

How to study consciousness as a natural phenomenon

An interview with
Tim Bayne

by Alessio Bucci & Matthieu Koroma

Tim Bayne
timothy.bayne@monash.edu
Department of Philosophy
Monash University,
Melbourne, Australia

Alessio Bucci
alessio.bucci@unito.it
Department of Philosophy
University of Turin, Italy

Matthieu Koroma
mkoroma@ens-paris-saclay.fr
Ecole Normale Supérieure
PSL University, Paris, France

Abstract

The field of the scientific study of consciousness has seen a flourishing of methodologies and theories. The debate over what defines consciousness and how we should study it is, however, yet to be settled. Philosopher Tim Bayne has proposed the “natural kind” approach, suggesting that consciousness properties should be empirically informed rather than defined a priori. Relying on the cross-talk between philosophy and empirical science, he proposes a cautious and integrative outlook that takes into account the diversity of the conscious phenomenon, defending a multidimensional model of conscious states.

keywords: altered states of consciousness, global state of consciousness, levels of consciousness, multidimensional model, natural kind approach.

Consciousness remains one of the most puzzling phenomena for both philosophers and scientists. Could you explain what triggered your interest in this topic? What are the questions about consciousness you have been addressing with your research? Do you think that there are questions about consciousness that cannot be answered?

My initial interest in consciousness focused on the unity of consciousness. I was struck by the fact that philosophers have traditionally regarded unity as one of the central features of consciousness, whereas many consciousness scientists (and a significant number of philosophers!) seem to assume that consciousness is anything but unified. I began to wonder whether those who were “pro” unity had the same conception of “the unity of consciousness” as those who were against it, and the more I read on the topic the more convinced I became that there were many conceptions of “the” unity of consciousness. One of the first papers that I published—co-written with my PhD supervisor David Chalmers—was largely devoted to distinguishing different conceptions of the unity of consciousness. Since then I’ve gone on to

consider the ways in which the various aspects of the unity of consciousness might be explained, and what implications the unity of consciousness has for theories of consciousness. I've also written on various other facets of consciousness, such as the relationship between consciousness and agency and the nature of conscious thought.

Are there questions about consciousness that cannot be answered? Quite possibly. Many theorists believe that scientific advances will enable us to bridge the explanatory gap between the physical/functional properties of the brain on the one hand and conscious experience on the other but I have my doubts. I'm sure that there is much about consciousness that we will succeed in understanding, but I suspect that a residue of mystery will always remain.

“ I'm sure that there is much about consciousness that we will succeed in understanding, but I suspect that a residue of mystery will always remain. ”

In “Modes of Consciousness” (Bayne and Hohwy, 2016), modes of consciousness are defined as global states of consciousness which are mutually exclusive. In previous works, you argued against the possibility for a single creature to have different streams of consciousness at the same time, discussing for example the case of split-brain patients (Bayne, 2008). Is there a link between modes being mutually exclusive and the unity of consciousness you defended previously in Unity of consciousness (2010)? Still, some animals like dolphins or birds exhibit patterns of sleep in some restricted parts of the brain while having other parts awake (Rattenborg et al, 2000). Without falling in the mental error fallacy, which would infer a mental property directly from patterns of neuronal activity, how do you deal with such cases from the perspective of modes of consciousness?

Interesting question! Let me start by stepping back a bit. The main point here was to sharpen the contrast between conscious modes (or what I now prefer to call “global states of consciousness”) and conscious contents. This contrast is quite intuitive, but few theorists have provided an analysis of it. One of the suggestions that Jakob Hohwy and I made is that one of the things that distinguish modes from contents is that modes tend to exclude each other whereas contents don't. For example, you can hear a trumpet whether or not you are tasting ice-cream, but if you are (say) in the state of alert wakefulness then you can't also be mildly sedated.

Now, what should we say about cases in which an organism might seem to be awake in one hemisphere but asleep in another? It's a good question, and one that I'm not entirely sure how to answer. One option would be to resist the idea that modes of

consciousness should be ascribed to hemispheres, and to insist that it's only the organism as a whole that is in a particular mode of consciousness. The other option would be to allow that modes can be ascribed to hemispheres, but to insist that no single hemisphere can be in more than one mode at a time. I can see problems with both options. The worry with the first is that it threatens to ignore what we seem to be learning about the neural basis of some modes; the worry with the second is that it threatens to sever the conceptual link between modes and the functional capacities that govern the behavior of the organism as a whole. It's certainly a topic that requires more thought!

