

Is Consciousness Primary?

Michel Bitbol

Abstract

Six arguments against the view that conscious experience derives from a material basis are presented, none of which is entirely new taken in isolation but whose conjunction is compelling. These arguments arise from epistemology, phenomenology, neuropsychology, and philosophy of quantum mechanics. It turns out that any attempt at proving that conscious experience is *ontologically* secondary to material objects both fails and brings out its *methodological* and *existential* primacy. No alternative metaphysical view is espoused (not even a variety of Spinoza's attractive double-aspect theory). Instead, an alternative *stance*, inspired from F. Varela's *neurophenomenology* is advocated. This unfamiliar stance involves (i) a complete redefinition of the boundary between unquestioned assumptions and relevant questions; (ii) a descent towards the common ground of the statements of phenomenology and objective natural science: a practice motivated by the quest of an expanding circle of intersubjective agreement.

Key Words: Consciousness; epistemology; phenomenology; quantum mechanics; neurophysiology; neurophenomenology

NeuroQuantology 2008; 1: 53-71

Introduction

The view according to which consciousness derives from a material basis (that it is so to speak *secondary* to special arrangements of material entities) turns out to be much weaker than what is currently believed. I will try to show that its scientific and philosophical credentials are indeed highly disputable. But in the end, I will not try to endorse any alternative metaphysical view such as "consciousness is some self-existent stuff independent from matter", or "consciousness and matter are aspects of a

common underlying stuff", which would not be easier to support than their opposite. In line with Francisco Varela, I will rather advocate a radical change of *stance* regarding objectivity and subjectivity.

As a preparation, let me try to circumscribe what we tend to *mean* by "consciousness", taken as immediate experience rather than self-awareness. To give a flavour of the very unusual signification of this word, the quickest way is to paraphrase Wittgenstein's remark on sensations: "(Consciousness) is not a *something*, but not a *nothing* either!" (Wittgenstein, 1983)².

Corresponding author: Michel Bitbol

Address: Director of Research at CNRS, (Centre National de la Recherche Scientifique), Associate Lecturer at University Paris I, Panthéon-Sorbonne, Centre de Recherches en Epistémologie Appliquée, (CREA/Ecole Polytechnique), 1, rue Descartes, 75005 Paris, FRANCE

e-mail: michel.bitbol@polytechnique.edu

² The complete quotation (which is about sensations rather than consciousness as a whole) reads thus: "And yet, you again and again reach the conclusion that the sensation itself is a *nothing*" – not at all. It is not a *something*, but not a *nothing* either!" (Wittgenstein, 1983 §304)

Consciousness is not *something*. If we wish to understand this short but puzzling statement, it is expedient to make an inventory of the kind of “things” we are able to indicate by means of language.

- We can first use a *noun* in order to refer to some manipulable or abstract *object*. But an object is an entity which is supposed to exist independently of situations and subjective states. By contrast, consciousness as experience is *situated*; it is *what it feels like to be a subject*.
- In our language, we can also use a *predicate* in order to ascribe a *property* to an object, after having ascertained by reliable criteria that it indeed belongs to that object. Now, it is tempting (and commonplace) to think that consciousness is a property of very special objects such as living human bodies. But do we have reliable *criteria* to ascribe consciousness to somebody else? I’ll argue later that any criterion, such as presence or absence of bodily movement, presence or absence of verbal report, or even presence or absence of certain shapes on an electroencephalogram, is weak and ambiguous. In practice, we forget criteria and tacitly encompass other beings who are reasonably similar to us in the field of experience. We replace questionable criteria with empathy, and property ascription with unquestioned presupposition³.
- Finally, we can use *special restrictive predicates* in order to point towards a *phenomenon* rather than a property. For instance, we may say that an object *appears* to be red instead of saying that it *is* red. But conscious experience is no phenomenon in this sense. Indeed, a phenomenon as indicated by a special restrictive predicate is a *content* of experience; *it is by no means experience as*

a whole. If anything, then, consciousness is the very fact of phenomenality rather than some particular phenomenon. Even less can consciousness be called an *illusion*, as it was implicitly suggested by some philosophers (Dennett, 1991). For after all, illusion is still a state of consciousness in which certain judgments or interpretations have led us astray. As Susan Blackmore rightly pointed out, several features of our experience, such as believing in a permanent self subject to a unified and continuous stream of consciousness are likely to be “illusory”; but she also insisted that this cannot mean that consciousness *per se* is only an illusion (Blackmore, 2005).

So, consciousness as experience *is not something*; it is neither an object, nor a property, nor even a phenomenon.

But it is not nothing! For us, consciousness might even be *everything*. It is not something that we *have*, but it identifies with what we *are* in the first place. It is not something that can be known or described by us in the third person as if we were separated from it; but it is what we dwell in and what we live through *in the first person*.

I would recapitulate this by saying that consciousness is *existentially* primary.

With this all-pervasive kind of primacy in mind, I will now present some arguments against the view that conscious experience derives from a material basis, namely against the view that consciousness is *ontologically* secondary to matter. I will successively develop one argument from epistemology, one argument from phenomenology, three arguments from neurophysiology and neuropsychology, and one argument from the philosophy of physics. None of these arguments is uncharted, but the devastating power of their conjunction is seldom appreciated.

1) Argument from Epistemology

An analysis of ordinary and scientific knowledge shows us that objective domains of knowledge are elaborated in two steps, with conscious experience as an implicit departure point. Firstly, one progressively pushes aside any feature of experience on

³ As J.-P. Sartre cogently pointed out, there is no *conceivable* “proof” that there *are* other beings endowed with experience. Our only way to confute skepticism is to realize that the problem of “other minds” does not concern knowledge but existence. We must then turn the problem upside down: start from the fact that we normally do not question other conscious experiences, and analyze what in the (normal) course of our own conscious existence makes virtually impossible not to take them as granted (Sartre, 1943).

which conscious subjects cannot agree, such as individual tastes, community values, or the emotional tinge which is associated by individuals and communities with particular situations. Secondly, one only retains a sort of structural residue of conscious experience that can be the object of a consensus, and of a collectively efficient use as a predictive tool. At its most abstract, this structural residue is mathematical; but it can also consist of general propositions stating various types of relations between entities and predicates, such as “brains are bioelectrical organs made of neurons and glial cells”.

Let me give an illustration of this process of objectification, borrowed from the dawn of thermodynamics. The long and difficult process by which the thermodynamic variables such as temperature, pressure, and even volume (though at a much earlier period of history) have been extracted from their experiential basis is a *locus classicus* of the philosophical history of science (Bachelard, 1938, 1973; Mach, 1986). In the beginning, there were bodily “sensations”, ordinary practices, and an overabundance of qualitative observations about color of metals, fusion or ebullition of materials, expansion of liquids according to whether they are cold or hot etc. Heat and temperature were hardly distinguished from one another, and from the feeling of hotness. As for pressure, it was little more than a name for felt strain on the skin. But, progressively, a new network of quantitative valuations emerged from this messy experiential background, together with the laws that connect them (such as the ideal gas law). Even though sensations of hotness and strain still acted as a root and as a last resort for these valuations, they slipped farther and farther away from attention, being the deeper but less reliable *stratum* in a growingly organized series of criteria for assessing thermodynamic variables. At a certain point, the sensation of hotness no longer played the role of an implicit standard at all; it was replaced by phase transitions of water taken as references for a scale of variable dilatations in liquid thermometers. This scale, which posits a strict order relation of temperatures, replaced the mixture of

non-relational statements of hot or cold and partial order relation of hotter and colder which tactile experience together with qualitative observation of materials afford. Accordingly, the visual experience of graduation readings, or rather the *invariant* of many such visual perceptions, was given priority over the tactile experience of hotness. Later on, when the function “Heat” was clearly distinguished from the variable “temperature”, and its variation defined as the product of the “heat capacity” times the variation of temperature, tactile experience was submitted to systematic criticism: the feeling of hotness was now considered as a complex and confused outcome of heat transfer between materials of unequal heat capacities and the skin, and also of the physiological state of the subject. From then on, declarations about tactile experience, which had acted initially as the tacit basis of any appraisal of thermic phenomena, were pushed aside and locked up in the restrictive category of so-called “subjective” statements. In view of their genesis, subjective statements are then endowed with very special features: they are completely independent of statements of other types; they hold irrespective of the truth or falsity of statements of other types; they are like isolated cogwheels that do not engage with the rest of language (Wittgenstein, 1983). For instance, the subjective statement “I feel hot” is admittedly indisputable, even when confronted with the thermometer-reading statement “the temperature of this room is 5°C”. By contrast, the statement “the temperature of this room is higher than the boiling point of alcohol” clearly conflicts with the former thermometer-reading statement. The metaphysical translation of this process of isolation of certain “subjective” statements with respect with the other types of statements is either dualism or reductionism: dualism if one projects the two-realms organization of statements onto a two-realms organization of entities or properties; or reduction of subjectivity to a network of objective entities if one takes the criticism of experiential expressions on objective grounds as a sign of subordination of the former to the latter. But if one looks back at the whole

cognitive process by which the two-realms organization of statements was established, it clearly appears that the very alternative of dualism and reductionism is flawed. Indeed,

