Trends in Cell & Molecular Biology
Vol. 15, 2020

Review

Testing for consciousness in cerebral organoids

Stuart Hameroff[1,*,#] and Alysson R. Muotri[2,*,$]

[1]Departments of Anesthesiology and Psychology, Center for Consciousness Studies, The University of Arizona, Tucson, Arizona; [2] Departments of Pediatrics and Cellular & Molecular Medicine, University of California, San Diego, School of Medicine, Center for Academic Research and Training in Anthropogeny (CARTA), Archealization Center (ArchC), Kavli Institute for Brain and Mind, La Jolla, CA 92037, Sanford Consortium for Regenerative Medicine, La Jolla, California 92093, USA.

ABSTRACT

Consciousness cannot be directly measured, but there exist in the brain putative 'neural correlates of consciousness' ('pNCCs'). We describe four novel electrophysiological pNCCs found in humans and animals. pNCC1 is 'high gamma' frequency (~85-165 Hz) electro-encephalography ('EEG'), selectively abolished by anesthesia, and enhanced in meditation. The second, pNCC2, is related to 'phase-amplitude coupling', or 'nesting' between different EEG frequencies, e.g. between delta/theta waves (1-10 Hz), and 'low gamma' EEG waves (25-65 Hz). EEG nesting correlates with cognitive short-term memory and perception, but paradoxically is enhanced by anesthesia, suggesting that deviation from nested EEG correlates with consciousness (pNCC2). Similarly, pNCC3 involves 'non-computable' deviation from 'integrate-and-fire', Hodgkin-Huxley neuron behavior, as found in brain pyramidal neurons in awake animals. pNCC4 is a generalized reverberation, or EEG 'chiming' in response to a focused stimulus. pNCC2 and pNCC3 suggest consciousness supervenes on non-conscious cognitive ('autopilot') mechanisms, seen as deviation from algorithmic processing as an observable 'shadow' of consciousness. Invasive pNCC testing in humans and animals is limited for ethical reasons, but simpler biological systems with pNCCs are potentially conscious, with fewer ethical concerns. Recently, 'high gamma' and nested EEG were observed in cerebral organoids, stem cell-derived cooperative assemblies of ~2.5 million cortical neurons. We propose to perturb cerebral organoids in ways known to affect consciousness, and evaluate perturbing effects on pNCCs1-4 in the context of various theories of consciousness. Organoids will be perturbed by 1) anesthesia, 2) psychedelics, 3) electromagnetic and ultrasound energy, 4) drugs which impair membrane receptors and ion channels, and 5) drugs which impair cytoskeletal microtubules inside neurons. We hope to establish criteria for assessing consciousness in biology, evaluate contributions from membrane/synaptic and intra-neuronal cytoskeletal processes to consciousness, find support for one or more theories of consciousness, and reach an informed opinion on whether cerebral organoids are conscious.

KEYWORDS: consciousness, cognition, cerebral organoids, phase-amplitude coupled EEG, Hodgkin-Huxley neuron, neural correlates of consciousness (NCCs), supervenience, anesthesia, theories of consciousness, pyramidal neurons.

Introduction: The mystery of consciousness

Despite ever-increasing detailed knowledge about the brain, it's most important function – consciousness - remains scientifically unexplained. We define consciousness as any phenomenal experience, including awareness, external and internal sensory

*Corresponding authors
[$]muotri@ucsd.edu; [#]SHameroff@anesth.arizona.edu

perceptions, feelings like pleasure and pain, emotions, sense of self, thought, memories, planning, dreams, and/or conscious choice. The essential distinction between conscious and non-conscious processes is the ineffable quality of experiential existence, composed of what philosophers term 'qualia'. Pondered since ancient times, the nature of phenomenal experience, of qualia, is currently framed philosophically as 'what it is like to be' [1], the 'explanatory gap' [2], and/or the 'hard problem' [3]. The nature and origin of conscious experience have eluded scientific explanation. Questions abound, for example:

1) *Are only humans conscious?* What about primates, mammals, all other animals, and all living systems? Did consciousness emerge during the course of evolution as is commonly assumed, or has consciousness 'been here all along'? And what about inanimate objects? Panpsychists assert conscious experience is a property of reality, of the universe [4], and many artificial intelligence ('AI') proponents argue that computers can be conscious. How do we decide?

2) *Is the brain a computer of neurons?* Brain function has always been likened to contemporary information technology, from the ancient Greek idea of memory encoding as a 'seal ring in wax', to the mind as a telegraph switching circuit, and now computers and artificial intelligence [5]. The brain-as-computer analogy is based on the Hodgkin-Huxley ('H-H') 'integrate-and-fire' neuron formulation, which casts neurons as threshold logic 'input-output' computational devices [6]. According to H-H, dendrites and cell bodies/soma of neurons receive synaptic inputs as membrane potentials, integrating them to a specific threshold potential which, when reached, triggers axonal firings, or action potentials ('spikes', Figure 1). These discrete all-or-none spike propagations lead to synaptic transmissions, actions and behavior, and are often taken as 'bits' in computational views of consciousness. Completely algorithmic H-H computation would appear to limit and possibly preclude conscious free will, insight and creativity.

There are flaws in the algorithmic H-H neuron, brain-as-computer approach.

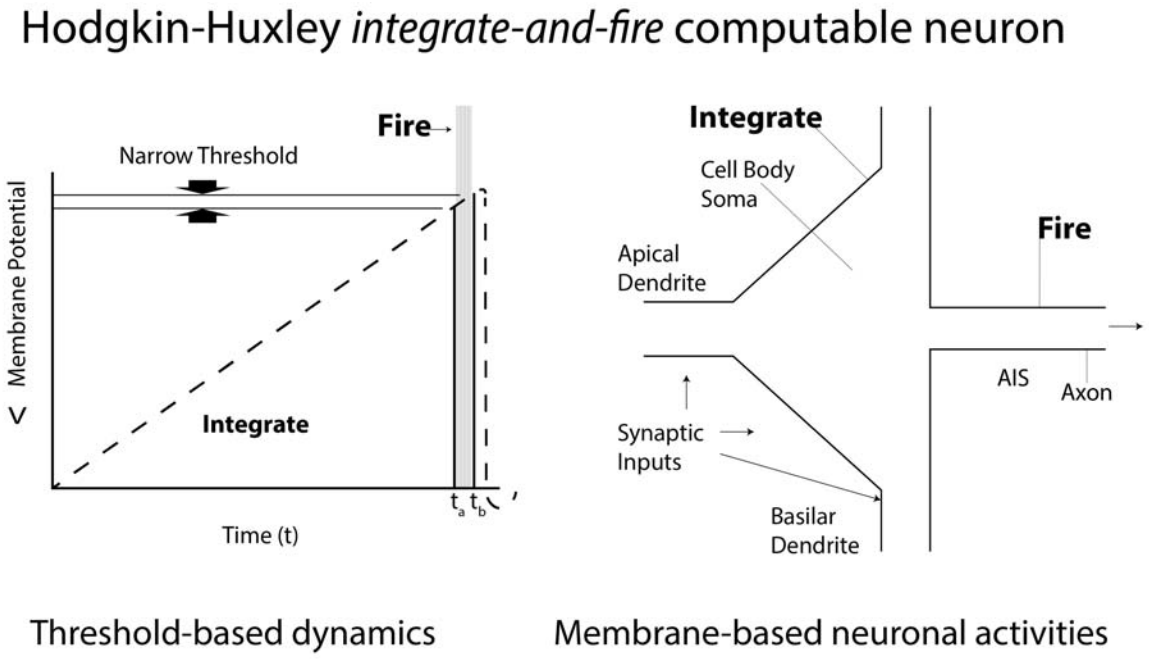

Figure 1. Left: Schematic diagram of Hodgkin-Huxley ("H-H") integrate-and-fire neuronal activity. A linear process is assumed, and the narrow firing threshold potential and small temporal firing windows reflect algorithmic "computable" behavior. Right: Cartoon version of pyramidal neuron (apex pointing left, axon to right) utilizing only its surface membrane potential for integration, signaling and firing threshold, as prescribed by Hodgkin-Huxley. AIS: Axon initiation segment.

