

Commentaries on the Monograph

CITATION: Pereira Jr A. A Commentary On De Sousa's "Towards An Integrative Theory Of Consciousness". *Mens Sana Monogr* 2013;11:210-29

A Commentary On De Sousa's "Towards An Integrative Theory Of Consciousness"

Alfredo Pereira Jr*

ABSTRACT

De Sousa's comprehensive two-part review of a diversity of contemporary approaches to the study of consciousness is highly welcome. He makes us aware of a proliferation of theoretical and empirical approaches targeting a common theme, but diverging in many ways. He skilfully accomplishes a classification of kinds of approach, identification of the main representatives, their contributions, and respective limitations. However, he does not show how the desired integration could be accomplished. Besides summarising De Sousa's efficient analytical work, I make critical comments and briefly report my contribution for the integration project.

Key Words: Astrocytes; Consciousness; Feeling; Integration; Neurons; Representation

Introduction

I comment on the sections of De Sousa's critical reviews (De Sousa 2013a^[9], 2013b^[10]), highlighting the main contributions and limitations of the reviewed literature. I jump only a couple of sections, because they are not central for the project of theoretical integration. Having discussed definitional

*Ph.D. Adjunct Professor, São Paulo State University (UNESP), Institute of Biosciences, 18618 - 970, Botucatu, SP, Brazil

Address correspondence to: Dr. Alfredo Pereira, Universidade Estadual Paulista (UNESP), Department Educação, Instituto de Biociências, 18618 - 970, Botucatu, SP, Brasil. E-mail: apj@ibb.unesp.br

Received 12 Oct 2012. Revised 15 Oct, 19, 21 Nov. 6, 9 Dec 2012. Accepted 10 Dec 2012.

Access this article online	
Quick Response Code: 	Website: www.msmonographs.org
	DOI: 10.4103/0973-1229.104495

issues elsewhere (Pereira and Ricke, 2009^[21]), I do not find myself in a serious disagreement with his statements in the first section. Although not being the main focus of the reviews, definitional issues are important to establish the conceptual basis on which the scientific, philosophical, artistic, religious and/or commonsensical approaches to consciousness are constructed.

Here is a summary of the concept of consciousness that I assume in this commentary. Conscious processes are dynamic phenomena that occur to living individuals, when

- They process information about what happens in their bodies and the world;
- They attribute meanings to the information contents;
- They produce a feeling about what happens;
- This feeling modulates the processing of the information, influencing action and memory formation; and
- During their lives, internalised information kept in memory and recurrent feelings form endogenous feedback cycles, allowing them to construct conscious episodes autonomously, for example, in dreams and imagination.

The above statements imply that information processing and attribution of meaning are two necessary, but not sufficient, steps in the formation of conscious episodes. Only with the addition of feelings these processes do become conscious; otherwise, they remain unconscious, as in the cases of the operations of the immune system—the distinction of what belongs to the system and what does not—and the mechanisms of physiological stress, for example, the increase of cortisol, leading to diabetes and/or other unnoticed tissue damages.

The latter kind of process can be classified as an “unconscious emotion”, revealing an important difference between the concepts of *feeling* and *emotion*: Feelings are subjective states considered to be always conscious to some degree, while emotions are physiological and behavioural processes that can precede the occurrence of feelings and/or persist after the corresponding conscious state/process has faded away. For instance, a traumatic psychological event is likely to elicit a conscious feeling, which often leads to an unconscious emotional process that persists after the conscious state/process disappeared.

In the Western philosophical tradition, the concept of knowledge and the respective knowing subject were emphasised. This conceptual bias was inherited by cognitive neuroscience, with the consequence of conceiving consciousness as the faculty of forming representations about the world or about the cognitive subject himself or herself. Accordingly, such representations were related to patterns of activity of neuronal assemblies, conceived as the *neural correlates of consciousness*. Binary patterns of axonal conduction (i.e. states of “firing” or “not firing”) were further related to binary codes of digital computers; this kind of discrete physical medium was assumed as adequate to the instantiation of representations.

More recently, attention has been called to the nonrepresentational side of consciousness, composed of feelings, affects, emotions, and the like (e.g., Panksepp, 2005^[20], 2007^[19]). At the same time, brain scientists have called attention to the “other brain” (Fields, 2009^[12]), the networks of glial cells that process information parallel to neuronal networks, possibly supporting aspects of conscious experiences that neurons cannot instantiate. This progress in psychophysiological research opens new possibilities for an integrative theory of consciousness, advancing one step further the project raised in De Sousa’s papers. In my concluding remarks, I summarise this new approach towards an integrative theory of consciousness.

Neuroanatomical models: Too many correlates

Neuroanatomical models of consciousness are based on the premise that the identification of the location of brain modules or networks—evolutionary specialised for a specific conscious function—would help to explain the relation of consciousness to the brain or the mechanisms that support conscious processing of information. Based on this premise, or just for the progress of empirical knowledge, several authors have investigated conscious functions related to the reticular activating region, amygdala, insula, thalamus, precuneus, parietal and/or prefrontal cortex, as reviewed by De Sousa (2013a^[9]).

The use of functional magnetic resonance imaging is adequate for this method of investigation, leading to the discovery of the so-called “neural correlates” of conscious processes. There is a problem of abundance: Too many correlates have been found, preventing simplistic conclusions about the correspondence of structure and function (e.g., for visual consciousness, see Pereira Jr., 2009^[23]).

Researchers often speak of brain structures as if they had powers to produce conscious experiences; for instance, the insula would generate pain or pleasure; the amygdala would generate fear, and so on. One of the most invested regions is the prefrontal cortex (e.g., Del Cul *et al.*, 2009^[11]), often taken as the locus of our higher cognitive capacities and moderator of the interplay of cognitive with affective and emotional factors.

As the identification of single regions has been insufficient to establish a unitary (one-to-one) relation between activations and functions, recently the focus has reasonably changed to networks and systems, such as the *default*, *executive* and *salience* networks and the *reward system*. A search in PubMed with these terms retrieves hundreds of papers worth reading, but this surely needs a supplementary integrative effort to draw the relevant conclusions for a theory of consciousness.

