

DOSSIER

# The colored-brain thesis

## A tese do encéfalo colorido

Osvaldo Pessoa Jr.<sup>1</sup>

### ABSTRACT

The “colored-brain thesis”, or strong qualitative physicalism, is discussed from historical and philosophical perspectives. This thesis was proposed by Thomas Case (1888), in a non-materialistic context, and is close to views explored by H. H. Price (1932) and E. Boring (1933). Using Mary’s room thought experiment, one can argue that physicalism implies qualitative physicalism. Qualitative physicalism involves three basic statements: (i) perceptual internalism, and realism of qualia; (ii) ontic physicalism, characterized as a description in space, time, and scale; and (iii) mind-brain identity thesis. In addition, (iv) structuralism in physics, and distinguishing the present version from that suggested by H. Feigl and S. Pepper, (v) realism of the physical description. The “neurosurgeon argument” is presented, as to why the greenness of a visually perceived avocado, which (according to this view) is present in the brain as a physical-chemical attribute, would not be seen as green by a neurosurgeon who opens the observer’s skull. This conception is compared with two close views, Russellian (and Schlickian) monisms and panprotopsychnism (including panqualityism). According to the strong qualitative physicalism presented here, the phenomenal experience of a quale  $q$  is identical to a physico-chemical quality  $q$ , which arises from a combination of (1) the materiality  $\omega$  associated with the brain, and (2) the causal organization or structure of the relevant elements of the brain  $\Sigma$ , including in this organization the structure of the self:  $(\Sigma\omega)_q$ . The “explanatory gap” between mental and physical states is shifted to a gap between the physico-chemical qualities  $q$  and the organized materiality of a specific brain region  $(\Sigma\omega)_q$ , and is seen as being bridged only by a set of non-explanatory postulates.

**Keywords:** Colored-brain thesis, qualitative physicalism, mind-brain identity thesis, qualia, panprotopsychnism, sensorium.

### RESUMO

Discute-se em um viés histórico e defende-se a “tese do encéfalo colorido”, ou fisicismo qualitativo forte. Esta tese foi proposta por Thomas Case (1888), em um contexto não materialista, e é próxima a visões exploradas por H. H. Price (1932) e E. Boring (1933). Usando o experimento mental do quarto de Mary, pode-se argumentar que fisicismo implica fisicismo qualitativo. O fisicismo qualitativo envolve três afirmações básicas: (i) internismo perceptivo e realidade dos qualia; (ii) fisicismo ôntico, caracterizado como uma descrição no espaço, no tempo e na escala; e (iii) identidade mente-encéfalo. Além disso, adiciona-se (iv) o estruturalismo na física e, distinguindo a presente versão daquelas sugeridas por H. Feigl e S. Pepper, (v) o realismo da descrição física. Apresenta-se o “argumento do neurocirurgião”, de porque a verdidão de um abacate percebido visualmente, que estaria presente no encéfalo como um atributo físicoquímico, não seria visto como verde por um

<sup>1</sup> Universidade de São Paulo  
– USP – Programa de Pós-  
Graduação em Filosofia. Av. Prof.  
Luciano Gualberto, 315. Cidade  
Universitária, 05508-010. São  
Paulo, SP, Brasil. Email: opessoa@  
usp.br.

neurocirurgia que abra o crânio do observador. Compara-se esta corrente com duas visões próximas, os monismos russellianos (e schlickianos) e o pamprotopsiquismo (incluindo o panqualitativo). Segundo o fisicismo qualitativo forte aqui apresentado, a vivência fenomênica de um quale  $q$  é idêntica a uma qualidade físico-química  $q$ , que surge de uma combinação da (1) materialidade  $\omega$  associada ao encéfalo e (2) da organização ou estrutura causal  $\Sigma$  dos elementos relevantes do encéfalo, incluindo nesta organização a estrutura do *self*:  $(\Sigma\omega)_q$ . A “lacuna explicativa” entre estados mentais e físicos migra para uma lacuna entre qualidades físico-químicas  $q$  e a materialidade organizada de uma região específica do encéfalo  $(\Sigma\omega)_q$ , e é vista como podendo ser coberta apenas por postulados não explicativos.

**Palavras-chave:** Tese do encéfalo colorido, fisicismo qualitativo, identidade mente-encéfalo, qualia, pamprotopsiquismo, sensório.

## 1. History of the view

The *colored-brain thesis* is the name given by Leopold Stubenberg (1998, p. 169) to the view that subjective phenomenal qualities, or qualia, are “properties of the brain”. Henry H. Price (1932) referred to this thesis as the “hypothesis that sense-data are cerebral”:

*To say that when a man looks at a tomato he is acquainted with a reddened portion of his own brain, or with a sounding tract of it when he hears a noise, is very singular. And others besides Bradley find it hard to believe that ‘when I smell a smell I am aware of the stinking state of my own nervous system’ (Price, 1932, p. 127).*

Price points out, in the above quotation, that the Hegelian philosopher Francis Herbert Bradley criticizes the theory proposed by the Oxford philosopher Thomas Case (1888), who characterized sense perception as the “the immediate apprehension of an internal physical object inside the nervous system of a sentient being” (Case, 1888, p. 33).