1) *The notion of neurons as simple 'bit states' in computer-like networks is belied* by non-neuronal single cell organisms like *Paramecium* which swim, learn, avoid obstacles, find food and mates, and have sex. *Paramecium* uses its own internal cytoskeletal microtubules, including those in cilia, to sense surroundings and move purposefully. Single cells can have significant cognition, regardless of whether they may be conscious. Equating neurons with bit states is an insult to neurons.

2) According to Gödel's theorem, mathematical proof of any algorithmic computation requires a 'non-computable' factor outside, or apart from the algorithmic computation itself. Nobel laureate Sir Roger Penrose [7] applied Gödel's theorem to 'understanding', suggesting a non-computable effect outside, or within the brain's recognized neuronal computational system is required in consciousness.

3) *Recordings from cortical pyramidal neurons in brains of awake animals show wide variability in firing threshold on a spike-to-spike basis* [8] (Figure 2). These deviations from H-H behavior suggest a non-algorithmic, non-computable factor may modulate and supervene to cause deviation in 'integrate-and-fire' processing. The source of H-H deviation may be close to the origin of consciousness.

4) *At what scale levels of brain activity does consciousness occur?* Most view consciousness as a higher order 'emergent' effect of computation among neurons in hierarchical networks. But are neurons the bottom level, the fine scale origin, the fundamental level of the hierarchy? If so, does consciousness correspond with neuronal dendritic-somatic integration, or axonal firing? Or does physiology relevant to consciousness extend inside neurons to smaller, faster activities in microtubules, for example? Are quantum activities, e.g. collective terahertz oscillations in microtubules [9], entangled to give unified, brain-wide states necessary for consciousness?

5) *Is consciousness fundamental?* Unable to account for consciousness strictly through neuronal computation and emergence, many philosophers and neuroscientists rely on panpsychism, in which consciousness is a ubiquitous property of reality, or pan-protopsychism, in which precursors of consciousness are omnipresent and available for conversion to consciousness [10]. If so, how does the brain access and connect to fundamental aspects

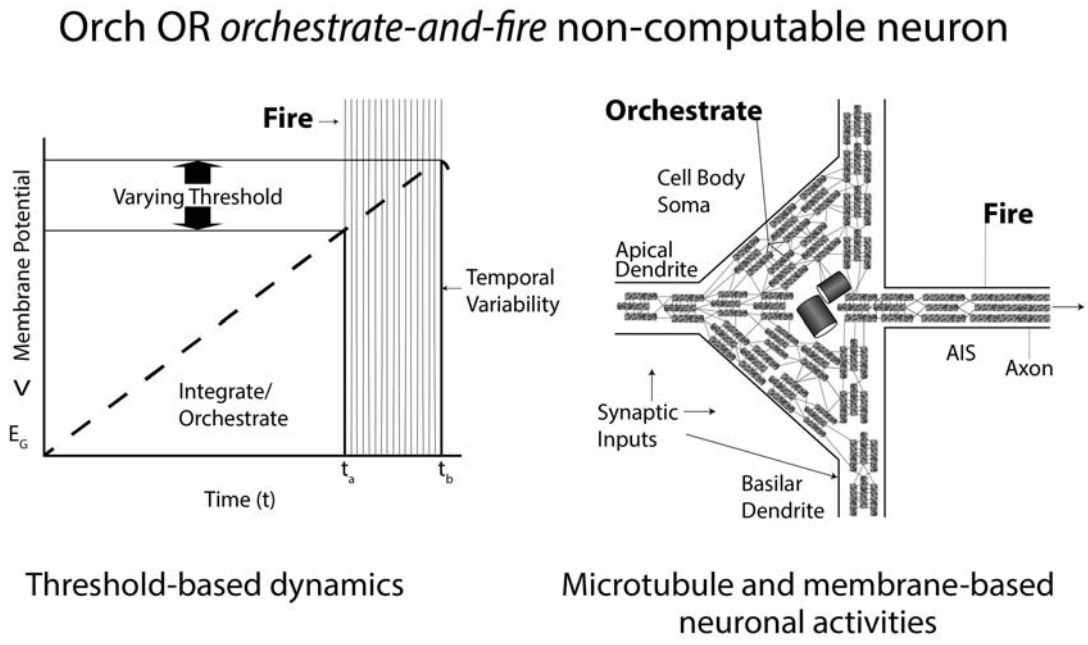

Figure 2. Left: Recordings from pyramidal neurons in awake cats show widely variable membrane potential thresholds and time windows (Naundorf *et al.*, 2006), i.e "non-computable". Right: Pyramidal neuron with interior microtubules, Orch OR non-computable influence on axonal firing.

of the universe? Does consciousness somehow extend downward to the basic level of spacetime geometry, as Penrose has suggested? Are quantum brain activities required for the connection to fundamental reality [11-13]?

Putative neural correlates of consciousness ('pNCCs')

Consciousness occurs in biological brains, but cannot be directly measured or observed other than by one's own consciousness. However there are measurable putative 'neural correlates of consciousness' ('pNCCs') observable in humans and other animals (Table 1). If one or more pNCCs are observed in a biological brain-like system, the possibility of consciousness cannot be excluded, and should be considered [14]. Some pNCCs indicate consciousness may "cast a shadow", causing deviation from non-conscious 'autopilot' cognition and algorithmic neural-level processing. Do pNCCs converge on a common mechanism, e.g. as proposed by any particular theory of consciousness?

pNCCs include the following types:

1) *Electro-encephalography* (*'EEG'*): Among physiological functions, the most widely recognized pNCCs are related to electro-encephalography (EEG), collective electrical activity of brain neurons measured from cortical surface or scalp. EEG consists of 5 fairly distinct frequency bands: delta: 0-4 hertz (Hz), theta: 4 to 7 Hz, alpha: 8 to 15 Hz, beta: 16 to 31 Hz, gamma: > 32 Hz, including 'low gamma' (~33 to 85 Hz), and 'high gamma' (difficult to detect, > 85 Hz).

Under anesthesia with loss of consciousness, EEG generally slows, with particular inhibition of gamma, especially high gamma. Perhaps the earliest correlation of consciousness with gamma EEG came in visual cortex of awake monkeys [15], with further elaboration in cats [16]. Crick and Koch [17] suggested low gamma synchrony ('coherent 40 Hz') was a neural correlate of consciousness, but later reversed their opinion [18], and indeed there are discrepancies in low gamma synchrony as a proper NCC [19]. However under anesthesia with loss of consciousness, and in dreamless sleep, low gamma oscillations (~40 Hz) generally disappear, and other EEG frequencies slow. In 'rapid eye movement' 'REM' sleep associated with dreaming, EEG signals resemble those of awake, fully conscious subjects.

When it can be measured, 'high gamma' EEG (~85 to 155 Hz or higher) is dampened or abolished

Table 1. Putative neural correlates of consciousness, perturbations, and theories of consciousness.