The discovery of involvement of specific networks and systems in specific conscious functions has heuristic value, for example, medical intervention and

electromagnetic stimulation procedures, but does not add to the understanding of the mechanisms of consciousness, for a simple reason: Nobody has found notable differences between these regions and circuits at the cellular and tissue levels. There are some details that differ, but knowledge about them does not help to explain why these regions and circuits instantiate different conscious functions and contents.

The physiology of the neuron is almost the same in all parts of the brain. Intracellular signal transduction pathways and intercellular connectivity patterns are also similar. The result is that knowledge of functional neuroanatomy explains the workings of consciousness as much as knowing the map of a country explains how people live there. This limitation leads us to the next section of De Sousa's first review paper, focussing on molecular and cellular approaches to consciousness.

Neuronal models: One step forward, one backwards

In this section, De Sousa discusses molecular and cellular processes involved with consciousness. While attention to the micro- and mesoscopic levels of activity surely increases our understanding of brain mechanisms underpinning conscious processes, there is also a limitation of this part of the review, due to the "neurocentric" approach adopted by him. This term refers to the "Neuron Doctrine" advanced by Ramon y Cajal, considering the neuron as the structural and functional unit of the nervous system (Bullock *et al.*, 2005^[6]).

In the first given example, the participation of glial cells is completely neglected: "Often it is assumed that consciousness emerges not before several hundreds milliseconds after stimulus onset...given the short time constants of membranes of neurons, recurrent connections are obviously necessary to store and process the stimulus before consciousness is reached" (De Sousa, 2013a^[9], p. 100-50). The above quotation can be contrasted with a recent publication (Pereira Jr., 2012^[27]), where I argue that the timing of consciousness is generated by a combination of neuronal *and* astroglial timings, the first in the scale of milliseconds and the latter in the scale of seconds.

Like many neural network modellers, De Sousa believes that "complexity is for sure of primary importance for consciousness because networks with the same number of neurons can create trivial as well as complex behaviour depending on their connectivity" (De Sousa, 2013a^[9], p. 100-50). However, it is not clear how the quantity of connections or the patterns of feed-forward or feed-back (recurrent) connections would make a great difference, turning ordinary electric signals into conscious contents of experience. Too many recurrences would be needed to explain, for instance, event-related potentials that take more than half of a second to produce a conscious outcome. A recurrent signal is just another action potential, so there is no reason to expect that the magic of consciousness could

derive from it. In any case, what could one thousand recurrent signals produce that one hundred do not?

Molecular and cellular hypotheses about consciousness are invariably based on important aspects of brain physiology, which are inflated in the attempt to explain conscious processing. In this regard, for example, “the apical dendrite activity theory takes...the apical dendrite part of thalamocortical circuits as the generator of consciousness”, for the reason that “the length of the apical dendrite increases the stability of the thalamocortical circuit activity and thereby increases the stability of the apical dendrite wave activity itself” (De Sousa, 2013a^[9], p. 100-50). The NMDA hypothesis assumes that “occurrence of states of consciousness critically depends on specific class of computational processes that are mediated by the NMDA synapse...NMDA receptor activity is capable of forming representational states in the brain and all general anaesthetics ultimately inhibit NMDA receptor activity” (De Sousa, 2013a^[9], p. 100-50). For the more famous neural assembly hypothesis, consciousness is believed to be based on the synchronised activity of neurons. One version of this hypothesis is the thalamocortical system functioning “on the basis of temporal coherence embodied by simultaneity of neuronal firing based on passive and active dendritic conduction along the apical dendritic core conductors. This results in the thalamocortical resonant column, which compromises the basic functional unit of consciousness” (De Sousa, 2013a^[9], p. 100-50).

Needless to say, all these hypotheses grasp an important aspect of brain physiology supporting consciousness, but hardly explain conscious processes as a whole-if only because their theoretical framework is limited to physiological language and consciousness cannot be fully described in these terms.

Altered states: A multiplication of theoretical problems?

The section on altered states in De Sousa’s first paper begins with an efficient review of the very important work made by Steven Laureys, Melanie Boly and collaborators on reduced states of consciousness (De Sousa, 2013a^[9], p. 100-50). Their results point towards a new paradigm for interpretation of fMRI and EEG data, focussing on transient coupling of activity of several brain regions. This approach is in accordance with Giulio Tononi’s Information Integration, as well as with Bernard Baars’ Global Workspace theory of consciousness (both properly referenced in De Sousa, 2013a^[9]). Not surprisingly, Laureys co-authored with Baars some years ago (Baars *et al.*, 2003^[3]; Baars and Laureys, 2005^[2]) and Tononi co-authors a recent publication from the group, one that has a potential to impact the field of consciousness science (Rosanova *et al.*, 2012^[29]).

Research on anaesthesia also has a positive contribution to consciousness theorizing. Although the mechanisms of anaesthetics are still largely unknown, a partial consensus was formed about the “anaesthetic cascade” of events reviewed

by De Sousa. However, Stuart Hameroff and colleagues have argued that volatile anaesthetics have a different pathway of action, resulting in the perturbation of quantum processes in microtubules, which they consider to be necessary for conscious processing (Hameroff *et al.*, 2006^[14]).

Other kinds of altered states, such as under the effect of psychoactive drugs, may bring new pieces of evidence to solve the consciousness puzzle, but De Sousa completes his section with cases of “split-brain” patients. These cases are interesting because of the consequences of the surgery on phenomenal consciousness, making it possible to establish a correspondence between split brain hemispheres and split conscious processes, thus favouring a monist conception-for which brain and mind are two aspects of the same system.

Although the interpretation of altered states is itself a theoretical problem, the combination of their related evidences has proven to be useful for the progress of knowledge about consciousness.

