



A systematic approach to brain dynamics: cognitive evolution theory of consciousness

Sergey B. Yurchenko¹

Received: 29 March 2022 / Revised: 29 June 2022 / Accepted: 21 July 2022 / Published online: 10 August 2022
© The Author(s), under exclusive licence to Springer Nature B.V. 2022, corrected publication 2022

Abstract

The brain integrates volition, cognition, and consciousness seamlessly over three hierarchical (scale-dependent) levels of neural activity for their emergence: a causal or ‘hard’ level, a computational (unconscious) or ‘soft’ level, and a phenomenal (conscious) or ‘psyche’ level respectively. The cognitive evolution theory (CET) is based on three general prerequisites: physicalism, dynamism, and emergentism, which entail five consequences about the nature of consciousness: discreteness, passivity, uniqueness, integrity, and graduation. CET starts from the assumption that brains should have primarily evolved as volitional subsystems of organisms, not as prediction machines. This emphasizes the dynamical nature of consciousness in terms of critical dynamics to account for metastability, avalanches, and self-organized criticality of brain processes, then coupling it with volition and cognition in a framework unified over the levels. Consciousness emerges near critical points, and unfolds as a discrete stream of momentary states, each volitionally driven from oldest subcortical arousal systems. The stream is the brain’s way of making a difference via predictive (Bayesian) processing. Its objective observables could be complexity measures reflecting levels of consciousness and its dynamical coherency to reveal how much knowledge (information gain) the brain acquires over the stream. CET also proposes a quantitative classification of both disorders of consciousness and mental disorders within that unified framework.

Keywords Brain dynamics · Consciousness · Metastability · Criticality · Complexity · Bayesian brain · Mental disorders

Introduction

Recent advances in artificial intelligence (AI), inspired by neurobiology, support the idea that consciousness could arise from machine learning in exclusively computational ways without requiring any kind of freedom from artificial neural networks, if these were endowed with the global architecture for self-monitoring and metacognition (Lake et al. 2017). Machine consciousness might progress by investigating the architecture, then transferring those insights into computer algorithms (Dehaene et al. 2017). However, what criterion of conscious experience should be reliable there? How can we say with confidence that a machine is or is not conscious, if the machine operates like humans? The existing tests for machine consciousness under criteria, such as flexibility, improvisation,

spontaneous problem-solving (Pennartz et al. 2019), are widely practicable in neuroscience, AI studies, and robotics, but they depend ultimately on our subjective interpretation of behavior (Elamrani and Yampolskiy 2019). Consciousness as a phenomenon that is self-evidential through Descartes’ cogito remains elusive. How could consciousness be certified in any particular case beyond reportability?

To answer these questions, the science of consciousness has to explain how the brain integrates volition, cognition (including perception and memory), and consciousness seamlessly over three hierarchical levels of brain dynamic respectively: (i) a causal or ‘hard’ level, (ii) a computational (unconscious) or ‘soft’ level, and (iii) a phenomenal (conscious) or ‘psyche’ level. Although this schema produces an intuitive analogy with a computer’s hardware and software, the analogy is rather supportive for indicating the absence of a psyche level, i.e., consciousness, in computers by a reason unknown to us. Nonetheless, the division would still be trivial, unless it was put upon a strict physical foundation. The cognitive evolution theory (CET),

✉ Sergey B. Yurchenko
s.yucko@gmail.com

¹ Independent Research Center of Brain and Consciousness, Andijan, Uzbekistan

outlined here, argues that such a foundation can indeed be proposed by assuming that the levels for the emergence of volition, cognition, and consciousness are scale-dependent. Each of these phenomena can be best accounted for at a separate scale of emergence.

In the literature, dividing a system of interest into different spatiotemporal scales is typically defined across micro-, meso- and macroscales. Their further specification depends on the size and nature of the system of interest (e.g., the Solar system vs. the cell). Accordingly, CET relates volition, cognition (absorbing perception and memory), and consciousness to these three scales of neural activity—neuronal, modular, and whole-brain dynamics.

Volition

Generally, volition is always concerned with internally-generated or self-initiated (consciously or unconsciously) action. It must necessarily have causal power. Although causation can be described at various spatiotemporal scales, depending on the size and relevant dynamics of a system of interest, a microscale proposes always a more rigorous and detailed fine-grained picture than those of a meso- or a macroscale. Because every scale of description is biased by averaging many variables to a single one, the micro-causation is, in fact, only responsible for causal processes at all coarser scales. Ignoring this fact generates fallacious concepts in cognitive neuroscience such as ‘downward causation’, which should be more correctly called ‘correlation’ (Atmanspacher and Rotter 2008). In particular, it is now commonly acknowledged that statistical dependencies based on functional connectivity patterns extracted from neuroimaging data can produce spurious causation which is only correlation (Reid et al. 2019; Mehler and Kording 2018; Weichwald and Peters 2021). Accordingly, CET argues that, on this theoretical account, volition can genuinely be accounted for only at a causal (hard) level of physically interacting neurons. Thus, a hard level is to be associated exclusively with a microscale of brain dynamics (including atomic and even a quantum scale if only it cannot be explicitly neglected as noise at larger scales).

Cognition

In a simplest formulation, cognition (including perception and memory) is learning. This must be causally (biologically or artificially) implemented. It goes in line with the fact that cognition, in contrast to volition, occurs not at a microscale of single neurons as binary input–output devices but at a mesoscale of networks of such devices, i.e.,

anatomical brain regions exchanging information at a computational (soft) level. Of course, it can be noted that cognition and sentience occur yet at a level of unicellular organisms (Torday and Miller 2016; Baluška et al. 2021), but it is not of interest here. Importantly, in neural networks the computations cannot have causal power (allegedly via downward causation) to volitionally influence brain dynamics at the ultimate microscale. Otherwise, we should agree that computers or, at least, learning AI systems already have their share of free volition.

What is of interest here, as mentioned above, is that these systems have neither consciousness nor volition but may implement some sort of cognition and even overcome brains in performing some tasks. What is then special in brain cognition to generate conscious experience? CET argues: it is free volition. However, if volition is physically predetermined from the past, superdeterminism comes into play. Superdeterminism argues that the brain is exactly it, a leaning machine indiscernible from those AI systems. To prevent superdeterminism, quantum randomness must somehow trespass into classical neural activity at a microscale. This is a logical way that leads many physics-oriented researchers to try to reconcile consciousness with quantum effects which clearly lack in modern computers and AI systems.

Consciousness

After all, conscious states emerge globally at a macroscale of the whole-brain network. This scale corresponds to a phenomenal (psyche) level, to which many volitional and cognitive systems contribute, thereby generating what is viewed as the neural correlates of consciousness (NCC) (Crick and Koch 2003). This psyche level is neither volitional nor cognitive but only representative of these both. It is self-evidential experience of a specious (Varela 1999) or remembered (Edelman 1989) present over which a person’s ‘way of being’ (Tononi 2008) unfolds as the stream of consciousness.

These levels of emergence can, at first sight, have something to do with Marr’s (2010) tri-level explanation: computation level (why), algorithmic level (what), and implementation level (how), each suggesting its own context-dependent explanation of the same phenomenon (vision). Instead, in CET the levels are spatially scale-dependent, each being physically responsible for the emergence of the brain’s separate features—volition, cognition, and consciousness respectively. This must not be also confused with Zeki’s (2003) three spatiotemporal levels, each being hierarchically nested within a larger one. Although scale-conditioned, these are again proposed for the same phenomena: for micro-consciousness, for macro-

consciousness, and for unified experience of a person, composed of all those. In fact, Zeki's approach does the opposite to CET by rendering consciousness ubiquitous and scale-independent. For example, Hunt and Schooler (2019) go further and suggest extending Zeki's levels over the evolutionary timeline beginning with a rudimentary form of consciousness in non-organic matter at an atomic scale. Going this way, one might then come to postulating proto-consciousness at a quantum scale (Hameroff and Penrose 2014). CET does not consider this conjecture.

This suggests a dynamical model based on a framework, drawn over diverse neuroscientific domains, with contributions from classical and quantum physics, critical dynamics, predictive processing, information theory, and evolutionary neuroscience to approach a general theory of consciousness. The approach is based on three general prerequisites: physicalism, dynamism, and emergentism. These entail five consequences about the nature of consciousness: discreteness, passivity, uniqueness, integrity, and graduation. CET starts from the assumption that brains should have primarily evolved as volitional subsystems of organisms at a hard level of a microscale, not as prediction machines at a soft level of a mesoscale (Knill and Pouget 2004; Clark 2013). Only then these two levels might account for the emergence of consciousness at a psyche level at a macroscale. This also argues that consciousness is a process that can be consistently described only as a temporal stream of discrete states.

There are now a number of dominant theories of consciousness, each identifying consciousness with something else: integrated information, global workspace, predictive processing, or self-monitoring. It does not them rival, rather, fragmentary in explaining how the brain integrates consciousness (psyche), cognition (soft), and volition (hard) seamlessly across the three hierarchical levels. Many authors attempt to compare the theories (Shea and Frith 2019; Hohwy and Seth 2020; Mashour et al. 2020; Doerig et al. 2020; Sattin et al. 2021; Del Pin et al. 2021; Signorelli et al. 2021; VanRullen and Kanai 2021), or even converge them to one or another dynamical framework (Northoff and Lamme 2020; Chang et al. 2020; Cofré et al. 2020; Safron 2020).

Unlike static theories above, CET shares certain features with two dynamical theories of consciousness—Operational Architectonics (OA) (Fingelkurts et al. 2010) and Temporo-spatial theory (TTC) of Northoff and Huang (2017): they all describe the stream of consciousness, yet, doing it over critical brain dynamics, also called often scale-free dynamics (Stam and de Bruin 2004; He et al. 2010; He 2014; Fields et al. 2021). It is well-known that self-organized criticality and scale-free topology facilitate each other (Heiney et al. 2021). CET refers to the concept of criticality (Bak et al. 1987; Kelso, 1995; Blanchard et al.

2000; Chialvo 2010) as it is well grounded in thermodynamics and leads naturally to entropy-based concepts such as order, disorder, and complexity, which all are fundamental in CET. Instead, scale-freeness originates from network science where it is linked to small-world organization and self-similarity upon which AO and, especially, TTC are based.