Putative neural correlates of consciousness (pNCCs)
pNCC1 'High gamma' frequency (~ 100 to 400 Hz) electro-encephalography ('EEG')
pNCC2 Deviation from 'phase-amplitude coupling' between EEG rhythms
pNCC3 Deviation from 'integrate-and-fire', Hodgkin-Huxley neuronal behavior
pNCC4 Chiming response (neuronal reverberation in EEG) to focused stimulation
Perturbations
Anesthesia
Psychedelics
Membrane receptor and ion channel disrupting drugs
Cytoskeletal microtubule disrupting drugs
Theories of Consciousness
GNW Global neuronal workspace
HOT Higher order
PC/RP Predictive coding/recursive processing
IIT Integrated information theory
Orch OR Orchestrated objective reduction

with anesthesia and loss of consciousness. Specifically, cortico-cortical coherence and frontal-parietal 'entropy transfer', both in the high gamma band (85 to 155 Hz) are markedly decreased during anesthetic-induced unconsciousness [20]. High gamma EEG is a valid pNCC.

The nature and origin of EEG remain unclear, usually attributed to membrane and synaptically mediated feedback loops, pacemaker circuits, network reverberation, and, for gamma oscillations, inter-neuron gap junctions [21, 22]. At the neuronal level, EEG is often assumed to derive from coherent axonal firings, or 'spikes', but EEG signals are primarily post-synaptic 'local field potentials' ('LFPs') from neuronal dendrites and cell bodies/soma, i.e. occurring during integration phases, rather than axonal firings. Similarly, many assume consciousness derives from axonal firings, or spikes in a computational brain, but anesthetics act post-synaptically, erasing consciousness without direct effects on spiking capabilities. EEG detected at the scalp is generated mostly from apical dendrites which ascend from cortical layer V pyramidal neurons.

It is generally assumed that EEG rhythms and coherent EEG derive completely from collective effects of neuronal membrane activities, be they axonal firings, or dendritic-somatic local field potentials. But EEG may actually derive from deeper levels, for example cytoskeletal microtubules inside neurons have spontaneous electrical oscillations in gamma synchrony frequency ranges [23], and higher frequency oscillations in microtubules have been discovered as well [24]. Despite its historical prevalence and clinical utility, the underlying origin, mechanism and significance of EEG remain unknown [25].

2) *Phase-amplitude coupled, 'nested' electro-encephalography*: EEG rhythms are sometimes phase-amplitude coupled, or 'nested', meaning the amplitude of a faster rhythm is coupled to the phase of a slower rhythm. Series of low gamma waves (~32 to 85 Hz) are often coupled with phase of slower delta, theta, alpha or beta waves, resulting in distinct 'wave packets' to which have been ascribed cognitive functions. For example Lisman and Idiart [26] suggested that nested theta-low gamma oscillations underlie the specific capacity limits of working memory. The approximately seven gamma cycles nested in a theta cycle are proposed to correspond to the seven, plus or minus two, items that can be stored in working memory. Van Rullen and Koch [27] suggested nested alpha-gamma cycles produce 'discrete perceptions', alpha wave ensembles of gamma wave 'snapshots'. Phase-coupled, nested EEG have also been linked to communication between brain areas [28], and feature-based attention [22]. Sequences of beta-gamma coupling process feature detection, such as orientation or color, in different regions of visual cortex [29]. EEG nesting, or phase-amplitude coupling may be involved in cognition, but what about consciousness?

Pal *et al*. [30] studied effects of 3 types of general anesthetics (propofol, sevoflurane, ketamine) on phase-amplitude coupling between delta or theta waves with low frequency gamma in rat frontal cortex, and found that all 3 types of anesthesia *increase* phase-amplitude coupling. This could imply that phase-amplitude EEG coupling is necessary for non-conscious cognition ('autopilot'), and that some other processes mediating consciousness supervene to cause deviation from, over-riding of, non-conscious cognitive functions as a pNCC.

From where and how could consciousness supervene on nested EEG, a manifestation of neuronal membrane and synaptic activities? The 'Orch OR' theory of consciousness [11-13, 31] suggests 'conscious' activities of microtubules, particularly those inside dendrites and soma of brain neurons in layer V cortical pyramidal cells, do indeed influence axonal firings, membrane, synaptic and network activities. Microtubules have conductance resonances ranging through terahertz, gigahertz, megahertz and kilohertz frequencies [32-35]. These activity patterns may be phase-amplitude coupled, or nested over 12 orders of magnitude, including EEG, shown to have scale-free, self-similar patterns in measurable EEG frequencies [36]. Orch OR proposes that consciousness occurs in this nested hierarchy, and that interference among higher frequency oscillations in microtubules gives rise to EEG as slower 'beat frequencies'. Nesting and coherence transfer between terahertz (10^{12} Hz) and petahertz (10^{15} Hz) frequencies have been observed in quantum biological systems [37].

3) *Non-computable deviation from Hodgkin-Huxley behavior*: The prevalent view in science

and philosophy is that the brain is a computer, based on algorithmic integrate-and-fire Hodgkin-Huxley ('H-H') model neurons, interacting via synapses. According to H-H, neuronal dendrites and cell bodies/soma receive synaptic inputs from other neurons as membrane potentials, which are integrated to a threshold potential at the axon-initiation segment ('AIS'). At that moment when threshold is reached, at that place on the proximal axon, a 'firing' occurs and an 'action potential' or "spike" is triggered which rapidly propagates the length of the axon to a synapse, exerting causal action and regulating behavior. In computer analogies and artificial intelligence ('AI'), firings or spikes are often taken as bit-like states, and the integration and firing of H-H neurons are considered to be 'computable', algorithmic, deterministic processes, with minimal random variability (Figure 1).

However, actual cortical neurons in awake animals appear to deviate from Hodgkin-Huxley 'computable' membrane-mediated algorithmic behavior. Naundorf *et al*. [38] placed electrodes in cortical pyramidal neurons of awake cats, and found that axonal firing threshold varies in individual neurons on a 'firing-to-firing', 'spike-to-spike' basis (Figure 2). The cause of this deviation was characterized as synaptic 'noise' [39], but as the 'noise' is found only in cortical pyramidal neurons of awake, presumably conscious animals, the deviation may reflect processes related to consciousness – the "shadow" of consciousness. Thus a 'non-computable' factor other than algorithmic synaptic inputs and membrane potentials appears to regulate axonal firings (and thus behavior) in pyramidal neurons of awake animals. What might that be?

The Orch OR theory suggests the non-computable factor may reflect consciousness from deeper level quantum processes in microtubules, especially in dendrites and cell bodies/soma, e.g. in layer V cortical pyramidal neurons.

Microtubules in dendrites and cell bodies/soma of neurons are uniquely arrayed in mixed polarity, anti-parallel networks, optimal for recursive processing and interference, leading to slower 'beat frequencies' (EEG) in cortical pyramidal neurons and other neurons. Non-computable deviation from H-H behavior, a possible "shadow of consciousness", is a pNCC.

4) *Chiming response*: Modulating, or 'zapping' the brain with a short pulse of magnetism (TMS, transcranial magnetic stimulation), or low intensity ultrasound can cause a prolonged reverberation and spatial spreading of brain activity ('chiming') as measured in EEG, and analyzed by data compression ('zipping', hence 'zap-and-zip, [40, 41]. In subjects or animals who are anesthetized, or have suffered brain-damage, the chiming is reduced, and thus chiming serves as a pNCC. We will test such techniques in cerebral organoids, and look for chiming responses. If chiming occurs, organoids would be considered potentially conscious.