Cognitive psychology: When top-down does not meet bottom-up

In the section on cognitive psychology, De Sousa first reviews a controversy about the relation of attention and consciousness, focussing on the CODAM model. Authors of CODAM propose that corollary discharges would have a role in attentional control. This model “has been interpreted as possessing the ability to create both the conscious experience of content, as well as providing a neural underpinning for the phenomenological experience of ownership” (De Sousa, 2013a^[9], p. 100-50). It is implied that a mechanism for motor control of attention would explain consciousness. Although the relation of movement, attention and consciousness is still an open issue, the CODAM model should contain a part of the truth, since our conscious contents are not about the brain itself but about the world we move in (in other words, they are “projected” into the world, as proposed by Velmans, 1990^[33]).

In the middle of the section, De Sousa reviews cognitive models of consciousness-some of the best approaches we have, but often unrelated to micro- or mesoscopic brain mechanisms. As the Global Workspace is well known, he reasonably spends more efforts to review the less known Cognitive Hierarchy model, which seems to be favoured by him for integrative purposes. Combining it with another well-known model (The “Two-Systems” approach, S1 being “heuristic, affective, and intuitive”, and S2 “deliberative, cognitive, and rational”), he suggests that mental phenomena could be categorised in three axes: Automatic/controlled, conscious/nonconscious and cognitive/emotional.

The rest of the section is devoted to the discussion of contemporary views of the automatic/controlled dichotomy in relation to perception and action, and the relation of consciousness with learning. An emergent theme is the

possibility of sophisticated modalities of unconscious learning, revealing that the “automatic versus controlled” dichotomy is too simple to account for our cognitive abilities. A final mention of the important ART model is well placed, since this model stresses the importance of “bottom-up” processes meeting the “top-down” ones, thus generating a resonance that somehow contributes to conscious processing. Although the authors of the theory did not clarify the nature of this resonance, it surely points towards the integrative approach desired by De Sousa.

Self-consciousness: Not a general feature of consciousness

The discussion of morality in relation to consciousness suffers a conceptual circularity: Self-consciousness is important for morality, while morality is based on the existence of self-consciousness.

To begin the discussion, it is necessary to make assumptions about the concept of *self* and how it relates to brain and behavioural functions. There is an interesting controversy about the location of the self: Is it based on vision (our dominant sense) or in evolutionarily older modalities such as touch? The authors reviewed by De Sousa present two possibilities: “In certain pathological conditions, as during an out-of-body experience, the self can be localized at the origin of the visual perspective even though this location is different from the seen location of one’s body”, but “the data suggest that participants localize their self where they perceive to be touched” (De Sousa, 2013b^[10], p. 151-209).

Another question that may be posed is whether the self should be related to sensation and perception or if it belongs to higher levels of elaboration, as interpersonal relations: “Consciousness is to some extent a social phenomenon. Though each individual has his own distinctive point of view on the world, a good deal of the content of individual experience is picked up from contact with others...just as consciousness depends on the wirings of the social mind, social mindedness may depend on the wirings of consciousness” (De Sousa, 2013b^[10], p. 151-209). Although the reasoning makes sense, it falls on circularity again: The social mind is called to explain individual consciousness, while the latter is called to explain the former.

One possible solution is to consider the mind as social (including a “collective unconscious” of the Jungian kind), while consciousness is to be conceived as individual. This solution would help to conceive the ground upon which is built the autonomy of the individual. His or her mind absorbs information from the physical and social environment, elaborating on it to construct a singular personality and self-consciousness. This solution implies an ontological difference between mind (conceived as a system of information) and consciousness (conceived as a subjective take on information, as discussed in my last section).

Artificial consciousness: Exploring possible worlds

One of the advantages of the artificial consciousness approach would be that “asking questions about phenomenal consciousness in machines and building models could improve our understanding of human consciousness and take us closer to a solution to the hard problem”. An open question that arises is whether “the reproduction of human behaviour, cognitive states, or internal architecture leads to real phenomenal experiences” (De Sousa, 2013b^[10], p. 151-209).

Decades of artificial intelligence studies taught artificial consciousness researchers how to reproduce cognitive functions, but not how to reproduce affects and emotions. In spite of this limitation, some authors claim to have reproduced them: “Haikonen’s architecture also includes emotions—for example, there is an analogue of pain, which uses information about physical damage to initiate withdrawal and redirect attention.” (De Sousa, 2013b^[10], p.151-209). However, in this case there is no real pain—only a simulation that has an effect on observable behaviour.

The deeper conceptual problem is that affective states cannot be reduced to representations; the very concept of feeling is missing in these approaches. In my last section, I additionally argue that the instantiation of feelings requires the consideration of wavelike dynamics.

Feelings are also absent in the new field of Synthetic Phenomenology: “To be synthetically phenomenological, a system S must contain machinery that represents what the world and the system S within it seems like, from the point of view of S. An unpacked version of this definition is used by Aleksander and Morton to argue that their kernel architecture is synthetically phenomenological, whereas the global workspace architecture is not” (De Sousa, 2013b^[10], p. 151-209). In this sentence, egocentric representations are taken as lived experiences, forgetting that the latter contain nonrepresentational aspects such as feelings and emotions.

There is a perspective of interdisciplinary collaboration. On the one hand, “Neurophenomenology...the description of human phenomenology from a third person perspective using measurements of brain activity gathered using techniques, such as fMRI, EEG or electrodes... is easier than synthetic phenomenology”. On the other hand, “both disciplines are attempting to use external data to identify phenomenal states in a system, and there is considerable potential for future collaboration between them” (De Sousa, 2013b^[10], p. 151-209). There is also a limitation for this collaborative effort, since AC is not a study of our actual world, where consciousness occurs to living individuals, but the exploration of possible worlds, where there may be different kinds of consciousness.

Miscellaneous facets: The background becomes a figure?

Some issues treated by De Sousa in a section entitled *Miscellaneous Facets and Approaches to the Study of Consciousness* turn out to be important for the desired theoretical integration. He begins by resuming the limitation of anatomical approaches: "There is no final integrator station in the brain, one which receives input from all visual areas; instead, each node has multiple outputs and no node is only a recipient" (De Sousa, 2013b^[10], p. 151-209). This reasoning leads naturally to the "binding problem": "If any binding occurs to give us our integrated image of the visual world, it must be a binding between micro consciousnesses generated at different nodes. Since any two micro consciousnesses generated at any two nodes can be bound together, perceptual integration is not hierarchical, but parallel and post conscious" (De Sousa, 2013b^[10], p. 151-209).