However, there are other distinctions between CET and these theories. Both AO and TTC seem to adopt the idea of James that the stream of consciousness should be continuous. Accordingly, conscious states are proposed to have duration with an abrupt transition from one to another in critical points of brain dynamics (Fingelkurts et al. 2013; Northoff and Zilio 2022). The basic states with duration about 200 ms will then be hierarchically nested within more and more extended temporal slices up to few seconds and further over long-range temporal scales, thereby generating self-similar patterns of 'operational' or 'intrinsic' spacetime. CET takes the opposite view similar to that of cinematic theory of Freeman (Freeman 2007; Kozma and Freeman 2017): the stream of consciousness consists of discrete states which are ignited transiently at a psyche level like momentary snapshots (VanRullen and Koch 2003; Herzog et al. 2020) just at moments of phase transitions while the brain processes information continuously at a soft level in unconscious ways. The stream will then be formalized on the time continuum as a transitive and irreflexive chain of point-like conscious states. This is a principled distinction between CET and both these theories.

The term 'conscious state' can have two, at least, very different meanings. On a strict physical account, the 'state of consciousness' is a state of the brain at a given moment, typically presented with some function $f(t)$. At that moment t the brain can or cannot hold a particular conscious state. The second meaning refers to consciousness in a general sense as if averaged over time, for example, when one says that a patient is in conscious state, i.e., in the state of permanent wakefulness. But this does not mean that the patient must be in state of awareness every moment, unless one neglects the rigorous notion of the mathematical continuum. To be continuous, consciousness should pervade every point on the time continuum, e.g., within a 'sliding window' (Fekete et al. 2018). In contrast, in a discrete stream, consciousness can occupy only separate points of the continuum that are exposed to a psyche level while not making the brain generally unconscious as it is, e.g., in sleep, in coma, or under anesthesia.

Historically, the hypothesis of temporally continuous consciousness is implicitly linked with another old belief of humans in the active role of consciousness, i.e., free will. Indeed, both assumptions need a physical model which would divide brain dynamics into two continuous parts: the

“underground” for unconscious processing, and the “highway” for conscious processing as if those were going in parallel, yet requiring their own separate NCC for corresponding dynamics. It should then be assumed that the most part of its continuous time consciousness can routinely control information presented by sensorimotor regions to go on in completely deterministic ways but sometimes intervene in unconscious processing (underground?) to make its own free choice.

While it remains unclear whether or not AO and TTC admit free will, in CET, consciousness will certainly be discrete and passive. This considers the ‘subjective feeling of continuity’ and the ‘experience of free will’ two inter-linked illusions of consciousness generated by its self-evidential and representative nature: consciousness cannot in principle detect its own absence in the brain. In the stream, conscious experience will always be available on introspection as if it was self-initiated there.

CET argues, each conscious state must be volitionally driven from oldest subcortical arousal systems in the brainstem integrating functions of many vital systems and containing numerous cranial nuclei and white matter tracts to higher thalamocortical areas (Parvizi and Damasio 2001; Merker 2007). Only then those areas might be involved in perception and cognition at a soft level (Fig. 1). This is a

reason why this theory is called CET. Here, the word ‘evolution’ combines two meanings—cognitive and biological. First, unlike ordinary (though complex) dynamical systems the brain is a system that learns and memorizes. This is not merely a dynamical but an evolutionary process. In other words, ‘cognitive evolution’ means *cumulative* cognitive neurodynamics which accumulate information (knowledge) over time. Only learning systems can evolve, and brains were created by biological evolution just for it. In Darwinist terms, they do it to promote their organisms’ adaptive success (fitness). This is a point where cognitive evolution of a particular brain over its lifetime (ontogeny) and biological evolution of the brain over species (phylogeny) converge to a timeline where they advance each other.

While random gene mutations are responsible for the variety of species (and their brains) these alone might not be sufficient for evolution. Cognition is needed for action and goal-directed behavior. Under selection pressure, a brain that is more successful in its cognitive evolution by minimizing prediction-error survives and spreads its genes over generations. Those are material for new gene mutations. However, CET argues, in both ontogeny and phylogeny, before cognition may come into play, volition must be at place. To put simply, an error must be volitionally

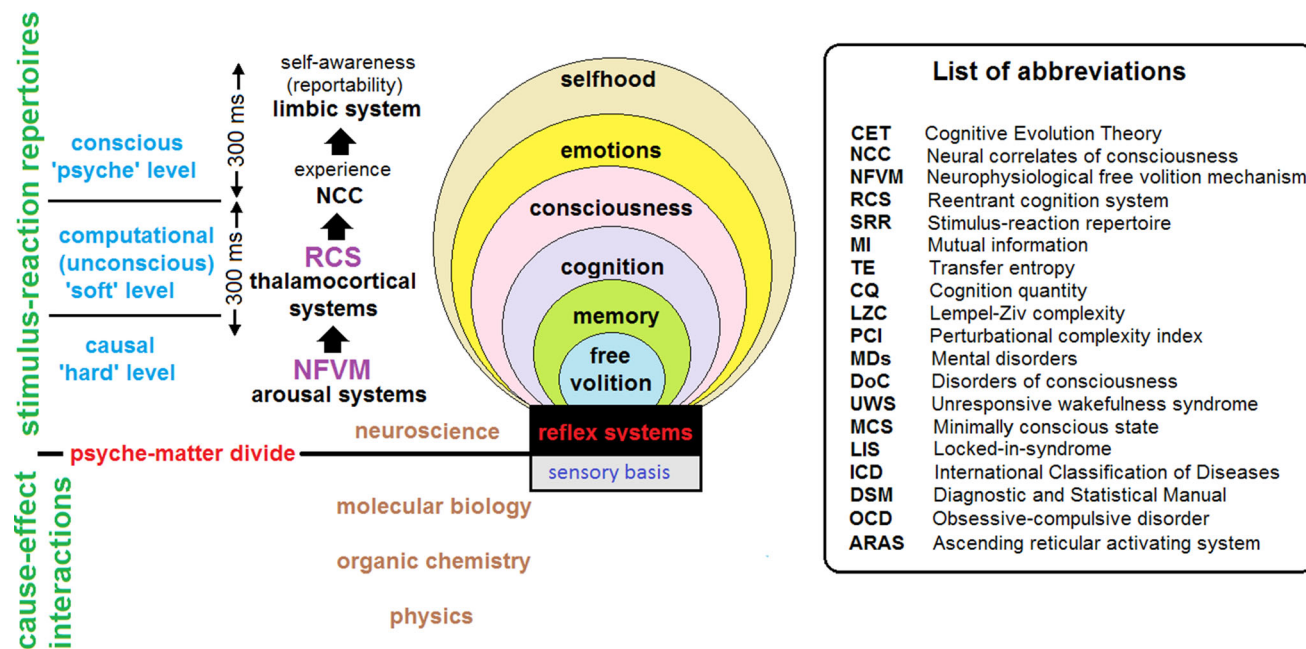


Fig. 1 The origin of consciousness. CET starts from the assumption that the brain should have primarily evolved as volitional subsystems of organisms from simplest neural reflexes. At a causal (hard) level, those should put a principled psyche-matter divide between organisms, exploiting their stimulus-reactions repertoires freely, and non-living systems, governed completely by cause-effects interactions. On the evolutionary scale, memory and cognition should evolve together,

thereby advancing each other. Their volition-driven unconscious cooperation would generate momentary conscious states over the stream at a phenomenal (psyche) level. Likewise, emotions can hardly be dissociated from self-awareness; their (limbic) neural substrates should evolve in parallel with conscious cortex-centered substrates, and motivate cognition by emotional decision-making in the functional integrity of the brain

initiated before being cognitively minimized. Hence, on the evolutionary timeline, the brain should primarily have evolved as volitional subsystems of organisms (rooted in the most ancient part of the brain, the myelencephalon). Thus, in CET, volition must be causally accounted for at a microscale, yet placed anatomically into the brainstem.

It is remarkable that just this ancient brain region, which contains nuclei of most primitive automatic functions for maintaining the body's physiological homeostasis, such as regulating blood pressure, heart rate, and breathing, combines them with arousal centers responsible for the highest phenomenon of brain activity—consciousness. Instead of speculating whether or not consciousness has free will, CET argues that in the stream of consciousness its every state has already to be volition-driven. In this sense, the brainstem not only regulates the sleep cycle for maintaining general states of consciousness (wakefulness) over extended periods of time, this also volitionally initiates 'pulsating' consciousness (Freeman 2007) like heartbeat within those periods of wakefulness. The evolution of this passive and discrete consciousness from simplest organisms to humans would then be a mere consequence of developing thalamocortical regions which might (i) modulate a volitional impulse from myelencephalon and (ii) enrich cognitive contents of consciousness over species.

After all, CET implies that not integrated information of irreducible causal mechanisms (Tononi 2008; Oizumi et al. 2014), architecture peculiarities of neural networks (Dehaene and Naccache 2001; Baars 2003), or predictive processing (Knill and Pouget 2004; Clark 2013; Seth 2014), but volition is a main obstacle that prevents computer scientists from making AI systems conscious. Just like 'consciousness', the concept of free will is far from obvious. While the Turing tests on machine consciousness can tell us nothing about the essence of consciousness, the free will tests lack a convincing theoretical paradigm for studying volition. Overall, free will is thought to be either illusory or hard to certify (Lavazza 2016). In general, the problem splits into two main accounts depending on how volition and consciousness are conceptualized.

In neuroscience, (i) consciousness is clearly taken to emerge from neural activity, and to be or not to be able to control unconscious processes. A typical conclusion is that consciousness has no control over neural processes, but the brain itself as a 'Bayesian optimal estimator' (Knill and Pouget 2004) can on its own part perform voluntary actions to minimize prediction-error or informational free energy (Friston et al. 2013). Thus, some kind of free volition is assumed to act on the authority of the brain not of consciousness itself. In this context, the volitional repertoires are usually viewed as a correction function in feedback circuitry (Clark 2013). The genuine causal freedom of such self-initiated actions is not questioned.

In physics, (ii) consciousness is presented mainly in the context of observer-dependent quantum phenomena. Superdeterminism is a hypothesis that all observed events and conscious observers themselves are completely controlled in the unambiguous causal ways from the Big Bang ('t Hooft 2016). This claim is irrelevant to the computational theory of mind from the position (i), assuming that the brain itself implements hierarchical predictive processing constantly by bidirectional causal cascades (Friston 2008; Pezzulo et al. 2018). Instead, superdeterminism holds that all neural processes are only a small part of the global causal process going over the whole universe by actions of law. As regards the brain's own counterfactual computations at a soft level, those cannot be dissociated from brain dynamics at a hard level to advance free volition despite determinism. Neither consciousness nor even the brain can have a bit of freedom. Bayesian active inference might be free of predetermination, if only its feedback circuits were closed causal loops, strictly forbidden in physics.