Case was led to this view by an application of the ancient principle of attraction of like to like (cf. Empedocles’ stanza in Aristotle, *De Anima*, Bk. I, Ch. 2, 404b 8): “The similar can be inferred only from the similar, therefore the physical can be inferred only from the physical” (Case, 1888, p. 23):

*If, then, natural science requires that the object of sense must be within my nervous system in order to be sensible, and logic that it must be physical in order to infer physical objects of science in the external world, how can the sensible object be at once physical and internal? I answer, it is the nervous system itself sensibly affected. The hot felt is the tactile nerves heated, the white seen is the optic nerves so coloured. (Case, 1888, p. 24).*

Case considers his “physical realism” to be a middle way between the direct or “intuitive” realism of the Scottish school (Thomas Reid, Dugald Stewart and William Hamilton) and the “hypothetical” realism of “cosmothetic idealism” (René Des-

cartes, John Locke and Samuel Clarke), which postulates an unknowable reality beyond the senses (in opposition to the absolute idealism of George Berkeley and Arthur Collier, which denies this reality) (compare this analysis with Hatfield’s, 2015, p. 111):

*Such is the realism proposed in this essay. It may be expressed in two propositions: there are physical objects of science in the external world; therefore there are, as data to infer them, physical objects of sense in the internal nervous system. It is the via media between intuitive realism and hypothetical realism of the cosmothetic idealist. As it recognizes physical realities, it is realism. As the objects, which it supposes to be sensible, are not external but internal, it is not intuitive realism. As the objects of sense, which it supposes to be the data of inferring an external physical world, are not psychical but physical, it is not hypothetical realism. As they are physical data within, to infer physical objects without, the realism which I advocate may be called Physical Realism (Case, 1888, p. 26).*

Case’s position, however, is not materialist or physicalist, first because he considers that God created and rules the world (Case, 1888, p. 20). Moreover, the internal object of sensation, which he postulates as being physical, is taken to be distinct from the “internal operation” that grasps it, which would be of the order of the “psychical”: “There is some plausibility in saying that the act of consciously touching is psychical, there is none at all in saying that the heat felt is psychical” (Case, 1888, p. 24).

It seems that the colored-brain thesis was not explicitly advocated by anyone else, although Price (1932) wrote that “philosophers have been accustomed to discuss the question whether sense-data are physical or mental” (Price, 1932, p. 127). In fact, in the interwar period, when sense-data theories and materialism coexisted in English speaking philosophy, the American psychologist and historian of psychology Edwin Boring came close to the colored-brain thesis, in his book *The physical dimensions of consciousness* (1933). It was this work that influenced U.T. Place (1956) to develop his version of the mind-brain identity thesis (Place, 2000). The mind-brain identity thesis has a history that is intertwined with materialism, and

was put forward by Boring in the following statement, quoted by Place: "To the author a perfect correlation is identity. Two events that always occur together at the same time in the same place, without any temporal or spatial differentiation at all, are not two events but the same event" (Boring, 1933, p. 16). Place next ponders on why Boring was ignored by philosophers:

*Boring moreover, was himself apparently committed to combining the identity theory with a phenomenalist account of sensory qualities which on Leibniz's principle of the Identity of Indiscernibles would commit him to the view that certain brain events are literally green, high pitched, warm, sour or putrid, which for a philosopher would constitute an immediate knockdown reductio ad absurdum of his position (Place, 2000, p. 1).*

We see therefore that Boring got close to the colored-brain thesis, in spite of not having mentioned it explicitly. Place had criticized this thesis in his seminal paper of 1956, in which he supported a reductionist materialism. In this work, there is not properly speaking a mind-brain identity thesis, but the thesis that the mind is "composed" of brain parts, and nothing else, which is a form of reductionism which comes close to eliminativism. The colored-brain thesis is mentioned and criticized in the following excerpt:

*This logical mistake, which I shall refer to as the 'phenomenological fallacy', is the mistake of supposing that when the subject describes his experience, when he describes how things look, sound, smell, taste or feel to him, he is describing the literal properties of objects and events on a peculiar sort of internal cinema or television screen, usually referred to in the modern psychological literature as the 'phenomenal field'. If we assume, for example, that when a subject reports a green after-image he is asserting the occurrence inside himself of an object which is literally green, it is clear that we have on our hands an entity for which there is no place in the world of physics. In the case of the green after-image there is no green object in the subject's environment corresponding to the description that he gives. Nor is there anything green in his brain; certainly there is nothing which could have emerged when he reported the appearance of the green after-image. Brain processes are not the sort of things to which colour concepts can be properly applied (Place, 1956, p. 49).*

## 2. The neurosurgeon argument

The usual reaction to the colored-brain thesis is to consider it absurd, as did Bradley, in the sense that it is an obvious

mistake to suppose that "to see red, there must be red neurons in the brain" (O'Regan and Noë, 2001, pp. 947, 1010, 1018).

However, the answer of strong qualitative physicalism to this objection is simple, involving a subtle Gestalt shift. The view associated with the colored-brain thesis is "internalist" in relation to colors and other qualia: the subjective greenness we experience as we look at an avocado is not in the fruit, but in our brain. The avocado is the *cause* of the greenness produced in us, and this cause is associated to the electronic properties of the pigment molecules in the skin of the avocado, that modulate the reflection of incident light.

Take a neurosurgeon, to be named Wilder, who has opened the skull and brain of patient Ullin and observes it in the appropriate part of the "sensorium" (the place where qualia allegedly exist). Suppose that Ullin contemplates an avocado and is having the subjective experience of greenness: it is clear that Wilder (looking only at his patient's brain) would not also have the subjective experience of greenness. In the words of Price, who imagined "if a living brain were cut up by a physiologist", "those colour-experiences would not be the ones which the owner of the brain was aware of, but would differ from them completely in shape, colour and position" (Price, 1932, p. 128).

The reason for this difference is simple: Ullin's brain in the state of greenness does not have the electronic properties (mentioned above) that might, after being illuminated, selectively absorb light and cause in Wilder the appearance of the quale of greenness. *Subjective color has nothing to do with light* (except for the meticulous causal connection between the two): our brains are dark (Dennett, 1992, p. 28).

The Gestalt shift involved is the change from common sense externalism ("greenness is in the avocado") to the internalism of views such as sense-data theory or the thesis of the reality of qualia ("greenness is in the mind"). Adding to this the mind-brain identity thesis, one arrives at: "greenness is in the brain". The idea that the colored-brain thesis would imply that Wilder would see green neurons inside Ullin's brain (or that the mad scientist who licks the patient's brain, while the latter is eating chocolate, would taste the chocolate - Nagel, 1987, p. 29) stems from an implicit adoption of the externalist view (Stubenberg, 1998, p. 173-74 misses this point, together with Dennett and Nagel).