5) *Neuroanatomical*: Where in the brain does consciousness arise? Neuroanatomical approaches aim to identify the locus of consciousness through activity revealed by brain imaging, primarily functional magnetic resonance imaging ('fMRI'). Results point to anatomical pNCCs in cerebral cortex, thalamus [42], thalamo-cortical loops [43], visual cortex connections to prefrontal cortex [44], posterior cortical 'hot zones' [45], pre-frontal cortical projections [46], recursive networks [47], brain stem [48], claustrum [49], and globally broadcast networks [50, 51].

However we don't know if activity correlating with consciousness occurs in any one localized area all the time, in global networks including these areas, among cortical pyramidal neurons, or whether conscious processes might literally move around the brain [52]. fMRI depends primarily on blood flow related to metabolism, and may not necessarily correlate with consciousness at all, at least in some cases. For example, studies of subjects "highly conscious" on psychedelic psilocybin show very low fMRI activity [53], suggesting that while cognition may depend on high energy-dependent membrane processes, consciousness may derive from deeper level, lower energy (possibly quantum) processes inside neurons.

Anatomy within cerebral organoids is not well maintained, especially in long-term organoid cultures, and would not necessarily distinguish among theories of consciousness. Thus, anatomy will not be included in this study description, although it would be interesting to look at polarity, e.g., development along a frontal-posterior axis in cerebral organoids using future protocols that allow for better cortical layering definition.

We propose to detect and observe pNCCs 1-4 as described above in cerebral organoids, and perturb them in ways known to affect consciousness or particular neuronal cell structures.

Perturbations

1) *Anesthesia*: Consciousness is selectively and reversibly inhibited by anesthesia, which largely spares non-conscious brain activities. In the 19th century a group of gases were discovered which, when inhaled at low concentrations, caused giddiness and euphoria. But when inhaled and equilibrated at moderate and higher concentrations, animals and humans were rendered unresponsive and apparently unconscious. Potencies of the different anesthetic gases vary, but for each gas there is an average inhaled concentration which, when equilibrated with lungs, blood and brain, effectively 'anesthetizes' half of animals or human subjects. This anesthetic 'effective dose' varies among gases, the lower the necessary concentration, the more potent the anesthetic gas, but is the same for each gas to 'anesthetize' protozoa, insects, mice or humans.

At the turn of the 20th century, Hans Meyer [54] and Charles Overton [55] independently found that potencies of all anesthetic gases correlated over orders of magnitude with their gas solubility in a specific non-polar, lipid-like solvent akin to olive oil (the 'Meyer-Overton correlation'. There are such regions in fat stores, lipid membranes, and inside proteins all over the body. Anesthetics bind and dissolve there by weak, quantum interactions called van der Waals London forces ('dipole dispersion' forces).

Paradoxically, anesthetics bind, dissolve and act non-specifically in large amounts by weak quantum interactions, and yet affect only consciousness. One possible explanation is that consciousness depends on highly organized quantum effects, e.g. dipole oscillations, easily disrupted by anesthetic 'dipole dispersion'.

In the 1980s [56], anesthetics were found to act in tiny non-polar, olive oil-like regions inside certain brain proteins, regions containing 'pi electron resonance' rings of aromatic amino acids phenylalanine, tyrosine and tryptophan. These molecules have the same 'lipid-like' solubility as fat stores and membranes, but in tiny water-aversive ('hydrophobic') pockets, and support quantum effects in their delocalized electron clouds. In which specific proteins do anesthetics act? These were initially presumed to be membrane receptor, and ion channel proteins, and anesthetics do bind within $GABA_A$, serotonin, acetylcholine and glycine post-synaptic receptors. But after decades of research, consistent effects of anesthetics on these or other neuronal membrane proteins were not found; there was nothing which matched the Meyer-Overton correlation with potency [57]. There was no 'Meyer-Overton correlation' for any membrane protein, or group of proteins.

Systematic studies showed anesthetic gases bind to over 70 different proteins in brain neurons, about half in membrane and half in cytoplasm [58]. Genomic, proteomic and optogenetic experiments then suggested anesthetics act inside neurons on tubulin, the component protein of microtubules [59-61]. Computer modeling further suggested anesthetics act in tubulin/microtubules by dampening collective terahertz quantum vibrations [62]. The site and selective mechanism of action of anesthetic gases remains an open question, and a potential clue to the origin of consciousness.

2) *Psychedelics*: Psychedelic molecules such as lysergic acid diethylamide ('LSD') and psilocybin cause perceptual distortion, enhanced experiential processing and content capacity, 'hallucinations', and increased brain-wide entropy of neuronal electrical activity, the latter suggesting a finer, more detailed grain of information processing [63]. Psychedelic molecules are largely pi resonance rings, similar to aromatic amino acid rings to which anesthetics bind within proteins to prevent consciousness. LSD and other psychedelics bind to membrane receptors including the 5HT2A receptor on pyramidal neurons, but the mechanism by which receptor binding induces altered mental states is unknown [64]. 5H2A receptors are coupled to G-proteins which convey signals to the cytoskeleton inside neurons, and non-polar psychedelic molecules (like anesthetics) may enter the neuron directly. The mechanism of psychedelic action remains an open question but has been suggested to increase pi resonance dipole oscillation frequencies in tubulin and/or other proteins, the opposite of anesthetics which dampen or disperse such dipoles.

3) *Membrane ion channel and receptor-perturbing drugs*: We will expose cerebral organoids to ion channel blockers, such as lidocaine, and receptor channel blockers such as bicuculline and others, and see effects on pNCCs. Significant dampening would tentatively support GNW, HOT, PC and IIT

4) *Microtubule-perturbing drugs*: We can test effects of drugs like epothilone D which destabilizes microtubules, and nocodazole which excessively stabilizes them, making microtubules less dynamic and responsive on cerebral organoids, and see how they affect pNCCs.

Theories of consciousness

A proper theory of consciousness should account for neural correlates of consciousness including the 4 pNCCs listed above. A recent review [14] compared 17 theories of consciousness, and from among them, we consider five representatives, the first 5 theories considered in the Templeton World Charity Foundation program 'Accelerating research on consciousness'. They are 1) Global neuronal workspace: 'GNW', attributed to Bernard Baars, Stanislaus Dehaene and Jean-Pierre Changeux, 2) Higher order theory: 'HOT', David Rosenthal and Hakwan Lau, 3) Integrated information theory: 'IIT', Giulio Tononi, Christof Koch, 4) Predictive coding: 'PC', or Recurrent processing: 'RP', hence 'PC/RP', Victor Lamme, Karl Friston, 5) Orchestrated objective reduction 'Orch OR', Sir Roger Penrose, Stuart Hameroff.

1) *Global neuronal workspace*: GNW emulated early efforts in artificial intelligence ('AI') regarding central processing units in computers, areas where information becomes available to various subsystems for integration and analysis. Baars [50] adapted the widespread availability of central processing units to the idea of a cognitive broadcast in the brain, the global workspace (GW). Dehaene, Changeux and others [51, 65] portrayed a more anatomical global neuronal workspace (GNW) based on non-linear, large scale ignition of global neuronal activity from pre-frontal and frontal cortex, broadcasting 'top-down' to temporal, parietal and occipital cortical areas, with concomitant recurrent feedback and recursive processing. Ignition and recursive processing are presumed to be mediated by Hodgkin-Huxley 'integrate-and-fire' neurons, with action potential firings, or 'spikes' propagating along long axons, most specifically from pyramidal neurons in layers 2, 3 and 5 of cortex. Consciousness is presumed to occur primarily in the broadcast ignition phase, primarily from frontal and pre-frontal cortex, and also parietal and cingulate cortex to more disparate areas.