In this sense, the problem of free will put to the dichotomy between conscious volition and unconscious decision-making becomes unessential. The only scientifically legitimate question one can ask follows from the position (ii). Can free volition (causally independent of the past) be in principle feasible in the brain as a physical body governed by deterministic laws? A more profound conceptualization of free volition should be suggested under a criterion applicable universally to various biological and artificial systems. It happens that the problem of free will, outstanding over centuries, is intrinsically coupled with another general problem of conscious presence in those systems. Even if consciousness as a special state of matter can be measured unambiguously like mass or charge in physics (Tononi 2008; Oizumi et al. 2014; Tegmark 2015), the amount of information the system is able to integrate is secondary to the main question. What is special in this state of matter to discriminate exactly between conscious brains and non-conscious systems, integrating information as well?

But if volitional mechanisms of a system can be certified, this reveals a very special behavior, which could evolve to conscious properties by providing the system with computational power. Unlike machines, brains consist of neurons which are themselves living systems not merely binary devices (Signorelli and Meling 2021). A natural phenomenon that may then account for their autonomy lies in the quantum domain despite the fact that all significant neural processes occur apparently at classical spatiotemporal scales. On this assumption, primitive neural networks should have primarily evolved as free-volitional (quantum in origin) subsystems of organisms, not as deterministic prediction machines, requiring larger biological resources.

Accordingly, their conscious properties, typically related to higher animals, should have appeared much later than their unconscious functions as, for example, in invertebrates (Brembs 2011).

To put it sharply from the position (ii), can invertebrates have some volitional mechanism, evolutionarily embedded in their neural networks to make a choice? How might the brain make a genuinely free choice not only accessible to an organism's adaptive behavior but also evolved over species to conscious properties of higher animals? Conscious experience would then emerge as a byproduct of volitional mechanisms and cognitive thalamocortical computations based on oscillatory neural synchronization and complex patterns of brain dynamics (Ward 2011). In other words, the way Nature had chosen to evolve biological brains under natural selection can diverge significantly from the way of computer scientists in making AI systems.

The article is organized to cope more or less consistently with the multiple levels, aspects, and approaches in studying consciousness. After discussing the free will problem at a causal (hard) level and its relation to the active role of consciousness, we introduce the concept of the stream of consciousness into the framework of critical dynamics. The next section incorporates the volitional mechanism into brain dynamics to account for conscious states at a phenomenal (psyche) level. The gap between the two levels must be then filled with cognitive function at a computational (soft) level. CET adopts predictive processing as a strong candidate to provide a basis for explaining conscious contents processed unconsciously. Complexity measures are suggested in the next section to explain how consciousness might be statistically estimated. Then CET suggests the 'cognition quantity' measure that should account for the (algorithmic) coherency of cognitive processes and their impairments in mental disorders. The discussion section is given to the biological function of consciousness compared to machine consciousness.

Free will problem

The aim of this section is to account for volition at a causal (hard) level of brain dynamics.

In neuroscience, the free will problem is traditionally conceived from the position (i) to a trial whether consciousness makes a choice at will or the brain itself decides it covertly in unconscious ways (Haynes et al. 2007; Guggisberg and Mottaz 2013; Schultze-Kraft et al. 2016). Since Libet's (1985) findings, experiments were clearly put to the question: Can consciousness let an action emerging from the motor area go on or block it with the explicit veto on the movement, implemented by the prefrontal areas?

However, since any kind of intentional veto has to be also neurally processed, it can be noted that the very awareness comes after the decision was made by the brain (Soon et al. 2013).

To put the problem at the fundamental physical level from the position (ii), CET will follow Bell's approach in his famous no-go theorem (Bell 1993) and its modified version, the Free Will Theorem of Conway and Kochen (2008). Without entering into details of these theorems, the assumption of our interest here is one that concerns free will. Its conceptualization is postulated as the ability of agents to decide freely, for example, how to prepare an experiment or which measurement to perform. This is then presented by conditional probability as

$$p(A|\lambda) = p(A) \quad (1)$$

Here A is an experimenter's actual choice, and λ stands for the hidden deterministic variables, conditioned on our incomplete knowledge about the dynamics of a system. Note the system of our interest is the experimenter's brain, making a choice, not anything else. The variables λ are assumed to embrace all necessary information about the past of both the experimenter and the environment. For clarity, this conceptualization does not discriminate between consciousness and the brain as being neutral to an initiator of volition. The Bell's assumption given by Eq. (1) states that the experimenter's actual choice A has to be independent of the past (or, more exactly, of its past lightcone), just as the probability $p(A)$ holds.

Clearly, if those hidden variables could exist in principle to enable us to make the exact predictions ahead of time by $p(A|\lambda)$ for a certain choice A of the experimenter, the choice should be given up to a subjective illusion as it is usually reported in Libet-type experiments, conceived from the position (i). Moreover, the variables would dismiss any even unconscious volitional mechanism from the position (ii) as well. This would generally mean that by uncovering those hidden deterministic variables and applying them to artificial neural networks, the subject's consciousness might be copyable to run automatically on many digital clones, clearly, with no freedom available there. Thus, the relevance consciousness and volition will become obvious by noting that it seems impossible to give any operational difference between a perfect machine predictor of a subject's states, and a machine copy of the subject's consciousness regardless of their nature (Aaronson 2016).

In contrast to classical information, however, quantum information cannot be uncovered due to a random wavefunction collapse not controlled by λ . An important consequence of it comes from the No-cloning theorem that states that it is fundamentally impossible to make a perfect copy of an unknown quantum state because of its 'privacy' (Wootters and Zurek 2008). At the level of neuroscience

and computer science, it also makes impossible to clone a particular consciousness, if all its *private* states are quantum-triggered. On this condition, some free-volitional mechanism can be the only scientifically legitimate obstruction to machine-cloned consciousness.

The hypothesis that the brain is totally controlled by the hidden deterministic variables λ , contrary to Eq. (1), is called superdeterminism. Although many physicists (including Bell) find the hypothesis implausible, some still advocate it ('t Hooft 2016; Hossenfelder and Palmer 2020). Superdeterminism argues that brains are just deterministic (hence copyable) machines: humans do what the universe wants them to do while keeping in mind the illusion of free will. The striking conclusion of it is that the outcome of a subject's particular decision at any time should have been predetermined long before the subject's birth. In fact, superdeterminism amounts to a scientifically rigorous version of old-fashioned fatalism (Gisin 2013). Fortunately, while conceived to banish any sort of mysticism from quantum mechanics ('t Hooft 2016) such as nonlocal (faster than light) correlations or the observer-dependent wavefunction collapse, superdeterminism leads inevitably to much more mystical consequence such as 'cosmic conspiracies' that should violate standard statistical inequality of the Bell theorem in precisely prepared quantum experiments (Gallicchio et al. 2014).

CET puts the Bell's condition given by Eq. (1) to its foundations as a mathematically rigorous formulation of free will: *no hidden deterministic variables can have full control over brain dynamics*. In the physical framework, however, there is simply no other legal way to account for volitional mechanisms besides quantum randomness because classical processes in the brain rule out any other kind of genuine freedom (Yurchenko 2021). The probabilities in quantum mechanics are fundamentally different from those in statistical mechanics. In fact, statistical mechanics dealing with big data is still a deterministic theory. In contrast, quantum entanglement and superposition are widely used in cryptographic applications to generate the so-called Bell-certified random numbers that could not be prepared classically (Pironio et al. 2010). Hence, if we want to account for free volition, we need to admit quantum effects in the brain.

Modern theories of consciousness can be divided into two camps—classical and quantum-inspired—depending on how they decide the free will problem. Most of dominant theories belong to the first camp. Although the free will problem is largely ignored there, they implicitly rely on classical statistical physics. Thus, all those theories are prone to superdeterminism. For example, to rescue conscious will within that classical account in the context of Libet-type experiments, some neuroscientists assume that noisy neural fluctuations can be involved in self-initiated

actions (Schurger et al. 2016). They find that the key precursor process for triggering internally generated actions could be *essentially random* in a stochastic framework (Khalighinejad et al. 2018). Indeed, as the brain contains a huge number of neurons, causal neural processes can be estimated there mainly with the help of statistical descriptions. These descriptions, however, reflect the state of our knowledge that by itself does not violate determinism. This is just the reason why Bell-certification was suggested in cryptographic applications of quantum mechanics, for example, for generating a string of random numbers a Turing machine could not compute. Analogously, Bell-certification should be applicable to the volitional mechanisms of the brain in contrast to a mere reduction in classical stochastic noise.¹

In contrast, the second camp pays much attention to quantum effects which have now be well confirmed in biological systems as opposed to the expectations that those should be rapidly thermalized as noise in the warm and wet environment (O'Reilly and Olaya-Castro 2014; Chenu and Scholes 2015). It was proposed that large-scale quantum entanglement across the brain due to microtubules could endow consciousness with an active role in brain dynamics at a psyche level (Hameroff 2012; Hameroff and Penrose 2014) or be involved in long-lasting quantum cognitive processing due to spin-entangled Posner molecules at a soft level (Fisher 2015 2017). These and other quantum-inspired models of consciousness and cognition (Sabbadini and Vitiello 2019; Georgiev 2020) are beyond the scope of this paper. Most importantly, CET does not belong to any of those two camps.

First of all, CET, as stated above, rejects the possibility that consciousness might somehow be active at psyche level of brain dynamics. Second, to account for the brain's free volition at a hard level, CET proposes to do it with a minimal use of quantum randomness at a microscale of neural activity, where volition causally originates, without resorting to much more mysterious macroscopic quantum effects. Thus, while dismissing any kind of conscious will at a psyche level of a macroscale, CET assumes that quantum randomness can influence brain dynamics at a hard level of a microscale. To solve this problem, CET will recruit a molecular machinery of exocytosis according to the hypothesis of Beck and Eccles (1992, 1998) that the

¹ In principle, Bell-certification can be applicable to all quantum-inspired models in neuroscience, including quantum computing, which CET does not adopt. Quantum computing is usually suggested to account for the binding problem: how different brain regions produce a unified conscious state at a given moment. Instead, CET relies on self-organized criticality accompanied with almost instantaneous phase transitions and spontaneous scale-free signatures to solve this problem at a soft level.

brain could utilize a quantum trigger of exocytosis in a synaptic cleft. Such a micro-event might be Bell-certified.

