existence of external bodies, we are merely contemplating our own ideas. But the mind, which has no regard for itself, deludes itself into thinking that it can conceive, and does conceive, that bodies exist without being thought or without the mind [4].

Even Berkeley himself was aware of several objections to his idealism: for example, it makes real things no different from imaginary things. It also seems absurd to suppress natural causes and attribute everything to the immediate operation of the mind. We could no longer say that fire heats, or that water cools, but that the mind heats or cools. On the other hand, we have the persistence of objects: do things continue to exist when no one perceives them? Another objection is the distinction between error and truth: since we judge the reality of things by our senses, how does one distinguish error from truth in situations such as when one thinks that an oar is crooked because one end is underwater? Finally, it also seems difficult from idealism to explain the similarity of specific objects of perception: why do certain things seem the same to all of us? [4].

#### 4.2.4 Neutral monism

According to neutral monism ultimate reality is intrinsically neither mental nor physical but neutral. For neutral monists the difference between the physical and the psychological lies not in the object but in the direction of investigation. Thus, for example, colour is a physical object insofar as we attend to its dependence on the light source, its relation to other colours, to temperature, to space, and so on. However, when we look at its dependence on the retina, etc., it is a psychological object, a sensation. [32]. Ernst Mach (1838-1916), the father of modern neutral monism, calls neutral entities *events/sensations*. For him, reality consists of a viscous mass of events, which in some places (as in the ego) is more firmly coherent than in others [32]. The paradigm of neutral monism is represented by William James (1842-1910), for whom consciousness is a non-entity, a mere echo of the archaic concept of the soul. The only thing that exists is *pure experience* in each present instant [25]. In the case of perceptual knowledge, the perceived object and its perception are only two names for an indivisible fact: experience. The object is in the mind, and the mind is around the object. Experience is part of a wider world and its connections can be traced in different directions, which are known as the physical and the mental [24]. In the case of conceptual knowledge, two experiences are interrelated in the same subject, where the second piece is representative of the first in the practical sense of substituting it in various operations, sometimes physical and sometimes mental, that lead to its associates and results [25]. The last representative of classical neutral monism is Bertrand Russell (1872-1970) for whom the *sensation* we have when we see an object is simply that object. The object and our sensation when we perceive it are the same thing [45]. For Russell, all that physics gives us are certain equations which give abstract quantitative properties of their changes. The qualitative aspect of mental objects stems from the fact that they are but sensations which reveal their intrinsic character and which offer the most indubitable knowledge of the world. On the contrary, our knowledge of the physical world is purely abstract, since we know only certain logical features of its structure, but nothing of its intrinsic character. [46]. As to what changes, what it changes from and what it changes to, physics is silent on this point [47].

One objection to neutral monism is that we have no indication of what these neutral entities actually are. In some versions of neutral monism, neutral entities seem to have a mixture of physical characteristics - such as being located in space - and phenomenological characteristics - qualitative character. And this makes them appear

to be entities that are both physical and mental, rather than neither physical nor  
mental. As neutral entities little can be said about them, and to the extent that their  
qualities are described they appear to be either physical or mental. Moreover, their  
supposedly neutral elements can be interpreted as mental because the way in which  
physical objects are constructed from neutrals is reminiscent of Berkeley's subjective  
idealism. [39]. The fact that there are intrinsic properties that explain the phenomenal  
and extrinsic relationships that construct the physical can be seen as a metaphysical  
speculation that is difficult to prove [12]. Ordinary material objects must be  
constructed from sensations. However, neutral monism was never able to show the  
method of construction and produced no more than sketches of how it should proceed,  
but never a set of working plans [51]. Chalmers has objected that even if fundamental  
neutral entities had constitutively phenomenal qualities there need not necessarily be  
conscious experience of those qualities. He relies on the quality/awareness gap,  
analogous to the physical/awareness gap when attacking physicalism. No instantiation  
of qualities requires awareness of them. It is conceivable that all those qualities and  
properties are instantiated without any awareness of them. And this leads us to doubt  
that there is room for consciousness in a neutral universe [9].