2) *Higher order theory*: HOT suggests that mental representations occur at the top of a hierarchy of sensory inputs and integration, i.e. the pre-frontal cortex, which exerts top-down effects to other brain regions [48]. In HOT, a higher order system is conscious of a lower order system (as opposed to GNW is which the broadcast itself is conscious). Both GNW and HOT rely solely on membrane and synaptic activity in the form of axonal firings in algorithmic H-H 'integrate-and-fire' neurons as primary mediators of activity most related to consciousness.

3) *Predictive coding/recurrent processing*: PC/RP suggests consciousness results from continuously comparing and updating mental models of the world with incoming information, and depends critically on recurrent, or recursive processing, i.e. feed-forward and feed-back processes [66-69]. Global, top-down projections, as suggested in GNW and HOT, are unnecessary, and recursive processes at many levels are proposed in PC/RP to compare previous hypothetical, or remembered reality with sensory inputs. Predictive coding PC is consistent with 'First order theories', e.g. Ned Block's [70] idea which suggests consciousness can occur due to local recursive processes, e.g. in primary cortex.

4) *Integrated information theory*: IIT proposes that consciousness depends on maximal integration of information in any physical substrate [71]. The integration is irreducible, has maximal cause–effect structure, and trans-scalar, optimizing integrated cause-effect power ϕ, or phi. Thus IIT is scale-independent, ϕ manifesting at multiple spatiotemporal hierarchical scales, with consciousness attributed to the level of maximal ϕ. What that level might be, and what the most basic, fundamental (i.e. small, fast) level might be, are not specified in IIT. Based on brain anatomy, maximum ϕ is hypothesized to reside primarily in posterior cortical posterior "hot zones". IIT does not stipulate neuronal processes relevant to phi, although Koch [72] has maintained that axonal firings mediate consciousness.

5) *Orchestrated objective reduction*: Orch OR attributes consciousness to quantum computations in microtubules inside brain neurons, with upward effects on neurons and networks. Quantum computers process information as quantum superpositions of multiple co-existing possibilities (quantum bits, or 'qubits'). These superposition states are proposed to unify by entanglement, evolve and compute until reduction, or 'collapse' occurs, selecting definite output states. Sir Roger Penrose [7] proposed reduction occurred due to an objective threshold in the fine scale structure of the universe ('objective reduction, 'OR'), and that the OR event itself resulted in proto-conscious experience. Subsequently, Penrose and Hameroff suggested that microtubules were quantum computers, with collective quantum dipole oscillations acting as qubits [11-13]. Evolution of entangled qubits among many neurons is 'orchestrated' during integration phases of integrate-and-fire neurons, reaching 'Orch OR' threshold for full, rich moments of conscious experience. Each Orch OR event selects microtubule states which 'non-computably' modulate axonal firings, controlling behavior (and thus cause deviation from H-H 'integrate-and-fire' behavior, and from phase-amplitude coupling in EEG). Orch OR is most likely to occur in dendrites and soma of brain neurons, e.g. layer 5 cortical pyramidal neurons in which microtubules are uniquely arranged as mixed-polarity, anti-parallel networks, and thus suitable for recursive processing, nested phase amplitude coupling and interference 'beats' seen as EEG.

Overview: GNW and HOT are essentially architectural schemes of brain information flow, mediated by 'firings' of 'integrate-and-fire' neurons, and consciousness is asserted to 'emerge' from particular flow patterns. PC/RP is also concerned with information flow, specifically in recursive, or recurrent, feed-forward and feedback pathways at any scale, with recursive processing the essential feature leading to consciousness. IIT sees consciousness as arising from maximal integration which can occur at various levels and scales, including neurons, networks and potentially inside neurons in microtubules. But IIT has yet to identify specific biological measures of integration correlating with consciousness. In Orch OR, consciousness is also seen to occur in a multi-scale hierarchy, but dependent on both 1) information processing at a deeper level, in microtubules inside neurons, and 2) quantum computing 'orchestrated' in microtubules terminated by Penrose OR, a specific proposal for (proto-) conscious experience. The current proposed study addresses the first difference: Is consciousness fully accounted for by collective membrane and synaptic effects of brain neurons and networks, or does it also require intra-neuronal activities in microtubules? Whether such processing in microtubules is quantum in nature is being addressed elsewhere. See http://osf.io/zqnjd/

Cerebral organoids as model systems to test for consciousness

Consciousness cannot be directly measured or observed, but detectable putative neural correlates of consciousness ('pNCCs') may point to the underlying mechanism of consciousness. However, 'obviously conscious' biological systems present difficulties. Humans are conscious, but ethical considerations preclude invasive procedures, and the blood brain barrier blocks certain drugs. Animals have pNCCs and are likely to be conscious, but also have ethical considerations and a blood-brain barrier. Brain slices and neuronal cell cultures are accessible, and ethical concerns are negligible, but generally lack pNCCs potentially suggestive of consciousness. But a more advanced, new type of brain-like biological model derived from human pluripotent stem cells – 'cerebral organoids' – has properties suggestive of pNCCs, and are relatively easy to study.

Induced pluripotent stem cells ('iPSCs') can be generated from reprogramming somatic human cells, and then induced to form particular types of cells which self-assemble and cooperate in miniaturized organ-like tissues ('organoids'). Human brain cortical neurons, including pyramidal cells associate and develop as tiny, 3-dimensional 'cerebral organoids', generated by patterned or non-patterned protocols. The latter, non-patterned organoids, rely on spontaneous morphogenesis, and intrinsic differentiation which generate structures with the most freedom for self-organization and heterogeneity of discrete regions with cooperative electrophysiology. Cerebral organoids grow over the course of about 10 months to roughly 2.5 million neurons, have receptors and neurotransmitters including $GABA_A$ and NMDA, mimic developmental features at cellular and molecular levels in human fetal brain cells, and exhibit EEG-like activity similar to pre-term human babies [73].

Recently, non-patterned, self-organizing cerebral organoids recorded on multi-electrode arrays were discovered to have local field potentials ('LFPs') with nested EEG-like activity. During the course of organoid development, low frequency 1 to 4 Hz delta EEG-like activity occurs at about 2 months, whereas 'high gamma' 100 to 400 Hz EEG-like activity phase-coupled to slow delta occurs at 4 to 6 months, suggesting an increasing multi-scale hierarchy in organoid development. As described above, this is cross-frequency phase-amplitude coupling in which high-frequency content of the LFP is entrained to the phase of slow oscillations [74].

In humans and animals, 100-400 Hz would be 'high gamma', a correlate of consciousness, but whether this applies to organoids is unclear. Also, phase-amplitude coupling, or nesting between delta (or theta) and low gamma in animals occurs with unconsciousness (under anesthesia), but again it is unclear whether this applies to 400 Hz in organoids. Cerebral organoids manifest potential putative neural correlates of consciousness (pNCCs).

Testing for consciousness in cerebral organoids

We propose to seek and monitor pNCCs in cerebral organoids, and measure their responses to various forms of perturbation in terms of specific theories of consciousness.

Discussion - Could cerebral organoids be conscious?

There are reasons to be skeptical of the idea that cerebral organoids may be conscious. But we don't really know how consciousness occurs in the human brain, nor how far 'down' in biological evolution consciousness might go.