By solving this problem, CET meets a new obstacle, which, however, can naturally be overcome within the three levels of brain dynamics. It was pointed out many times that randomness alone has nothing to do with free actions caused by a reason not randomly (Koch et al. 2009; Aaronson 2016). Only two ultimate explanations are possible there. First, if volition emerges unconsciously from completely deterministic neural processes to awareness, as it is typically reported in Libet-type experiments, there is no genuine freedom in it, and this kind of volition can be ascribed to machines as well. Second, if a subject's action would indeed be free of the past, then it was difficult to find a testable difference between physical randomness and behavioral freedom, albeit uncontrolled (Conway and Kochen 2008).

To reconcile causal freedom with cognitive control, a volitional mechanism, responsible for random quantized events at a hard level of a microscale, should be classically amplified and modulated across a mesoscale at a soft level of brain dynamics. A particular conscious state generated by the brain would then be passive at a psyche level but not predetermined from the past at a hard level. Thus, admitting quantum randomness via some neurobiological free volition mechanism (NFVM)—like the Beck-Eccles quantum trigger—is rather of logical necessity to certify the brain's genuine freedom to act against a superdeterministic and/or classical stochastic account of its dynamics (Jedlicka 2017). After all, CET places the NFVM into arousal centers of the brainstem, a phylogenetically oldest part of the brain. We will return to this issue and incorporate the NFVM into brain dynamics after introducing the stream of consciousness.

Stream of consciousness in brain dynamics

The aim of this section is to formalize the relation between brain dynamics at a causal (hard) level, and consciousness at a phenomenal (psyche) level.

The notion of the stream of consciousness has been pervasive in the literature but never properly defined. Formalizing the stream could make our understanding of consciousness, associated with multiple meanings (see e.g., Sattin et al. 2021), more operational and distinguishable from different brain processes in the same way as developing classical mechanics and thermodynamics had allowed physicists to distinguish weight and mass, or heat and temperature. Consciousness will remain elusive until we introduce a unified framework for brain dynamics, then separating conscious experience from all concomitant and

overlapping neural processes maintained by different systems.

The basic prerequisites of CET are these.

Physicalism (causality): consciousness depends entirely on neural activity governed by natural laws at a hard level, not on anything else;

Dynamism (temporality): consciousness not only requires the neural correlates of consciousness (NCC), it also needs time to be cognitively processed at a soft level;

Scale-dependence (emergentism): neural activity at micro- and mesoscopic scales cannot account for its subjective, internally generated mental phenomena at a psyche level without resorting to large-scale brain dynamics.

Physicalism, also called the mind-brain identity, deprives consciousness of any active role in brain dynamics. Or, speaking in philosophy terms, CET adopts epiphenomenalism by rejecting the idea that consciousness can have causal power over the brain. Dynamism makes consciousness the discrete stream of states like momentary snapshots (VanRullen and Koch 2003; Herzog et al. 2020), which cannot control brain dynamics at a soft level as well. After all, scale-dependence excludes multiple conscious entities in the brain like those admitted in Integrated Information Theory (Tononi and Koch 2015) or Resonance Theory of Consciousness (Hunt and Schooler 2019).

Based on the three prerequisites, CET will model consciousness as the stream of macrostates at a hard level, each specified by a particular structural–functional configuration of the whole-brain network \mathcal{N} , with $\text{NCC} \subseteq \mathcal{N}$. Here \mathcal{N} stands for a graph $G = (N, E)$, where N is the set of nodes (ideally, neurons), and $E \subseteq N \times N$ is the set of edges representing synapses. The configurations with each node's own dynamics averaged over spontaneous fluctuations in neural activity are typically presented with network statistics, extracted locally from various neuroimaging data.

The large-scale brain dynamics can then be approximated in terms of stochastic non-equilibrium systems by Coordination dynamics of coupled phase oscillators (Tognoli and Kelso 2014) or, more generally, in the Langevin formalism, a mixture of deterministic γ and stochastic ω contributions to the motion of a system,²

² Recall, 'stochasticity' that is often associated with 'randomness' is not synonymous with indeterminism. This kind of (epistemic) randomness depends on the state of our knowledge expressed in probabilistic descriptions that by itself does not violate determinism. For example, deterministic chaos cannot be Bell-certified, unless it is effectively influenced by quantum (ontic) randomness. Thus, CET explicitly assumes that the stochastic term ω in Eq. (2) is affected by the NFVM. In other words, the NFVM is a source of irreducible

$$\frac{d\psi}{dt} = -\gamma\psi + \omega(t) \quad (2)$$

Here $\psi(\mathcal{N}, t)$ is a descriptive function whose representation by the order parameter in a phase space O should account for metastability, scale-free avalanches, and self-organized criticality in brain dynamics (Bak et al. 1987; Blanchard et al. 2000; Beggs and Plenz 2003; Hesse and Gross 2014). In complex neural processes, metastable states are multiple near criticality in the activities of neurons that make up a system, thereby enlarging network repertoires for flexible behavioral outcomes (Deco and Jirsa 2012; Cocchi et al. 2017; Dahmen et al. 2019). Criticality is of crucial importance in neural activity as enhancing information processing capabilities of the brain, poised on the edge between order and disorder (Chialvo 2010; Beggs and Timme 2012).

As consciousness cannot be detected directly because of its subjective nature, accessible experimentally via a subjective report, the only way the science of consciousness is left with is to postulate its emergence from neural activity. Of course, it makes its falsification problematic (Kleiner and Hoel 2021), and explains why there are now a bewildering number of very different theories, each suggesting its own account of conscious presence in brains and other systems (Doerig et al. 2020).

In CET, the stream $S(\tau)$ of consciousness will be formally defined as a derivative of brain dynamics in *discretized* time τ ,

$$S(\tau) \stackrel{\text{def}}{=} \frac{d\psi}{d\tau} \quad (3)$$

In effect, this equation should capture the instantaneous transitions from the continuous brain dynamics to discrete conscious states, each identified with a single point $o \in O$ in a phase space, where the response of the brain to external stimuli is maximized (Shew et al. 2011; Tagliazucchi et al. 2016). A similar approach to studying consciousness over critical dynamics had been proposed by Werner (2009). Accordingly, the transitions over metastable brain states can then be viewed as neural correlates of *pulsating* conscious experience in the framework of the cinematic theory of cognition (Freeman 2007; Kozma and Freeman 2017). This approach finds now experimental evidence in many studies (Lee et al. 2010; Haimovici et al. 2013; Mediano et al. 2016; Tagliazucchi 2017; Kim and Lee 2019) exhibiting that only states integrated near criticality can ignite consciousness.

Formally, consciousness can be viewed like the physical force, derived from the momentum in Newtonian mechanics, $F = dp/dt$. Although the force is measurable and calculable, it is not a real entity but only a dynamical property of a moving system. Seeing consciousness as ‘mental force’ of brain dynamics seems to be more accurate and insightful than the view that consciousness is an intrinsic property of matter like mass (Oizumi et al. 2014; Tononi and Koch 2015). Speaking in terms of physics, there is a principled ontological difference between the force and mass in $F = ma$, where m is a scalar quantity, which is indeed intrinsic to a system constantly, whereas F is a vector quantity of a system’s action that can be zero sometimes. Similarly, consciousness can trivially lack in the brain, not to mention other material (biological or artificial) systems.

According to Eq. (3), the brain has no mental force if its dynamics depart from criticality as it occur in unconscious states such as coma, sleep, or general anesthesia (Hudetz et al. 2014; Tagliazucchi et al. 2016; Lee et al. 2019). This also says us that even in critical dynamics the brain lacks the mental force during some interval Δt while reaching dynamically a next critical point. There are two complementary ways to estimate Δt —either by monitoring brain dynamics to calculate phase transitions or due to a subjective report. The problem of monitoring, however, is non-trivial because of heterogeneous timescales involved there (see e.g., Golesorkhi et al. 2021). An optimum timescale in resting state and in task data is usually reported to be around 200 ms (Kozma and Freeman 2008; Deco et al. 2019).

Here we follow the second way and assign the interval to a wide temporal window $\Delta t \approx 100 - 450$ ms, comprising many experimental findings from video sequences of intelligible images about 7–13 per second (VanRullen et al. 2014) to the attentional blink on masked targets separated by 200–450 ms (Shapiro et al. 1997; Drissi-Daoudi et al. 2019). Yet, an important neurophysiological aspect of brain dynamics is that the stream cannot be normally delayed for a period longer than about 300 ms because this timescale is crucial for the emergence of consciousness (Dehaene and Changeux 2011; Herzog et al. 2016). Consciousness spontaneously fades after that period, for example, in anesthetized states (Tagliazucchi et al. 2016).

In CET, consciousness and unconsciousness do not cooperate in parallel as if advancing each other in two separate dynamics (highway vs. underground). Conscious states appear instantaneously as a snapshot accompanied with a phenomenal percept of the ‘specious’ (Varela, 1999) or ‘remembered’ present (Edelman 1989) at a particular moment of time. Consciousness is not an independent observer of how the states were prepared so that awareness requires no time to ignite. Since the ignition across the

Footnote 2 continued

uncertainty in the initiation of every conscious state in brain dynamics due to which the privacy of a subject’s stream cannot in principle be uncovered and cloned.

whole brain's workspace at a psyche level occurs phenomenally due to self-organized criticality (Friston et al. 2012), it is not dynamical process that should transmit information into a special site of the network \mathcal{N} to conscious experience. Experience is just the information the brain has processed at a moment τ . Although conscious experience emerges only at critical points like a snapshot (VanRullen and Koch 2003; Herzog et al. 2020), subjects can feel the continuity of being as if they were conscious all the time.