#### 4.2.5 Anomalous monism

Mental events resist being explained by physical theory. How can this fact be reconciled  
with the causal role of mental events in the physical world? On the assumption that  
both the causal role and the anomaly of mental events are undeniable facts, Donald  
Davidson's aim as the creator of anomalous monism was to explain how they can be  
compatible with the physical world. [16]. Davidson formulates this apparent  
contradiction by considering three principles: (i) at least some mental events interact  
causally with physical events, (ii) where there is causality, there must be a law: events  
related as cause and effect fall under strict deterministic laws, and (iii) there are no  
strict deterministic laws on the basis of which mental events can be predicted and  
explained. How can the three principles be reconciled? Causally interacting mental  
events (first principle) must instantiate some property of strict law (second principle),  
but mental properties are not suitable for inclusion in strict laws (third principle).  
Therefore, mental events must instantiate some other property that is suitable for such  
inclusion, and this other property must be physical. Consequently, causally interacting  
mental events must be identical to physical events. The conclusion Davidson reached is  
that a distinction had to be made between *type identity* and *token identity*: although  
the class or type of mental events cannot be reduced to the class of neural events, each  
individual mental event - each case or token- is nevertheless identical to a physical  
event [16].

It is often objected to anomalous monism that the identity of two individual events is  
not compatible with the types or classes by which they are characterised being  
irreducibly different [29]. It has also been blamed on anomalous monism, which implies  
an absence of causal power of mental properties. If we assume that a given event, by  
virtue of its mental property, causes a physical event, the causal closure of the physical  
domain says that this physical event must also have a physical cause. We can consider  
the possibility that each of them is only a partial cause, and that the two together  
constitute a complete or sufficient cause. But this violates the principle of causal closure  
of the physical, since a complete causal story of how this physical event comes about is  
at least partially outside the physical realm. Could it be that the mental cause and the  
physical cause are sufficient? In that case, the physical effect is overdetermined.  
Moreover, the idea of overdetermination also seems to violate the principle of causal  
closure of the physical [26].

## 4.2.6 Dual-Aspect theories

457

According to dual-aspect monism, the single substance of the world has a mental (experiential, intentional) aspect, just as it has a physical aspect. Dual-aspect monism respects both the physical and mental dimensions of existence equally [49]. Some of the dual-aspect theories combine an epistemic dualism with an ontic monism that suggests an alternative to the conventional physicalist programme of naturalising the mind. [1]. The origin of this school of thought is usually traced to Spinoza (1677), for whom thinking substance and extended substance are one and the same substance, which is now understood under this attribute, now under that one [17]. In the 20th century, Julian Huxley defended monism against dualism on the basis of the progress of science and the theory of evolution. But the substance of which the world is made reveals material or mental properties depending on the point of view: when the world is seen from the outside we have matter and when it is seen from the inside we have mind [21].

458

459

460

461

462

463

464

465

466

467

468

469

But if there is an underlying reality that we can understand as either mental or physical, depending on the point of view from which we observe it, neutral monism and dual-aspect theory share a central claim: there is an underlying reality that is neither mental nor physical. If the dual-aspect theory insists that the two aspects are fundamental and irreducible to each other, we would fall into panpsychism (See 4.3.2). If not, it would be closely associated with emergentism (See 4.3.3). In each case, the most challenging criticisms would be those of neutral monism, panpsychism or emergentism, respectively [1, 49].

470

471

472

473

474

475

476

477

## 4.3 Beyond monisms and dualisms

478

### 4.3.1 Functionalism

479

According to functionalism, certain functional states are invariably correlated with mental states. Putnam, for example, takes as a model for the mind a probabilistic automaton in which a program specifies, for each state and set of inputs, the probability with which the automaton will transition to each possible subsequent state and produce some particular outcome [40]. In general, functionalism can be seen as an extension of behaviourism. In *behaviourism*, all mentalistic language was eliminated because of its inherent subjectivity and replaced by a language of mere behavioural dispositions that simply correlate inputs-stimuli and outputs-behaviours of the system. But the functional states to which functionalism refers are not mere behavioural dispositions, since they are specified in terms of their relations not only to inputs and outputs, but also to the state of the machine at the time. Thus, functionalism separates itself from behaviourism by including internal states as propositional dispositions, i.e. beliefs and desires.

480

481

482

483

484

485

486

487

488

489

490

491

Internal states:

492

-(i) can be considered representations and serve to explain the representational character of mental states,

493

494

-(ii) are not tied to any particular physical realisation since the same program can run on different types of hardware,

495

496

-(iii) can be fully described in terms of their relationships to input, output and themselves; and

497

498

-(iv) can be included in descriptions and predictions of the output-behaviour of a system.