- (i) At the first stages of the process, there is simply *no* clearcut distinction between the two realms. As long as Touch still partakes of the basis for thermic judgments, a statement such as “this liquid is hot” made on the basis of mere tactile feeling can still be disputed on the ground of other pieces of evidence, just as much as the other pieces of evidence can sometimes be disputed on the ground of a set of experiential statements. At this point, experiential judgments have no reason to be called “*only* subjective” ; they are not isolated in a sort of “true-for-me” trap ; they are still pieces of the overall machinery of language and public debate instead of being isolated cogwheels.
- (ii) At the latest stages of the process of objectification, experience has been eliminated nowhere, even in the purest theoretical or experimental statements. There is still judgment in theoretical statements, and there is still perception in experimental statements. But judgments and perceptions tend towards an ideal of invariance across situations and persons.

We can sum up what has been said until now by pointing out that consciousness is and remains *methodologically primary*. Any objective descriptions arises, in history and on a day-to-day basis as well, as an invariant structural focus for subjects endowed with conscious experience (Bitbol, 2002).

Now, the problem is that the very success of this procedure of extracting invariants yields a sort of *amnesia*. The creators of objective knowledge become so impressed by its efficacy that they tend to *forget* or to minimize that conscious experience is its starting point and its permanent requirement. They tend to forget or to minimize the long historical process by which contents of experience have been carefully selected, differentiated, and impoverished, so as to discard their personal or parochial components and to distillate

their universal fraction as a structure. They finally turn the whole procedure upside down, by claiming that experience can be explained by one of its structural residues. Husserl severely criticized this forgetfulness and this inversion of priorities, that he saw as the major cause of what he called the “crisis” of modern science (Husserl, 1970). According to him, it is *in principle* absurd to think that one can account for subjective conscious experience by way of certain objects of science, since objectivity has sprung precisely from what he calls the “life-world” of conscious experience.

One might suspect that this is only the old-fashioned opinion of some philosophers of the past who knew virtually nothing about modern neurophysiology. But, interestingly, the same remark was stated in several texts of modern scientists, as an elementary truth one is bound to rediscover after a long wandering in the labyrinth of naturalism. One finds it, *inter alia* :

- in many articles of Francisco Varela, according to whom “Lived experience is where we start from and where all must link back to, like a guiding thread” (Varela, 1998).
- In a lucid paper of Piet Hut and Roger Shepard, who urge us to “set aside the presuppositions that have been foisted upon us by the standard scientific view and build, instead, upon the foundation of what is indubitably given in experience” (Hut & Shepard, 1998).
- And, also, at the end of Adam Zeman’s book entitled *Consciousness*: “I have described consciousness as a ‘further fact’, but it might be described more accurately as the fundamental fact of our human lives” (Zeman, 2002). The only objection which may arise against this latter remark is that, in virtue of what it says about the existential and methodological primacy of conscious experience, it should have been taken as the background of the whole book, instead of its being shyly formulated as an additional assertion in the last pages.

This preliminary argument is considered by some phenomenologists (Henry, 2001) as sufficient to declare that naturalism is faulty

from the outset. The fact that some scientists finally relinquish their naturalist dogma when they have carefully pondered about this argument could be taken as a further reason to stop the enquiry at this point. But, in view of the remarkable development of natural science and of the implicit adoption of naturalism that goes along with it for a vast majority of scientists, we cannot content ourselves with an argument completely external to science. We must try to adopt for a while the standpoint of the scientists who (*pace* Husserl) tend to forget or neglect the *existential and methodological* primacy of consciousness, and who turn its effective order of priorities upside down. We must take seriously these scientists who, being fascinated by their own mental power in manipulating objects of knowledge and using them for technological mastery of nature, presuppose that consciousness is *ontologically* secondary to such objects. We must follow carefully their project of identifying *in the objective world*, as described by physics or physiology, *where* and *how* consciousness arises. The problem is that, even if one does not retain Husserl's warning that it is manifestly invalid, this question (especially the question "how?") turns out to be immensely challenging.

2) Argument from phenomenology

First of all, the question about the origin of consciousness in the objective world is challenging *in principle*, as shown by further reflections derived from phenomenological analysis. In the previous section, we have seen that the end-product of the procedure of objectification is structural (or abstractly *relational*). But there is a consensus in phenomenology (and beyond) that if anything can be called "absolute", it is conscious experience. Not *ontologically* absolute, of course (since consciousness is not "something"), but *self-evidentially* absolute. How can ordinary naturalism reconcile these conflicting characterizations of the objective domain and of conscious experience?

To start with, E. Husserl pointed out that there is a huge gap, a major difference of "essence", between natural bodily objects

and "lived experience". Natural objects are given through mere "adumbrations" (*abschattungen*). They are given through sensed facets or profiles, connected with expectations about other profiles that could be sensed in the future. These expectations are themselves organized according to a law-like network of mutual connections. In other terms, natural objects are always *incompletely* given; and this defective *given* is embedded in a systematic pattern of anticipated *relations* which, unlike the actual sensory "intuition" of their present quality, can be shared intersubjectively. As a consequence of their mode of being that incorporates incompleteness in it, the existence of natural objects is "essentially" subject to doubt. Indeed, future contents of experience can always disconfirm previous conjectures about them. By contrast, lived experience is immediately and completely given. Future experience can by no means *disconfirm* its existence here and now, but only take its place. It may therefore be called an "absolute", in so far as its existence is indubitable whenever it is present (Husserl, 1993). For it, the distinction between reality and appearance is flatly irrelevant. As J. Searle after Husserl cogently pointed out, "where consciousness is concerned, the existence of the appearance *is* the reality" (Searle, 1997). Consciousness is the very fact that there is appearance; appearance is the reality of consciousness.

The claim that conscious experience is to be taken as "absolute" was further argued by Husserl because (i) any ascription of existence *presupposes* the existence of conscious experience (endowing conscious experience with a "transcendental" primacy), and (ii) confuting expectations about the existence of anything else leaves it untouched (Husserl, 1993), as can be seen in the experience of "epoché", "phenomenological reduction", or cartesian "hyperbolic doubt".

J.-P. Sartre concurred with this analysis, apart from some qualifications that will not concern us here. Sartre especially agreed with Husserl that consciousness can be taken as an absolute in virtue of the identity between being and appearing which is proper

to it. But he added two related considerations.

- Firstly, Sartre insisted that the mode of existence of consciousness is “circular”, that it necessarily involves a sort of loop: “any conscious existence exists *qua* conscious of existing” (Sartre, 1943). In the latter sentence, Sartre meant that consciousness is inherently consciousness of something *and* consciousness of itself *at the same time*. Yet, he noticed, the expression “consciousness is consciousness of itself” is misleading. Consciousness comprehends self-consciousness in a very peculiar, “non-positional”, way. It does not imply being conscious of the object “itself” besides other objects (which would mean some internal subject-object duality), but it involves perfect co-extensivity between being conscious of some object and being conscious of this being-conscious.
- Secondly, Sartre pointed out that consciousness cannot arise as the outcome of some “ontological proof” (that he takes as a derivation from essence to existence, from a certain stringent definition to an ascription of being). Indeed, consciousness is never merely possible apart from existing; it is no possible instantiation of a definition apart from being actual, since its existence is the (“transcendental”) precondition of any ascription of possibility. In phenomenological terms, the existence of consciousness is thus by no means contingent. It is absolute also in this sense.

