

Tying Phenomenal Consciousness and Quining the Gap*

Maria Clara Dias

Centro de Ética e Filosofia da Mente/UFRJ

In his article about the problem of consciousness and the explanatory gap¹, J. N. Kaufmann presents interesting contributions to the discussion of the qualitative characteristics of subjective (or first-person) experiences of mental phenomena. Basically, the problem emerges from the fact that; in one hand, those characteristics seem to play a causal role in human actions and, in the other hand, no conceptually satisfactory reference to those elements seem to fit (due to their subjective character) into explanatory structures (due to the demands of objectivity associated to these structures). Therefore, there is a gap between the qualities of subjective experience and acceptable psychological explanation of human action, in a physicalist framework. Kaufmann accepts the formulation of the problem in J. Levine's terms.²

If there is such a gap, it would consist in a serious problem to any naturalization project of psychological explanation. The problem would be a serious one, because it is not simply a matter of fact, relative to the present state of our science

*I would like to express my thanks to Carlos Eduardo Nogueira Loddó (UQTR) for his extremely useful comments on earlier versions of this paper. His stimulating observations helped me to present my points in a clearer way, especially as far as conceptual matters are concerned. He also helped essentially in formulating a precise interpretation (in what is relevant for this discussion) of the perspective proposed by J. N. Kaufmann, with whom he worked closely, translating many of his articles (including the one here commented) to Portuguese. For the English revision of the present text, I would like to thank Mário Nogueira and André Fuhrmann.

Of course, my use of the neologism "to quine" is taken from Dennett, D. "Quining Qualia", in A. Marcel and E. Bisiach (eds.) *Consciousness in Contemporary Science*. Oxford: Oxford University Press, 1988). I am also coining another new term, "to tye," taking advantage of the phonetic coincidence between Tye's name and the English verb, "to tie."

¹Kaufmann, J. N., "A Problemática da Consciência e a Lacuna Explicativa." *Princípios. Revista de Filosofia*, vol. 6, n.7 (1999), pp. 87-120.

²J. Levine, "Materialism and Qualia: the explanatory gap". *Pacific Philosophical Quarterly*, 64 (1983).

(which could change in the future), but a *matter of principle*. Being the explanatory gap problem formulated as a matter of principle, Kaufmann observes that it should be also seen as a *conceptual* matter.

Because of the conceptual character of the problem, Kaufmann starts out his analysis by attempting to circumscribe precisely the *concept* of consciousness. He first presents a series of distinctions and conceptual oppositions concerning consciousness, and establishes which concept of consciousness is the relevant one for this discussion. Specially important for Kaufmann is N. Block's opposition between *access consciousness* and *phenomenal consciousness*.³ According to such opposition, *access consciousness* is representational and definable functionally. *Phenomenal consciousness*, on the other hand, cannot be defined as a general concept, although it is possible to indicate which objects would fall under it. Kaufmann observes that the list of such objects vary, from author to author. But this last concept is, anyway, the notion of consciousness that is involved in first-person or subjective experience. Therefore, Kaufmann locates here the concept of consciousness around which the problem of the explanatory gap is formulated, and relatively to which it should be treated.

Next, Kaufmann goes on to examine two contrasting perspectives, relatively to the explanatory gap problem. First, he examines Colin McGinn's position, which includes arguments favorable to the idea that there is a gap.⁴ Kaufmann examines two of McGinn's arguments for the gap, using Owen Flanagan's reconstructions of those arguments.⁵ Kaufmann retains Flanagan's criticisms to the arguments and adds to those criticisms some others, coming from Husserl's phenomenology.

But it is Kaufmann's examination of David Chalmers's position the most relevant one to what we will discuss in this text. Chalmers's position is in favor of "closing the gap", this is to say, he tries to provide a solution to the problem. As Kaufmann observes, Chalmers presents different strategies for this, in different texts. In the early texts⁶, Chalmers attempts to supplement physicalist explanations with conceptual analysis typical from phenomenology.