499

For functionalism, the mind is explained by these internal functional states and not by a certain physico-chemical state of the brain or a behavioural disposition. In other words,

500

501

functionalism is opposed to physicalism and behaviourism, and functionalists put forward empirical reasons for this. Pain, for example, would not be a physical-chemical state of the brain but a functional state of the whole organism. The brain state corresponding to a sensation of pain would depend on the evolutionary details of each phylogenetic lineage of each species. However, its functional character could be independent of such details [40].

The best-known objection to functionalism is Searle's Chinese room argument [48]. Let's imagine that I am locked in a room. Following an instruction book in English, I am able to answer by means of Chinese symbols to questions asked by means of Chinese symbols about the script of a story that I do not know at all. I never understand any of the story or the questions and answers, but I am able to answer correctly because the instruction book only refers in English to the manipulation of Chinese symbols in order to answer according to the Chinese symbols of the questions. Evidently for me the symbols are simply meaningless pictograms. From the point of view outside the room, I seem to understand Chinese perfectly well. Searle thus dismantles the idea that following a set of syntactic rules can be equated with thinking. The problem of *qualia* is also objected to functionalism. Qualia are the qualitative components of conscious experiences, such as those we experience when we see the colour red, for example. Critics of functionalism argue that a system could be functionally equivalent to the human brain even with a total absence of qualia. To this end, Block proposed to imagine the individuals of the Chinese nation working together in a way that is functionally equivalent to that of a human brain [5].

### 4.3.2 Panpsychism

According to panpsychism, elemental entities have their own basic forms of conscious experience, and in the brains these conscious elemental entities somehow coalesce to constitute human and animal consciousness. Although panpsychism literally means that everything has a mind, in practice, panpsychists are not committed to the thesis that every inanimate object has a mind. For them, it is sufficient that some fundamental physical entities (e.g., quarks or photons) have mental states, i.e., conscious experiences. The best arguments for panpsychism are actually arguments against dualism and physicalism, its main alternatives. Panpsychism claims to have the virtues of both views and the vices of neither [9]. We have previously seen arguments against dualism such as the causal argument, and arguments against physicalism such as Chalmers' zombies and the epistemological argument. These same arguments would also support *panprotopsychism*: roughly speaking, the view that fundamental entities have certain special properties that are precursors of consciousness and that can collectively constitute consciousness in larger systems.

Emergent panpsychists argue that macro-experiencing is strongly emergent from micro-experiencing (See 4.3.3). However, *emergent panpsychism* inherits many of the problems of dualism. Its opposite, *constitutive panpsychism*, is the thesis that our macro-experience is based on the micro-experiences of our constituent elements. Intuitively, constitutive panpsychism holds that the micro-experiences somehow add up to produce the macro-experience. The less problematic constitutive panpsychism is the one that holds that there is an *a priori* linking of microphenomenal truths to macrophenomenal truths [9].

Another important variety of panpsychism is *Russellian panpsychism*, or a version of Russell's neutral monism, which holds that physics reveals the relational structure of matter but not its intrinsic nature. Russellian panpsychism is the view that some intrinsic properties are microphenomenal properties. Russellian panpsychism addresses

two metaphysical problems - what is the place of phenomenal properties in nature, and what are the intrinsic properties underlying physical structure? -And, in fact, he answers both at the same time: Fundamental phenomenal properties play fundamental microphysical roles and underlie fundamental microphysical structure. There is a non-Russellian panpsychism that claims that there are microphenomenal properties that do not play microphysical roles, but it would run into problems with mental causality. According to Chalmers, the least problematic version of panpsychism would be a Russellian constitutive panpsychism [9].

The main objection to panpsychism is *the problem of combination*. It is very difficult to make sense of the conscious micro-subjects of experience with their micro-experiences coming together to form a conscious macro-subject with its own macro-experience. For William James, for example, even if we group conscious experiences together, each will remain enclosed in its own 'skin', windowless, ignorant of what other experiences are and mean. In the same way, private minds do not agglomerate into a higher composite mind [23].

For Chalmers, the problem of combination is actually a set of seven different problems (i) *The anti-aggregation argument*, or that aggregates have no objective existence, but exist only for observers who perceive them as such; (ii) *The subject-summing argument*, or that the existence of a number of subjects with certain experiences does not necessitate the existence of a distinct subject, and, in particular, the existence of a number of micro-subjects does not necessitate the existence of a macro-subject; (iii) *A conceivability argument* that it is possible to conceive of zombies that are microphysically and microphenomenally the same as us but do not have our macrophenomenal experiences; (iv) *An epistemological argument*, assuming that inside her black and white room, Mary is told all the microphysical facts, and also learns all the microphenomenal facts (what is like to be a quark, a photon, etc.); (v) *The palette argument*, if Russellian panpsychism is true, we can only expect a handful of microqualities, corresponding to the handful of fundamental microphysical properties, but how can this limited palette of microqualities combine to give rise to the vast array of macroqualities? (vi) *The revelation argument*, or that the vast array of micro-experiences that supposedly constitute our macro-experience is not revealed to us in introspection; and (vii) *The structural mismatch argument*, or that the macro-phenomenal structure of consciousness seems quite different from the macro-physical structure of the brain, when constitutive Russellian panpsychism should demand that the structures be the same [10].