Now let $M : O \rightarrow V$ be a mapping from the phase space O onto a vector space V over the product $N \times N$ of all nodes (neurons) of the brain. The map returns $S(\tau)$ from a point $o \in O$ to an N -dimensional vector $x = [n_1, \dots, n_N]$ where $n_i = 1$ or $n_i = 0$ for neurons active or inactive at a given time. Thus, each state $S(\tau)$ can be represented by x as a certain structural–functional configuration of \mathcal{N} at a moment τ responsible for subjective experience at a given critical point. This is also a particular NCC (see the next section). We write,

$$S(\tau) \xrightarrow{M} x \quad (4)$$

The discreteness of the stream means that all conscious states can, at least in principle, be naturally enumerated from a subject's birth, not merely by a lag in experimental settings. Let the brain bring consciousness to a state x_i at a moment $\tau = t$. We can return Eq. (4) into the continuous time description (omitting details),

$$\psi(\mathcal{N}, t) = x_i \quad (5)$$

The next conscious state will emerge over the time interval as

$$\psi(\mathcal{N}, t + \Delta t) = x_{i+1} \quad (6)$$

In a timeless description, the stream $S(\tau)$ is a discrete chain $(X, <)$, where $x_i \in X$ and whose relation $<$ standing for temporal/causal order is transitive and irreflexive. Here the irreflexivity means $(\forall i)x_i \not< x_i$ that forbids closed causal loops and, in particular, instantaneous feedback circuitry in brain dynamics that might enable consciousness with causal power over the brain, for example, due to the quantum temporal back-referral of information (Hameroff and Penrose 2014). In CET, consciousness is neither active nor continuous so that it cannot—classically or quantum-mechanically—choose its own way in brain dynamics. How can then the stream be free of predetermination?

Consciousness NFVM-driven in brain dynamics

The aim of this section is to incorporate free volition into the stream of consciousness. To do it consistently, consider the concept of NCC in the framework of the stream. The

NCC has been traditionally defined as the *minimal neural substrate* expressed by specific signatures that are necessary and sufficient for any conscious experience (Crick and Koch 2003). This is based on the assumption that a key function of consciousness is to produce the best current interpretation of the visual scene and to make this information available to the planning stages of the brain (Rees et al. 2002). In general, the empirical search for NCC is implicitly based on the idea that that consciousness plays an active role in presenting a subject with a multimodal, situational survey of the environment, and in subserving their complex decision-making and goal-directed behavior (Pennartz et al. 2019). This idea is widely accepted in the neuroscientific community. When unfolded over an evolutionary scale, the idea, thus, leads to a scenario that there should have been developed special neural networks which might enable consciousness with mental power to control cognition at a computational (soft) level, yet picking up free decision-making from the neural computations at a causal (hard) level.

On the other hand, the importance of dissociating the true NCC from the variety of neural processes which underpin conscious experience has been often stressed, and the role of different areas of the brain in specifying conscious contents is debated over decades (Noë and Thompson 2004; de Graaf et al. 2012). However, the extensive cortical and subcortical networks involved in integrity of large-scale brain dynamics make it difficult to precisely identify the contribution of individual brain regions to NCC (Mashour and Hudetz 2018). Most importantly, the problem of defining the true NCC is tightly intertwined with the problem of defining consciousness and its biological function. How can we study the correlates of consciousness without knowing what the consciousness is and how had it evolved?

CET allows specifying the problem more operationally by decomposing the concept of NCC into certain neural configurations that are responsible for different conscious states. In principle, we can uncover NCC for any particular state x_i by merely detecting activity patterns in \mathcal{N} at that moment τ . Then we can define the minimal neural substrate by the intersection of all those states over the stream, or, more generally, as

$$\text{NCC}^{\min} = \bigcap_{i=1}^{2^N} x_i \quad (7)$$

Here 2^N is a set of all possible states from full vigilance to sleep to coma a subject might have during lifetime. Thus, to identify which minimal correlates should be necessary for consciousness, we need to associate NCC^{\min} with the most primitive core of conscious presence presented clinically in brain-injured patients with the

unresponsive wakefulness syndrome (Giacino et al. 2014) or by subcortical consciousness in infants born without the telencephalon, e.g., in hydranencephaly or anencephaly (Merker 2007).

Otherwise, if one wishes to assume active (and continuous) consciousness thought to be involved in attentional effort, active inference, decision-making, planning, goal-directed behavior, and other functions, it can be that NCC would comprise the most part of the brain for its own “highway” to produce these activities,

$$\text{NCC}^{\text{active}} = \bigcup_{i=1}^{2^N} x_i \quad (8)$$

Thus, it can be futile to try to identify neural correlates of consciousness without having confidence in its role in brain dynamics. On the other hand, because conscious contents cannot be evaporated from conscious states to account for empty or contentless phenomenal experience (Hohwy 2009; Bachmann and Hudetz 2014), neural correlates thought to be involved in conscious experience have already been involved in volitional and cognitive processes (Naccache 2018; Aru et al. 2019). Of course, this fact by itself is indifferent to the question “what causes what?” Namely, it does not explain whether or not consciousness is active there.

Contrary to the idea of active consciousness, CET takes the ‘inverted perspective’: consciousness is a passive phenomenon, ignited at critical points of brain dynamics and resulting completely from unconscious computational processes at a soft level. Instead of discussing different brain regions with their contributions to subjective experience, e.g., the prefrontal cortex vs. posterior ‘hot zones’ (Koch et al. 2016; Boly et al. 2017), or involving pre-stimulus and post-stimulus activity correlates (Northoff and Zilio 2022), CET argues: there is no special NCC that might causally influence brain dynamics at a hard level. Conscious states emerge transiently at a psyche level as neural configurations $x = [n_1, \dots, n_N]$ triggered by the NFVM and classically amplified by bottom-up causation via scale-free avalanches that are intrinsic to and ubiquitous in critical dynamics (Beggs and Plenz 2003; Hahn et al. 2010; Shew et al. 2011).

Another supportive evidence comes from experiments showing that perturbations or nanostimulations in vivo of a single neuron can cause those avalanches and induce phase transitions of cortical recurrent networks, thereby modifying global brain states (Fujisawa et al. 2006; Cheng-Yu et al. 2009; London et al. 2010; Houweling et al. 2010) with a marked impact on conscious states (Tanke et al. 2018; Knauer and Stüttgen 2019). Of course, such experiments, if viewed in the context of Libet-type experiments (Fried et al. 2011) to account for volition, do not provide

direct evidence to the NFVM because the amplifications are mainly detected on cortical neurons.

Meanwhile, CET places the NFVM into the brainstem to account for internally-generated quantized neuronal events that might generate scale-free avalanches across the brain. It is based on the fact that just the brainstem is responsible for spontaneous arousal and permanent vigilance conducted through the ascending reticular activating system (ARAS) to thalamocortical systems (Parvizi and Damasio 2001). Although the cortex is mostly responsible for elaborating conscious contents, only the ARAS and intralaminar nuclei of the thalamus can abolish consciousness. Moreover, due to the brainstem’s anatomical location in the neural hierarchy, its neuromodulatory influences, acting as control parameters of criticality, are capable of moving the whole cortex through a broad range of metastable states, responsible for cognitive processing in brain dynamics (Bressler and Kelso 2001).

On the evolutionary timeline, brains had evolved gradually as multilevel hierarchical systems consisting of anatomical parts that were selection-driven as adaptive specialized modules for executing one or another function. Any brain function requires an appropriate neuronal structure for generating various dynamical patterns to carry out it optimally. It is well known that the global architecture of the brain is not uniformly designed across its anatomical parts, which structural features are specialized under corresponding functions. For example, the cortex and the cerebellum exhibit various network properties. Possibly, the network characteristics of the brainstem with its reticular formation were developed to be especially conducive to small neuronal fluctuations that might be amplified across many spatiotemporal scales to account for reflexes and primary volitional reactions projected afterwards to higher thalamocortical systems (see “Discussion”).

How could then conscious states be causally free from the past? CET takes for an illustration the proverbial coin toss scene. If someone, say, Alice tosses a coin at her truly free will resulting from a micro-event in her brain, and amplified in spontaneous neural activity, the action is independent of the past, and, hence, the macro-event caused by Alice occurs genuinely random (not predetermined by the entire previous history of the universe). Although the outcome of tossing is typically probabilistic with a corresponding distribution, the trajectory of the coin is completely deterministic. The random outcome is thus *epistemic*, i.e., related to the state of our knowledge about the coin’s behavior, not of the behavior itself. Nevertheless, this can be Bell-certified, if Alice’s conscious states (coupled with corresponding actions) were indeed NFVM-triggered in her brainstem in the same way as, for instance, quantum effects can participate in bird navigation based on

the interaction of electron spin with the geomagnetic field (Ritz 2011; Hiscock et al. 2016).

To put the stream $S(\tau)$ and the NFVM together, CET will replace the fundamental cause-effect framework by a behavioristic stimulus-reaction space. In general, all physical interactions of any sort throughout the world can be viewed in the behaviorism language insofar as any physical system from a particle to a planet depends on its environment. Instead of using the cause-effect language, one can assume conversely that all physical systems ‘respond’ on the environmental ‘stimuli’ by adding nothing to a standard physical theory, for example, by saying that planets behave adaptively to gravitational fields in spacetime. Clearly, no behavioral freedom could be possible there.

In contrast, CET assumes that the brain has some freedom to respond to stimuli, and introduces a stimulus–response repertoire (SRR). Considering brain dynamics within the SRR results in information not about what possible causal mechanisms should lead the brain to its current state, but how the brain could arrive at a certain state among many possible responses from a given state (stimulus). Integrated Information Theory, for example, stresses the cause-effect repertoire of brain dynamics to compute the information generated when the system transitions to one particular state out of a repertoire of *statistically* possible (counterfactual) states (Tononi 2008). In reality, however, every next state of the brain emerges from the previous state (a particular NCC) that has been already actually defined in the past. Moreover, the SRR may be

physically possible just due to metastability in critical dynamics that provide variability and perceptual transitions (Haldeman and Beggs 2005), thus, leaving room for volitional responses there (Fig. 2a).

The responses should then be presented by the probability p , related to our incomplete knowledge about a system’s behavior that, however, could be completely predetermined at a hard level. The probability distribution behind an SRR would thus amount to n counterfactual outcomes the brain might arrive at a moment τ . However, if we adopt the Bell’s assumption, no hidden deterministic variables λ can control the NFVM. Now we can formally define the mechanism by translating Eq. (1) into CET:

$$\text{NFVM} : (\forall i)p(x_i|\lambda, x_{i-1}, y) = p(x_i|x_{i-1}, y), \quad (9)$$

where x_{i-1} and y stand for the previous state in the stream $S(\tau)$ and the environmental variables respectively.

Equation (9) returns CET to standard stochastic descriptions of brain dynamics, but now the descriptions can be Bell-certified, not merely *statistically* independent from the environment. On this condition, the probability of a subject’s choice could not be refined to unlimited precision in principle for lack of such variables. Overall, the stream $S(\tau)$ evolves as a chain $(X, <)$ of separate conscious states, each computed unconsciously at a soft level within a given SRR (Fig. 2b).