Nor do we know how the experiential nature of consciousness (Chalmers' 'hard problem' [3]) relates to cognition, or the so-called 'easy problems' like sensory discrimination, information integration, purposeful movement and focusing attention, all of which may be either conscious, or non-conscious 'autopilot' functions. Purposeful, cognitive functions occur in single cell organisms such as *Paramecia* that swim about, find food and mates, and have sex, all without synapses, neurons or neuronal networks. Even plants respond purposefully, their actions dampened by anesthetics. This is not to say these cognitive functions in protozoa and plants are conscious, but demonstrates that cognition extends to cellular and intra-cellular activities.

Human cognition utilizes computational auto-pilot functions upon which consciousness may supervene, as a somewhat separate activity. For example, we may walk, or even drive on autopilot non-conscious cognitive mode, but then suddenly consciousness takes control – supervenes – as something novel happens, requiring adaptive behavior. The Orch OR theory suggests conscious supervenience occurs from non-computable quantum processes in microtubules in dendrites and soma within brain neurons, e.g. cortical pyramidal cells, producing moments of consciousness, and regulating axonal firing patterns.

Since we don't really know what causes consciousness in the brain, the presence of putative neural correlates of consciousness (pNCCs) in cerebral organoids raises the possibility that they may be conscious. To probe further, we will see how pNCCs in organoids respond to various perturbations known to affect consciousness in humans and animals. Responses will be evaluated in the context of various theories of consciousness (Table 1, Figure 3).

Theories of consciousness described above are not necessarily mutually exclusive. For example GNW, HOT and PC/RP may be correct, but also require Orch OR inside neurons. Anti-parallel, mixed polarity microtubules in dendrites and cell body/soma may perform PC/RP-like recursive processing at a deeper level in microtubules inside neurons. In IIT, maximal Phi may optimize in Orch OR-based processes in microtubules.

What would it be like to be a cerebral organoid? With limited external sensory receptors, and no peripheral 'body', they would likely not feel 'embodied'. But organoids could be conscious of their internal states, as we often are of ours. They could have a series of disconnected experiences, or a rudimentary conscious self through memory and context. One might imagine that cerebral organoid (proto-) consciousness would differ from ours, with organoid consciousness of lower intensity and content of experience.

Among theories of consciousness, Orch OR predicts a spectrum of conscious events, with experiential intensity and content bandwidth related

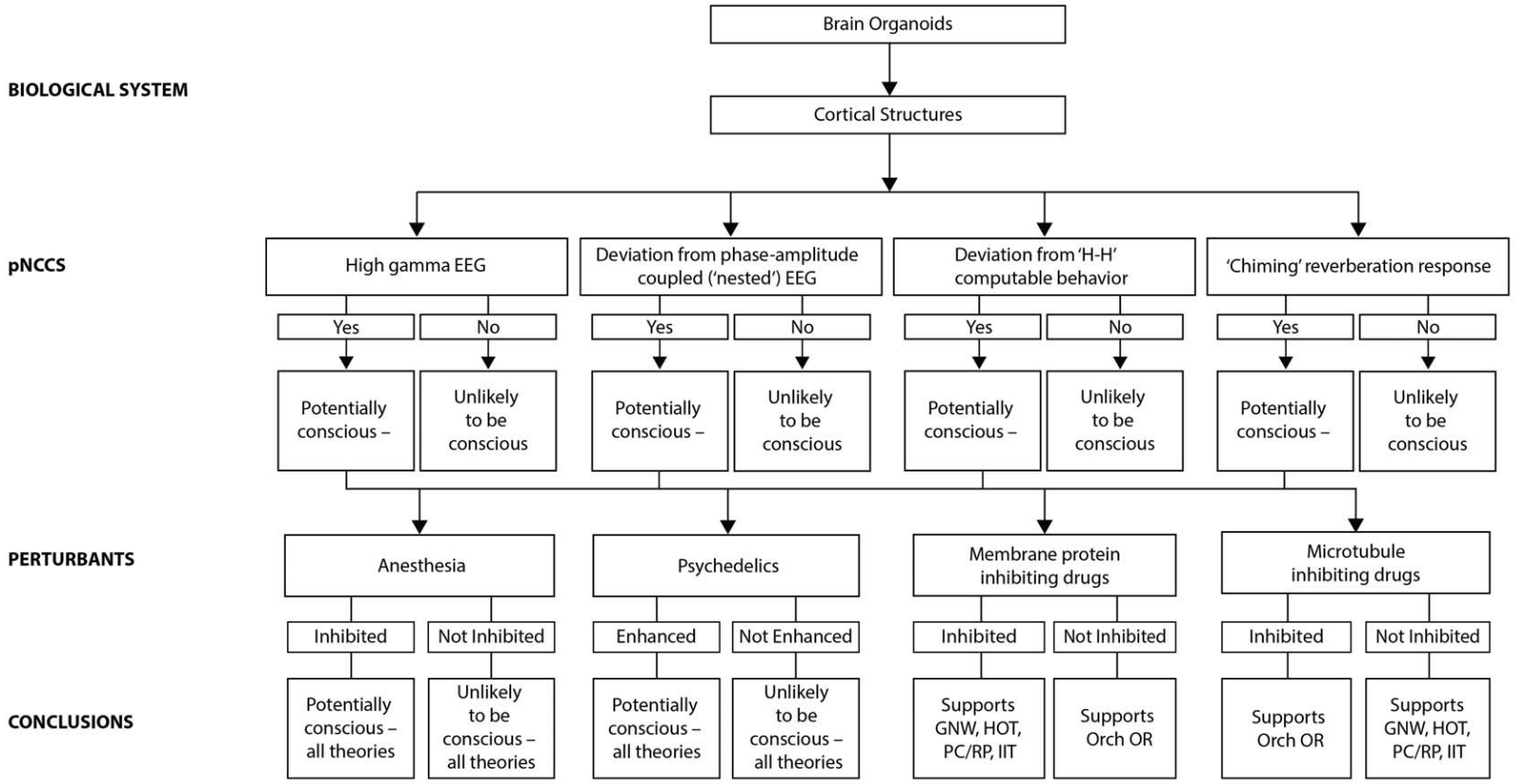

Figure 3. Schematic flow chart of experimental approaches to consciousness in cerebral organoid cortical structures. 3rd row: Putative neural correlates of consciousness (pNCCs): 1) High gamma (100 – 400 Hz) EEG-like oscillations, 2) Deviation from phase-coupled EEG (e.g. coupled 4 Hz and 400 Hz), 3) Deviation from H-H computable behavior, and 4) "Chiming", reverberative EEG response to brief TMS or ultrasound stimulation. If present, these pNCCS will be observed when perturbed in 4 different ways (7th row, Perturbants): 1) anesthesia, 2) psychedelic molecules, 3) membrane receptor and channel blocking drugs, 4) microtubule stabilizing/destabilizing drugs. Implications for each outcome in terms of possible consciousness in cerebral organoids, and particular theories of consciousness are shown as possible outcomes.