### 4.3.3 Emergentism

Already in classical Greece Aristotle claimed that the whole is greater than the sum of its parts, but it was John Stuart Mill who exploited the idea to propose the existence of heteropathic laws that do not comply with the principle of composition of causes [35]. It was finally his disciple George Henry Lewes who introduced the term *emergent* to refer to heteropathic effects [31]. *Prima facie*, emergence occurs when a complex system is observed to have properties or behaviours that its components do not have on their own, i.e. they only emerge when the parts interact as an overall complex system. In philosophy of mind, emergentism has been used to interpret the mental as an emergent property of the human brain, in which the components are clearly physical [6].

However, due to the vagueness of the *prima facie* definition of emergence, different interpretations arise. Two of them are paradigmatic: Strong and weak emergence. *Strong emergence* means that the causal power of the emergent property is irreducible to that of the micro-properties in which it supervenes (See Subsection 4.1.2). Strong emergence exerts its influence directly *downwards*, in contrast to the functioning of a

simple structural macro-property, whose causal influence is produced through the activity of its constituent micro-properties [38]. In contrast, *weak emergence* occurs when the macrostate can be derived from the microdynamics and external conditions by simulation [3]. From a philosophical point of view, strong emergence and weak emergence are diametrically opposed. Strong emergence, if it really exists, would presuppose the incompleteness of physicalism. In contrast, weak emergence itself supports physicalism by showing how all emergent phenomena are based on underlying laws [11].

Therefore, I will henceforth focus on strong emergentism since weak emergentism can be assimilated to physicalism. When the question arises as to whether there really is strong emergence in nature, the answer is usually that the best candidate for it is human consciousness. But an emergence beyond the weak implies that high-level facts and laws are not deducible from low-level laws. Simulations would be unable to deduce facts about some high-level phenomena. And this, in turn, implies an inability to deduce even all low-level facts from low-level laws, since if all low-level facts were derivable it would be possible to deduce high-level facts from them due to supervenience. Therefore, strong emergence implies incompleteness of physical laws even in the characterisation of low-level processes. This feature of strong emergence can be called top-down or downwards causality and means that the higher level is not only irreducible but also causally effective. A consequence of this is that low-level laws are incomplete as a guide for the evolution of both low-level and high-level processes in the world. It must be emphasised that the causal impact of a high-level phenomenon on low-level processes is not deducible even in principle from the initial conditions and the low-level laws [11].

Precisely, the main objection to strong emergence is related to the downwards causal powers of emergent properties. Kim's argument is based on three principles: (i) Emergent properties supervene on microphysical properties (See Subsection 4.1.2 for the definition of supervenience), (ii) emergent properties are neither reducible nor identical to microphysical properties, and (iii) mental properties have causal efficacy. If we add to these the principle of the closure of the physical domain (iv), and the principle of causal exclusion (v) according to which no event can have more than one sufficient cause, the conclusion is that all five principles cannot be true simultaneously, so we have to give up something. For Kim the only renounceable point is the causal power of emergent properties and the conclusion would be that if we use strong emergence to explain consciousness it would be an epiphenomenon [27].

## 5 Discussion

The relationship between the mental and the physical remains a fascinating mystery. As a philosophical problem it left us perplexed, and despite the many proposed solutions, the objections they all raise may make us doubt that we will ever solve the question. We have seen that each answer to the problem has consequences that seem unacceptable. Consciousness is a major challenge to science as it is known today. The objections raised to physicalism by Nagel, McGinn, Chalmers and others seem insurmountable on the near horizon.

To conclude, let us highlight two significant facts. The first is that the origin of the mind-body problem and the origin of the scientific method as we know it today coincide around the figure of Descartes. Descartes precisely developed the concept of Cartesian coordinates that allow us to characterise the positions of bodies in space, and this allows us to mathematise the conception of the physical as that which occupies an extension in space. It is also significant that McGinn's ideas about the insolubility of

the problem date from 1989, just before the neuroscientific revolution. What is surprising is that the enormous development of neuroscience and neuroimaging techniques since that year has not shed significant light on the philosophical problem associated with the phenomenon of consciousness. Thus, time seems to have proven right, so far, those who saw a conflict between the objective scientific view of reality and consciousness as a subjective experience.