At this point he makes an unjustified assumption, that the nodes of a neuronal network generate "micro-consciousnesses": "Visual consciousness consists of many, functionally specialised, micro-consciousnesses which are spatially and temporally distributed if they are the result of activity at spatially distributed sites" (De Sousa, 2013b^[10], p. 151-209). He neglected the possibility of distributed neuronal fields being unconscious, and their information content becoming conscious only at the moment of perceptual binding. In this case, there would be no separate *micro-consciousnesses* and the binding process would be pre-conscious.

As proposed by Gestalt theories long ago, we are always conscious of episodes as a whole. In other words, *we are not conscious of our micro-consciousnesses*. This contradiction destroys any argument for micro neuronal fields or assemblies being conscious by their own powers. As if he was aware of the contradiction, De Sousa argues for an approach to perceptual integration compatible with his assumptions: "If integration occurs between different nodes, the communication between them must influence the micro-consciousness that each creates in a consistent way, leading to consistent, integrated percepts" (De Sousa, 2013b^[10], p. 151-209). However, one can ask: Why are we conscious only of the integrated percepts? We perceive integrated scenes; even in pathological cases, the dissociation is always partial, and the gap in conscious content is restricted to the missing contributions of the lesion regions.

De Sousa is inclined to agree with authors who deny the binding problem. Assuming the existence of *micro-consciousnesses*, he questions, "whether they are bound at all, given what appears to be the nonunitary nature of conscious experience" (De Sousa, 2013b^[10], p. 151-209). There is an epistemological mistake here, since the argument was based on the distributed parallel character of neuronal processes, not on the phenomenology of conscious experiences. What could be a first-person, phenomenological evidence for consciousness being nonunitary? The author also quotes Anne Treisman, who made a critical analysis of binding issues, but does not mention that her landmark paper (Treisman,

1996^[32]) argues for a diversity of binding processes, many of which are relevant to a theory of consciousness.

After discussing micro-consciousness, De Sousa turns to what he calls “computational neuroscience”, but his usage of this expression is not standard (the one that refers to ordinary computational mechanisms of the brain, such as sparse population coding and frequency modulation of spike trains). He states that “the theory described by computational neuroscience suggests that it feels like something to be an organism or machine that can think about its own thoughts...raw sensory and emotional subjective feelings arise secondary to having evolved such a higher order thought system, and that sensory and emotional processing feels like something...when there are thoughts about the system, i.e. higher order syntactic thoughts (HOSTs), and the system is reflecting on its first-order thoughts...then it is a property of the system that it feels conscious” (De Sousa, 2013b^[10], p. 151-209).

There is a category mistake in the above reasoning: How could operations of some kind of entity on itself produce another kind of entity, that is, how could thoughts about thoughts generate feelings? This mistake only stresses the need to look for the specificity of affective and emotional processes. After proposing his three axes for categorisation of mental processes, De Sousa does not deepen the analysis of subtle differences between the concepts of feeling, sensation, affect, emotion, mood and others, and remains attached to the view that affects and emotions depend on representational processes: “Both emotion and consciousness depend on neural representations of the subject’s own body” (De Sousa, 2013b^[10], p. 151-209). This statement seems to imply that when we focus on the external world we develop cognitive processes, and when we focus on our body we develop emotional processes. The difference between cognitive and emotional would be given by the referent of the intentional act, but this view is not plausible in light of well-accepted philosophical theories of intentionality, claiming that the referent does not determine the meaning of a mental operation.

De Sousa’s assumption about importance of self-representations for consciousness leads him to speculate on the possibility that invertebrates elaborate on self-representational states: “Do all invertebrates have explicit central interoceptive representations and can this criterion be used to determine which species might be capable of conscious experience?” In spite of these doubtful speculations, a reasonable test for his idea is formulated some paragraphs later: Does an “alteration to structures that represent physiological changes in one’s own body...alter or destroy conscious experience” (De Sousa, 2013b^[10], p. 151-209)? This is a good question for experimental research. Although the structures that represent body, physiological changes may not be the places where feelings and emotions are generated or instantiated, they may indeed be necessary to trigger them.

After apparently hitting the finger two times, some paragraphs later De Sousa hammers the nail on the head with this question: "Is there an aspect of emotion experience that is relatively independent of thought and reflection, and an aspect that depends on it" (De Sousa 2013b^[10], p. 151-209)? One answer appeared in the proposal of *Affective Consciousness*, recently discussed in another *Mens Sana Monographs* paper by Almada *et al.* (2013^[11]). De Sousa's summary is very close to the main idea advanced in that paper: "Affective processes are supported by brain structures that appeared earlier in the phylogenetic scale (such as the periaqueductal grey area)...run in parallel with cognitive processes, and can influence behaviour independently of cognitive judgments" (De Sousa 2013b^[10], p. 151-209). Making common cause with a growing consensus led by Damasio and Bechara, he comments later in the paper that "ethical judgments are always emotional and conscious, but they can also have a cognitive-appraisal component that complements the somatic signalling that is also part of the account" (De Sousa, 2013b^[10], p. 151-209).

Finally entering the domain of affective and emotional consciousness, De Sousa requires that "a theory of emotional consciousness should be able to explain how...different experiences are generated by neural operations...A theory of emotional consciousness should provide a mechanism for explaining such differences in intensity. It should also provide a mechanism for valence, the positive or negative character of emotions. Positive emotions like happiness and pride have very different qualitative feel from negative ones like fear, anger, and disgust" (De Sousa, 2013b^[10], p. 151-209). These excerpts definitely moved the paper in a productive direction, except for the "neurocentric" assumption that relates affects and emotions only to "neural operations", thus neglecting other cells and informational mediums of the body.