You proposed multi-dimensional modeling as a tool for the study of consciousness that better accounts for the differences among global states of consciousness (Bayne, Hohwy & Owen, 2016). What are the specific epistemic merits of this approach that you see for the field of consciousness studies? An issue concerning this approach is that it opens a Pandora's box: the dimensions can vary in number, nature (e.g., they can be functional or content-based) and thus far there are no clear guidelines on how to identify the most relevant dimensions to classify states of consciousness. Do you have a pluralist view on the dimensions that should be used to distinguish states of consciousness or do you think some dimensions would emerge as the most relevant? What would be the significance of one or the other finding for the conceptualization of consciousness?

It's true that we didn't propose any clear guidelines on what the relevant dimensions are for characterizing global states of consciousness, but whether that's a bug or a feature seems to me to depend on one's perspective. When it comes to global states, theorizing begins with cases. For example, the vegetative state and the minimally conscious state involve global states that are distinct from the ordinary neurotypical state and from each other. The state of consciousness associated with mild sedation is also a distinct global state, as is the conscious state associated with epileptic absence seizures. Psychedelics induce yet another kind of global state. After this—it's all theory. The main point of Bayne et al. 2016 was to show that the dominant approach to global states—the levels account—won't work. According to that account, consciousness comes in degrees (you can be more or less conscious), and your global state of consciousness just is a matter of how conscious you are. This proposal is pretty obviously hopeless, for it's clear that two people can be in different global states without differing in the degree to which they are conscious. But since the levels account is in effect a unidimensional account, rejecting it requires moving to a multidimensional account.

That, of course, doesn't answer the question that you posed about what the dimensions of these global states are and how they might be identified. A good place to begin in addressing these issues is to ask why we assume that (say) the vegetative state and the minimally conscious state are associated with distinct global states of

consciousness, or why we assume that REM sleep is associated with one global state of consciousness and psychedelics with another. I suspect that any plausible answer to this question will have to consider two things: what cognitive and behavioral capacities are in place, and what kinds of contents are (and are not) permitted by or associated with the global state in question.

In a recent article (Bayne, 2018), you argued that the scope of conscious phenomena covered by the axioms of the Information Integration Theory (IIT) (Oizumi et al., 2014) is limited to a restricted set of conscious phenomena or philosophical positions about consciousness. IIT proponents defend the view that axioms should be considered as working hypotheses that allow the theory to be coherent. Presumably, one advantage of such an approach is its explicitness about the assumptions endorsed by the theory. As an alternative, you propose the natural kind approach. What are the respective roles of theoretical thinking and empirical investigation in the natural kind approach?

I have no problem with the appeal to working hypotheses when it comes to developing theories of consciousness, but on my reading of IIT that's not the role that they ascribe to the axioms (the label "axioms" here is not an accident—working hypotheses are typically called "hypotheses" rather than "axioms"). As I read them, Tononi and the other advocates of IIT take an axiom to be a self-evident claim about the essential phenomenological features of consciousness. I have doubts about whether there are any axioms in this sense of the term, and even if there are any axioms, I doubt that it's possible to ground a theory of consciousness on the basis of an appeal to them in the way that the advocates of IIT attempt to do.

The natural kind approach doesn't start with alleged self-evident truths, but instead begins with defeasible claims about consciousness. These claims need not hold universally (for example, they might apply to human beings but not to the members of other species), and they need not be restricted to the phenomenology of consciousness (for example, they might also include the functional role of consciousness). The idea is to use these defeasible claims—"working hypotheses", if you will—as tools for homing in on the mechanisms that underlie consciousness.

“ The natural kind approach doesn't start with alleged self-evident truths, but instead begins with defeasible claims about consciousness. ”

When it comes to the natural kind approach theory-building and empirical investigation go hand-in-hand: we need theory for hypotheses formation and we

need empirical work in order to test the hypotheses. The situation is the same here as it is for any other domain in which we're looking for natural kinds.

In a recent article (Bayne & Carter, 2018), you argued against the characterization of psychedelic states as higher states of consciousness. The line of argument you develop revolves around cognitive functions associated with conscious processing. You reviewed evidence that some cognitive functions are enhanced during psychedelic experience (e.g. imagery and sensory salience), while others are impaired (e.g., decision making). You conclude that this argues against unidimensional accounts of consciousness that rank consciousness levels on a unique scale. As a consequence, in your view, this provides a challenge to theories like IIT and complexity-based approaches. Nevertheless, in such approaches, consciousness is not ordered according to cognitive or behavioral dimensions, but rather as the amount of integrated information accounting for phenomenal experience (Oizumi et al, 2016). Thus, it seems that your criticism is based on different conceptions about what dimensions are relevant for capturing the conscious phenomenon. In fact, a recent application of the IIT framework to the study of psychedelics suggests that even IIT might contest general claims about psychedelics being 'higher' states of consciousness (Gallimore, 2015). Could you comment on how your critique of the notion of levels of consciousness depends on theoretical assumptions guiding consciousness research?