Of course, these conclusions are likely to be challenged on the ground that their phenomenological source is not the sort of things that can be exhibited and publicly examined. Rather, it relies on an implicit demand of prompting the same experience as the author of the phenomenological description, and then recognizing in this description a reasonably accurate expression of one’s own experience. However, I have already displayed in section 1 an equivalent of the former statements in the “grammar” of our language, along with Wittgenstein’s urge to deflect the questions of phenomenology on the plane of language. Indeed, the connection between

absoluteness and self-evident indubitableness is still perceptible in our use of sentences including special restrictive predicates such as “appears to me as red”. These statements are either taken to be indisputable or subject to a classification in terms of *sincerity* rather than *truth*. They are sheltered from the standard procedure of empirical test by which we assess the truth or falsity of propositions about nature. *They are functionally absolutized*.

But wasn’t functional absolutization restricted to the so-called “subjective statements”? Isn’t this way of construing certain propositions as “freely moving cogwheels” which fail to engage with the rest of language and discourse the clearest symptom of the great “bifurcation of nature” (Whitehead, 1979), to wit of the dualist divide which was instituted by the completion of objectification? The latter is true only to a certain extent. Functional absolutization is in fact all-pervasive as an inherent aspect of any statement claiming an experiential basis, irrespective of whether the objectification process has been completed or not. Upstream from the “great divide” of objectification, a statement such as “this liquid is hot” was both functionally absolutized *qua* expressive of experience, and disputable *qua* partaking of a quest of interpersonal agreement. What objectification did was only to separate and freeze these two functional aspects of experiential statements.

In sharp contrast with the monadic / absolute character of conscious experience, the *content* of objective scientific knowledge is relational / structural throughout⁴, as I already mentioned and as I will now develop. Reduction of knowledge to structure is in fact the price which has to be paid for intersubjectivity or intersituationality. Karl Mannheim, one of the founding fathers of the sociology of knowledge, was aware of that. According to him, coordinating the variety of individual or collective perspectives entails an ever increasing formalization of

⁴ The (experimental and theoretical) *content* of scientific knowledge is relational/structural, but not its *experiential basis*, of course, as it has just been pointed out.

knowledge (K. Mannheim, 1997). Individuals and social groups can only understand each other at the high stage of abstraction provided by structures. And the residual part of the phenomena which does not fall under any structural frame expresses nothing more than the specific (not to say parochial) components of individual or collective situations. Before that, Kant and the neo-Kantian school of philosophy built their entire theory of knowledge on this premise. According to Kant, the properties of manifested objects of collective empirical study are only relational, and the so-called substance itself, construed as the permanent nucleus of a complex of phenomena, is "totally and fully a set of pure relations" (Kant, 1996). Similarly, Cassirer pointed out that the history of science as a whole tends towards relinquishment of substantial concepts and research of "invariant relations" instead (Cassirer, 2002). If properties are referred to, Cassirer writes in a Kantian spirit, it is only after the concept of property has been (re)defined in such a way that "it includes in itself the concept of relation" (Cassirer, 2004). Every single property referred to in scientific theories and discourse fall under this analysis, be it physical, chemical or biological. Little effort is needed to realize that "electrical charge", "valence", or "function" are thoroughly relational concepts. The quantum paradigm only reinforced this realization by including *contextuality* (and incompatibility) of observables in its own foundations, through its canonical commutation relations.

At this point, the project of accounting for the self-evidentially *absolute* conscious experience in terms of the *relational* concepts of objective science sounds utterly implausible. Actually, it might well represent a major aspect of the "category mistake" that was pointed out by G. Ryle (Ryle, 1949) when he discussed the mind-body problem. Such conflict between absolutist and relational concepts in the philosophy of mind has been documented by D. Chalmers (Chalmers, 1996) and evoked by G. Strawson (Strawson, 2006). Both authors formulated very similar suggestions aiming at overcoming this conflict in naturalist terms. Both of them

argued that there must exist in the world some hidden *non-relational* property (of either information processing devices or matter in general) which has no counterpart in the purely structural laws of physics, chemistry, and biology, but which is able to account for the absolutist status of conscious experience. However, when these authors try to characterize the hidden intrinsic property they refer to, they seem to be short of inspiration. In view of their construal of objective science as structural / relational, they can by no means make use of physical concepts for this characterization. They are then bound to use experiential concepts instead; those very experiential concepts whose origin they sought to identify in the physical world. In a circular way, they account for conscious experience by invoking the "psychical", or "proto-experiential" properties of information processing devices or material entities. Experiential facts being the only available model of intrinsic features, they project this model on the material process they hold responsible for the origin of conscious experience. They appear to be caught in the same *petitio principii* as Leibniz, who could think of the intrinsic properties of material entities only by assimilating these entities to "monads" endowed with (more or less refined) conscious experience. As Kant cogently adverted, the basis of the Leibnizian *Monadology* is the remark that "we can attribute to substances no other intrinsic state than that whereby we ourselves inwardly determine our sense" (Kant, 1996).

True, another, diametrically opposite, strategy, to overcome the gap between absolutist and relational concepts seems to be available at this point. It relies on a variety of structural realism, according to which *all there is* (and not only what science can describe) reduces to structure. This view has been advocated as a response of scientific realism against the challenge of paradigmatic changes and of the correlative instability of the set of entities that are accepted as existent at each stage of research. Several authors (Worrall, 1989, Ladyman, 1998) pointed out that structures are good candidates for the role of transparadigmatic invariants which entities can no longer play.

Unlike entities such as phlogiston, or caloric, or the individual corpuscles of classical physics, many laws survive across scientific revolutions, as limiting cases for the new theory. And some other structures (such as principles of invariance of laws with respect to any change of place, time, and orientation) are even more perennial, since they retain their full generative aptitude irrespective of the paradigmatic changes in concepts. But if indeed *all there is* is structure, it would be pointless to invoke further intrinsic properties of matter unaccounted by a purely structural physics (Seager, 2006). It would even be tempting to claim, in the name of the new structuralist ontology, that consciousness itself *is* pure structure; and that, therefore, the explanatory gap is (or at least can be) overcome.

But doesn't this mean ignoring the lesson of phenomenology we previously reminded? Isn't the deliberately non-phenomenological claim that consciousness itself *is* pure structure an ontological projection of a methodological choice in favor of objective science as the only source of knowledge? It seems that many scientists are satisfied with statements like: "with our structural objective science, we can only account for the structural features of the contents of conscious experience; we then declare that there *must* be nothing more in conscious experience than structure". However, this kind of statement is utterly doubtful from a philosophical standpoint, since it springs from an almost implicit epistemological bias. It remains doubtful even though it may be acceptable as a simplification for research and as a good shortcut to obtain interesting results without bothering about troublesome conceptual issues.

The two antagonistic strategies to overcome the explanatory gap between a relational objective realm and an absolute experiential realm thus appear to have failed. Neither absolutizing some properties of matter, nor relativizing / structuralizing experience, works in principle.

3) Argument from the mixing up of the "easy problems" of neurophysiology with the

"hard problem" of the origin of consciousness

As suggested in section 1, the question about *where* and *how* consciousness arises in the objective world as described by physics or physiology indeed turns out to be very challenging. It is challenging for reasons of principle as we have just seen in section 2. But these reasons of principle also have a practical counterpart. In a few words, the "category mistake" which is involved in a question about the origin of the absolute from the relational, of the subjective from the objective, manifests itself in practice by a *regressio ad infinitum*. Let us see how this *regressio* arises.

Those who think that it is possible to provide the question about the material origin of consciousness with an answer ask for patience. They point out that many easier problems can be and have been solved in the past, such as the problem of finding neurophysiological *correlates* of certain structural features of experience. And, they go on, one may hope that if enough of these easier problems are indeed elucidated in the future, a solution of the harder problem of the physical or biological origin of consciousness will automatically arise. This hope, or this belief, has proved to be a powerful incentive for research in neuropsychology; and the outcomes of research developed under its impulse are impressive. But there are good grounds to think it is *illusory*.