In a later text⁷, in view of correlating consciousness to neuronal processes (the conclusion of his argument), Chalmers presents what he calls "bridging princi-

³Block, N. "On a Confusion About a Function of Consciousness." *Behavioral and Brain Science* **18** (1995), pp. 227-247.

⁴McGinn, C. *The Problem of Consciousness*. Oxford: Blackwell, 1983.

⁵Flanagan, O. *Consciousness Reconsidered*. Cambridge, Mass.: MIT Press, 1992.

⁶Chalmers, D. J. *The Conscious Mind: In Search of a Fundamental Theory*. New York: Oxford University Press, 1996; Chalmers, D. J. "Facing up to the Problem of Consciousness" – "Response". In J. Shear (ed.) *Explaining Consciousness—the Hard Problem*. Cambridge, Mass.: MIT Press, 1997, pp. 9-33.

⁷Chalmers, D. J. "On the Search for the Neural Correlate of Consciousness". In S. R. Hameroff, A. W. Kaszniak and A. Scott (eds.). *Toward a Science of Consciousness II: The Second Tucson Discussions and Debates. Complex Adaptive Systems*. Cambridge, Mass.: MIT Press, 1998.

ples". Functionally definable states are linked to neuronal structures, because the first might be realized in the last (the empirical premise of Chalmers's argument). We might have access to consciousness by means of reportability or by other behavioural modes of expression (here is the bridging principle of Chalmers's argument). Kaufmann analyses Chalmers's argument, presenting a series of criticisms to it, in some cases, making appeal, once more, to Husserl's phenomenology.

It is not my purpose here, neither to examine Chalmers's argument, nor to provide a full account of Kaufmann's criticism to it. However, I will retain a single point observed by Kaufmann. He notices that Chalmers's bridging principles may only have success in linking physicalist explanations to consciousness, *if* consciousness is conceptualized as what Block calls *access*-consciousness (the one which is representational and definable functionally), but *not at all* if it is conceptualized as *phenomenal*-consciousness.

This last concept, as seen, is where the explanatory gap problem lies, according to Kaufmann's approach. Even though this observation is critical to Chalmers's attempt to close the gap, Kaufmann's attitude is more optimistic than his evaluation of Chalmers's suggests, as far as the project of closing the gap is considered. Kaufmann observes that, *if* Chalmers would show (contrarily to what Block supposes) that a concept of functionally defined consciousness is systematically correlated with a concept of phenomenal consciousness, *then* his argument could perhaps stand, and the gap would be closed.

In the following lines, I will examine an alternative perspective that seems to be able to close the gap. I will argue that there is no gap between physicalist explanations and phenomenal consciousness, as a matter of principle. This perspective is the one suggested by Michael Tye.⁸ In a sense, what Tye proposes corresponds exactly to Kaufmann's suggestion to save Chalmers's argument, since Tye does correlate systematically a concept of phenomenal consciousness with a concept of functionally definable consciousness. According to Tye, more radically, the concept of phenomenal consciousness is functionally definable.

Based on an analysis of certain *sui generis* properties of phenomenal consciousness, I aim to show that the apparent explanatory gap between the mental and the physical can be eliminated. More specifically, I want to defend the claim that the so-called *sui generis* qualities of mental events, referred to as *qualia*, are physical properties. These properties would constitute, as Tye puts it, the non-conceptual representational content of our conscious experiences. Moreover, they are the input of physical processes whose output would be intentional states with propositional content.

We can understand physicalism, in general, as the perspective according to which everything that exists or that may be known belongs to the physical world,

⁸Tye, M. *Ten Problems of Consciousness*. Cambridge, Mass.: MIT Press, 1995.

i.e. to the domain of objects and laws of an elementary microphysics, possibly to be developed in the future.

