

Letter

Islands of Awareness or
Cortical Complexity?

Benedetta Cecconi,^{1,2}
Steven Laureys,^{1,2} and
Jitka Annen^{1,2,*}



Bayne *et al.* (2020) [1] envision the fascinating possibility that disconnected hemispheres (through hemispherotomy), *ex cranio* brains, and cerebral organoids may be islands of awareness (IOAs). They defined IOAs as systems fully disconnected from the external surrounding, both in terms of input (i.e., sensory information) and output (i.e., motor responses), yet capable of conscious experience (i.e., aware). In order to test whether these islands are effectively aware, the authors propose to use measures of human/animal consciousness that have been well validated in human/animal settings by correlation with pretheoretical measures of consciousness (e.g., behavioural reports) or subjective reports (in the case of humans). In particular, Bayne *et al.* [1] mention: (i) measures of dynamical complexity, capturing both integration and differentiation of the detected brain signal, such as the perturbational complexity index (PCI); and (ii) measures of algorithm complexity of EEG (i.e., entropy measures), quantifying the randomness of brain dynamics and capturing mainly the differentiation of the recorded brain activity. Two questions come to mind: one of a theoretical and one of a methodological kind. (i) Can subjective experience be collapsed to a 1D entity (i.e., a certain measure of cortical complexity)? (ii) If we accept such reduction, can measures validated in one context generalize to different species and neural systems?

If we accept that consciousness, now generally defined in the literature as subjective experience, can be assessed in isolated systems by the mere application

of a certain measure of cortical complexity, consciousness is reduced to a 1D construct, in which an index is equated to (the capacity for) subjective experience. For example, let us consider the above-mentioned dynamic complexity (DC). DC operationalises a (phenomenological) structural property of consciousness, the integration and differentiation of every conscious scene (structural properties of consciousness are those properties common to all conscious experiences). While sufficiently high values of such measure appear necessary for consciousness, evidence that it is sufficient is lacking, as high DC could be present in unconscious networks [2]. In fact, consciousness would seem to possess other structural properties necessary for it to arise: for example, the features of perspectivalness (subjective first-person perspective), presentationality (to be in the present, the experience of 'nowness'), and transparency (attentional unavailability of earlier processing stages) [3]. Perhaps Bayne *et al.* [1] do not refer to subjective experience when speaking about consciousness in IOAs. In this case, it is legitimate to wonder what kind of consciousness would be identified if these systems were to be found dynamically complex: for example, will they possess a sense of phenomenal ownership, which seems necessary for the experience of suffering [3]? This is crucial for determining their moral status and regulating our behaviour toward them legally. Moreover, DC and entropy measures, quantifying different aspect of cortical complexity, may operationalise different phenomenological properties of conscious experience, potentially leading to conflicting interpretations of IOAs' state of consciousness. It is therefore necessary to specify which measure is to be considered to define an IOA as such.

Even if we accept the possibility of reducing consciousness to a certain measure of cortical complexity, we are still faced with the problem of defining a set of values consistent with

consciousness. For example, electrically stimulated slices of the occipital cortex show a slice PCI score of 0.2 [4]; considered to index unconsciousness in the human PCI context [5]. Can we conclude that these slabs are unconscious? To help answer this question, let us imagine that human body temperature would be proposed as a universal index of health. A specific range of normal temperature has been set and validated as an index of human health by correlating it with human behaviours (i.e., sickness), such that if a woman at rest has a temperature of 40°C, we can conclude she is ill. We now apply this well-validated index to a bird. Since birds have an average body temperature of approximately 40°C, applying our index would make us conclude that it is sick. Similarly, the thresholds for (un)consciousness as defined by objective measures might not retain the same meaning when applied to other systems, without prior calibration. This calibration, requiring some form of subjective report/pretheoretical measure of consciousness, is by definition impossible to obtain in true islands of awareness. This raises the question whether consciousness can be quantified without relying on subjective/behavioural reports in systems very distant from us. Currently, human (or animal) consciousness is univocally probed through the subjective/behavioural reports of the person (or animal). Bayne and colleagues concluded that 'for the first time we might be able to tell that there is something it is like to be a disconnected brain without being able to tell just what it is like'[1], suggesting the presence of a subjective experience in IOAs. Or might we instead have reached the ultimate challenge of consciousness science, that is, to account for first-person phenomenology from a third-person perspective?

Acknowledgements

The authors thank the University and University Hospital of Liège, the GIGA Doctoral School for Health Sciences, the Belgian National Funds for Scientific Research (FRS-FNRS), the European Union's Horizon 2020 Framework Programme for Research and

Innovation under the Specific Grant Agreement No. 945539 (Human Brain Project SGA3), the European Space Agency (ESA), the Belgian Federal Science Policy Office (BELSPO) in the framework of PRODEX, the fund Generet and the King Baudouin Foundation. SL is research director at FRS-FNRS.