Let me try to show that advances concerning the easy problems of consciousness afford us *no clue whatsoever* about the harder problem of its physical origin.

To that purpose, I'll consider the example of perception of *color*, showing that, even though more and more easy problems about it are solved, the hard problem remains exactly as elusive as ever.

Two centuries ago, physicists could claim to have clarified this issue by displaying a relation between the *wavelength* of light radiation and the perception of color as reported by human subjects. Couldn't one

call this a true “reduction”⁵ of an experiential feature to physics? Couldn’t one say that color *is* wavelength of electromagnetic radiation, full stop? The answer is “No”, for two types of reasons :

- A. The correlation between wavelength and perceived color is imperfect.
 - i. Electromagnetic radiation with a given wavelength (or “monochromatic” light) gives rise to a well-defined class of experiences of color only under *normal physiological circumstances* ;
 - ii. No straightforward juxtaposition of the experiences of colors is obtained when several fluxes of monochromatic light of various wavelengths are combined. For instance a combination of monochromatic red and monochromatic blue does not give rise to “blueish red” but to *violet*, which corresponds to a monochromatic radiation with wavelength shorter than *both red and blue* ;
 - iii. Moreover, similar perceptions of color can be associated to very different mixtures of light of different wavelength and intensities.
- B. There is no conceptual connection, no way to figure out a passage, between the numerical value of a wavelength and *what it is like to experience* redness or blueness.

The first kind of reason proved accessible to further research. Maxwell modeled the system of relations between perceived colors by means of a triangle of three primaries with chromatic coordinates. This model was soon related to the presence of three types of cone-cells in the retina, with different photopigments. One could then understand some of the physiological reasons why the

connection between wavelength and perceived color is not straightforward, especially when several fluxes of light of various wavelengths are combined. In other terms, one could account for certain *structural* features of reported experience of color in terms of elaborate *structural* features of relevant objects (Petitot, 1999). Much later, in the second half of the twentieth century, one could even claim knowledge of the place where retinian information about impinging electromagnetic radiation is processed and articulated to other pieces of information. Cone-cell outputs are first sharpened by the neuron-system of the retina and, after a complex path in the brain (through the lateral geniculate nucleus), they project on the V1 area of the occipital lobe of the cortex. This part of the brain cortex in turn involves “color columns” of cells that only fire according to a certain chromatic signal. Doesn’t this mean, at last, “reduction” of an experiential feature to physiology? Here again, the answer to this question is “No”, for the two *same* reasons we already documented. No strict correspondance between reported experience of color and activation of the color columns of V1, and no conceptual bridge between experience and physiology. The correspondance was soon improved by further physiological work that disclosed the existence of the V4-V8 area, whose activity is in closer correlation with the subjective appreciation of color, including when shade and diurnal variations of light occur. But, here again, the perception of color may involve many additional ingredients such as motion, orientation, and activity, that are processed in other parts of the brain (Varela et al., 1991 ; Cohen, 2006 ; Churchland, 2007) ; and moreover we still do not have the slightest clue about *how* the rythmic firings of a bunch of neurons of V4-V8 can give rise to *experienced* red or blue (not any more than about how experience could arise from activity in any other part of the living body). And so on, and so forth... No end is in sight for this research about how experience of color arises, despite its remarkable fruitfulness.

To sum up, one witnesses an open-ended progress towards closer and closer

⁵ The analogy between reducing experience to neurophysiology and reducing heat to average molecular kinetic energy is highly disputable, and it is partly based on an ambiguity. What has been subject to a partial intertheoretical reduction is the objective thermodynamical function Q (heat), and by no means the *quale* “hotness”. In fact, the whole process by which the objective function Q has been extracted and the quale “hotness” has been progressively pushed aside as purely “subjective” is usually ignored or overlooked (Peschard & Bitbol, 2008). Hence the spurious alternative between reductionism and dualism (see section 1 of this paper).

correspondance between the structure of brain processes and the structure of conscious experiences as expressed by reports. But this open-endedness is associated with *incompleteness*. Scientists who believe that solving many such “easy problems” about consciousness will finally clear up the harder problem of its physical origin, look like somebody who believes one can finally reach the horizon by walking far enough. In the same way as the walker ignores the *category gap* between a line *in* space and an apparent line seen *through* space, these scientists ignore the category gap between the exclusively structural connections provided by science and the absolute of experience analyzed through a structured framework. In fact, these scientists come up against the unavoidable consequences of the initial decision to extract a structural residue of experience in order to build an objective science (what I called the epistemological argument); but since they have overlooked or forgotten this decision, they are unable to recognize the persistence of the “hard problem” as the most glaring symptom of the bounds which are inherent to a knowledge based on it. They have forgotten that objective knowledge is made possible by carving the lacuna of first person experience within it; and they cannot recognize that, therefore, the lines of progress of objective science can only have an *apparent* convergence towards consciousness in the infinite. Being oblivious of the blindspot of knowledge, they disregard the inaccessibility of the perspectival eye-point that represents its projection in the field of the known.

True, some scientists are more prudent. They content themselves with pointing out that very often science dismisses problems rather than trying to solve it, and concentrates on other problems, more tractable by its methods (Mills, 1996). According to them, science does not have to tackle every problem that can be formulated. Instead, its major task is to tell apart those problems that are liable to a scientific study and banish any other problem as “unscientific”. And, they claim, unlike the issue of isomorphism between objective

structures and experiential structures, the “hard problem” of the origin of consciousness is likely to be among those “unscientific” problems.

But even their attitudes and their values tend to maintain our culture as a whole in a sort of schizophrenic state. On the one hand, we are so impressed and enthusiastic about the remarkable results of neurophysiology that we are prone either to ignore the yawning gap it leaves behind it, or to dismiss it so strongly as “unscientific” that we finally *think* it does not really exist. On the other hand, however, we *experience* the gap in our conscious lives, we realize that “describing is not living” (Edelman & Tononi, 2001), and we partake of the existential crisis of our civilization as a whole.

4) Argument from neuropsychology against functionalism

One glaring symptom of our radical inability to figure out a passage from objective properties to lived experience, is the ongoing debate about the proper *level of organization* at which this passage must be sought.

Some claim that the relevant level concerns generic *functions* in a network of information processing. This is the so-called functionalist thesis, which implies that, provided certain organized informational fluxes are in place, experience may arise irrespective of the material basis on which these fluxes are implemented. In other terms, functionalists imply that even *computers* could be endowed with conscious experience, provided they have a certain functional structure, imposed by some appropriate *software*. But many arguments have weakened the functionalist thesis.

One such argument has been drawn from the well-known and hotly debated experiments performed by B. Libet in the 1970's. Libet's experiments tend to show (at least under one of their most popular interpretations) that conscious experience of having decided to move is delayed by several hundreds of milliseconds with respect to the corresponding electroencephalographic “readiness potential” in the brain. Therefore, they irresistibly suggest that consciousness is at most *epiphenomenal* with respect to any

informational function (Gray, 2004, Flanagan, 1997). This suggestion is all the more likely when we consider the fact that these functions (including the higher-order functions of “*binding*” which is often seen as a typical task of consciousness) can be described in physiological terms throughout, irrespective of whether the material process involved are associated or not with conscious experience. At this point, the question about how conscious experience arises looks more mysterious than ever, since the issue of informational functions appears irrelevant to it.

By the way, the insistent comeback of epiphenomenalism might also challenge a major resource of those who claim the physical origin of consciousness. This resource is the now commonplace idea that consciousness have a survival value of its own, presented as an explanation of its alleged progressive *emergence* from a physical-physiological substrate during biological evolution. How can one claim that consciousness have a *specific* survival value, if every single function it fulfills can be accounted for in narrowly physiological terms, irrespective of whether the described physiological processes are associated with conscious experience or not ?