Quantum consciousness: A necessary, but possibly an insufficient approach

In the beginning of this section in De Sousa's second review paper, he unsuccessfully attempts to cover the basics of the epistemology of physics, the nature of time and even implications of second law of thermodynamics. Considering that consciousness is "transcendental and immanent" (what does it mean?), he assumes a radical empiricist view of the epistemology of physics ("physicists realize that science can account for and explain only a small part of reality that is the part that we see and perceive"), and discusses some aspects of time and entropy without precision, adequate references or an explicit relation with the main theme of the paper-except for brief mention in places of the Kantian idea that time is the "innermost dimension of consciousness" and the highly controversial conclusion of a few authors that "without consciousness, there would be no time" (De Sousa, 2013b^[10], p. 151-209). In both cases, however, it is an understanding of consciousness that would be required to understand time.

Before entering the quantum approaches that give the section its name, De Sousa briefly discusses electromagnetic field and tensor network theories of consciousness, which are not properly in the quantum consciousness theoretical domain. Another imprecision appears here, when he states: "We have still not managed to describe a single empirical example of a spatial electromagnetic pattern that covaries with a particular kind of *human* conscious experience" (De Sousa, 2013b^[10], p. 151-209). He could be referring to token-token identity, but if the word "kind" is not misplaced in the above sentence then he could account for Microstate Theory, proposed by Dietrich Lehmann back in the 1990s (see Lehmann, forthcoming^[17]).

I am not commenting on the issues of gravity or three separate worlds-briefly reviewed by De Sousa-since they are poorly justified philosophically and only loosely connected to scientific consciousness issues. In a brief mention of the Anthropic Principle, on the other hand, he poses an important question about the place of consciousness in our universe: "By the use of this principle we show that consciousness is inevitable by virtue of the fact that sentient beings, that is 'we', have to be around to observe the world" (De Sousa, 2013b^[10], p. 151-209). However, this argument operates under a subjective idealist, Berkleyan assumption that the world exists only if observed by a conscious subject.

The next remarks are about the differences between the brain and the computer. De Sousa reasons that the difference would lie on quantum principles such as superposition and entanglement, and cogently suggests "there also seems to be some relation between this oneness of consciousness and quantum parallelism. In quantum theory, different alternatives at the quantum level are allowed to co-exist in linear superposition. Thus a single quantum state could in principle consist of a large number of different activities all occurring simultaneously" (De Sousa, 2013b^[10], p. 151-209). Here I call attention to the fact that many authors-including myself-have thought that the "oneness" of quantum entangled states would help to solve the binding problem (Pereira Jr., 2012^[27]), but De Sousa does not develop this possibility, probably because he assumed the micro-consciousness hypothesis.

More to the point, quantum theory seems to be important to brain sciences and to a scientific theory of consciousness because our cognitive and affective states are embodied in ionic dynamic patterns, such as those registered by the EEG. Their microscopic dynamics cannot be explained by classical electromagnetism alone, since biological ions are moving charged atoms. There is no reason to rule out "a priori" the existence of ionic entanglement in the living brain, or even in other parts of the living body.

Philosophical approaches: Crucial when not generating additional confusion

In the history of philosophy, the main agenda was to account for human knowledge, not for consciousness in general or affective states in particular.

Contemporary philosophy of mind inherited this habit, resulting in what De Sousa appropriately calls “representationalism”: “If qualia cannot be dismissed as unreal then how does an intentionalist theory of consciousness deal with them? The answer is a philosophy called *representationalism*, which is the view that qualia are nothing more than representational properties of conscious experiences” (De Sousa, 2013b^[10], p. 151-209).

“Representationalism” helped the progress of cognitive neurosciences in the 1990s, but became an obstacle in the next decade when the paradigm entered a period of scientific crisis (in the Kuhnian sense of the word). A legion of graduate students made their way in the scientific community using functional magnetic resonance imaging to highlight brain regions where one or other kind of “representation” is instantiated. Instead of understanding how the brain works, the goal of brain science became the mapping of cognitive functions, in the same way that—in the same period—the goal of genetics became the mapping of the genome of several species (instead of understanding the mechanisms of information transmission and metabolic control, which were finally brought to attention in the 2000s with the discovery of new roles of the RNA and the then called “junk DNA”).

After presenting the main philosophical tendencies in the study of consciousness, such as Materialism and Idealism, De Sousa mentions one alternative view that could encompass the best of both, without the worse of each: “There is a sense in which qualitative conscious states may be identified with states of the brain. Perception of a brain state and introspection of a mental state may be seen as two different ways of representing the same thing” (De Sousa, 2013b^[10], p. 151-209). If we replace “representing” by “perceiving”, we arrive at the concept known as *Dual-Aspect Monism* (DAM; Velmans, 2009^[34]; Pereira Jr. *et al.*, 2010^[24]). Although there is a diversity of conceptions in Indian philosophy (aptly reviewed by De Sousa), the prevalent conception seems to be close to the above-mentioned DAM.

Mistaken and/or complicated philosophies explicitly assumed by professional philosophers or implicitly by productive scientists are an actual obstacle to the development of consciousness science, since they add conceptual confusion to a field that deals with a complexity of phenomena and explanatory factors. Materialism, the conception that holds that brain physical and chemical processes generate mental states, and Idealism, the converse doctrine that understands that mental states produce physical and chemical effects, are good examples of philosophies that increase conceptual confusion in consciousness studies, mostly by inadequate usages of the concept of “causation” (the physical causing the mental or the mental causing the physical). DAM has the advantage to avoid this kind of problem, pointing to the necessity of considering physical and psychological processes on their own, as well as their correspondences. For DAM, conscious processes are always psychophysical: Both the causes and the effects are psychophysical.

The ending of the section is a bit frustrating, because De Sousa presents his sketch for a scientific methodology for consciousness research, instead of further developing the interesting philosophical issues he had raised. The proposed methodology is again based on the assumption of conscious modules, to be studied in their physical-biological and psychological aspects. Although consistent with DAM, no argument is presented in favour of the existence of psychological modules; the reader is left with an impression that the impossibility to overcome neural modularity is imposed on psychological phenomena. However, considering that brain networks are made not only of neurons, but also glial cells (forming a syncytium where information is processed by ionic waves), it is possible to conceive a physical-biological continuum that-according to the monist assumption-would correspond to integrated conscious episodes; in this case, we could have DAM without the modularity assumption.