The colored-brain thesis leads to the idea that the experienced greenness is a real physico-chemical *attribute* not yet recognized by contemporary physics, a real quality of matter. The term "attribute" is used in a generic sense, as it is not clear whether it would be a property, a relation, a process or another metaphysical category.

## 3. Other criticisms

The colored-brain thesis is naturally rejected by any dualistic view, such as that of John Stuart Mill (1843, bk. I, ch. III, § 4): "These are states of my body; but the sensation of blue, which is the consequence of these states of body, is not

a state of the body: that which perceives and is conscious is called Mind" (Mill, 1843, p. 68).

According to some of its critics, the colored-brain thesis is subject to the so-called "sense-data fallacy" of Harold Prichard and John Austin: "(...) the inference from the fact that one perceives a tomato that appears red to the fact that one perceives a red appearance" (Den Otter, 2005, p. 160). Considering that perception involves a causal chain that starts from things in the external world, passes through the sense organs, and enters the brain, one may consider (as is done in the causal-pluralist theory of observation, in Pessoa, 2019b) that one is observing either the external tomato, or a pattern on the retina, or the most proximal cause to perception in the sensorium (which would be the material part identical to the subjective perception of the tomato – what could be called the "red appearance"). Thus, there would be no fallacy. Another version of this same point (with which we also disagree) is the "phenomenological fallacy", mentioned above in the quote by Place (1956), which denies the reality of the green appearance experienced in an after-image (due to looking, for example, at a white wall after looking at a red tomato).

We have already quoted excerpts from Price (1932), who discusses the thesis that "sense-data are cerebral" over five pages (pp. 127-131). The Welsh philosopher comes to interpret the thesis with sympathy, provided that the brain is not considered just a material or physico-chemical system, but also "an organ of a living thing", so that the sense-data "are neither physical nor mental but *vital*, in the sense in which breathing and digesting are vital" (Price, 1932, p. 127).

The approximation between "psychosubstantialism" (or "quidism", the thesis that the mind involves qualities or inscrutable, which should be distinguished from mechanical working of the parts of the brain, as will be reviewed in section 6)) and vitalism (in a broad sense, the thesis that life is something beyond the mechanical working of the living parts) leads to a view that there are "vital qualities" that would distinguish for example a flatworm from a robot that simulates it cell by cell. Vitalist conceptions were common at the turn of the 19th to the 20th centuries, but they usually postulated causal powers above the causes described by physiology. The position suggested here does not conceive of additional causal powers, but only the concomitant existence of "qualities", which would constitute an epiphenomenalist view if it were not for the identity thesis.

Another suggestion that Price makes is to interpret the problem following the theory of emergence of Lloyd Morgan and C.D. Broad (among others), thus stating the colored-brain thesis in the following way:

*In so far as cerebral processes have this sensuous aspect, the brain, one would say, is also the sensorium; if so, the thesis is that it is the sensorium which is sonorous when we hear a bell, and red when we see a tomato. Or one might use the language of the Emergent Theory, and hold that sensual qualities*

*such as red and loud emergently qualify certain physico-chemical processes in the brain when these reach a certain degree of complexity (Price, 1932, p. 128).*

Price then discusses some difficulties that he sees in this view, concerning especially the localization of sense-data in space. We will return to this discussion in section 8.

## 4. What is "physical"?

The question of whether qualia are physical or not assumes that we have a definition of "physical". Any physicalist position must define what it means by "physical", but there is no consensual definition of the term.

A common definition in the philosophy of mind is associated with Mary's room thought-experiment (Jackson, 1982). Inside her black and white room, Mary would have access to "all physical knowledge" about colors, but she would never have experienced colors (other than those in the scale of gray). When Mary finally sees a color for the first time, one concludes that she acquires some "non-physical knowledge" about the world (if not before, at least after learning how to name the color), i.e., acquaintance with qualia involves "non-physical knowledge".

In this thought experiment, physical knowledge is defined as theoretical statements and models, involving language and mathematics, which lead to the experimental capacity for manipulating nature. The "physical" would then be what is described by the linguistic and mathematical structures, which amounts to the causal structure of the world, and which is amenable to experimental manipulation.

And qualia, would they be "physical" entities or not? Jackson's argument is limited to concluding that qualia are not known by "physical knowledge". To conclude that qualia are non-physical entities, it would be necessary to add one more hypothesis to the argument, that "if something is knowable and if it is physical, then it is physically knowable" (this proposition is denoted by "L4" in Pessoa, 2019a, p. 189). With this hypothesis, one infers that there is something non-physical that is knowable, which in the thought-experiment are qualia. Therefore, physicalism would be false, since not everything would be physical.

On the other hand, the rejection of the aforementioned proposition L4 that is, the acceptance of the thesis that there is something physical that is knowable without being "physically knowable" (in Jackson's sense), is consistent with the thesis that qualia are physical, as held by qualitative physicalism.

In short, the acceptance of Jackson's argument forces the physicalist to accept the colored-brain thesis. In other words, accepting the definition of "physical knowledge" from the Mary's room argument, it follows that *physicalism implies qualitative physicalism*.

But the physicalist may find Jackson's definition too restrictive, and consider that acquaintance with qualia consti-

tutes direct physical knowledge, as opposed to the indirect physical knowledge typical of the theoretical, linguistic-quantitative knowledge which Mary has within her room with respect to colors. The thesis that all theoretical knowledge in physics refers to causal structures in the world (and to other structures, such as symmetry relations) is known as “structuralism in physics.” The thesis that physical knowledge is limited to the description of these structures may be called “mechanicism,” in a broad sense. The qualitative physicalism maintained here shares with Mach, Schlick, and Russell the thesis that these structures are relations known in an indirect and inferential way, through the immediate knowledge of sensations or qualia.