Another argument against functionalism has been drawn from experiments about word-color *synaesthesia* (perception of a well-defined color when a word is told or read). It has indeed been found that the experience of color does not depend at all on whether it is triggered by direct visual perception or by indirect induction by a word. Since vision and verbal processing have very different *functions* in the informational network of a human being, this strongly suggests that function cannot determine qualitative experience by itself (Gray, 2004).

Therefore, many authors now tend to think that the right level of passage from an objectified nature to conscious experience is located *below* the level of functions. This is made likely by the outcome of the experiment we evoked above : even though a direct perceptive experience of color and a synaesthetic experience of color do not involve the same function, they are both

correlated with an activation of the same cortical area V4-V8. The passage might then be located at the level of brain tissue, at the cellular level, or even at the level of certain elementary physical processes in the molecular content of neurons. Yet at this point, there is still absolutely no clue about the nature of the passage, namely about *how* conscious experience is generated. The authors who support the “tissue” view, still claim that consciousness must “emerge” from some neural basis because experience reports are more closely correlated to well-defined brain processes than to abstracted functions. But they have no inkling whatsoever about how to go beyond such an improved correlation.

True, some of them claim that they can establish a *causal* link between brain tissue excitation and experience, since electrical stimuli of (say) V4-V8 can trigger experience of (say) color⁶. This kind of evidence, they insist, goes clearly beyond mere correlation, and this supports the idea of an ontological primacy of brain processes over conscious experience. But any lopsided conception of the connection between brain tissue and experience, any conceptual asymmetry in favor of physiology, is still precluded by the fact that an influence can be exerted in the opposite direction as well. After all can't one alter the electrophysiological or chemical configuration of parts of the brain (say V4-V8) by orienting one's experience voluntarily (say by asking someone to imagine a colored surface) ?

This possibility of “downward causation” from experience to physiology could be taken by some as mere evidence that conscious experience emerges from a neurophysiological basis in the “strongest” sense of the concept of emergence, as opposed to “weak” emergence (Bedau, 1997). But suppose we add a further constraint. Suppose we adopt a very strict criterion of emergence. Galen Strawson recently submitted this type of criterion : “For any feature of E (anything that is considered

⁶ Salvatore Maria Aglioti, private discussion after the conference “Festival della scienza”, Genoa, Italy

to be Emergent from the Basis B), there must be something about B and B alone in virtue of which E emerges, and which is sufficient for E" (Strawson, 2006). The problem is that, as I have suggested more and more insistently, there is *nothing specific* about functions, neural tissues, or molecular structures in virtue of which conscious experience should emerge. Any loose talk of emergence of consciousness from brain processes in the name of mere *correlations*, or even mere experiments of mutual triggering, then appears to be ruled out by this strong criterion. From the latter negative statements, Strawson infers that conscious experience is nothing emergent at all. Combining this inference with a materialistic monistic principle, he concludes in favor of *panpsychism*, or rather pan-experientialism⁷. But, then, his problem is to explain how micro-experiences "add up" to full-fledged human consciousness. Moreover, it is not easier to understand why and how an atom has elementary experience than to understand why and how a living human brain has an elaborated consciousness. Ascribing micro-experiences to atoms just seems an *ad hoc* additional postulate about matter.

So, at this point, we are still completely stuck, with no idea whatsoever about how to handle the "hard problem" of the origin of conscious experience in an objectified nature.

5) Argument from fine-grained analysis of the neural correlates of deprivation of consciousness

Even the modest results about the Neural Correlate of Consciousness (NCC) turn out to be partly questionable due to a methodological constraint. In order to assess the presence or absence of conscious awareness in a subject, a neuropsychologist has to rely on quite complex and integrated behavior. He/she relies on *language*, or,

⁷ An intermediate position is held by Francisco Varela, who thinks (on the ground of an analysis of first-person experience) that consciousness arises when there is elementary "concern", namely polarization towards conditions of survival. In that case, the most elementary conscious beings would be unicellular living organisms, with their polarization towards food (or threats).

more generally, on the ability to *report* experience by way of signs or gestures. Ability to report in turn implies ability to *discriminate* between specific events, a high level of *reflectivity*, and a reasonable amount of *memorization*. So, how can we be sure that, when no report can be obtained, there *is no experience at all*? Couldn't it happen that certain elaborated abilities needed to *make the report* are defective, although some sort of experience was indeed lived by the subject⁸?

Let's then look at the current neurological theories of the *locus* of conscious experience with a critical eye, and with this methodological problem in view. Here is a non-exhaustive list of four contemporary neurological views of consciousness :

- Edelman's theory of signal "reentry" states that conscious experience requires feedback loop interaction of regions of the cerebral cortex with nuclei in the thalamus ;
- Dehaene's and Changeux's theory stipulates that for conscious experience to arise, a "Global workspace" connected with many specialized areas of the brain must be activated ;
- Rosenthal's theory says that higher-order thoughts (and associated recursive activity of the neurons) are indispensable for experience ;
- Finally, Crick's and Koch's theory claims that an area can be a NCC only if it is connected to the frontal executive cortex, and if it is thus able to partake of the process of "binding" various fluxes of information by way of a synchronized firing of neurons at 40-70 Hz.

All these views rely heavily on the subject's ability to discriminate, to memorize, and to *report*, which is used as the ultimate experimental criterion. Can we preclude the possibility that the large-scale

⁸ Even Libet's experiments have been criticized on this ground. Indeed, in order to assess the moment of their conscious decision to move, the subjects were asked to notice the position of the needle of a sophisticated clock at the very instant when their decision was taken. Couldn't it be that the time interval between the "readiness potential" and the so-called "decision" to move, is in fact a time interval between the potential and the ability to notice the position of the needle discriminately enough to give a reliable report about it?

synchronization of neural activity deemed indispensable for *consciousness*, is in fact only required for *interconnecting* a number of cognitive functions including those needed for *memorizing* and *reporting* ?

Many experimental results feed this doubt. Some results of this sort concern split brain patients (Sperry, 1983). In these patients, the left hemisphere of the brain may have all the areas corresponding to a full visual perception activated, whereas the right hemisphere is at rest. Relying on the ability to report by language would yield the conclusion that the left hemisphere's activation is correlated to no visual experience at all. But there are other, non-linguistic, ways to obtain reports from the left hemisphere (e.g. by asking answers by means of a lever activated by the *right* hand), and in this case the conclusion is that its activity is indeed correlated to experience. Can we preclude that smaller areas of the brain are correlated to experience, although one can hardly figure out how to get any report from them?

Other experiments reinforcing this doubt can be found in the literature about deprivation of consciousness by sleep, fainting, anesthesia, coma, epileptic seizure etc. Let me take the example of general anaesthesia, whose depth can be modulated according to the dose of drug which has been administrated. Growing doses of a certain class of anaesthetic drugs were tested on voluntary subjects, and they were correlated with progressive drop of the average coherent frequency of electroencephalographic waves (Zeman, 2002). When the doses of drugs increased and coherent frequency decreased, mental abilities were lost step by step, one after another. At first, subjects lost some of their appreciation of pain, but could still have dialogue with the experimenter and remember every event. Then, they lost their ability of later recalling explicit memories of what was going on, but they were still able to react and answer demands on a momentary basis (this latter finding can be connected to situations where, with another class of anaesthetic drugs, patients could react during anaesthesia by pressing a lever, and could

even report pain *at that moment*, but did not remember anything *after the event*.) With higher doses of drugs, patients also lost ability to respond to requests, in addition to losing their explicit memory ; but they still had "implicit memories" of the situation (namely, their reactions after anaesthesia were modulated by certain events during anaesthesia, although no explicit report of these events could be obtained). Experiments of this kind display a case of mutual dissociation of the faculties that are usually taken *together* as necessary to conscious experience. They open the further possibility that some sort of conscious experience can be present, although none of the usual criteria can be ascertained ; not even the most elaborate electroencephalographic criteria of consciousness since the latter were based *in fine* on an ability to elaborate reports, either immediately or retrospectively (Zeman, 2002 ; Van Lommel et al. 2001).