Whichever definition of the “physical world” we would adopt, it would seem that it should incorporate a set of properties applicable in fact to a specific group of objects that, at least on a first impression, cannot be imputed meaningfully to physical objects. For instance, we say that “no one else can feel my pain” and also that pain, sadness, gladness and other mental states just exist while lived by a subject or while belonging to a subject. My consciously perceptible experience, my body sensations, emotions, passions and moods seem to be related to me in a way that my car and my clothes, for instance, cannot be. Another person could have my car and my clothes and those items would not cease to exist if not possessed by someone.

The “monopolization” of experienced mental phenomena by a subject, by a self, seems to characterize a sort of necessarily inner belonging, which is alien to the objects that only contingently belong to me. Should we not distinguish the objects to which I have a privileged access, that is, the objects that can only be “possessed” and “known” in my experienced perspective and that cannot exist unless they are related to the subject of the experience from the other objects to which I simply relate to, because I am the one that contingently possesses them? If the characteristic properties of the objects of first group—that is, the property of privacy, the property of being necessarily owned and the property of perspectivity⁹—cannot be attributed to physical objects, should we not then admit the existence of nonphysical entities?

We would run into difficulties if we adopt such an assumption. For it seems to violate a widely accepted principle, namely that the physical world is causally closed. This principle says that for any physical event whatsoever its causes are other physical events. Nonphysical entities have, therefore, no causal role to play in the processes by which physical events are brought about. Phenomenal events could of course be still effects, but never causes of physical events. This is the epiphenalist solution to problem of *qualia*. But are we really prepared to admit that phenomenal states are forever causally irrelevant?

Anyway, many people share the intuition that there is an unbridgeable between the physicalist explanation of our mental life and what we innerly believe to happen in some contexts. Would it be possible to explain in physicalist terms, that is to say, in terms of elementary microphysics, even if only in principle, what it means to be the subject of some experiences, or to express somehow the privileged point of view that we assume towards our sensations, emotions, passions and moods?

Faced with this problem, we may find three alternatives:

⁹On these properties, see: Michael Tye (1995).

1. to accept the *sui generis* character of phenomenal objects and their irreducibility to the domain of the physical world;
2. to show that phenomenal objects do not really have these *sui generis* properties; or to show that these objects do not exist and, thus, that we cannot apply these properties; or
3. to acknowledge the existence of phenomenal objects and their characteristic properties denying, however, that they do not apply to physical objects;

The first alternative is clearly dualist. It reinstates an ontological dualism in which the relation between the mental and the physical becomes contingent, arbitrary or even impossible. At best, it would be possible to restore an epiphenomenalism, according to which the mental would be causally inert. Our 21st century materialist beliefs, however, make it difficult to understand what it means to think about objects that do not produce effects on the physical world. How could these objects ever be perceived if they do not leave traces in the physical world? To claim the existence of something seems to imply causal relations with other physical objects or events.

The second alternative is either reductionist or eliminativist. Respectively, it either reduces the phenomenal objects to physical ones, thereby denying their peculiar characteristics; or it denies the existence of these objects and consequently the realization of *sui generis* properties. This alternative leaves physicalism untouched, but not our ordinary practices of attributing causal powers to mental states.

The third alternative seems to combine physicalism with some kind of dualism (or pluralism) of properties, thus affording some understanding of the peculiar character of some of our experiences. The second and the third alternatives share the kernel of the physicalist intuition. The difference between them is subtler than it seems to be at first sight. Some sort of reductionism or eliminativism seems unavoidable even to the third alternative. The difference between them lies only in the attempt of the third alternative to conciliate physicalism with some of our most valuable intuitions about our subjective states. But which subjective states, specifically? Or, in other words, which of their peculiar characteristics can plausibly be combined with the kernel of our physicalist beliefs?