Although we do not explicitly define what is “physical,” we will characterize a physical process with three properties: (a) localization in space and time; (b) occurrence in a scale (micro, macro, etc.); and (c) absence of final causes, i.e., at the elementary level there is no purpose or intentionality (for a naturalistic approach to intentionality, see Millikan, 2000). A consequence of this characterization of “physical,” for strong realist qualitative physicalism, is that every mental process (a visual quale, a theoretical idea) must be associated with a spatial localization (in addition to temporal and scalar localizations), even if such localization is distributed over a network in the central nervous system (more on this in section 8).

## 5. Assumptions of the view

In defending the colored-brain thesis, the following assumptions were implicitly assumed:

(i) *Perceptive internalism and reality of qualia*. Internalism in the philosophy of perception is the doctrine of the primary and secondary properties of Galileo, Descartes, Locke, etc., according to which the phenomenal qualities (qualia) exist only in the mind, not in the external world. When looking at a green avocado, or smelling the scent of sandalwood, we can focus our attention on the greenness or the sandalwoodness. The reality of qualia thesis states that such phenomenal qualities do in fact exist, as does a green afterimage that arises after staring at a tomato, and a female voice that one hallucinates after a night of exhausting work. At the beginning of the 20th century, qualia were described as “sense-data,” and conceived as something exclusively mental, separate from the thing that causes the sensation (e.g. the avocado in the fruit bowl, the perfume in the bottle) and from the material processes occurring in the brain.

(ii) *Ontic physicalism*. According to this thesis, everything is material, i.e., everything that exists can be conceived from within the worldview of physics and chemistry. Furthermore, living beings are considered as having evolved and emerged

from a physico-chemical universe lacking intentionality, within the framework of space-time. Contemporary physics may be insufficient for the understanding of consciousness, but consciousness must have a material basis located spatio-temporally. As indicated in the previous section, a quale or an idea must have spatial localization, size, temporal coordinate and duration.

(iii) *Mind-brain identity thesis*. The most direct way of connecting theses (i) and (ii) is through the thesis that qualia are identical to brain processes. With Place (1956), the identity thesis started being associated with the version that privileges the material ontological status of brain processes, a status conceived in linguistic-quantitative terms. But Bor-ing’s (1933, p. 16) approach, mentioned in section 1, placed qualitative states on equal terms as brain states. Qualitative physicalism’s use of the identity thesis should amount to considering qualitative mental states as identical to qualitative physico-chemical states (see more in section 9). One may also consider Spinoza’s attribute dualism as a version of the identity thesis, a form of monism that also appears in Gustav Fechner (1966, p. 3) and in Thomas Nagel (1987, chap. 4), views which privilege neither matter or mind, but a more fundamental substance.

In section 8 we will complete this list of assumptions of qualitative physicalism.

## 6. Russellian type monisms

In his book *Analysis of matter* (1927), Bertrand Russell arrived at a position close to qualitative physicalism, a monism constituted of sensorial elements or “percepts.” Such a domain would be “neutral” in relation to matter and mind, as Ernst Mach had formulated, in a view that had a good reception in the United States, being adopted by William James (Banks, 2014, p. 1-3). Traditional neutral monism’s claim is that both the external world and the self are collections or bundles of sensations.<sup>2</sup> Identity applies to the sensorial element, not to the distinct configurations of elements that receive the names of mind and body. Thus, one concludes that the physical world consists of the same elements that we consciously experience.

Russell emphasized that physics only captures the relations between things, and not the things in themselves (not the *relata* of the relations), and this in an indirect, inferential way. As mentioned in section 4, such a position is known as “structuralism in physics,” or “structural realism.” Russell stressed that science is built upon inferences, and the only intrinsic reality to which we have access (besides the relations between sensations, which reveal the extrinsic relations between things) are the sensations themselves: “Percepts are the

<sup>2</sup> More recently, the term “neutral monism” has come to refer to any monism that is neutral in relation to the poles of mind and matter, such the attribute dualism of Spinoza mentioned in section 5. Here we are referring to the “traditional” neutral monism of Mach, Avenarius, Petzoldt, James, Perry, Holt, Schlick, and Russell (Stubenberg, 2016).

only part of the physical world that we know otherwise than abstractly" (Russell, 1927, p. 402).

Russell's position has influenced several recent authors, culminating in the collection edited by Alter and Nagasawa (2015). Authors like Grover Maxwell, Michael Lockwood, and David Chalmers developed views inspired by the English philosopher, generating a class of conceptions they called "Russellian monisms". What they put forth that is important is the notion that there are "qualities," "quiddities" or "inscrutables" (the things in themselves) that constitute the entire physical world, and to which we do not have direct access, except the sensations that we experience subjectively (qualia). Alter and Nagasawa (2012, pp. 432-6) present four conceptions about the nature of these inscrutables: (a) phenomenal properties, as suggested by Russell; (b) non-physical "protophenomenal" properties, as suggested by Chalmers (1996, pp. 126-7); (c) neither physical nor mental properties, as in aspect dualism (Nagel, 1987, chap. 4); (d) physical properties of a special sort.

Qualitative physicalism falls into this last category, taking subjective sensation as identical to a real physico-chemical quality. Given that a quale is an inscrutable, by extension (that is, by inductive inference or analogy) one may assume that matter in general has similar qualities or quiddities, and that these combine in some way in the brain tissue to generate the complex of qualities that we experience daily. This possibility is considered by Herbert Feigl (1971), commenting on the "pan-quality-ism" proposed by the philosopher Stephen Pepper (see section 8), conceding that it "is not unreasonable provided that the 'intrinsic qualities' of inorganic things or systems are conceived as incomparably more 'colorless' than the qualities of human experience" (Feigl, 1971, p. 308). Feigl also stresses that a view similar to Russell's had been proposed by Moritz Schlick in 1918:

*The world is a variegated structure of connected qualities. Some of them are given to my (or to some other person's) consciousness and these qualities I call subjective or mental; others are not given directly to any consciousness and these I designate as objective or extra-mental (Schlick, 1974, p. 292).*

Feigl (1979, p. 325) points out that the monisms of Schlick and Russell are not, strictly speaking, materialist. One should add that Ernst Mach also combined neutral monism with structuralism in physics.