¹Coma Science Group, GIGA-Consciousness, University of Liège, Liège, Belgium

²Centre du Cerveau, University Hospital of Liège, Liège, Belgium

*Correspondence:

jitka.annen@uliege.be (J. Annen).

<https://doi.org/10.1016/j.tins.2020.05.007>

© 2020 Elsevier Ltd. All rights reserved.

References

1. Bayne, T. *et al.* (2020) Are there islands of awareness? *Trends Neurosci.* 43, 6–16
2. Seth, A.K. *et al.* (2006) Theories and measures of consciousness: an extended framework. *Proc. Natl. Acad. Sci. U. S. A.* 103, 10799–10804
3. Metzinger, T. (2003) *Being No One: the Self-Model Theory of Subjectivity*, MIT Press
4. D'Andola, M. *et al.* (2018) Bistability, causality, and complexity in cortical networks: an *in vitro* perturbational study. *Cereb. Cortex* 28, 2233–2242
5. Casali, A.G. *et al.* (2013) A theoretically based index of consciousness independent of sensory processing and behavior. *Sci. Transl. Med.* 5, 198ra105

Letter

From Complexity to Consciousness

Tim Bayne,^{1,2,*}
Anil K. Seth,^{2,3,4} and
Marcello Massimini^{1,2,5,6}



We thank Cecconi and colleagues for their probing questions about our opinion article ‘Are There Islands of Awareness?’ [1], and we are grateful for the opportunity to return to these important issues.

Our primary aim in writing the aforementioned paper was to give the notion of an island of awareness (IOA) some precision, and to argue that the possibility of IOAs should not be dismissed as a

philosophical fantasy but should be taken seriously by the science of consciousness.

The fundamental challenge, concerns the detection of IOAs. What would count as good evidence for the presence of an IOA? We suggest that the most promising approaches involve complexity measures, such as the Perturbational Complexity Index (PCI), and other measures, which are sensitive to the coexistence of functional differentiation and functional integration in neural systems. We point to these particular metrics because they go beyond mere correlation to operationalize explanatory relationships between neural dynamics and phenomenological properties [2,3], and because they have already been shown to perform well when assessing consciousness in patients in whom devastating injuries often result in disconnection from inputs and outputs [4,5]. Thus, we propose to use a ‘radar’ that we have learned to trust to explore the unknown landscape of completely disconnected brains.

Cecconi *et al.* [6] raise two interesting issues in connection with this proposal. The first is that these measures treat consciousness as a one-dimensional phenomenon, which, Cecconi and colleagues argue, is mistaken. We agree that consciousness is multidimensional [7], but we don’t agree that the use of complexity measures to detect consciousness presupposes that it is unidimensional. The key here is to realize that markers or indicators of consciousness need not be treated as definitions of consciousness. (Compare: crying is a marker of pain, but pain is not to be defined in terms of crying.) Consciousness could be multidimensional even if certain markers of it are unidimensional. (Cecconi and colleagues ask whether we mean to refer to subjective experience when speaking of consciousness. We do.)

The second issue that Cecconi *et al.* raise concerns the use of complexity measures

to identify IOAs. There are two aspects to this challenge. The first is whether it is possible to directly compare values of complexity that have been derived from different types of systems. For example, one cannot directly compare the values of the slice-adapted version of PCI (sPCI) to those validated in humans, in part because they are currently obtained with different techniques at different scales [8]. Determining what counts as the ‘most appropriate’ scale for a given system is challenging, but recent advances in the mathematical analysis of multiscale systems have made progress in addressing it [9,10].

Accompanying these technical considerations is the more general question of whether it is ever legitimate to apply markers of consciousness that have been validated in one population (intact human beings) to the members of other populations, such as ex-cranio brains, disconnected hemispheres, and cerebral organoids. Cecconi *et al.* appear to answer this question in the negative. We will confine our response to three points. First, anyone who is willing to ascribe consciousness to non-human animals faces a version of this challenge, for the behavioural and neurophysiological markers of consciousness that we employ to ascribe consciousness to non-human animals must themselves be validated with reference to human beings [11]. Second, the systems we are considering share many of the neurophysiological features of the brains of intact mammals; in particular, they have neurons and associated biological infrastructure. Concerns about the possible substrate dependency of consciousness can therefore be set aside [compare, for example, silicon-based artificial intelligence (AI) systems]. Third, as we note above and in [1], our validation strategy appeals to the capacity of the property in question to explain features that are associated with consciousness, starting from its unity and richness. This approach to validation (i.e., the ‘natural