Here we need to clarify our position in relation to the *explanandum* of the problem. What seem to resist to a physicalist reduction are not mental states in general, but some of their properties: privacy, necessary ownership and perspectival subjectivity. These properties reflect the subjective character of our experience; they are connected to what is lived only and exclusively from the perspective of the experienced subject. To use Nagel's phrase, those properties constitute the "what

it is like to be”¹⁰ character of our experience. In Nagel’s view, these properties are present in every intentional event, i.e., in every kind of relation that would involve consciousness and its objects. They constitute the irreducible subjective character of intentional events.

Here I sustain a more restricted position, a position that excludes from this domain our mental states such as beliefs, desires and intentions whose relevant content is expressed by a that-clause. In order to justify such a restriction, I intend to show that it is possible to provide, at least *prima facie*, a functional description of our propositional attitudes and that such a description seems not to leave out any relevant element of our comprehension of these mental states. The irreducible subjective character of experience, or *qualia*, on the other hand, will be seen as a peculiar quality of some mental states. Due to these peculiar qualities, the perspective of the experiencing subject seems to those who defend the explanatory gap to add something irreducible to the experience itself. I would, therefore, like to characterize the problem of the explanatory gap as one directed specifically to mental events that does not seem to be sufficiently described without a direct reference to the agent.

Before addressing the problem of the gap itself, I shall give a brief exposition of what could be a functional explanation of our propositional attitudes, thus justifying the restriction mentioned above.

The functionalist perspective in general identifies mental states either with the function accomplished by brain states or with the brain states that accomplish that function. What grounds the attribution of a certain content *p* to a state of belief, for instance, is relation of this state with certain sensorial stimuli, with other inner states and with behavior. Therefore, mental states are both inner causes of behavior and effects produced by the external world. In the words of Jackson and Braddon-Mitchell:

A functionalist theory of mind specifies mental states in terms of three kinds of clauses: input clauses which say which conditions typically give rise to which mental states; output clauses, which say which mental states typically give rise to which behavioral responses; and interaction clauses, which say how mental states typically interact.¹¹

To avoid circularity, the functionalist characterization of any intentional state must be made exclusively in nonintentional, physical terms. In particular, the intentional states that we intend to describe in physical terms must not be surreptitiously reintroduced in the characterization of the output. As with inputs, outputs

¹⁰Nagel, T. “What it is like to be a bat,” in: *Mortal Questions*. Cambridge: Cambridge University Press, 1979.

¹¹Braddon-Mitchell, D. and Frank Jackson. *The Philosophy of Mind and Cognition*. Oxford: Blackwell, 1996, p. 40.

must be also described entirely in physical terms. The functional description of mental states specifies a net of physical relations, involving dispositions to behavior, typical causes, as well as other mental states as far as these are also characterized in functional terms.

The most obvious applications of the theory follow the model of a machine programmed to perform specific tasks. This includes very simple machines like vending machines that provides us with Coke when we introduce a coin of a certain value, but also highly complex ones like connectionist machines of the PDP kind or neural nets. Evidently, the complexity of the functions realized by mental states requires flexible and very complex models. In many cases the complexity is so as that we are as yet unable to specify the models in question. This, however, is no objection to the claim that, in principle, there is nothing to prevent relevant aspects of mental events from being described in physical terms.

The human mind described in a functional way thus operates like a computing machine composed of several modules. There would be a scanner in the first module, which is responsible for the reception of the inputs. From that point on, we can imagine several modules. Among them, there would be an evaluator module, responsible for the selection of the pieces of information generated by an alternative generator. Finally, some motor module outputs a specific behavior. The peculiarity of such an assembly of modules is its capacity to learn, that is, to alter the final result, due to new inputs reflecting the effects produced by the agent's behavior in the previous stages. In other words, the produced output would promote external answers, which, in their turn, would be introduced as input and evaluated by the program, so that it can provide a new result. Thus, e.g., a program able to recognize alphabetic characters, like the ones successfully used in postal services, transforms and improves its performance when different tokens of the same character are provided and the manager answers affirmatively or negatively to its recognition. Such programs are capable of organizing the information so that a character previously misinterpreted can later be recognized. Based on the functional analysis of flexible program models, we could therefore account for one of the main characteristics of the human case: our capacity of learning. Similarly, other typically human characteristics, such as our capacity to deliberate and to intervene in the course of our actions—on which depends our attribution of responsibility and freedom to an agent—could also be recovered successfully.¹²