Concluding remarks: How to make the integration? [See also Figure 1: Flowchart of the paper]

After reading the useful reviews written by De Sousa, I am more convinced that the integration of scientific theoretical hypotheses about consciousness requires a sound philosophical framework. Some attempts have been made in this direction, in collective publications, meetings, online discussions and e-mail lists. A consensus seems to be emerging that there is little chance of a revolutionary empirical discovery in this field, and that scientific progress will derive mainly from an adequate theoretical framework to interpret the thousands of published results and to guide the planning of pertinent new experiments.

There are relevant publications absent in De Sousa's huge reference lists (on consciousness theory-e.g., Jaynes, 1976^[16]; philosophical systematisation-e.g., Seager, 1999^[30]; consciousness conferences-e.g., Rakić *et al.*, 1997^[28]; philosophy encyclopaedia entries-e.g., von Gulick, 2004^[35]; encyclopaedia of consciousness-e.g., Banks, 2009^[4] and integrative project-e.g., Brook and Raymont, forthcoming^[5]), but I will not attempt to complete it; his reviews already cover the main topics to be addressed and provide a good sample of the most relevant publications. My disagreements concern the interpretation of some of the topics and how they could contribute to an integrative scientific theory of consciousness.

In this commentary I have expressed some points of disagreement; now is the time to show how my interpretations of the literature could aggregate to form a different ontological picture. In my own concept of "Triple-Aspect Monism" (TAM; in reference to three aspects of reality, physical-chemical-biological, informational and conscious; see Pereira Jr., forthcoming^[22]), I have defended the following ideas:

- Consciousness is more than information; it is *the feeling* of informational content (see, for instance, the conception presented by Harnad and Scherzer,

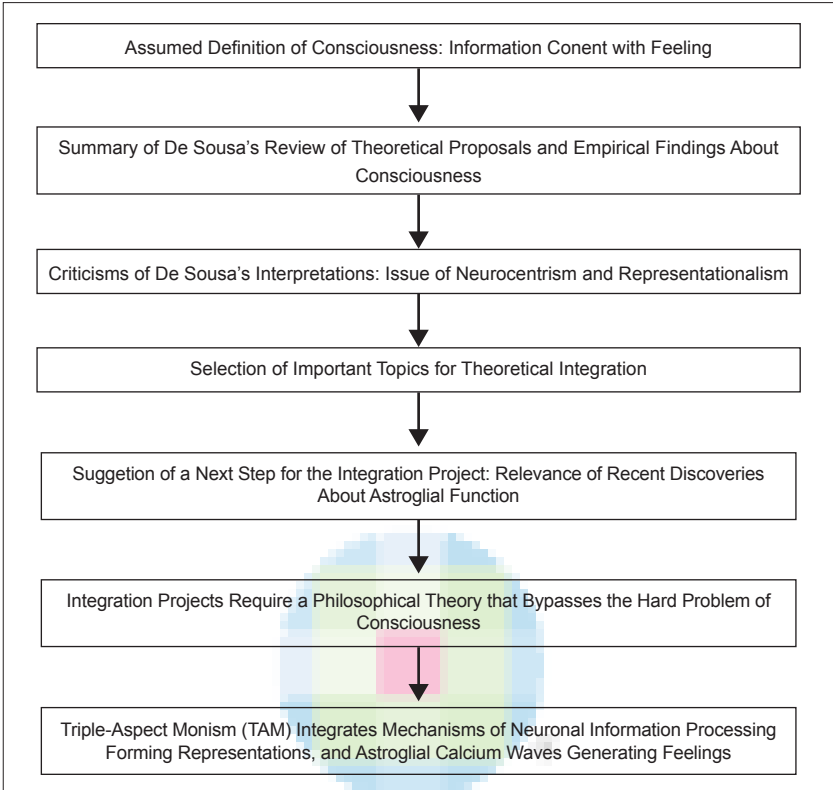


Figure 1: Flowchart of paper

2008^[15]). Feelings are considered to be typically conscious (while emotions can be unconscious). There are two major classes of feelings. What I call a “sensitive feeling” refers to the experience of states of the body, for example, feeling hunger and thirst, heat or cold, pain or pleasure. What I call an “affective feeling” refers to experiences elicited by the content of information, for example, feeling happy or sad about something, interested or bored of something, loving or hating something (Pereira Jr., forthcoming^[22]).

- Consciousness, understood as “the feeling of what happens”-title of a famous Damasio (2000^[8]) book-is a fundamental aspect of the universe, although not primitive. It exists as a potentiality of nature and requires specific mechanisms to be actualised;
- Feelings are actualised in wavelike media, such as the astroglial network in the brain, and can occur-with different degrees of actualisation-in any continuous waveform phenomenon found in nature, but probably not in

neurons-because of their structural (basically, their separation by the synaptic cleft) and physiological (as described by the classical Hodgkin-Huxley equation) properties;

- Neuronal and astroglial networks interact, forming a larger *brain endogenous network* (Carrara-Augustenberg and Pereira, 2012^[7]); this is the network that supports conscious processing. Neurons process information and construct representations, while astrocytes instantiate feelings about them, thus modulating neuronal activity according to the valence attributed to the information content;
- The conscious focus of attention is determined by the matching (or “resonance”) of affective and cognitive processes in the domain of neuro-astroglial interactions; therefore, attention is closely related but not synonymous to consciousness, since there are several modalities of peripheral consciousnesses.

The resulting concept of consciousness is a *proto-panpsychist one*, in the sense that consciousness is conceived as a fundamental aspect of nature, inseparable from the others (physical and informational), but depending on the operation of specific mechanisms present in some kinds of systems (typically, living systems), but not in others (e.g., rocks) to become actualised.