## 7. Panprotopsychism

Feigl (1979, pp. 326, 335-6) also mentions the "panpsychism" of American neorealists Charles A. Strong, Durant Drake and Roy Wood Sellars. This particular panpsychist tradition originated with the work of the English mathematician and philosopher William Clifford, who in 1878 defined

a certain *mind-stuff* that would be present at the atomic level, and the composition of which would generate the conscious mind of humans: "(...) a moving molecule of inorganic matter does not possess mind or consciousness; but it possesses a small piece of mind-stuff" (Clifford, 1878, p. 65). William James analyzed this view in his *Principles of psychology* (1890, chap. VI), rejecting it due to the difficulty of explaining how the myriad of mind-stuff grains would combine to generate a continuous phenomenal image (this "grain problem" or "combination problem" was analyzed more recently by Lockwood, 1993, and by Chalmers, 2013).

We highlight Drake's view, which arrived at a clearly panqualityist position:

*Your brain-processes are you, the conscious part of you. All the rest of the restless, moving world has its own inner nature; but it has no mechanism of introspection, or of speech. It cannot tell us, it does not know, what its own inner nature is. Because it has no mechanism of introspection, or of perception, or memory, we cannot properly call these other bits of matter minds. But we can say that their inner nature is homogeneous with that of our brain-processes. Ordinary matter is the stuff out of which brains develop; its inner nature we may call 'mind-stuff' (Drake, 1933, pp. 381-2).*

Drake regrets having called his position "panpsychist" in his previous book, *Mind and its place in nature*:

*The term 'pan-psychism' has proved misleading. It suggests a denial of the existence of matter, and its replacement by a different sort of reality, an airy, insubstantial world. But the theory does not deny any of the teachings of physics, it merely rounds it out by telling us what matter is. It says that the world of matter-in-motion is the real world, but that physics has no means of discovering its inner nature, and needs to be supplemented by an inference drawn from the special knowledge we have of our little corners of the world.*

*The term 'pan-psychism' also suggests that all the world is conscious. But, though, according to this theory, the whole universe is of the same substance as our conscious life, it lacks – except where brains have developed – the peculiar organization of elements which make up a mind. The brain is not a special sort of stuff, it is a very special sort of mechanism. So the mind, which is the brain, considered in its inner nature, is not a different sort of stuff from the rest of nature, but is a highly complex and delicately adjusted mechanism. It is a mechanism which carries on the processes of perception, memory, imagination, thought, and emotion. These*

*are all extremely complex processes; and where they are absent, we cannot properly speak of a 'mind' (Drake, 1933, p. 382).*

In the previous section, we mentioned that qualitative physicalism can adopt the “panqualityist inference” and assume that matter in general (outside the brain) also has qualities. This inference, adopted by the Russellian-Schlickian monisms, leads to the notion that there is a “protopsychism” in nature, as expressed by Clifford and more clearly by Drake, resulting in the “panprotopsychoist” view (term coined by Chalmers, 2013). These inscrutables would combine in a mechanically complicated way in the brain tissue to generate the qualitative complex that we experience.

## 8. Herbert Feigl and Stephen Pepper

In a paper published in 1963, Feigl discussed a “familiar objection” to the mind-body identity theory, which addresses a version of the “explanatory gap” (Levine, 1983), to be further examined in sections 9 and 10: “how could directly experienced qualities such as colors, sounds, smells, pains, emotions, or the like, be identical with neural processes whose properties are so fundamentally different?” (Feigl, 1963, p. 257). In presenting his view, Feigl appears to accept the three assumptions of qualitative physicalism (section 5), contrary to the materialists Place and Smart, who tend to reject the thesis that qualia are real. Still, Feigl does not derive the colored-brain thesis, probably due to its counterintuitive character, so his view may be characterized as a weak form of qualitative physicalism (and ours a *strong* form).

But there is also another difference between our approach and that of Feigl, concerning the spatial localization of qualia. To begin with, Feigl explicitly adopts “structuralism in physics,” as we do, following Schlick and Russell. But he contends that the intuitive models we form of the brain, in a spatiotemporal background, cannot be considered real, but only heuristic and didactic models: “The geometry employed in the description of physical space is a conceptual system which, though based upon the evidence of the sensory kind of spatiality, is itself not adequately intuitable (visualizable, etc.)” (Feigl, 1963, p. 331). This is an anti-realistic stance on physics (in this case, classical physics). In contrast, the qualitative physicalism presented here takes a *realist* view of the picture of the physical world, at least within the limits of validity of classical physics. Both versions of materialism take seriously the localization in physical space of mental properties, but such physical space is considered abstract by Feigl, and he attributes reality only to perceptual space: “Hence there is no conflict and no incompatibility in regard to the ‘location’ of, e.g., a directly experienced patch of color. It is where we ‘see’ it in phenomenal space. The systemically identical cerebral process is assigned a place in the abstract 3-dimensional manifold of physical space [...]” (Feigl, 1963, p. 259).

In the view maintained here, the patch of color is in a region of the real physical brain, and when we see the color or patch, we are observing directly this region of the optical sensorium (and only indirectly observing an avocado on the table). Thus, we add two theses to the previous list of section 5: (4) Structuralism in physics. (5) Realism of the structural (functional, mechanistic) description of physics. This last item defines our *strong realist qualitative physicalism*.