This lends some credit to Semir Zeki's alternative thesis in which every single region of the brain and every level of neural activity, weak or strong, can be correlated to something like "micro-consciousness". According to this theory, each micro-consciousness just happens to be amplified, stabilized, and interconnected to other micro-consciousnesses when larger and larger neural assemblies are activated, thus reaching the pathways of report. Along with this view, ability to report, as well as massive, correlated, and distributed neural activity, should not be taken as a sign of consciousness, but only of interconnection and synergy of micro-consciousnesses. As for *absence* of report, either immediate or retrospective, it could in many cases be caused by a lack of interconnectedness of microconsciousnesses rather than by the absence of any experience whatsoever. A similar thesis was formulated by Susan Greenfield's, according to whom consciousness is a "continuously variable" feature, growing with brain complexity, and progressively amplified in each given brain according to its ability to recruit larger and larger neural assemblies. Since no lower bound is fixed for the presence of micro-consciousness, we are once again slipping on

the slope that leads inexorably to panpsychism, but with no additional clue about the material origin of consciousness (Greenfield & Collins, 2005).

6) Argument from physics against materialism

The thesis that matter is primary and consciousness secondarily emergent from it, relies on physics in the last resort.

To begin with, this thesis may need physics as a provider of the hoped for (though till now inexistent) ultimate explanation of the material source of consciousness. Some researchers, such as Roger Penrose and Stuart Hameroff, thus tentatively ascribe the origin of consciousness to putative quantum processes going on in the brain. Their proposal is based on the following concatenation of assumptions :

- State vectors or wave functions undergo an “objective reduction (‘OR’)” ;
- OR is triggered by quantum gravitational processes ;
- Gravitational processes are non-computable ;
- Quantum coherence takes place in neuron microtubules, and this is the basis of a quantum information processing in the brain ;
- OR occurs in microtubules, thus suppressing coherences ;
- Conscious thought is non-computable (in view of anti-mechanicists arguments based on Gödel’s theorem) ;
- Consciousness therefore arises from microtubular OR (this is the final claim, submitted to experimental test by various ways of acting on microtubular coherences).

Unfortunately, *none* of these assumptions can be considered, by far, as unproblematic. They remain highly controversial, for several good reasons. Let me consider two of them.

The first group of reasons is internal to Penrose’s and Hameroff’s framework of assumptions. M. Tegmark thus challenged the view that the central feature of quantum processes, namely state superposition, can have a sufficiently long life-time in the brain

to fit the characteristic time scales of reportable experiences (which is at least 25ms). Indeed, he claimed, the decoherence time in microtubules is shorter than this characteristic time scales by several orders of magnitude (Tegmark, 2000). But a reply from Hameroff and his group (Hagan et al., 2002) showed that this issue of decoherence time is far too uncertain, in view of the various molecular environments that may surround microtubules, to be taken alone as a fatal blow against their quantum theory of consciousness.

The second group of reasons is much more serious, because it challenges the very initial premise of Penrose’s and Hameroff’s proposal. Their root assumption is that quantum mechanics is a theory describing the objective world, and that every single alteration of the formal elements of the theory is to be ascribed to a change in the objective world. In particular, any reduction of the wave function is descriptive of an objective process called “the objective reduction”. But this interpretation is only one among many, and it is not even the most credible one in view of its speculative “surplus” of hypothesis. By contrast, there are several interpretations of quantum mechanics in which wave packet reduction plays no role (or only an ancillary role as a practical tool for simplifying probability assessment). Some of the latter interpretations, from Bohr’s to Quantum Information, even challenge the underlying belief that quantum mechanics is aimed at describing anything “out there”. Instead, they point out, many features of this theory automatically lose their paradoxical flavor if one accepts that quantum mechanics might only be aimed at predicting the effects of our intricate relation (or “interface”) with the environment (Fuchs & Peres, 2003). In this case, far from being a picture of the world construed as detached from us, quantum mechanics would be a picture of the *bounds* of detachment in physics. This view is further supported by the fact that one may derive many formal features of quantum theories by assuming that they have indeed a purely predictive status, and that they express the bounds of the procedure of detaching

properties from phenomena (Bitbol, 1998). But in this case, it looks plainly absurd to entertain the hope that quantum mechanics will do the job of ordinary naturalism, namely the job of accounting for conscious experience by describing relevant parts of an objectified nature.

This does not mean, however, that quantum mechanics interpreted in this non-realist way can teach us *nothing* about the issue of conscious experience. In fact, this theory, and the way it turns a limitation of objective knowledge into an advantage, have important conceptual clarifications in store for the philosophy of mind as Bohr himself realized very early (Bohr, 1997). The first clarification comes from the essential feature of contextuality of phenomena that quantum mechanics has to cope with (Bitbol, 2000, 2002). Microphysical phenomena adhere to the contraptions in which they arise ; they are not independent of the experimental situation which makes them manifest ; accordingly, they cannot be said to “reveal” an underlying independent property, unless one accepts the “surplus” speculative scheme of hidden variable theories. In this very special situation, quantum physicists were nevertheless able to build an intersituationally and intersubjectively acceptable theory. They obtained intersubjective consent without detachment of an object in the usual sense of the word. They reached this aim by elaborating universally valid rules for *predicting* contextual phenomena (or “values of observables”), and by stating universally efficient prescriptions for mastering *directly* the *technological* implementations of the predictive rules, without the help of a fully consistent model of objects.

But this is exactly what is needed for a proper science of mind. Indeed, the epistemological situation of this science is exactly isomorphic to the epistemological situation of quantum physics. Just as microphysical phenomena adheres to experimental device and cannot be detached from them, conscious experience adheres to conscious beings and cannot be detached from them (as stressed in the introductory section, consciousness is not “something” but

what is lived by somebody). Just as there are no true “quantum properties” but only “observables”, there are no experiential properties, but only “livables”. In such situation, it is a bad strategy to look for a scientific naturalist account of the integrality of mind in the ordinary sense of defining and characterizing objects (either neural objects or mentalistic objects). Rather, one should look for ways to obtain intersubjective consensus about “livables”, and try to formulate prescriptions for acting on contents of experience without the help of a model of objects.

The second conceptual clarification that quantum physics can offer deals with the conception of *matter* it tends to promote. After all, the view that consciousness emerges from matter is bound to rely on physics to disclose *what is (or at least what is not) matter*, and in which sense *matter* can reasonably be taken as the basis of everything else including consciousness. Unfortunately, the answer modern physics has elaborated about this issue is highly unsettling for the materialist thesis (Bitbol, 2007). To sum up, this answer is that matter in the classical sense (namely a set of bodies extended in space and enduring in time) is nothing else than an *appearance* shaped by the coarse sense-organs or experimental apparatuses we use to explore our environment. According to standard quantum mechanics, the very features that define the identity of a material body (such as the position and the velocity) are no more than values of mutually exclusive observables, which can be assessed only by using macroscopic instruments. Quantum *decoherence* adds theoretical details about how material bodies at our scale only *seem* to be localized, *seem* to occupy a region space at a given time, and more generally *seem* to have the properties we ascribe to them (Joos et al., 2003). As for quantum field theories, they tend to go even further, since they do not only deprive elementary particles of matter of intrinsic properties but also of intrinsic existence. According e.g. to Paul Teller, particles are no longer “(...) the sort of things that are either There or Not There” (Teller, 1995). They are only the old-

fashioned name we give to a *potentiality* of quantized events of detection, embedded in certain group-structures (Wigner, 1939). To borrow a beautiful metaphor formulated by Jean-Marc Lévy-Leblond, they have the “mode of existence of rainbows”, because they depend of a relational network of conditions. What is left is only an abstract pattern of field-like dispositions holding for any experimenter, at any time, at any place, out of which matter-like *appearances* may emerge at our scale. As a consequence, matter as we know it, the matter of which our brains appears to be made out, is no more fundamental than anything else. Matter can then hardly be taken as the real stuff out of which everything else emerges, including consciousness.