This picture is in agreement with many neuroscientific findings, firstly, on bistable perception within a constant SRR. Binocular rivalry is a phenomenon of visual

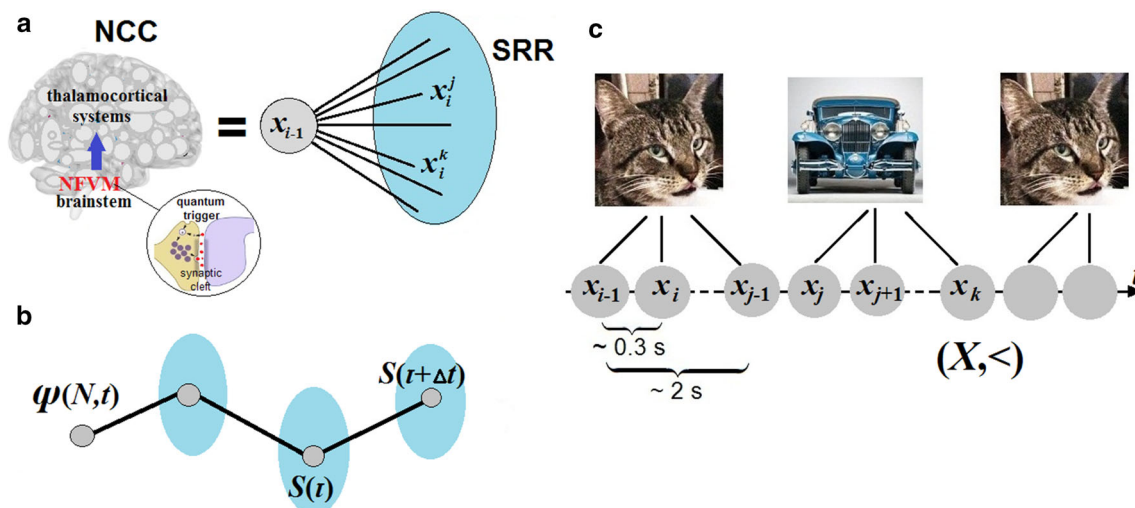


Fig. 2 The stream of consciousness. **a** In brain dynamics, every conscious state evolves from the previous one as a schematic bunch of all possible metastable states x_i^j , processed within a current SRR in a state-space and collapsed after Δt to a certain conscious state balanced at criticality. Placing the NFVM into the brainstem responsible for arousal and vigilance guarantees that each conscious state will initially be free from predetermination. **b** The stream $S(\tau)$ is shown as a broken (bold) line, running over bunches of different SRRs, each

triggered by the NFVM. A state $\psi(\mathcal{N}, t) \rightarrow x_i$ emerges instantaneously as the ‘winner-take-all’ coalition that does not transmit information to special NCC. **c** In binocular rivalry, while the incoming signals remain constant, the percept switches to and fro over a temporal period about 2 s during which many states x_i are to be processed in $S(\tau)$. Instead of visually experiencing a confusing picture of two images (a cat and a car) simultaneously, subjects report a perceptual alternation in seeing only one of those at a given time

perception, placed between different images, which are presented separately to each eye but at the same time. Instead of the two images being seen superimposed, only one single image is consciously perceived at a time. After a few seconds, while the brain has processed many states x_i focused on the same image, there is a switch to perceiving the other image, after which the cycle repeats (Fig. 2c). Binocular rivalry appears between two competing hemispheres beyond any conscious volition, and can be an example of how the NFVM affects perceptual switches passively observed by consciousness.

How should the brain unconsciously arrive at a certain state x_i by deciding between two equivalent stimuli during a bounded interval Δt ? Commonly accepted approaches to binocular rivalry stress just the role of randomness that should account for alternating conscious scenes within a constant SRR. Data of various experiments characterize the alternation by a crucial influence of noise in neural activity mediating deterministic dynamics (Brascamp et al. 2006). A similar explanation is given in Hohwy et al. (2008) on predictive processing as a competition of priors between two error-minima, each per image in a free-energy landscape with bistability in stochastic resonance. First, any of mentioned explanations of binocular rivalry in terms of classical stochastic processes does not contradict CET. We only ask how the conscious states alternated within the same SRR might be free in brain dynamics. According to Eq. (9), the principled premise here is the NFVM, which guarantees that the very arousal underlying each conscious state in the stream $S(\tau)$ will already be free of predetermination.

The NFVM can well be reconciled with some mental diseases such as obsessive–compulsive disorder, accompanied with distortions of the sense of agency, when patients fail to respond whether or not they were responsible for a particular action. The experience of free will is reported to be (often painfully) affected (Oudheusden et al. 2018) by the presence of intrusive recurrent thoughts and unwanted urges with compulsively repetitive acts. Such distortions must be directly related to cognitive function: if the process of unconscious control is violated at a soft level, a subject can experience distortions of the sense of agency at a psyche level as if someone else had dictated the subject's choice. But the NFVM is intact. In other words, the NFVM is just the invisible one that initiates at a hard level those decisions, internally-generated at a soft level and exposed then to a psyche level.

Volitional-cognitive complex

The aim of this section is to provide discretized stream of consciousness at a phenomenal (psyche) level with the multitude of cognitive (unconscious) processes at a

computational (soft) level, imposed upon brain dynamics at a causal (hard) level and implemented by various functional systems, thereby connecting consistently all the hierarchical levels across the three spatial scales of neural activity.

Many proponents of the active role of consciousness suggest that free will can trespass computationally into brain dynamics, but only under the set of special circumstances. For example, higher-order thoughts can involve the use of language when planning future actions at a soft level. Yet, the soft level should require its own causal explanation beyond a hard level (Rolls 2020). CET rejects this hypothesis as physically implausible. First of all, Eq. (3) does not discriminate between different kinds of conscious states. All the states must be uniformly processed, yet Bell-certified at a hard level so that some states cannot be causally freer than others. On the other hand, CET recognizes that only rapid and random reflexes might benefit from the NFVM because the stream of consciousness, consisting of completely random states, would be cognitively disconnected and, thus, ill-adaptive. The mechanisms of control are also necessarily for acquiring knowledge and understanding experience through coherent predictive processing. This implies a two-stage model in which random neural events, initiated by the NFVM from arousal nuclei at a microscale, will be unconsciously constrained by cognitive thalamocortical systems at a mesoscale before reaching conscious states at a macroscale.

To do it, CET adopts the predictive processing theory (PPT) as a strong candidate for a soft level that can bridge the gap between brain dynamics at a hard level and phenomenal experience at a psyche level. PPT postulates that brains should have evolved mainly as prediction machines (Knill and Pouget 2004; Friston 2008; Clark 2013) which minimize prediction-error to support best adaptive responses within alternating SRRs. Hohwy and Seth (2020) argue that PPT—precisely because it is a theory of perception, cognition, and action along which the stream $S(\tau)$ unfolds dynamically—could provide a systematic basis for a complete theory of consciousness. Such a theory needs to incorporate volition, consciousness, and cognition seamlessly into a general framework. There were proposed various combinations of PPT with known theories of consciousness such as Integrated Information Theory or Global Workspace Theory (Safron 2020; VanRullen and Kanai 2021). The main advantage of PPT before these static theories is its intrinsically dynamical nature. Another way to introduce dynamics into the theories is self-organized criticality (Tagliazucchi 2017; Kim and Lee 2020). However, in the framework of CET criticality and predictive processing are well compatible in describing brain dynamics: the former is about causation at a hard level, and the latter is about computation at a soft level. Moreover,

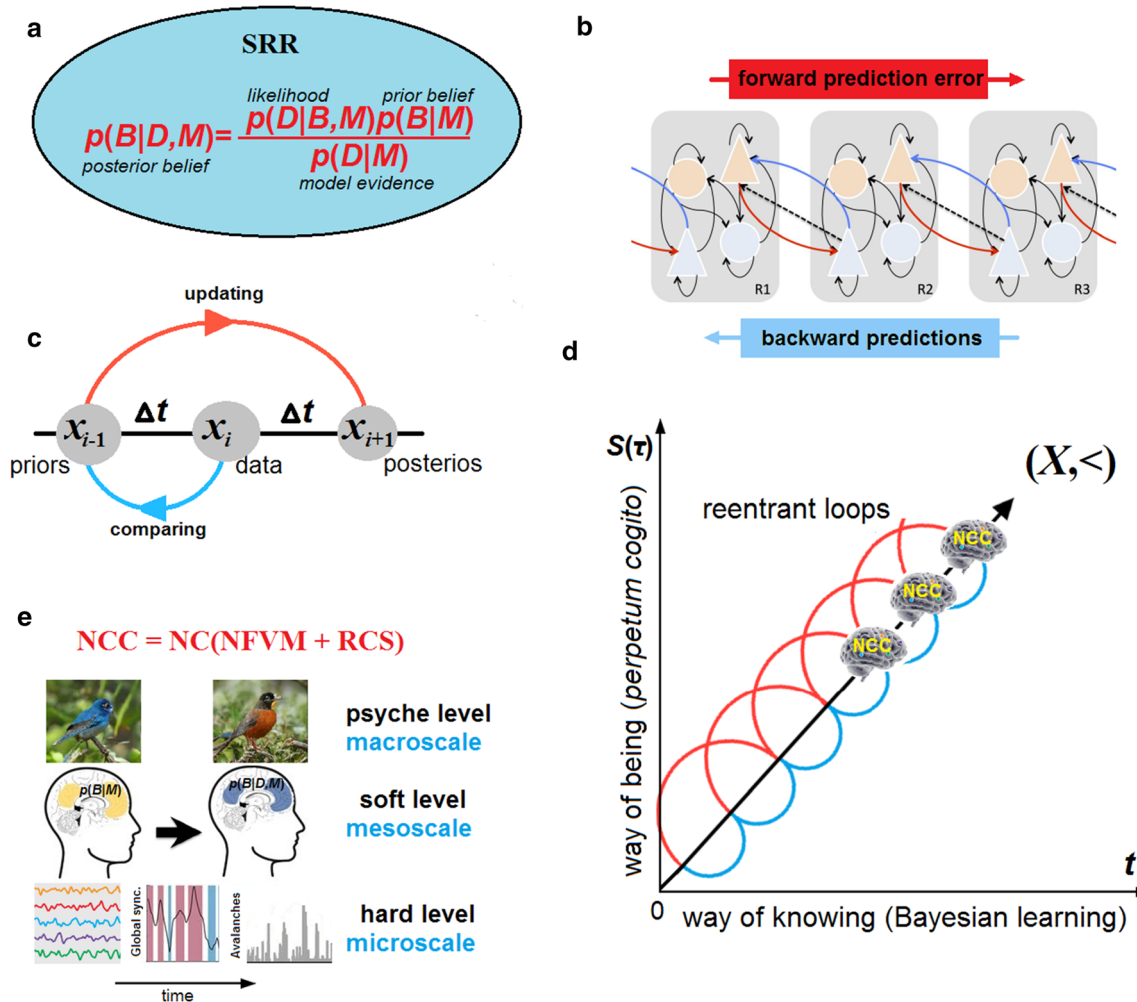


Fig. 3 The reentrant cognition system in predictive processing. **a** The Bayes theorem describes how the prior belief B (expectation) based on the brain's generative model M is transformed into the posterior belief over data acquisition D , all placed into a state space of a given SRR. **b** Hierarchical predictive processing across three cortical regions with feedforward and feedback information flow (adapted from Friston 2008). **c** Triplets of successive states, accompanied with self-organized recurrent neural activity across hierarchically distributed brain areas, are connected by the RCS over two Δt intervals as an unclosed causal loop. Consciousness at the present state x_i (data