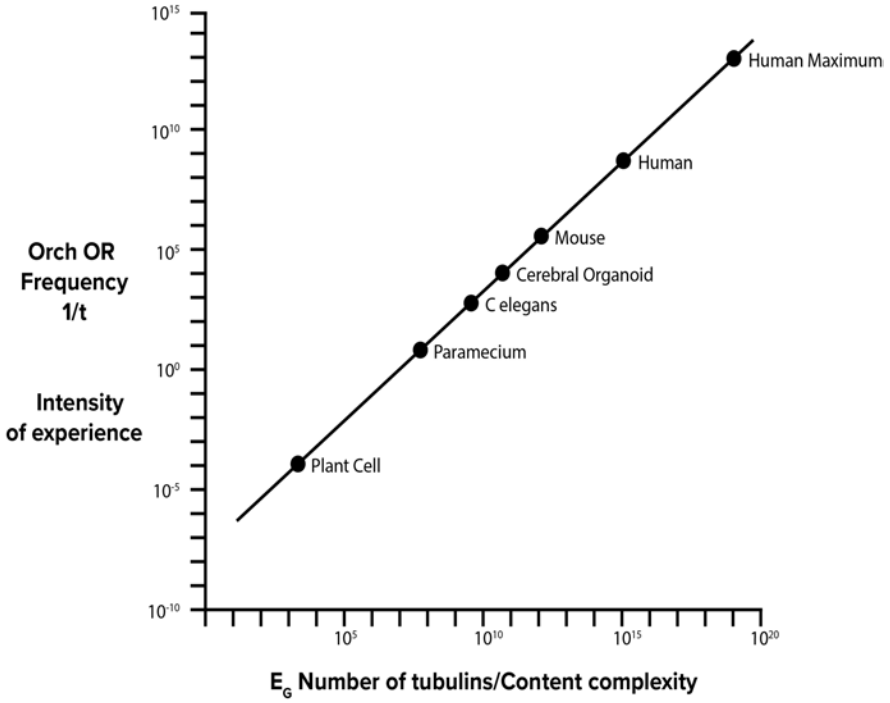

Figure 4. Orch OR events occurring at time $t = \hbar/E_G$ give a spectrum of possible levels of intensity and content complexity based on number of tubulins in brains of particular organisms, and other settings.

to frequency and amount of microtubule protein in quantum superposition states. The frequency and intensity of Orch OR conscious moments would depend on the number of microtubule subunits available for quantum computation, in turn related roughly to brain mass. Figure 4 shows a plot of a spectrum of Orch OR frequency and intensity of conscious moments for various organisms and organoids.

CONCLUSION

The scientific basis for consciousness remains unknown, and difficult to study, as consciousness cannot be directly measured or observed. However there are measureable putative 'neural correlates of consciousness' ('pNCCs') in biological brains, perturbations known to affect consciousness, and theories of consciousness (Table 1).

Recently, a relatively simple biological system of cerebral organoids, stem cell-derived 'mini-brains', has shown evidence of one pNCC (pNCC1), namely, high frequency gamma EEG, and other activity related to pNCCs including phase-coupled, nested EEG, deviation from which correlates with consciousness (pNCC2). In cerebral organoid cortical neurons, we will test for another pNCC, non-computable deviation from algorithmic Hodgkin Huxley ('H-H') 'integrate-and-fire' behavior (pNCC3). Finally, pNCC4 involves reverberative ("chiming") response in EEG. We propose to measure these pNCCs with 4 methods of perturbation: 1) anesthesia, 2) psychedelics, 3) drugs which affect neural membrane proteins, 4) drugs which affect microtubules inside neurons. In all cases one can monitor effects of these perturbations on NCCs, treat the organoids ethically, and evaluate them in the context of different theories of consciousness: GNW (global neuronal workspace), HOT (higher order theory), PC/RP (predictive coding/recursive processing), IIT (integrated information theory), and Orch OR (orchestrated objective reduction).

As summarized in Figure 3, we hope to determine mechanisms for pNCCs and their perturbations, narrow the field of theories of consciousness, come to an informed opinion on whether cerebral organoids are conscious, and establish a method for approaching consciousness in biological systems.

ACKNOWLEDGEMENTS

Thanks to Edgar Mendoza for artwork and Abi Behar-Montefiore for manuscript preparation.

CONFLICT OF INTEREST STATEMENTS

S.R.H declares no conflict of interest.

A.R.M. is a co-founder and has an equity interest in TISMOO, a company dedicated to genetic analysis and brain organoid modeling focusing on therapeutic applications customized for autism spectrum disorder and other neurological disorders with genetic origins. The terms of this arrangement have been reviewed and approved by the University of California San Diego by its conflict of interest policies.

REFERENCES

1. Nagel, T. 1974, The Philosophical Review, 83(4), 435-450.
2. Levine J. 1983, Pacific Philosophical Quarterly, 64, 354-61.
3. Chalmers, D. J. 1996, The Conscious Mind: In Search of a Fundamental Theory, Oxford University Press. New York.
4. Skrbina, D. 2005, 2017 rev, Panpsychism in the West, London, England; Cambridge, MA:MIT Press.
5. Jaynes, J. 1976, The Origin of Consciousness in the Breakdown of the Bicameral Mind. Boston: Houghton Mifflin.
6. Hodgkin, A. L. and Huxley, A. F. 1952, J Physiol 117(4), 500-44.
7. Penrose, R. 1989, The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics. Oxford; New York: Oxford University Press.
8. Naundorf, B., Wolf, F. and Volgushev, M. 2006, Nature, 440, 1060-63.
9. Craddock, T. J. A., Kurian P., Preto, J., Sahu, K., Hameroff, S. R., Klobukowski, M. and Tuszynski, J. A. 2017, Sci. Rep., 7(1), 9877.
10. Kastrup, B. 2019, The Idea of the World: A Multi-Disciplinary Argument for the Mental Nature of Reality" John Hunt Publishing.
11. Hameroff, S. and Penrose R. 1996a, Math Comput. Simul., 40, 453-80.
12. Hameroff, S. and Penrose R. 1996b, J. Conscious Stud., 3(1), 36-53.
13. Hameroff, S. and Penrose, R. 2014, Physics Life Reviews, 11(1), 39-78.
14. Doerig, A., Schurger, A. and Herzog, M. H. 2020, Cogn. Neurosci., 1-22.
15. Hughes, J. R. 1964, Int. Rev. Neurobiol., 6, 99-152.
16. Singer, W. and Gray, C. M. 1995, Annu. Rev. Neurosci., 18, 555-86.
17. Crick, F. and Koch, C. 1990, Seminars in the Neurosciences, 2, 263-75.
18. Crick, F. and Koch, C. 2003, Nat. Neurosci., 6(2), 119-26.
19. Li, D., Hambrecht-Wiedbusch, V. S. and Mashour, G. A. 2017, Frontiers in Systems Neuroscience, 11, 16.
20. Pal, D., Li, D., Dean, J. G., Brito, M. A., Liu, T., Fryzel, A. M., Hudetz, A. G. and Mashour G. A. 2020, Journal of Neuroscience, 40(3), 605-618.
21. Buzsáki, G. 2006, Rhythms of the brain. Oxford; New York: Oxford University Press.
22. Buzsáki, G. and Wang, X. J. 2012, Annu. Rev Neurosci., 35, 203-25.
23. Cantero, M. R. and Cantiello H. F. 2020, J. Neurol Neuromedicine, 5(3), 1-5.
24. Saxena, K., Singh, P., Sahoo, P., Sahu, S., Ghosh, S., Ray, K., Fujita, D. and Bandyopadhyay, A. 2020, Fractal and Fractional, 4(2),11.
25. Cohen, M. X. 2017, Trends Neurosci., 40(4), 208-18.
26. Lisman, J. E. and Idiart, M. A. 1995, Science, 267(5203), 1512-5.
27. VanRullen, R. and Koch, C. 2003, Trends Cogn Sci., 5, 207-213.
28. Bonnefond, M., Kastner S. and Jensen, O. 2017, eNeuro, 4(2), 0153-16.
29. Pagnotta, M. F., Pascucci, D. and Plomp, G. 2020, NeuroImage, 223, 117354.
30. Pal, D., Silverstein B. H., Sharba, L., Li, D., Hambrecht-Wiedbusch, V. S., Hudetz, A. G. and Mashour, G. A. 2017, Frontiers in Systems Neuroscience, 11.
31. Hameroff, S and Penrose R. 2014, Physics Life Reviews, 11(1), 39-78.
32. Sahu, S., Ghosh, S., Hirata, K., Fujita, D. and Bandyopadhyay, A. 2013, Appl. Phys. Lett., 102, 123701.
33. Sahu, S., Ghosh, S., Ghosh, B., Aswani, K., Hirata, K., Fujita, D. and Bandyopadhyay, A. 2013, Biosens Bioelectron, 47, 141-48.