There are a couple of issues here. One point concerns the criticisms that Olivia Carter and I make of the idea that the psychedelic state is a "higher" state of consciousness. Here, it's important to keep in mind that our focus was on the claim that the psychedelic state is a higher state of consciousness in the same sense in which disorders of consciousness (such as the vegetative state or epileptic absence seizures) are lower states of consciousness. We argue that if this were right, then one would expect the psychedelic state to be associated with general improvements in cognitive function, but that seems not to be the case. Denying that the psychedelic state is "higher" in this sense is of course compatible with accepting that it is higher in some other sense of the term.

As you suggest, the advocates of IIT propose that an individual's level of consciousness corresponds to the amount of integrated information that it involves. But here we need to ask how this proposal should be understood. If, on the one hand, it is to be understood as a proposal for how the term "level" should be used in consciousness science then that's fine, but it's not clear to me why anyone would use the term in this way unless they were already committed to IIT. If, on the other hand, it is to be understood as an account of levels of consciousness as that term has been used within consciousness science for the last decade or more than it's clearly unsatisfactory, for given current usage an improvement in an individual's level of consciousness requires a concomitant improvement in their general cognitive and behavioral capacities.

Your article on psychedelics fits within the growing trend of research that has been labeled the “psychedelic renaissance” (Sessa, 2012). How do you see this trend in consciousness studies? More generally, do you see any advantage in incorporating so-called altered states or fringe phenomena (like trance, meditational states, out-of-body experiences, lucid dreaming, etc.) in the study of consciousness?

As far as the study of consciousness is concerned, the psychedelic renaissance is part of a wider movement to take altered states of consciousness more seriously. Of course, altered states of consciousness have always had a place within the study of consciousness, but for the most part they've been relegated to the fringes of the discipline, and mainstream consciousness science has been dominated by the study of consciousness as it occurs in adult human beings in the state of ordinary wakefulness. (Indeed, for the most part it's been dominated by the study of visual consciousness as it occurs in the state of ordinary wakefulness!) I see the move towards taking psychedelic experience more seriously as motivated by a more general sense that consciousness really is multifarious in its expression, and that models which account for adult human experience in the state of ordinary alert wakefulness might not apply so well when it comes to other forms of experience.

Some recent theoretical papers focused on empirical studies (Gerrans & Letheby, 2017; Millière et al., 2018) suggest that through some alterations of consciousness (induced by certain classes of drugs or by meditational techniques) we can achieve a state of “ego dissolution”, where consciousness persists without the experience of being or having a self. In the past, you proposed to see the self as a “virtual center of gravity” for the unity of consciousness (Bayne, 2010). In relation to this, what do you think about empirical cases of ego dissolution? More generally, do you think that the study of these kinds of alterations can tell us something new about the structural features of consciousness?

In my view this is one of the most interesting aspects of psychedelic experience. It's tempting to suppose that a sense of being a single subject of experience is essential to human experience, and that no matter how seriously impaired consciousness is the “I” is always attached to each and every experience. But ego dissolution tells us that that assumption is false, and that the subjectivity of experience is not a necessary feature of human experience.

The multiple metaphysical options on the nature of consciousness discussed in philosophy (e.g. the hard problem of consciousness, the problem of mental causation, etc...) still seem to pose a problem for a unified scientific theory of consciousness. Should science take into account these philosophical debates, or should science rather focus on the “real problem” of the empirical investigation of the conscious phenomenon, as formulated by Anil Seth (2016)? Does consciousness raise, in your opinion, specific conceptual issues regarding its scientific study compared to other mental phenomena?

It's interesting to note that although philosophers have a fairly limited role when it comes to the study of most phenomena—for example, I suspect that few zoologists are motivated to keep up with developments in the philosophy of biology!—there are other domains in which philosophers and scientists have a great deal to learn from each other, and the study of consciousness is certainly one such domain. Having said that, it seems to me that the value and importance of dialogue between science and philosophy depends very much on the question at issue. I'm not sure that the scientific study of consciousness is likely to be enhanced by engaging with philosophical reflection on the nature of the hard problem or mental causation, but there are a number of other topics on which dialogue really is crucial. For example, there are important questions about what a scientific explanation for consciousness should look like, and whether it should have the notion of a neural correlate of consciousness at its heart. There are also deep and difficult questions about how to measure consciousness, and whether the subjectivity of consciousness poses any kind of principled difficulty for its scientific study. There are important treatments of these issues within both philosophy and the sciences, and it's a shame that all-too-often they are limited by disciplinary boundaries.