Some philosophers who defend a “physicalist” view may be tempted to skip this difficulty about the traditional components of the definition of matter by saying something like this: “Consciousness is matter-based in a very general sense: it emerges from whatever physics *describes* as fundamental, be it a quantized field”. But the problem is not solved by this further flexibility. For, as I mentioned previously about the interpretation of quantum mechanics, modern physics cannot even be said to “*describe*” anything completely *independent* of the experimental and intellectual tools of investigation: it just affords a way of systematic *prediction* of what occurs *if* this investigation is carried out; and it establishes reproducible relations between these predictions. What is taken as objective by modern physics is no longer a conception of the ultimate stuff of which the world is made, but the very network of mathematical tools by which we can collectively anticipate the outcome of our most refined actions.

We then discover that, far from displaying the solid basis we were looking for, modern physics emphasizes *the very procedure by which objective knowledge is elaborated*. As we sketched it, the procedure consists in trying first to push aside any situated component in knowledge, then formulating a group of transformation, and finally singling out the universal invariants of

this group. But, reminding Husserlian phenomenology and the reflection of section 1, we realize that this procedure is universal, and that it stems from the most primitive operations one has to perform on the undisciplined conscious experience in order to achieve the carefully organized perceptive *stratum* of objective knowledge. At the root of this stratified protocol, which enables the constitution of more and more abstract objects (perceptive objects, classical bodies, and finally quantum group-theoretical invariants), we thus find conscious experience. Conscious experience is and must be taken as methodologically primary, including when the research which is meant to throw light on its so-called “material basis” is concerned.

At this point, we have almost reached our conclusion. Consciousness is *existentially, transcendentally, and methodologically* primary. Any attempt at showing how it can be *ontologically secondary* to material objects both fails and draws us back to its methodological primacy.

Final Remarks : Towards A New Stance

Can we go further? Should we say that conscious experience is *ontologically* primary? Or should we hold something like the highly seductive double-aspect theory formulated by Spinoza, in which conscious experience and material appearances are two facets of the same unfathomable stuff coarsely figured out by (say) the dispositional quantum field? I am reluctant to take this additional step. Indeed, the negative arguments I adduced are not sufficient to support any such thesis. And since no further positive evidence appears to be available even in principle (e.g. due to the supposedly pre-mental and pre-material nature of Spinoza’s putative “third stuff”), one can assert a double-aspect theory only at the cost of adding speculative elements.

Instead of an alternative *thesis*, what is then needed is an alternative framework of thought, or even better an alternative *stance* (Van Fraassen, 2002) of which a framework of thought is only a facet. The alternative stance involves a radically new approach of subjectivity: instead of being underrated,

neglected, or even despised, situated experience becomes a field of exploration on the same footing as the triangulated structural residue favored by the objective sciences (Wallace, 2000); instead of remaining like freely rotating cogwheels with no engagement with the rest of language, statements expressing first-person experience are reintegrated into the whole system of discourse, thus compensating for the spurious effects of the great dualist divide triggered by objectification (see section 1).

Two components of the new stance are especially relevant here. It is convenient to formulate them in a prescriptive form :

- The scope of the deepest question scientists usually accept *not* to ask should be expanded. Instead of renouncing to answer the question “why is there a physical universe rather than none?”, they should renounce to answer the broader question “why is there experience-of-a-physical-universe rather than nothing at all?”. Not because any question about the origin of consciousness is “un-scientific” (with a pejorative undertone), but because conscious-experience-of-something is itself the all-pervasive origin and presupposition of any endeavor, including the scientific enquiry about “things” and about a physical universe. After all, the most primitive “given”, the “world as I found it” (Wittgenstein), is neither an external universe nor a purely internal world : it is an inextricably united experience-of-a-world, out of which the poles of the usual duality are differentiated (Petitmengin, 2007).
- Our neuropsychological research should be more balanced : instead of focusing almost exclusively on careful elaboration of neurophysiology and physics, with the project of reducing conscious minds to brain processes, we should put exactly as much attention on cultivating the experiential side of the embodied mind, and try to establish what Francisco Varela referred to as “mutual generative constraints” between the mental and physiological domains. This would contribute to the dawn of another science

he called “neurophenomenology” (Varela, 1998).

At this point I will try to unfold the meaning of the latter prescription of what we can call the “Varelian stance” or the “neurophenomenological stance”, together with its potentially devastating impact on the debate about consciousness.

We have seen in section 2 of this article that, by contrast with the self-evidential *absoluteness* of conscious experience, any attempt at accounting for the latter by way of the law-like network of *relations* posited by an objective science is bound to fail. Conscious experience *as a whole* can by no means be reduced to structure. Yet, it is well known since Fechner’s *psychophysics* and Maxwell’s triangle of colors at least that (i) certain *contents* of conscious experience as reported by human or animal subjects are liable to systematic interconnection, and (ii) the nodes of this interconnected mesh can themselves be connected with objective phenomena and laws. Hence the program of the “Varelian stance” to amplify and systematize these results of psychophysics :

- (a) Do not try to *absorb* contents of experience or phenomenological reports into the structural network of objective science, either by elimination, or by reduction, or by a statement of *identity*. Rather, strive towards *embedding* phenomenological reports within a broader relational network, of which the law-like structure of the objective domain is only a fraction.
- (b) Avoid mere juxtaposition of an objective science with a poorly studied subjective realm. Instead, cross the threshold of a new amplified science with its own unprecedented structures. The following prescription is precisely aimed at reaching this aim.
- (c) If you wish to formulate a new amplified science, if you wish to embed contents of experience or phenomenological reports into a generalized relational network together with the laws of an objectified nature, do not content yourself with looking for *correlations* between previously established categories and

structures. Instead, show that the very process of interconnection between experiential data and objective data may give rise to new categories and new unexpected structures. Varela's notion of "mutual generative constraints" precisely points towards a process of reciprocal alteration and enrichment of experiential and objective concepts :

- o Phenomenological reports may help to pick out and ascribe meaning to previously unnoticed neural configurations (Petitmengin et al., 2006) ;
 - o Conversely, neurological findings may become an incentive for re-categorization and further development in phenomenological research (Depraz et al., 2002).
- (d) Show how objectivity arises from a universally accepted procedure of intersubjective debate. Do not construe it as a transcendent resource of which intersubjective consensus is only an indirect symptom. Draw inspiration from a careful reflection about physics : either from the process of emergence of objective temperature valuations from an experiential underpinning (see section 1), or from the model of quantum mechanics construed as a science of inter-situational predictive invariants rather than a science of "objects" in the ordinary sense of the word (see section 6). Then, recognize that intersubjectivity should be endowed with the status of a common ground for both phenomenological reports and objective science. Start from this common ground in order to elaborate the amplified variety of knowledge that results from embedding phenomenological reports and objective findings within a unique structure.
- (e) Do not rely on a minimal and most elementary form of intersubjective consent, but try to amplify the criteria of intersubjective understanding by refining the stability and sharpness of subjective experience. After all, the reason why numerical values and ratios are privileged as objects of intersubjective agreement is likely to be the fact that they are not too difficult to be agreed upon, even among

subjects with a poorly cultivated experience⁹. But if experience is systematically trained and educated, either in the first person by meditation (Wallace, 2000, Lutz et al. 2004), or in the second person by making explicit unsuspected features of experience in dialogue (Depraz et al., 2002 ; Petitmengin, 2007), or in a combination of first- and third-person modes by bio-feedback, the basis of possible intersubjective consensus is likely to expand beyond recognition.

- (f) Think about the most basic presupposition of the process of objectification and of establishment of law-like relations between objective quantities : a system of socially regulated *practices*. A *practice* of measurement, calibration, and elimination of noise ; a *practice* of feed-back loop between experimental activity and tentative theories ; above all a *practice* of distancing from subjective connotations of the findings and theories of science, despite their being so crucial as a creative background. A new set of practices including various types of experiential training in the *curriculum of studies*¹⁰ would deeply alter the *substratum* of research, stabilize an expanded version of the regulative ideal of intersubjectivity, and favor the new generalized paradigm of science which neurophenomenology forecasts.

Clearly, this program made of a set of prescriptions rather than theoretical statements (i.e. made of "ought" rather than "is") does not *solve* the "hard problem" of the physical origin of conscious experience. However, the reason for this non-solution is not that the problem is too difficult, but that in the proper stance it does not even arise. It does not arise because the physical world is no longer the standard of being, and objectivity is no longer the ultimate standard

⁹ One must not forget, however, that even this minimal ground of agreement must be educated. Not everybody is able to perform accurate measurements, or precise observations under the microscope.