acquisition) self-refers (blue short arc) to its previous state x_{i-1} (priors) to arrive (red long arc) at the future state x_{i+1} (posteriors). **d** Here brain dynamics are mapped onto the irreflexive chain $(X, <)$ in a 2-dimensional space of a physical axis t (way of knowing) and a phenomenal axis $S(\tau)$ (way of being). The chain evolves by the RCS as the *perpetum cogito* process, running from a subject's birth moment. Unlike a statistic description of Bayesian learning above, in dynamics, the priors, posteriors, and data acquisition become a single process. **e** The schematic of neural activity over the causal, computational, and phenomenal levels of description

criticality is thought to optimize information computing (Shew et al. 2011). It turns CET to Bayesian learning as a core computational devise of predictive processing.

Bayesian learning is the transformation of priors about the parameters into posteriors via data, presented by stimuli within a given SRR (Fig. 3a). The posteriors become over updating the priors of the brain's generative model for future data in predictive processing over SRRs. Bayesian learning is often thought of as a single process implemented by means of top-down and bottom-up signal flow over hierarchical layers (Friston 2008; Seth 2013) in the brain (Fig. 3b). The same models are successfully exploited in deep machine learning. First, unlike brains, such

machines lack any conscious experience at a psyche level,—and it occurs by a reason unknown to us. Second, suppose the machine might be conscious in that single process. Did it mean that its conscious states should all emerge only as priors or as posteriors (related to the output layers of generative models in machine learning)? To translate Bayesian learning into the language of the stream $S(\tau)$, CET takes priors and posteriors to be separate conscious states, each unconsciously processed during an interval Δt . One more state must then be placed between them for perceived data. Laying now priors to the boundary conditions of Bayesian learning, its full cycle needs a triplet $\{x_{i-1}, x_i, x_{i+1}\}$ (Fig. 3c). Importantly, such a triplet

arises only in a static representation, requiring three successive conscious states. In brain dynamics, they become just mixed into a single process in which priors turn out into posteriors that serve for data acquisition in next states.

Importantly, by turning to Bayesian learning, CET comes naturally at the hard-soft duality between brain dynamics at a causal level, and predictive processing at a computational level, both expressed with the same statistical tools. Meanwhile, predictive processing is about *subjective* information the brain has computed from its own perspective in a given *objective* state, not about the state itself (a particular NCC). In other words, this reflects a cognitive (epistemic) aspect of neural activity, not its physical (ontic) aspect, presented by $\psi(\mathcal{N}, t)$, which encodes that information in neurons. The NCC can be uncovered by neuroimaging data, whereas its contents are accessible only due to a subjective report. Without realizing this hard-soft duality, a reader can be confused. In own framework, priors, data acquisition, and posteriors all refer to conscious contents the brain has learned from its own perspective, whereas $S(\tau)$ conforms to a certain NCC responsible for those contents at the physical level. Because of the duality, we can know everything about the NCC but be still unable to explain how subjective experience appears there.

Indeed, consciousness is subjective self-evidential experience. This is the essence of Descartes self-referential cogito “I think, therefore I am.” Its stream, Tononi (2008) argues, is a way of being rather than a way of knowing. Conscious experience cannot, however, be dissociated from its cognitive contents (Hohwy 2009; Aru et al. 2019; Naccache 2018), as well as from its introspective account, i.e., self-awareness (Lau and Rosenthal. 2011; Friston 2018). Now we argue that the *way of being* (consciousness) and the *way of knowing* (cognition) go side by side by imposing self-referential cogito upon Bayesian learning. On this condition, every conscious state in the stream $S(\tau)$ should self-refer. Recall, however, that the chain $(X, <)$ of conscious states is irreflexive since closed causal loops are forbidden there. Neurophysiologically, thus, self-reference cannot be made instantaneously but needs time to be causally processed in brain dynamics, with consequent subjective experience. When consciousness self-refers, it refers to its present state, while coming causally and computationally into the next updated state in the stream $S(\tau)$.

Henceforth in CET, self-referential cogito will follow Bayesian learning in every conscious state over the stream. Meanwhile, consciousness and cognition both depend entirely on brain dynamics: the way of knowing originates from metastability, and the way of being emerges near criticality. Critical dynamics allow thus to naturally

separate unconscious predictive processing from conscious experience, ignited instantaneously only at particular moments of time. Without the conceptualization, presented by Eq. (3), it would be hard to explain how conscious snapshots were separated from both continuous brain dynamics at a causal level, and unconscious predictive processing at a computational level.

If so, then from a perspective of neural circuitry, information flow in the brain should somehow embody Bayesian learning and cogito with corresponding neural mechanisms. Reentry is a typical neurophysiological device suggested by Edelman et al. (2011) for the binding problem: How do functionally segregated areas of the brain correlate their activities in the absence of an executive program or superordinate map? Reentry is viewed as an ongoing process among competing neuronal groups of the dynamical core, which is central to the emergence of consciousness in a particular state (Edelman 2003; Baars et al. 2013). This emphasizes the role of recurrent activities between cortical areas by feedforward and feedback connections (Mashour et al. 2020). It is also shown that critical dynamics are well compatible with learning in recurrent neural networks (Del Papa et al. 2017).

In CET, the system comprising all thalamocortical areas involved in perception and cognition, with the predominant role of the prefrontal cortex in cognitive control (Miller and Cohen 2001), will be called Reentrant Cognition System (RCS). The RCS has to capture the dual aspect of brain dynamics and provides both global and local dynamical binding of neural activity: while being a causal system, this is responsible for long-term cognitive coherency of the stream $S(\tau)$ over Bayesian learning that is schematically depicted by an unclosed temporal loop imposed upon brain dynamics with respect to causality (Fig. 3c). The RCS must be (i) autonomous, (ii) self-connected over $S(\tau)$, and (iii) applicable uniformly to every conscious state experienced and remembered along the way of being through self-referential cogito. In the stream, self-awareness emerges from recursive applications of primary perceptual experience at a moment τ to cognitive contents at the next moment $\tau + dt$. In this sense, self-awareness is what the brain has learned about its own representations of the world (Cleeremans 2011).

This is just the reason why the process can be called “perpetum cogito” (Yurchenko 2017) in which priors, data acquisition, and posteriors intertwine with each other into a single process by recurrent (causally unclosed) neuronal structural–functional loops over time. Thus, self-organized criticality at a causal (hard) level, predictive processing at a computational (soft) level, and self-evidencing conscious experience at a phenomenal (psyche) level all should be covered by the perpetum cogito (Fig. 3d). The brain does not store perceptual data and intermediate computations;

only its ultimate decisions (“best guess”) over Bayesian learning will be stored. It explains why brain states can be preserved when reach conscious experience, whereas unconscious information, underlying the decisions, quickly decays (Dehaene and Changeux 2011). While being ignorant of unconscious processing (e.g., in visual masking), consciousness remains well informed about the brain’s ultimate decisions (e.g., in binocular rivalry), and thus acquires an illusion of volitional and cognitive control. Thus, the *perpetum cogito* process provides the discrete stream of consciousness with the persistent and temporally extended sense of Self. Importantly, this must not be confused with *conscious processing* which covertly requires its own “highway” in brain dynamics to control the unconscious “underground” of neural activity. CET finds the very term fallacious as leading to the illusion of free will. What might be loosely called ‘conscious processing’ should ultimately be the *perpetum cogito* as a discrete reentry process (way of being) following passively unconscious predictive processing (way of knowing) and exposed near criticality to a psyche level as more or less coherent decisions of Bayesian learning at that time. Their adaptive success depends on the RCS.

The NFVM and RCS both form a volitional-cognitive complex, anatomically extended over the whole brain. While the RCS occupies mainly the thalamocortical regions, the NFVM is a key underlying mechanism placed in the brainstem to be responsible for bottom-up initiation of conscious states, each then processed by the RCS *within* a given SRR during Δt . The states have also to be modulated in sensorimotor systems to provide cognitive function with coherence maximization *between* SRRs in the ever-changing environment. Thus, to be cognitively connected under the way of being that makes a difference, the brain should have the volitional-cognitive complex fine-tuned and exploited entirely.

According to the inverted perspective, adopted by CET, consciousness is a passive snapshot ignited at a psyche level having neither causal nor computation power over neural activity at both hard and soft levels. There is no neural correlates of consciousness that might be responsible for its active role, $NCC^{\text{active}} = 0$. Momentary conscious states emerge phenomenally at critical points of brain dynamics as ultimate decisions of Bayesian learning. Their neural correlates are just the neural correlates (NC) of the complex (Fig. 3e). Heuristically,

$$NCC = NC(NFVM + RCS) \quad (10)$$

More exactly, the neural correlates of a particular conscious state $S(\tau)$, presented by the variable $x_i = (n_1, \dots, n_N)$, depend not only on a set of neurons recruited by the complex at that moment but also on how well that configuration of diverse structural–functional

networks can maintain self-organized criticality to provide large-scale brain dynamics with the mental force. Its magnitude, traditionally referred to as the level or ‘quantity’ of consciousness in a given state, varies across different states, including clinical ones. Now there are a number different quantitative measures proposed to estimate the level of consciousness in different states. We will consider most promising of them in the next section.