Stephen Pepper, a thinker close to Feigl, in his book *Concept and quality* (1967), developed a “*qualitative* neural identity theory” (p. 76), stronger than Feigl’s in the sense of being panprotopsychoist. In discussing the identity thesis, he starts from a version previously developed by Feigl (1958) but later abandoned, the “double language theory”. According to this view, there are two languages usually employed in psychology, the physical and the phenomenal, which can be *translated* into each other. Physical language can refer to various scales, such as the macroscopic or “molar” scale, which describes overt behaviors, or the microscopic scale, which describes cells, or even the nanoscopic one, which refers to molecular processes. The phenomenal language does not seem to encompass such a wide range of scales. An example of translation between the two languages is as follows (adapted from Pepper, 1967, pp. 76-77): a subjective visual flash (phenomenal language) can also be described as the reflection of light into the retina, followed by a winking reflex (physical language). Such a “linguistic version” of the identity thesis (i.e., the double language theory) thus assumes that the two utterances have the same referent, but one does not ask what this referent is (Pepper, 1967, p. 84). Feigl ended up abandoning this version because the referents of the utterances were not the same after all. The direct referent of the phenomenal language would be “raw feels” or qualia. And how about the physical language? In the new Feiglian version we saw above, it refers to “descriptive symbols” (not, as a realist might expect, to physical reality). Pepper inherits this version of the identity thesis that combines realism of qualia with structuralism in physics interpreted antirealistically. However, unlike Feigl, as we have already mentioned (section 6), he adopts the “panprotopsychoist inference” and considers that there are qualities outside the mind (panqualityism).

Pepper gave an example involving neurosurgeon Wilder Penfield, who directly stimulated with a pair of electrodes a certain region of a patient’s cortex, at a certain point, and the patient reported a “sensation in face”. If Penfield were to give a detailed account of the neural processes involved (something we do not yet have the ability to do fully), “what is the event pointed at by all the indirect evidence?” Pepper’s answer is that it would be the “qualitative experience of the patient” (1967, p. 85).

Pepper transforms the mind-body problem into the quality-concept (or quality-structure) problem, “a problem of qualitative actuality and various symbolic descriptions of it” (1967, p. 92). But he concludes with the more realist view that the qualitative process is located in the brain:

*Just how do physical terms make contact with an actual qualitative process? The identity theory brings this question to focus. Taking expert physiological and observational results seriously, it shows just where actual quality and physical concept meet. They meet in what is physiologically indicated as areas of man's brain, or, perhaps more narrowly, of his cerebral cortex (Pepper, 1967, p. 93).*

## 9. A model for strong qualitative physicalism

Qualitative physicalism claims that a quale  $q$  is identical to a physico-chemical quality  $q$ , still unrecognized by science. The problem now is how to “explain” such a quality in the physical world. The explanatory gap or hard problem of consciousness now migrates from the mental/physical interface to the domain of physics and chemistry. The path to be explored rejects functionalism of qualia, or the claim that the causal structure of a system (describable in linguistic-mathematical terms) is sufficient for establishing its qualitative properties, and adds the element of “materiality”.

In Pessoa (2015, p. 210) we called “ $\Sigma\omega$ ” (“sigma omega”) the specific part of the brain (included in the sensorium) directly associated with the experience of a quale such as greenness,  $(\Sigma\omega)_{\text{greenness}}$ , a part that could be spatially localized in a restricted way (localizationism), or could be distributed in more extended regions (holism). The term  $\Sigma$  refers to the *organization* or structure of the system, i.e., the network of causal relations (plus other relations) that is described in a mechanistic approach, which in the case of mental phenomena includes neuronal spikes and the specific network of feedback loops that underlie the self. The term  $\omega$  refers to the *materiality* of the system, i.e., the physico-chemical constitution of the brain tissue.<sup>3</sup> This is a modified form of hylemorphism, as it concatenates matter and form (organization). For the strong qualitative physicalism proposed here, (i) a quale of greenness is identical to a physico-chemical quality, and (ii) such physico-chemical quality of greenness arises (emerges) from a specific material organization  $(\Sigma\omega)_{\text{greenness}}$ .

It is plausible to assume that only a change in the organization of the system (change in the frequency with which electrochemical spikes affect the cells in the sensorium), maintaining its materiality  $\omega$  (that is, keeping the molecules of the cell and its surroundings constant), can change the subjective quality associated with the cells (in a localizationist model).<sup>4</sup> But the dynamic structure  $\Sigma$ , which includes any relation that can be measured and reproduced in another system, has com-

ponents that do not contribute directly to the generation of a quale. That is, there is a spatiotemporal domain in which the structure is *essential* for the constitution of a quale, which can be denoted by  $\Sigma_q$ , but there are measurable relations  $(\Sigma_q)$  in other spatiotemporal domains that are *not* directly involved in the generation of the quale, but are part of what constitutes the mental phenomena. These could include structures that participate in the self, besides causal structures of support for the working of the local cognitive system.

The question of “why” a quale  $q$  arises from process  $(\Sigma\omega)$  is an expression of the hard problem of consciousness, or the explanatory gap mentioned in section 8. In the future, “bridge laws” will stipulate which  $q$  is instantiated when brain matter is organized in a certain way  $(\Sigma\omega)_q$ . At least part of these bridge laws will have to be established as *new scientific principles*, and not as an “explanation” in the usual sense of the term. An explanation in the usual sense starts from a linguistic-mathematical structural description  $\Sigma$  (for example, the description of the sun’s composition and the laws of nuclear physics), from which another linguistic-mathematical statement  $\Sigma'$  is derived (for example, intensity of solar radiation reaching the Earth), which can be verified by observing the interaction of scientific instruments with this radiation. But (contrary to what some supporters of mechanicism and functionalism seem to claim), one cannot start from a linguistic-mathematical description and “derive” a quality (in the same sense as previously exemplified). In the future, we will be able to say: “subjective greenness arises from  $(\Sigma\omega)_{\text{greenness}}$ ,” but this phrase by itself does not capture what it is like to experience greenness.