We live in a universe full of potentialities that unfold in time. The first aspect that was actualised (e.g., starting with the “Big Bang” or a similar event) is the physical-chemical, culminating with the origin of life. The development of physical-chemical-biological organisation-ruled in complex ways by the second law of thermodynamics and other physical laws-made possible the stabilisation of a variety of forms, from the elementary forms classified in the periodic table to the complex forms of plant and animal species. The communication of forms-also called “information transmission”-constitute a second fundamental aspect of the universe.

The existence of individual, autonomous systems able to use information to control their activities makes possible the emergence of the third aspect, consciousness. It occurs only when these systems develop a sensibility to the content of information, and the capacity of modulating their constitutive processes according to their lived experiences. Before the existence of these systems, consciousness was just a potentiality present in wavelike patterns of nature, such as sound (acoustic waves), vibrational patterns in solids (e.g., metallic plates) and liquid (e.g., the sea) media.

According to TAM, conscious systems are those which contain two active interacting networks, one for the construction of knowledge, and the other for the generation of feelings. The novelty is how the feeling system is conceived: As a system that operates with wavelike patterns, each kind of wave corresponds to a feeling. The “master hub” that instantiates feelings is proposed to be the

astroglial network (Pereira Jr. and Furlan, 2010^[25]; Pereira Jr., Pereira and Furlan, 2010^[26]), reflecting states of the whole body in the process of interaction with the world. In this regard, recent results have related calcium waves in the astrocytic network with brain information integration (Navarrete *et al.*, 2012^[18]), the state of consciousness (Thrane *et al.*, 2012^[31]) and neurological diseases impacting consciousness, such as Alzheimer (Furman *et al.*, 2012^[13]).

The astroglial network receives signals from the whole body by means of neurons, blood flow and cerebrospinal fluid (e.g., in the vestibular system). This whole-body state is dynamically compared with the information about the body in the world, as represented by the neuronal network. The result of this comparison is a momentary feeling that expresses the valence of the information relative to the state of the system. This feeling impacts the whole-body state, by means of neuro-immune-endocrine signalling-including activities of the heart and gut nervous systems. Consciousness is therefore conceived as the processing of information about the body and the world, constructing knowledge and investing it with feelings that influence the processing and consequently the action of the living individual in the world.

I hope the above work could help to advance one step further De Sousa's valid suggestions towards theoretical integration. The consequences for psychiatric theorising and practice are relevant, from the possibility of identifying the aetiology of most mental diseases as anomalies of the astroglial system (of course, related to patterns of interaction with the world) to the discovery of new pathways for therapeutic medication in the domain of neuro-astroglial-immune-endocrine interactions.

Take home message

From a critical analysis of De Sousa's reviews, I make proposals for the advancement of the project of an integrative theory of consciousness:

1. I consider that conscious systems are those that are composed of two active interacting networks, one for the construction of knowledge, and the other for the generation of feelings.
2. I suggest that the feeling system operates with wavelike patterns, each kind of wave corresponding to a feeling. The "master hub" that instantiates feelings is the astroglial network of the brain. It receives signals from the whole body by means of neurons, blood flow and cerebrospinal fluid. This whole-body state is dynamically compared with the information about the body in the world, as represented by the neuronal network. The result of this comparison is a momentary feeling that expresses the valence of the information relative to the state of the system. Feelings impact the whole body, by means of neuro-immune-endocrine signalling.
3. The concept of consciousness assumed in this proposal is that it consists of the processing of information about the body and the world of living individuals,

constructing knowledge and investing it with feelings that influence the processing, and consequently the action of the living individual in the same world.

Conflict of Interest

None declared.

Declaration

This is my original unpublished work, not under consideration for publication elsewhere.

Acknowledgements

CNPQ, FAPESP (Brazilian funding agencies) and Ajai Singh for the invitation to write this commentary.

References

1. Almada LF, Pereira Jr. A, Carrara-Augustenberg C. What the Affective Neuroscience Means for a Science of Consciousness. *Mens Sana Monogr* 2013;11. Available from: <http://msmonographs.org/preprintarticle.asp?id=100409>. [Last accessed on 2012 Nov 19].
2. Baars BJ, Laureys S. One, not two, neural correlates of consciousness. *Trends Cogn Sci* 2005;9:269.
3. Baars BJ, Ramsøy TZ, Laureys S. Brain, conscious experience and the observing self. *Trends Neurosci* 2003;26:671–5.
4. Banks WP. *Encyclopedia of Consciousness* (2 vols). Oxford: Elsevier; 2009.
5. Brook A and Raymont P. A Unified Theory of Consciousness. Cambridge, MA: MIT Press Forthcoming. A summary is Available from: <http://plato.stanford.edu/entries/consciousness-unity/>. [Last accessed 2012 Nov 21].
6. Bullock TH, Bennett MV, Johnston D, Josephson R, Marder E, Fields RD. The neuron doctrine, redux. *Science* 2005;310:791–3.
7. Carrara-Augustenberg C, Pereira Jr. A. Brain endogenous feedback and degrees of consciousness. In: Cavanna AE, Nani A editors. *Consciousness: States, Mechanisms and Disorders*. New York: Nova Science Publishers; 2012.
8. Damasio A. *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. New York: Harcourt; 2000.
9. De Sousa A. Towards an integrative theory of consciousness: Part 1 (Neurobiological and cognitive models). *Mens Sana Monogr* 2013a;11:100-50.
10. De Sousa A. Towards an integrative theory of consciousness: Part 2 (An anthology of various other models). *Mens Sana Monogr* 2013b;11: 151-209.
11. Del Cul A, Dehaene S, Reyes P, Bravo E, Slachevsky A. Causal role of prefrontal cortex in the threshold for access to consciousness. *Brain* 2009;132:2531–40.
12. Fields RD. *The Other Brain*. Simon and Schuster, New York; 2009.
13. Furman JL, Sama DM, Gant JC, Beckett TL, Murphy MP, Bachstetter AD, Van Eldik LJ, Norris CM. Targeting astrocytes ameliorates neurologic changes in a mouse model of Alzheimer's disease. *J Neurosci* 2012;32:16129–40.
14. Hameroff SR. The entwined mysteries of anesthesia and consciousness: Is there a common