“ I'm not sure that the scientific study of consciousness is likely to be enhanced by engaging with philosophical reflection on the nature of the hard problem or mental causation, but there are a number of other topics on which dialogue really is crucial. ”

Different strategies in consciousness science can be distinguished according to the role they give to phenomenology in their research program. Integrated Information Theory uses phenomenology as a starting point to define consciousness through the formulation of axioms. Neurophenomenological approaches consider phenomenological reports as first-hand empirical data that should be investigated equally along physical processes. Reductionists consider phenomenology as an *explanandum* that must be accounted for in terms of physical explanations. Would the natural kind approach allow these strategies to be complementary rather than in competition, insofar as it considers definitional properties of consciousness as testable working hypotheses while avoiding metaphysical commitments? What role should phenomenology play in today's consciousness science in your opinion?

In a sense, everybody starts with the phenomenal character of consciousness—that's the thing that we want to explain. The question is how far you can get by simply focusing on the phenomenology. My view is: not all that far. One of the problems

here is that people often disagree about the phenomenology. For an example of such disagreement take a look at discussions of “cognitive phenomenology”. But leaving disagreement about phenomenal character to one side, the problem is that there is an explanatory gap between phenomenal character and neurofunctional structure. Thus, it seems to me, it’s difficult to build a conceptual bridge from descriptions of the phenomenology to descriptions of neurofunctional structure. The natural kind approach takes the phenomenology of consciousness seriously, but it doesn’t assume that explanations of the phenomenology can be deduced from descriptions of it.

You developed a substantial body of work on consciousness that identifies key components of the conscious phenomenon through conceptual analysis of empirical results. How do you consider the current progress on a definition of consciousness? How do you see the development of the field of consciousness studies in the future? Is there any avenue that you deem particularly promising?

I’m not sure that the study of consciousness has made much progress with respect to defining “consciousness”. Here, it seems to me that we have yet to improve on Tom Nagel’s characterization of consciousness—a conscious state is a state that there’s “something that it’s like” to be in—even though it’s generally recognized that this characterization has at best limited utility. But significant progress has been made on other issues, even if that progress has often involved the posing of more questions than answers. I think that the increased attention given to non-visual consciousness has been salutary, and it is good to see burgeoning literatures in both science and philosophy on such topics as agentive awareness, conscious thought, and multisensory integration. It is also good to see lots of activity in the development of theories of consciousness, although I wish that this activity was accompanied by more discussion about how to test theories of consciousness. Looking forward, I suspect that issues relating to the detection of consciousness in machines and non-human animals will loom large in the coming decade.

References

- Bayne, T. (2008). The unity of consciousness and the split-brain syndrome. *The Journal of Philosophy*, 105(6), 277-300.
- Bayne, T. (2010). *The unity of consciousness*. Oxford University Press.
- Bayne, T. (2018). On the axiomatic foundations of the integrated information theory of consciousness. *Neuroscience of consciousness*, 2018(1), niy007.
- Bayne, T., & Carter, O. (2018). Dimensions of consciousness and the psychedelic state. *Neuroscience of consciousness*, 2018(1), niy008.
- Bayne, T., & Hohwy, J. (2016). Modes of consciousness. *Finding consciousness: The neuroscience, ethics and law of severe brain damage*, 57-80.
- Bayne, T., Hohwy, J., & Owen, A. M. (2016). Are there levels of consciousness?. *Trends in cognitive sciences*, 20(6), 405-413.
- Gallimore, A. R. (2015). Restructuring consciousness—the psychedelic state in light of integrated information theory. *Frontiers in human neuroscience*, 9, 346.
- Letheby, C., & Gerrans, P. (2017). Self unbound: ego dissolution in psychedelic experience. *Neuroscience of Consciousness*, 2017(1).
- McGinn, C. (1999). *The mysterious flame: Conscious minds in a material world*. Basic Books.
- Millière, R., Carhart-Harris, R. L., Roseman, L., Trautwein, F.-M., & Berkovich-Ohana, A. (2018). Psychedelics, Meditation, and Self-Consciousness. *Frontiers in Psychology*, 9.
- Oizumi, M., Albantakis, L., & Tononi, G. (2014). From the phenomenology to the mechanisms of consciousness: integrated information theory 3.0. *PLoS computational biology*, 10(5), e1003588.
- Rattenborg, N. C., Amlaner, C. J., & Lima, S. L. (2000). Behavioral, neurophysiological and evolutionary perspectives on unihemispheric sleep. *Neuroscience & Biobehavioral Reviews*, 24(8), 817-842.
- Sessa, B. (2012). *The psychedelic renaissance: Reassessing the role of psychedelic drugs in 21st century psychiatry and society*. Muswell Hill Press.
- Seth, A. (2016) The hard problem of consciousness is a distraction from the real one. *Aeon*. <https://aeon.co/essays/the-hard-problem-of-consciousness-is-a-distraction-from-the-real-one> ◊