The introduction of non-explanatory postulates is common in science: for example, instead of asking why light has the same speed in all reference frames, Einstein in 1905 introduced this hypothesis as a *postulated* principle within his theory of special relativity, and not as an explanation. Analogously, we suggest that for a basic set of qualities, the thesis that a process  $(\Sigma\omega)_q$  gives rise to a quality  $q$  should be taken as a postulated principle.

What would be the nature of this real physical materiality, which, along with its dynamical organization, give rise to physico-chemical qualities which are identical to subjective qualities? Considering the four fundamental physical interactions, it seems reasonable that it is part of the electromagnetic interactions, since it does not seem to involve gravitational and nuclear forces. There are several types of electrochemical interactions recognized at the molecular level: in addition to the stronger covalent bond, there are various types of non-covalent interactions. Another important question is the scale at which these subjective qualities emerge; a localizationist guess would be in the order of the size of a cell ( $\sim 5$  micrometers). As for the localization of the sensorium, a reasonable

<sup>3</sup> The question of whether the materiality of a cell can be reduced to lower level structures is open to investigation. The scale that seems crucial here is the quantum level, which governs chemical reactions.

<sup>4</sup> In Pessoa (2019a), this model is used to explore two different situations involving qualia inversion.

guess for visual qualia would be in subcortical regions such as the pulvinar, in the thalamus. Given this material substrate  $\omega$  in the sensorium, the hypothesis is that incoming electrochemical spike trains  $\Sigma$  from the neocortex generate different physico-chemical qualities.

## 10. The explanatory gap

While commenting on Price's analysis, we already mentioned that Charlie Dunbar Broad articulated an emergentist materialist view for the mind-body problem, in his book *Mind and its place in nature* (1925), the same title as Drake's book of the same year. He criticized the attempt to reduce a mentalist statement, such as "I am aware of a red patch," to a behavioral statement, such as "a part of this body is behaving in such and such a way" (Broad, 1925, p. 622). Taking further this criticism to reductionism, he extended the definition of behaviorism to include statements that describe the "behaviors" of microscopic parts of the body, as in "a molecular movement is going on in a certain part of my brain" (Broad, 1925, p. 622). Such a description is what I have called a "mechanicist" description. For Broad (1925), there is something which has the "characteristic" of being redness (in his words, of being "my awareness of a red patch"), and there is something which has the characteristic of being a molecular motion. These two "somethings" could even be identical, ponders Broad, but the two characteristics or properties are clearly different. He rejects the possibility that they be two identical characteristics, such as "rich" and "wealthy"; because for the qualitative awareness one can ask whether it is clear or confused, and for the molecular movement whether it is swift or circular, but not the other way around. He therefore concludes that *qualitative awareness cannot be reduced to molecular movement* (Broad, 1925, p. 623).

The qualitative physicalism presented here introduces the idea that the organized material states  $\Sigma\omega$  of the brain may give rise to physico-chemical qualities  $q$ , which are identical to subjective qualities  $q$ . However, the difference between the characteristics (properties) pointed out by Broad remains, although shifted to the physical world. The brain system has a microscopic dynamic organization  $\Sigma$  that may be described by the relational scientific language, and in addition has a specific materiality  $\omega$ , which combined give rise to qualities  $q$ . In the present approach, the mind-body problem, as posed by Broad, becomes an ontological issue regarding matter.

As a conclusion, consider the following comment by James Cornman (1962), which takes up Broad's argument (see Moravia, 1995, p. 115):

*We can talk about intense, unbearable, nagging, or throbbing pains. And yellow, dim, fading, or circular after-images. And dogmatic, false, profound, or unconscious beliefs. On the other hand we can also discuss publicly observable, spatially located,*

*swift, irreversible physical processes. Thus if the Identity Theory is correct, it seems that we should sometimes be able to say truthfully that physical processes such as brain processes are dim or fading or nagging or false [or yellow], and that mental phenomena such as after-images are publicly observable or physical or spatially located or swift* (Cornman, 1962, p. 490).

There are three cases to be considered here:

(i) Concerning qualia, such as yellowness, the colored-brain thesis argues that they are real physical properties localized in the properly organized brain matter.

(ii) Other attributes of mental objects, such as nagging (in the sense of persistent) or dimness, are mechanical properties ( $\Sigma$ , see section 9) identifiable both subjectively in ( $\Sigma \Sigma \omega$ ) and by neuroscientific measurement.

(iii) The truth value of an idea is a relational property between a mental state and a situation in the outside world. Thus, for the internalism of qualitative physicalism, being true or false is not an attribute of a mental state (contrary to what an externalist would say, like Kim, 1982, p. 57).

## References

- ALTER, T.; NAGASAWA, Y. 2012. What is Russellian monism? *Journal of Consciousness Studies*, **19**: 67-95. Reprinted in ALTER; NAGASAWA (Eds), *Consciousness in the Physical World: Perspectives on Russellian Monism*. Oxford, Oxford University Press, p. 422-51.
- ALTER, T.; NAGASAWA, Y. (Eds.). 2015. *Consciousness in the Physical World: Perspectives on Russellian Monism*. Oxford, Oxford University Press.
- ARISTOTLE. 1931. *On the soul (De anima)*, transl. J.A. Smith. Oxford, Clarendon Press. Available at: <http://classics.mit.edu/Aristotle/soul.1.i.html>.
- BANKS, E. C. 2014. *The Realistic Empiricism of Mach, James, and Russell: Neutral Monism Reconceived*. Cambridge, Cambridge University Press.
- BORING, E. G. 1933. *The Physical Dimensions of Consciousness*. New York, Century.
- BROAD, C. D. 1925. *The Mind and its Place in Nature*. London, Kegan Paul, Trench, Trubner & Co.
- CASE, T. 1888. *Physical Realism: Being an Analytical Philosophy from the Physical Objects of Science to the Physical Data of Sense*. London, Longmans, Green & Co.
- CHALMERS, D. J. 1996. *The Conscious Mind: In Search of a Fundamental Theory*. New York, Oxford University Press.
- CHALMERS, D. J. 2013. Panpsychism and panprotopsychism. *The Amherst Lecture in Philosophy*, **8**: 1-35. Available at: [http://www.amherstlecture.org/chalmers2013/chalmers2013\\_ALP.pdf](http://www.amherstlecture.org/chalmers2013/chalmers2013_ALP.pdf). Reprinted in ALTER; NAGASAWA (Eds), *Consciousness in the Physical World: Perspectives on Russellian Monism*. Oxford, Oxford University Press, 2015, p. 246-76.
- CLIFFORD, W. K. 1878. On the nature of things-in-themselves. *Mind*, **3**: 57-67.