¹⁰ As initiated, e.g., by the Mind & Life summer institute

of method. In the alternative stance, the standard of being is underpinned by a standard of self-evidence, and the methodological standard of objectivity is expanded into a more general standard of intersubjectivity. Then, in the same way as, according to Wittgenstein, “The solution of the problem of life is seen in the vanishing of the problem” (Wittgenstein, 1994), according

to Varela, the solution of the hard problem of consciousness is found in a certain stance and research program wherein the problem vanishes.

References

- Bachelard G. Etude sur l'évolution d'un problème de physique: la propagation thermique dans les solides. Vrin, 1973
- Bachelard G. La formation de l'esprit scientifique. Vrin, 1938
- Bedau M. Weak emergence. In: Tomberlin J (editor) Philosophical Perspectives: Mind, Causation, and World. Vol. 11, Blackwell, 1997
- Bitbol M. Schrödinger's Philosophy of Quantum Mechanics. Kluwer, 1996
- Bitbol M. Some steps towards a transcendental deduction of quantum mechanics. *Philosophia Naturalis* 1998;35:253-280
- Bitbol M. Physique et philosophie de l'esprit, Flammarion, 2000
- Bitbol M. Science as if situation mattered. *Phenomenology and the Cognitive Science* 2002;1:181-224
- Bitbol M. Materialism, Stances, and Open-Mindedness. in: B. Monton (ed.), *Images of Empiricism: Essays on Science and Stances, with a Reply from Bas C. van Fraassen*, Oxford University Press, 2007
- Blackmore S. *Consciousness, a Very Short Introduction*. Oxford University Press, 2005
- Block N, Flanagan O, & Güzelde G. (editors) *The Nature of Consciousness*, MIT Press, 1997
- Bohr N. *Atomic theory and the description of nature*. Ox Bow Press, 1987
- Cassirer E. *Philosophie der Symbolischen Formen*. Vol. 3, Chapter 5 (III), Meiner Felix Verlag, 2002
- Cassirer E. *Substance and Function & Einstein's Theory of Relativity*. Dover Publications, 2004
- Chalmers D. *The Conscious Mind*, Oxford University Press, 1996
- Churchland PM. On the Reality of Objective Colors. In : Cohen J and Matthen M (eds.), *Color Ontology and Color Science*. MIT Press, 2007
- Cohen J. Color and Perceptual Variation Revisited: Unknown Facts, Alien Modalities, and Perfect Psychosemantics. *Dialectica* 2006;60:307-319
- Cohen J and Matthen M. (editors) *Color Ontology and Color Science*. MIT Press, 2007
- Dennett DC *Consciousness Explained*. Little, Brown & Co., 1991
- Depraz N., Varela FJ, and Vermersch P. *On becoming aware: An experiential Pragmatics*. John Benjamins Press, 2002
- Edelman GM and Tononi G. *A Universe of Consciousness: How Matter Becomes Imagination*. Basic Books, 2001
- Flanagan O. Conscious inessentialism and the epiphenomenalist suspicion. in: Block N, Flanagan O and Güzelde G (eds.) *The Nature of Consciousness*, MIT Press, 1997
- Fuchs C and Peres A. Quantum Theory Needs No 'Interpretation'. *Physics Today* 2003;53:70-71
- Gray J *Consciousness, Creeping Up on the Hard Problem*. Oxford University Press, 2004
- Greenfield S and Collins TFT. A neuroscientific approach to consciousness", Laurey S (ed.), *Progress in Brain Research* 2005;150:11-23
- Hagan S, Hameroff SR, Tuszynski JA. Quantum Computation in Brain Microtubules? Decoherence and Biological Feasibility. *Physical Review E* 2002;65:061901
- Henry M. La Barbarie. P.U.F., 2001
- Husserl E. *Ideen zu einer reinen Phaenomenologie und phaenomenologischen Philosophie*. Max Niemeyer Verlag, 1993
- Husserl E. *The Crisis of European Sciences and Transcendental Phenomenology*. Northwestern University Press, 1970
- Hut P and Shepard RN. Turning the 'Hard Problem' Upside Down and Sideways. in: Shear J (ed.), *Explaining Consciousness, the Hard Problem*. MIT Press, 1998
- Joos E, Zeh HD, Kiefer C, Giulini D, Kupsch J and Stamatescu IO. *Decoherence and the Appearance of Classical World in Quantum Theory*. Springer-Verlag, 2003
- Kant I. *Critique of Pure Reason*. Hackett Publishing Company, 1996
- Ladyman J. What is structural realism? *Studies in History and Philosophy of Science* 1998, 29 : 409-424
- Mach E. *Principles of the Theory of Heat. Historically and Critically Elucidated*. Reidel, 1986
- Mannheim K. *Collected works volume 1*. Routledge, 1997

- Mills EO. Giving up on the hard problem of consciousness. *Journal of Consciousness Studies* 2006;3:26-32
- Peschard I and Bitbol M. Heat, Temperature, and Phenomenal Concepts. in E. Wright (ed.) *The Case for Qualia*, MIT Press (In Press), 2008.
- Petitmengin C, Navarro V, Baulac M. Seizure anticipation: Are neuro-phenomenological approaches able to detect preictal symptoms? *Epilepsy and Behavior* 2006; 9:298-306
- Petitmengin C. Towards the source of thoughts. The gestural and transmodal dimension of lived experience. *Journal of Consciousness Studies* 2007;14:54-82
- Petitot J, Varela FJ, Roy JM, Pachoud B. *Naturalizing Phenomenology: Issues in Contemporary Phenomenology and Cognitive Science*. Stanford University Press, 1999
- Petitot J. Morphological Eidetics for a Phenomenology of Perception. in: Petitot J, Varela FJ, Roy JM, Pachoud B. *Naturalizing Phenomenology: Issues in Contemporary Phenomenology and Cognitive Science*. Stanford University Press, 1999
- Ryle G. *The Concept of Mind*. Hutchinson, 1949
- Searle J. *The Mystery of Consciousness*. Granta Books, 1997
- Sartre J. *L'être et le néant*. Gallimard, 1943
- Seager W. The 'Intrinsic Nature' Argument for Panpsychism. *Journal of Consciousness Studies* 2006;13:129-145
- Shear J. *Explaining Consciousness, the Hard Problem*. MIT Press, 1998
- Sperry R. *Science and Moral Priority*. Basil Blackwell, 1983
- Strawson G. Realistic Monism: Why Physicalism Entails Panpsychism. *Journal of Consciousness Studies* 2006;13:3-31
- Tegmark M. The importance of quantum decoherence in brain processes *Physical Review E* 2000;61:4194-4206
- Teller P. *An Interpretive Introduction to Quantum Field Theory*. Princeton University Press, 1995
- Tomberlin J. *Philosophical Perspectives: Mind, Causation, and World*, Blackwell, 1997
- Van Fraassen B. *The Empirical Stance*, Yale University Press, 2002
- Van Lommel P, van Wees R, Meyers V, & Elfferich I. Near-death experience in survivors of cardiac arrest: a prospective study in the Netherlands. *The Lancet* 2001;358:2039 – 2045
- Varela F. Neurophenomenology: a methodological remedy for the hard problem. in: Shear J (ed.) *Explaining consciousness, the hard problem*, MIT Press, 1998
- Varela F, Thompson E, and Rosch E. *The Embodied Mind: Cognitive Science and Human Experience*. MIT Press, 1991
- Wallace A. *The Taboo of Subjectivity*. Oxford University Press, 2000
- Whitehead AN. *Process and Reality*. Free Press, 1979
- Wigner E. On unitary representations of the inhomogeneous Lorentz group. *Annals of Mathematics* 1939;40:149-204
- Wittgenstein L. *Philosophical Investigations*. Basil Blackwell, 1983
- Worrall J. Structural realism: The best of both worlds? *Dialectica* 1989;43:99-124
- Wright E. *The Case for Qualia*. MIT Press (In Press) 2008.
- Zeman A. *Consciousness, a user's guide*. Yale University Press, 2002.