- underlying mechanism? *Anesthesiology* 2006;105:400-12.
15. Harnad S, Scherzer P. First, scale up to the robotic Turing test, then worry about feeling. *Artif Intell Med* 2008;44:83-89.
 16. Jaynes J. *The Origin of Consciousness in the Breakdown of the Bicameral Mind*. Boston: Houghton Mifflin; 1976.
 17. Lehmann D. Consciousness: Microstates of the brain's electric field as atoms of thought and emotion. In: Pereira Jr., Lehmann Editors. *The Unity of Brain, Mind and World: Current perspectives on a science of consciousness*. Cambridge UK: Cambridge University Press.
 18. Navarrete M, Perea G, Fernandez de Sevilla, Gómez-Gonzalo M, Núñez A, Martin ED, *et al*. Astrocytes Mediate *In Vivo* Cholinergic-Induced Synaptic Plasticity. *PLoS Biol* 2012;10:e1001259.
 19. Panksepp J. Affective consciousness. In Velmans M, Schneider S, editors. *The Blackwell Companion to Consciousness*. Malden, MA: Blackwell; 2007.
 20. Panksepp J. Affective consciousness: Core emotional feelings in animals and humans. *Consciousness Cogn* 2005;14:30-80.
 21. Pereira Jr A, Ricke H. What is consciousness? Towards a preliminary definition. *J Consciousness Stud* 2009;16:28-45.
 22. Pereira Jr A. Triple-Aspect Monism: A conceptual framework for the science of human consciousness. In: Pereira Jr and Lehmann, editors. *The Unity of Brain, Mind and World: Current perspectives on a science of consciousness*. Cambridge UK: Cambridge University Press.
 23. Pereira Jr A. Visual Consciousness: Too many correlates. In: *Brain Physiology, Cognition and Consciousness*. 2009. Available From: <http://network.nature.com/groups/bpcc/forum/topics/1423> [Last accessed on 2012 Nov 21].
 24. Pereira Jr. A, Edwards J, Lehmann D, Nunn C, Trehub A, Velmans M. Understanding consciousness: A collaborative attempt to elucidate contemporary theories. *J Consciousness Stud* 2010;17:213-9.
 25. Pereira Jr. A, Furlan F. Astrocytes and human cognition: Modeling information integration and modulation of neuronal activity. *Prog Neurobiol* 2010;92:405-20.
 26. Pereira Jr. A, Furlan FA, Pereira, MAO. Recent Advances in Brain Physiology and Cognitive Processing. *Mens Sana Monogr* 2011;9:183-92.
 27. Pereira Jr. A. Perceptual information integration: Hypothetical role of astrocytes. *Cogn Computation* 2012;4:51-62.
 28. Rakić L, Kostopoulos G, Raković D, Koruga D. Proceedings of the First Annual ECPD International Workshop on Scientific Bases of Consciousness. Belgrade: European Centre for Peace and Development (ECPD) of the United Nations University for Peace; 1997. Available from: <http://www.dejanrakovicfund.org/knjige/1997-ECPD-Workshop.pdf>. [Last accessed on 2012 Nov 21].
 29. Rosanova M, Gosseries O, Casarotto S, Boly M, Casali AG, Bruno MA, *et al*. Recovery of cortical effective connectivity and recovery of consciousness in vegetative patients. *Brain* 2012;135:1308-20.
 30. Seager W. *Theories of Consciousness* London: Routledge; 1999.
 31. Thrane AS, Thrane VR, Zeppenfeld D, Lou N, Xu Q, Nagelhus EA, Nedergaard M. General anesthesia selectively disrupts astrocyte calcium signaling in the awake mouse cortex. *Proc Natl Acad Sci USA* 2012;109:18974-9.
 32. Treisman A. The binding problem. *Curr Opin Neurobiol* 1996;6:171-8.
 33. Velmans M. Consciousness, brain, and the physical world. *Philos Psychol* 1990;3:77-99.
 34. Velmans M. *Understanding Consciousness*. 2nd ed. London: Routledge; 2009.
 35. Von Gulick R. Consciousness. *Stanford Encyclopedia of Philosophy* 2004. Available from: <http://plato.stanford.edu/entries/consciousness/>. [Last accessed on 2012 Nov 21].

Questions that this Paper Raises

1. How do definitions of consciousness relate to projects of theoretical integration?
2. How to integrate cognitive and affective consciousness approaches?
3. How could the progress towards a science of consciousness be achieved?
4. Should the study of consciousness be based on properties of neuronal networks or should other cells and their respective networks be taken into consideration?

Is consciousness composed of integrated episodes or is it separable in micro modules? What are the phenomenological evidences for each view?

About the Author



Alfredo Pereira Jr. holds a degree in Philosophy from the Federal University of Juiz de Fora (1984), degree in Business Administration from the Machado Sobrinho Foundation (1983), MA in Philosophy from the Universidade Federal de Minas Gerais (1986) and PhD in Logic and Philosophy of Science from the State University of Campinas (1994). He was Postdoctoral Fellow on the Brain and Cognitive Sciences Department at the Massachusetts Institute of Technology (1996-1998). He is Adjunct Professor at Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP) since 1988, and teacher and counsellor accredited in master and doctoral programmes in Public Health (School of Medicine, UNESP-Botucatu) and Philosophy (Faculty of Sciences, UNESP-Marília). He is Visiting Scientist at the KEY Institute for Mind-Brain Research affiliated with the University of Zurich (2012), and Visiting Researcher at the Centre for Theoretical and Empirical Consciousness Studies, Department of Psychology, University of Copenhagen (2012). He is also Sub-Coordinator of the FAPESP Thematic Project "Systemics, Self-Organization and Information", linked to the Centre for Logic and Epistemology, University of Campinas (2011-2016). He has experience in the areas of Physiological Psychology (Cognitive and Affective Neuroscience), Mental Health and Epistemology, and does research and publishes on the following topics: Mind and Brain, Human Consciousness, Neuro-Astrocyte Interactions, Philosophy of Sciences of Life and Health, and the Biopsychosocial Model of the Health Disease Process.