- CORNMAN, J. 1962. The identity of mind and body. *Journal of Philosophy*, **59**: 486-92.
- DEN OTTER, S. M. 2005. Case, Thomas (1844-1925). In: BROWN, S. (Ed.). *Dictionary of twentieth-century British Philosophers*, vol. 1. Oxford, Oxford University Press, p. 159-63.
- DENNETT, D. 1992. *Consciousness Explained*. New York, Little, Brown & Co.
- DRAKE, D. 1933. *Invitation to Philosophy*. Boston, Houghton Mifflin.
- FECHNER, G.T. 1966. *Elements of Psychophysics*, vol. I. Transl. H.E. Adler. New York, Holt, Rinehart & Winston. (German original: 1860.)
- FEIGL, H. 1958. The 'mental' and the 'physical'. In: H. FEIGL; M. SCRIVEN; G. MAXWELL (Eds.), *Concept, Theories, and the Mind-Body Problem*. Minnesota Studies in the Philosophy of Science, vol. 2. Minneapolis, University of Minnesota Press, p. 370-497.
- FEIGL, H. 1963. Physicalism, the unity of science, and the foundations of psychology. In: P. A. SCHILPP (Ed.), *The Philosophy of Carnap*. La Salle (IL), Open Court, p. 227-68. (Originally written in 1954.)
- FEIGL, H. 1971. Some crucial issues of mind-body monism. *Synthese*, **22**: 295-312.
- FEIGL, H. 1979. Russell and Schlick: a remarkable agreement on a monistic solution of the mind-body problem. In: G. W. ROBERTS (Ed.), *Bertrand Russell memorial volume*. Oxfordshire, Routledge, p. 321-38.
- HATFIELD, G. 2015. Perception in philosophy and psychology in the 19th and early 20th centuries. In: M. MATTHEN (Ed.), *The Oxford Handbook of Philosophy of Perception*. Oxford, Oxford University Press, p. 100-117.
- JACKSON, F. 1982. Epiphenomenal qualia. *Philosophical Quarterly*, **32**: 127-136.
- JAMES, W. 1890. *Principles of Psychology*. 2 Vols. London, Macmillan.
- KIM, J. 1982. Psychophysical supervenience. *Philosophical Studies*, **41**: 51-70.
- LEVINE, J. 1983. Materialism and the explanatory gap. *Pacific Philosophical Quarterly*, **64**: 354-61.
- LOCKWOOD, M. 1993. The grain problem. In: H. ROBINSON, H. (Ed.), *Objections to Physicalism*. Oxford, Clarendon Press, p. 271-291. Reprinted in ALTER; NAGASAWA (Eds), *Consciousness in the Physical World: Perspectives on Russellian Monism*. Oxford, Oxford University Press, 2015, p. 143-160.
- MILL, J.S. 1843. *System of Logic, Ratiocinative and Inductive*, v. I. London, J.W. Parker.
- MILLIKAN, R. G. 2000. Naturalizing intentionality. In: B. ELVITCH, (Ed.), *Philosophy of Mind*. The Proceedings of the Twentieth World Congress of Philosophy, vol. 9. Bowling Green (KT), Bowling Green Philosophy Documentation Center, Bowling Green State University, p. 83-90.
- MORAVIA, S. 1995. *The Enigma of the Mind*. Transl. S. Staton. Cambridge, Cambridge University Press, 1995. (Italian original: 1986).
- NAGEL, T. 1987. *What Does It All Mean?* New York, Oxford University Press. 1987.
- O'REGAN, J. K.; NOË, A. 2001. A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, **24**: 939-1031.
- PEPPER, S. C. 1967. *Concept and Quality: A World Hypothesis*. La Salle (IL), Open Court.
- PESSOA JR., O. 2015. A consciência enquanto observação do cérebro. In: F. R. M. FERREIRA; H. HADDAD JR.; M. I. NOGUEIRA; O. PESSOA JR. (Eds.). *História e Filosofia da Neurociência*. São Paulo, Liber Ars, p. 201-219.
- PESSOA JR., O. 2019a. How to measure a quale. *Sofia (Vitória)*, **8**: 187-198.
- PESSOA JR., O. 2019b. A causal-pluralista metatheory of observation. *Open Philosophy*, **2**: 657-667.
- PLACE, U. T. 1956. Is consciousness a brain process? *British Journal of Psychology*, **47**: 44-50.
- PLACE, U. T. 2000. Identity theories. Unfinished manuscript edited by S. Schneider. Available at *A Field Guide to the Philosophy of Mind*: <http://host.uniroma3.it/progetti/kant/field/mbit.htm>.
- PRICE, H. H. 1932. *Perception*. London, Methuen.
- RUSSELL, B. 1927. *The Analysis of Matter*. New York, Harcourt Brace & Co.
- SCHLICK, M. 1974. *General Theory of Knowledge*. Transl. A. E. Blumberg. New York, Springer. (Translation of the 2nd German edition of 1925; 1st ed.: 1918).
- STUBENBERG, L. 1998. *Consciousness and Qualia*. Amsterdam, John Benjamins.
- STUBENBERG, L. 2016. Neutral monism. *Stanford Encyclopedia of Philosophy*, available at: <https://plato.stanford.edu/entries/neutral-monism/>.

Submitted on October 07, 2020.  
Accepted on December 07, 2020.