## References

1. Harald Atmaspacher, *20th century variants of dual-aspect thinking*, *Mind and Matter* **12** (2014), no. 2, 245–288. 4.2.6
2. Andrew Bailey, *Zombies, epiphenomenalism, and physicalist theories of consciousness*, *Canadian Journal of Philosophy* **36** (2006), no. 4, 481–509. 4.1.1
3. Mark A Bedau, *Weak emergence*, *Philosophical perspectives* **11** (1997), 375–399. 4.3.3
4. G. Berkeley, *Principles of human knowledge*, Start Publishing LLC, 2013. 4.2.3
5. Ned Block, *Readings in philosophy of psychology*, vol. 1, Harvard University Press, 1980. 4.3.1
6. C.D. Broad, *The mind and its place in nature*, International library of philosophy, Routledge, 2008. 4.3.3
7. David Chalmers, *The hard problem of consciousness*, *The Blackwell companion to consciousness* (2007), 225–235. 3
8. David Chalmers, *Panpsychism and panprotopsychism*, *Amherst Lecture in Philosophy* **8** (2013). 4.2.1
9. David Chalmers, *Panpsychism and panprotopsychism*, *Consciousness in the Physical World. Perspectives on Russellian Monism*, Oxford University Press, 2015, pp. 246–276. 4.1, 4.2.4, 4.3.2
10. ———, *The combination problem for panpsychism*, *Panpsychism: contemporary perspectives* **179** (2017), 214. 4.3.2
11. David J Chalmers, *Strong and weak emergence*, *The re-emergence of emergence* (2006), 244–256. 4.3.3
12. D.J. Chalmers, *The conscious mind: in search of a fundamental theory*, *Philosophy of Mind Series*, Oxford University Press, USA, 1996. 4.2.1, 4.2.1, 4.2.4
13. Paul M Churchland, *Eliminative materialism and the propositional attitudes*, *The Journal of Philosophy* **78** (1981), no. 2, 67–90. 4.2.2
14. K.C. Clatterbaugh, *The causation debate in modern philosophy, 1637-1739*, *Philosophy, science studies*, Routledge, 1999. 4.1.1
15. E. Craig and Routledge (Firm), *Routledge encyclopedia of philosophy*, *Routledge Encyclopedia of Philosophy*, no. v. 5, Routledge, 1998. 4.1.1
16. D. Davidson and Oxford University Press, *Essays on actions and events: Philosophical essays volume 1*, *Essays*, Clarendon Press, 2001. 4.1.2, 4.2.5

17. B. de Spinoza, *Ethics*, Penguin classics, Penguin Publishing Group, 1996. 4.2.6 687
18. D. C. Dennett, *Consciousness explained. boston (little, brown and co) 1991.*, 1991. 688  
3 689
19. Robert Francescotti, *Supervenience and mind*, Internet Encyclopedia of 690  
Philosophy (James Fieser and Bradley Dowden, eds.), 2011. 4.1.2 691
20. ALVIN I GOLDMAN, *In defense of the simulation theory*, *Mind & Language* **7** 692  
(1992), no. 1-2, 104–119. 4.2.2 693
21. Julian Huxley et al., *Evolution. the modern synthesis.*, *Evolution. The Modern* 694  
*Synthesis.* (1942). 4.2.6 695
22. Frank Jackson, *Epiphenomenal Qualia*, *The Philosophical Quarterly* **32** (1982), 696  
no. 127, 127–136. 4.2.1 697
23. W. James, *The principles of psychology*, ATLA monograph preservation program, 698  
no. v. 1, H. Holt, 1890. 4.3.2 699
24. William James, *The knowing of things together.*, *Psychological Review* **2** (1895), 700  
no. 2, 105. 4.2.4 701
25. ———, *Does 'consciousness' exist?*, *The Journal of philosophy, psychology and* 702  
*scientific methods* **1** (1904), no. 18, 477–491. 4.2.4 703
26. Jaegwon Kim, *The myth of nonreductive materialism*, *Contemporary Materialism*, 704  
Routledge, 2002, pp. 149–165. 4.2.5 705
27. ———, *Physicalism, or something near enough*, *Physicalism, or Something Near* 706  
*Enough*, Princeton University Press, 2007. 4.1.2, 4.3.3 707
28. Ran Lahav, *The amazing predictive power of folk psychology*, (1992). 4.2.2 708
29. Drew Leder, *Troubles with token identity*, *Philosophical Studies* (1985), 79–94. 709  
4.2.5 710
30. Joseph Levine, *Materialism and qualia: The explanatory gap*, *Pacific* 711  
*philosophical quarterly* **64** (1983), no. 4, 354–361. 4.2.1 712
31. G.H. Lewes, *Problems of life and mind*, Osgood, 1875. 4.3.3 713
32. E. Mach, C.M. Williams, S. Waterlow, and T.S. Szasz, *The analysis of sensations,* 714  
*and the relation of the physical to the psychical*, Dover Books S 525, Dover 715  
Publications, 1959. 4.2.4 716
33. C McGinn, *Consciousness and space*, *Journal of Consciousness Studies* **2** (1995), 717  
no. 3, 220–230. 4.2.1 718
34. Colin McGinn, *Can we solve the mind–body problem?*, *Mind* **98** (1989), no. 391, 719  
349–366. 1, 4, 4.2.1 720
35. J.S. Mill, *A system of logic, ratiocinative and inductive: Being a connected view* 721  
*of the principles of evidence, and methods of scientific investigation*, John W. 722  
Parker, 1843. 4.3.3 723
36. S. Nadler, *Occasionalism: Causation among the cartesians*, OUP Oxford, 2010. 724  
4.1.1 725