Supposing then that a functional description of our propositional attitudes is possible and that nothing essential to them is lost with a functional explanation, the problem of the explanatory gap becomes circumscribed to the question of *qualia*. How do they fit in a physical description of the world? What exactly do

¹²For a defense of compatibilism based on the exposition of the model of self-dominion, see Blackburn, S. *Think*. Oxford: Oxford University Press, 1999, p. 98f.

they refer to?

If we are not willing to accept a reductionist or eliminativist perspective towards *qualia*, we must be able to provide a physical description of them. In other words, we must explain in physicalist terms how the what-it-is-like properties may play a genuine causal role in our behavioral dispositions. In developing this possibility I draw on Michael Tye's PANIC theory of phenomenal character: "phenomenal character is one and the same as Poised Abstract Nonconceptual Intentional Content."¹³ Though Tye's theory distinguish between the nonconceptual aspect from the abstract one, I will consider exclusively the nonconceptual side of phenomenal experience, which is the relevant one for the present purposes.

The PANIC theory is a theory of phenomenal events or states as intentional but nonconceptual phenomena. These states may come together with other intentional states with propositional content. They can work as background elements to cognitive processes involving intentional states with propositional content. The crucial point here is that the intentional phenomenal states, as opposed to propositional attitudes, are not themselves conceptual. Bodily sensations, such as a toothache or an orgasm are representational, in such a way that they reflect causally, in normal circumstances, the alterations in specific body parts. These representations (1) are the output of nervous receptors responding to several stimuli in our body and (2) integrate the input of the cognitive processes that are for the first time cognitively revealed to the experience subject. Tye says further that through the action of cognitive processes operating in the nonconceptual representations the subject becomes aware of his phenomenal states. Becoming aware of his own experiences is a cognitive process that subsumes for the first time the phenomenal experience to concepts. Only at this moment, the experiential subject does adopt an epistemic (cognitive) attitude in relation to his phenomenal states. Below this level of awareness—traditionally called introspection—there is the phenomenal consciousness. The phenomenal content would therefore be better described as a nonconceptual representation of physical alterations, ready to integrate the input of processes that have as output intentional states with propositional contents (beliefs and desires). This way, to be the subject of a phenomenal experience would mean to possess certain nonconceptual intentional contents, which might be inputs of certain cognitive processes, that is, nonconceptual representations that could provide different responses.

The explanation provided by the PANIC theory allows us to make physicalism compatible with the intuition that the experience of *seeing red* and the experience of *knowing what it is like to see red* are distinct states irreducible to each other. Something definitely new (a new experience) occurs to someone that sees red for the first time even if this very same person already knows everything that

¹³Tye, M. (1995), p.137.

is possible to be known, in physicalist terms, about *what it is like to see red*. Here we face the case of one of the main arguments against physicalism, the knowledge argument. In what follows, I should like to show how the description of the phenomenal consciousness proposed by the PANIC theory allows us to account for Mary's new experience without jeopardizing physicalism.