34. Sahu, S., Ghosh, S., Fujita, D. and Bandyopadhyay, A. 2014, Sci. Rep., 4, 7303.
35. Saxena, K., Singh, P., Sahoo, P., Sahu, S., Ghosh, S., Ray, K., Fujita, D. and Bandyopadhyay, A. 2020, Fractal and Fractional, 4, 11.
36. He, B. J., Zempel, J. M., Snyder, A. Z. and Raichle, M. E. 2010, Neuron, 66(3), 353-69.
37. Rafiq, S., Fu, B., Kudisch, B. and Scholes, G. D. 2021, Nature Chemistry, 13, 70-76.
38. Naundorf, B., Wolf, F. and Volgushev, M. 2006, Nature, 440, 1060-63
39. Colwell, L. J. and Brenner, M. P. 2009, PLoS Computational Biology, 5, 1.
40. Koch, C. 2017, Sci. Am., 317(5), 28-33.
41. Sarasso, S., Boly, M., Napolitani, M., Gosseries, O., Charland-Verville, V., Casarotto, S., Rosanova, M., Girardi Casali, A., Brichant, J. F., Boveroux, P., Rex, S., Tononi, G., Laureys, S. and Massimini, M. 2015, Current Biology, 25, 3099-3105.
42. Redinbaugh, M. J., Phillips, J. M., Kambi, N. A., Mohanta, S., Andryk, S., Dooley, G. L., Afrasiabi, M., Raz, A. and Saalmann, Y. B. 2020, Neuron, 106, 66-75.e12.
43. Llinás, R. and Ribary, U. 1993, Proceedings of the National Academy of Sciences Mar, 90(5), 2078-2081.
44. Lamme, V. A. F., Supèr, H., Landman, R., Roelfsema, P. R, Spekreijse, H. 2000, Vision Research, 40(10-12), 1507-21.
45. Boly, M., Massimini, M., Tsuchiya, N., Postle, B. R., Koch, C. and Tononi, G. 2017, J. Neurosci., 37(40), 9603-13.
46. Lau, H. and Rosenthal, D. 2011, Trends Cogn. Sci., 15(8), 365-73.
47. Pollen, D. A. 2003, Cerebral Cortex, 13(8), 807-814.
48. Damasio, A. R. 1999, The Feeling of What Happens: Body and Emotion in the Making of Consciousness. Fort Worth, TX, US: Harcourt Brace New York.
49. Crick, F. C. and Koch, C. 2005, Philos. Trans. R Soc. Lond B Biol. Sci., 360(1458), 1271-1279.
50. Baars, B. J. 1988, A cognitive theory of consciousness, Cambridge [England], New York: Cambridge University Press.
51. Dehaene, S., Kerszberg M. and Changeux, J. P. 1998, Proc. Natl. Acad. Sci. USA, 95(24), 14529-34.
52. Hameroff, S. 2010, J. Biol. Phys., 36(1), 71-93.
53. Carhart-Harris, R. L., Erritzoe, D., Williams, T., Stone, J. M., Reed, L. J., Colasanti, A., Tyacke, R. J., Leech, R., Malizia, A. L., Murphy, K., Hobden, P., Evans, J., Feilding, A., Wise, R. G. and Nutt, D. J. 2012, Proceedings of the National Academy of Sciences, 109(6), 2138-2143.
54. Meyer, H. 1899, Naunyn-Schmied Arch. Exp. Path. Pharmakol., 42, 109-118.
55. Overton, E., Verlag von Gustav Fischer, 1901, Studien Uber Die Narkose Zugleich Ein Beitrag Zur Allgemeinen Pharmakologie Studies on Anesthesia, at the Same Time a Contribution to General Pharmacology, Originally published as Jena: Gustav Fischer.
56. Franks, N. and Lieb, W. 1984, Nature, 310, 599-601.
57. Eger, E. I. 2nd, Raines, D. E., Shafer, S. L., Hemmings, H. C. Jr. and Sonner, J. M. 2008, Anesth Analg., 107(3), 832-48.
58. Xi, J., Liu, R., Asbury, G. R., Eckenhoff, M. F. and Eckenhoff, R. G. 2004, Journal of Biological Chemistry, 279, 19628-19633.
59. Pan, J. Z., Xi, J., Tobias, J. W., Eckenhoff, M. F. and Eckenhoff, R. G. 2007, J. Proteome Res, Feb, 6(2), 582-92.
60. Pan, J. Z., Xi, J, Eckenhoff, M. F. and Eckenhoff, R. G. 2008, Proteomics, 8(14), 2983-92.
61. Emerson, D. J., Weiser, B. P., Psonis, J., Liao, Z., Taratula, O. A. Fiamengo, Wang, X., Sugasawa, K., Smith, 3rd A. B., Eckenhoff, R. G. and Dmochowski, I. J., 2013, J. Am. Chem. Soc., 135(14), 5389-98.
62. Craddock, T. J. A., Kurian, P., Preto, J., Sahu, K., Hameroff, S. R., Klobukowski, M. and Tuszynski, J. A. 2017, Sci. Rep., 7(1). 9877.
63. Carhart-Harris, R., Leech, R., Hellyer, P., Shanahan, M., Feilding, A., Tagliazucchi, E., Chialvo, D. and Nutt, D. 2014, Front. Hum. Neurosci., 8, doi.org/10.3389/fnhum.2014.00020
64. López-Giménez, J. F. and González-Maeso, J. 2018, Current topics in behavioral neurosciences, 36, 45-73.
65. Mashour, G. A., Roelfsema, P., Changeux, J. P. and Dehaene, S. 2020, Neuron Review.
66. Rao, R. P. and Ballard, D. H. 1999, Nat. Neurosci., 2(1), 79-87.

67. Lamme, V. A. F, Supèr, H., Landman, R., Roelfsema, P. R. and Spekreijse, H. 2000, Vision Research, 40(10-12), 1507-21.
68. Clark, A. 2013, Behav. Brain. Sci., 36(3), 181-204.
69. Friston, K. 2018, Nat. Neurosci., 21, 1019-1021.
70. Block, N. 2011, Analysis, 71(3), 419-31.
71. Tononi, G., Boly, M., Massimini, M. and Koch, C. 2016, Neuroscience 17(7), 450-61.
72. Koch, C. 2004, The Quest for Consciousness: A Neurobiological Approach.Roberts & Co.
73. Trujillo, C. A., Gao, R., Negraes, P. D., Gu, J., Buchanan, J., Preissi, S., Wang, A., Wu, W., Haddad, G., Chaim, I. A., Domissy, A., Vandenberghe, M., Devor, A., Yeo, G. W., Voytek, B., Muotri, A. R. 2019, Cell Stem Cell, 25(4), 558-569.e7.
74. Voytek, B. and Knight, R. T. 2015, Biol. Psychiatry, 77(12), 1089-97.