37. Thomas Nagel, *What is it like to be a bat*, Readings in philosophy of psychology **1** (1974), 159–168. 3, 4, 4.2.1 726  
727
38. Timothy O'Connor, *Emergent properties*, American Philosophical Quarterly **31** (1994), no. 2, 91–104. 4.3.3 728  
729
39. K.R. Popper and J.C. Eccles, *The self and its brain*, Springer Berlin Heidelberg, 2012. 4.1.1, 4.2.4 730  
731
40. Hilary Putnam, *The nature of mental states*, Art, mind, and religion (1967), 37–48. 4.3.1 732  
733
41. Victor Reppert, *Eliminative materialism, cognitive suicide, and begging the question*, Metaphilosophy **23** (1992), no. 4, 378–392. 4.2.2 734  
735
42. Howard Robinson, *Dualism*, Blackwell Guide to Philosophy of Mind (Stephen P. Stich and Ted A. Warfield, eds.), Blackwell, 2002, pp. 85–101. 4.1.1 736  
737
43. W.S. Robinson, *Epiphenomenal mind: An integrated outlook on sensations, beliefs, and pleasure*, Routledge Studies in Contemporary Philosophy, Taylor & Francis, 2018. 4.1.1 738  
739  
740
44. David M. Rosenthal, *A theory of consciousness*, The Nature of Consciousness (Ned Block, Owen J. Flanagan, and Guven Guzeldere, eds.), MIT Press, 1997. 3 741  
742
45. B. Russell, *The analysis of mind*, Library of philosophy, G. Allen & Unwin, 1921. 4.2.4 743  
744
46. B. Russell and J. Slater, *An outline of philosophy*, Routledge, 1995. 4.2.4 745
47. B. Russell and A. Wood, *My philosophical development*, Simon and Schuster, 1959. 4.2.4 746  
747
48. John R Searle, *Minds, brains, and programs*, Behavioral and brain sciences **3** (1980), no. 3, 417–424. 4.3.1 748  
749
49. David Skrbina, *Dualism, dual-aspectism, and the mind*, Contemporary Dualism, Routledge, 2014, pp. 228–252. 4.2.6 750  
751
50. D. Stoljar, *Physicalism*, New Problems of Philosophy, Taylor & Francis, 2010. 4.1.1 752  
753
51. Robert E. Tully, *Russell's neutral monism*, The Cambridge Companion to Bertrand Russell, Cambridge University Press, 2003, pp. 332–70. 4.2.4 754  
755
52. Robert Van Gulick, *Consciousness*, The Stanford Encyclopedia of Philosophy (Edward N. Zalta, ed.), Metaphysics Research Lab, Stanford University, Winter 2021 ed., 2021. 3 756  
757  
758
53. Elly Vintiadis, *Property dualism*, ch. 20, p. 26, Rebus Community, 2019. 4.1.2 759
54. J. Westphal, *The mind-body problem*, The MIT Press Essential Knowledge series, MIT Press, 2016. 3 760  
761