Mary is Frank Jackson's famous scientist who knows everything about the physical world. She has complete knowledge of the elements and the laws of the physics of the future. Mary's knowledge involves therefore everything that is possible to be known in the physical domain. After having introduced Mary, let's take a look at her story. Bright Mary has grown up and learned everything in a black-and-white room. All of her contacts with the external world have been through a black-and-white monitor. For the sake of the argument, we can even imagine that some complex device makes her see herself and everything else in her room in black-and-white. One day Mary leaves the room and sees for the first time a red tomato. She experiences the color red for the first time. Mary already knew not only everything about tomatoes but also about color experience. She knew, therefore, that ripe tomatoes, hydrants and British telephone booths are all red. She also knew everything about the physiological processes that underlie color experience. At this point, physicalism seems to be challenged: supposing that we share the intuition that Mary learned something new when she saw colors for the first time, should we not admit that there is something in our experience that cannot be described in physicalist terms? Mary's argument seems to indicate exactly the distinction between *to know something* and *to experience something*. Mary knows everything about colors, but she never had any experience of seeing colors, before leaving the room. This way, although she knows *what it is to see colors*, she has never known *what it is like to see colors*. What seems to be irreducible to physicalism is therefore this privileged character of the perspective of the person who is the subject of experience: the subjective character of the experience.

For the sake of the argument, I shall accept the elements of the story, although they may seem very implausible. I do not intend to deny the basic intuition that Mary experienced something new when she left the room. What I do want to deny is that this new something cannot be described in physicalistic terms or that it indicates the boundaries of physicalism. If Mary already knew everything about red tomatoes, obviously what happened to her when she first saw them could not have altered her knowledge. But we can say that something new is introduced in the generation of such knowledge. According to Tye's theory, new nonconceptual representations come to integrate the input of Mary's cognitive processes. Although they do not alter Mary's cognitive content, such representations can promote changes in her future behavior. They can, for instance, provide her with the skill to select more quickly the tomatoes she wants to put in a salad. The introduced elements are representations of elements that enter in the facts Mary

already knew, since she knew everything about the physical world. But, the introduction of these new elements allows a reorganization of the data available to Mary, which determines her future behavior. According to such an interpretation, what Mary acquires is not new cognitive content, but a new form of access to the cognitive content she already possessed. So, there would be no cognitive content that could not be covered by a physicalist theory of the world; neither there would be a nonphysical access to it. The physicalist hypothesis would be kept intact.

If the functional description of our mental states *in general* and the above description of the supposedly subjective character of our experience *in particular* are both plausible, would we then not only eliminate the explanatory gap but also the mentalist vocabulary and with it some widely accepted and successful explanatory practices?

This is the point where it seems to me reasonable to admit that we are facing a dilemma. Maybe it is also the point that separates a mere eliminativist perspective from a constructive functional explanation of mental states. We do not want to suppress the *sui generis* characteristics of our experience, but we want to explain them from a physical point of view. If we are successful in our attempt to offer a plausible explanation—which can be proven correct in the future—it may leave no room for other sorts of explanation. Indeed, our explanatory practices seem to enter in a progressive process of substitution. The events and states that we are now ready to consider as the real causes of the disorganized behavior of an epileptical patient exclude the explanations we used to give of this phenomenon. We take epilepsy itself in a causal story, conceptualized functionally, and associated by realization links to neurological processes. Devil possession, for instance, are now suppressed alternatives. It is not an acceptable hypothesis in the characterization of the *true* causes of epileptical behavior.

Some philosophers seem to see some definite limits to this progressive substitution of explanatory schemes.¹⁴ According to them, our successful explanatory practices cannot be improved by physicalist developments. Since this point, the present paper does not seem to threaten well-established explanatory practices. But here we cannot block the potential emergence of these substitutions. Can we accept as a real possibility the replacement of our established mentalistic way of describing and explaining aspects of our conscious life by a theory that exists only in the futuristic imagination of physicalistic oriented philosophers? Which sort of explanatory schema will succeed our daily view of the world of experiences? Our answer is: a *prima facie* physicalism, a bet in the future functional understanding of mental events that will fill actual explanatory gaps and will make our descrip-

¹⁴I think here about authors like Lynne Rudder Baker who, although an assumed materialist, intends to rescue the credibility of day by day explanatory practices which have in their causal antecedent psychological, economical and social facts or events. See Baker, L. R. *Explaining Attitudes*. Cambridge, Mass.: Cambridge University Press, 1995.

tion of the human design compatible with our materialist beliefs.