Complexity and transfer entropy in stream of consciousness

According to Eq. (3), conscious states emerge only near criticality where the brain is poised between order and disorder (Chialvo 2010). This provides an optimal state for dynamical variability and information storage, and has been suggested as a determinant for information-based measures of consciousness (Mediano et al. 2016; Tagliacuzzi 2017; Kim and Lee 2019). Indeed, these both are statistically relevant as describing neural activity at the same physical level (Werner 2009; Deco et al. 2015; Aguilera 2019). While the critical dynamics are characterized by the order parameter, for example, a mean proportion of activated neurons in \mathcal{N} , with the control parameter, depending on connectivity density over time (Hesse and Gross 2014), the information-based measures evaluate the degree of integration (order) of \mathcal{N} in a particular state. In CET, objective observables of consciousness at a moment τ will be complexity measures.

Here we consider only two measures that are most relevant to neural activity. In physics, a statistical complexity C_{LMC} was proposed to reflect a thermodynamic depth of physical systems with N accessible states arranged from an ideal gas in equilibrium to a perfect crystal, maximally ordered (López-Ruiz et al. 1995). This is a product of Shannon entropy H as the disorder measure, and the opposite measure D called “disequilibrium” as a distance between $H_{max} = \log N$ and H .

$$H = - \sum_{i=1}^N p(i) \log p(i) \quad (11)$$

$$D = \sum_{i=1}^N (p(i) - 1/N)^2 \quad (12)$$

In an ideal gas, $H = H_{max}$, and $D = 0$. Conversely, for a crystal, $H = 0$, and $D = 1$. Thus, the product $C_{LMC} = H \cdot D$ well captures the balance between order and disorder and becomes zero for both purely chaotic and purely crystalized systems. Nevertheless, it does not allow for complex non-ordinary systems that are themselves information-processing structures.

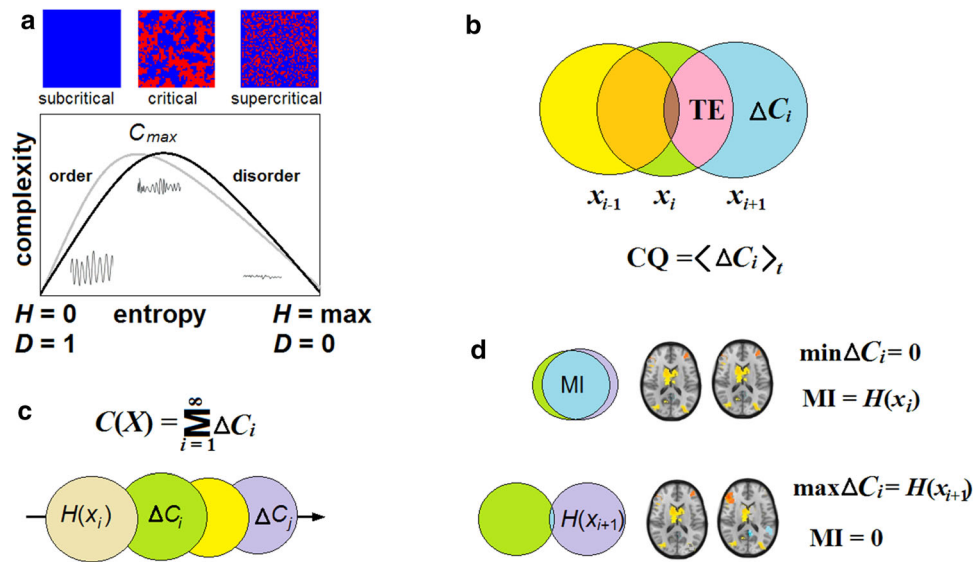


Fig. 4 Complexity measures and Cognitive evolution. **a** When extracted in time series of discretized measurements, C_{LMC} and $C_{\mathcal{N}}$ reflect a mixture of synchronization/desynchronization in brain dynamics with maximal values near criticality between subcritical and supercritical phases, presented here by the 2D Ising model (adapted from Tegmark 2015). **b** While TE depends on mutual

information, CQ can statistically reflect how much new information the brain has gained in time. **c** The $C(X)$ unfolds by formally summarizing increments ΔC over $(X, <)$. **d** If $CQ \rightarrow 0$, it makes brain dynamics functionally ‘crystallized’ in a subcritical regime. Conversely, $CQ \rightarrow max$ makes brain dynamics chaotic in a supercritical regime, thereby causing minimal coherency of $S(\tau)$

The neural complexity $C_{\mathcal{N}}$ is another measure (Tononi et al. 1994) focusing on structural–functional connectivity of the brain network \mathcal{N} . This is mathematically equivalent to the average information exchanged between subsets of a system and the rest of the system, summed over all subset sizes. The $C_{\mathcal{N}}$ can be calculated by mutual information (MI) obtained for all possible bipartitions of a system \mathcal{N} consisting of N elements,

$$C_{\mathcal{N}}(\mathcal{N}) = \sum_{k=1}^{N/2} \langle MI(\mathcal{N}_j^k; \mathcal{N} - \mathcal{N}_j^k) \rangle, \tag{13}$$

where \mathcal{N}_j^k is the j ’th bipartition running over all subsets of size k , and $\langle \cdot \rangle$ stands for their average integration. MI is defined as

$$MI(\mathcal{N}_j^k; \mathcal{N} - \mathcal{N}_j^k) = H(\mathcal{N}_j^k) - H(\mathcal{N}_j^k | \mathcal{N} - \mathcal{N}_j^k) \tag{14}$$

The $C_{\mathcal{N}}$ behaves like C_{LMC} : it is highest when segregation and integration are balanced in \mathcal{N} , and lowest under either total integration (order) or total segregation (disorder) of its elements (Fig. 4a). In CET, $C_{\mathcal{N}}$ should provide a measure for information that was integrated by the brain during a time interval Δt . This displays how well the brain is poised near criticality to gain the maximum information. On this condition, $C_{\mathcal{N}}$ refers to Bayesian active inference, inevitably coupled with self-awareness in the concomitant perpetuum cogito (Friston 2018). Thus, conscious states can emerge with different values of $C_{\mathcal{N}}$ reflecting the magnitude of the brain’s mental force at a given time.

An impressive review of complexity measures as reliable indices of the presence/absence of consciousness across many different conditions, such as sleep, anesthesia, meditation, drug-induced and hallucinatory states, epilepsy, and related disorders of consciousness, ranged clinically from coma and unresponsive wakefulness syndrome (UWS) to minimally conscious states (MCS) and locked-in-syndrome (LIS), is presented by Sarasso et al. (2021). Recall, the “state of consciousness” is defined there in a general sense as a state of wakefulness or vigilance averaged over time. Accordingly, complexity measures applied to the stream of many particular states over a slice $T = \sum \Delta t$ are also averaged over time. So, another line of experimental evidence, based on spatiotemporal scale-free signatures, cognate to complexity measures, can propose a more insightful picture of temporal variability in brain dynamics (Liu et al. 2014; Zhang et al. 2018; Huang et al. 2016) together with atypical intrinsic timescales (Watanabe et al. 2019; Golesorkhi et al. 2021) in altered states of consciousness.

On the other hand, it is shown that critical dynamics can also characterize human cognitive abilities and intelligence (Ezaki et al. 2020; Wang et al. 2021; Xu et al. 2022). It is therefore natural to ask how complexity measures of consciousness biased within experimental slices T can be applied to measuring cognition without losing fine-grained variability. In particular, can one show that a completely random or a completely periodic sequence of particular conscious states is not complex, while a sequence that

contains many different kinds of regularities does it? It seems obvious that such a kind of fine-grained variability must depend on cognitive processes implemented by the RCS. Thus, obtaining the cognitive variability measures could help in understanding mental disorders, which do not usually affect the general state of consciousness (wakefulness) but depend on how particular states of consciousness vary in their cognitive contents over the stream $S(t)$. In other words, mental disorders are a matter of unconscious predictive processing at a soft level not a matter of conscious states, which appear as ultimate decisions of Bayesian learning exposed to a psyche level. Neurologically, passive consciousness cannot be guilty there but its representative nature does it (yet fueled by the illusion of free will).

How might the entropy-based measures be usefully turned to studying mental (cognitive in origin) disorders which are then symptomatically detected in the stream of consciousness? Consider \mathcal{N} in brain dynamics over a particular segment $\{x_i, \dots, x_k\}$ of states, processed stochastically during a temporal slice $T = \sum \Delta t$ with corresponding probability distributions $p(x_i)$ over \mathcal{N} . The time-delayed mutual information MI_t between two nearest states is

$$MI_t(x_i; x_{i-1}) = H(x_i) - H(x_i|x_{i-1}) \quad (15)$$

MI_t is symmetric and upper-bounded by the entropy of both states but with no dynamical or directional information. It shows how good the brain is at predicting its own future state or, equivalently, how much information is inherited by the brain from its own past state. We expect all the states to be more or less connected in the sense that each future state x_{i+1} must somehow depend on the present state x_i , given the past state x_{i-1} , within Bayesian learning (Fig. 3b).

By applying MI_t to brain dynamics, we return to the notion of ‘information gain’ in Bayesian learning. The latter is typically defined by the Kullback–Leibler divergence between the prior and the posterior, both computed by the brain from its own perspective. However, we have no access to that information secluded in its generative model. For example, there is an obvious visual difference between seeing a cat and seeing a car, or even between seeing a cat and seeing a subject’s own cat (emotional difference). But how could we measure or, at least, identify the difference for testable predictions, if we are ignorant about the scenes and beyond a subjective report?

We need to learn how the brain itself gains information within its own way of knowing. In other words, the aim is to obtain more or less objective information about $(X, <)$, compared to exclusively subjective information computed by the brain from its own perspective. To do it, consider transfer entropy (TE), another statistical measure based on

MI_t and designed to detect the directed exchange of information between two variables, conditioned on common history and inputs (Schreiber 2000). TE holds a principled feature of neural complexity: it vanishes for both purely chaotic and purely ‘crystallized’ systems. Ideally, TE should be quantified over triplets $\{x_{i-1}, x_i, x_{i+1}\}$.³

$$TE_{x_i \rightarrow x_{i+1}|x_{i-1}} = MI_t(x_{i+1}; x_i|x_{i-1}) \quad (16)$$

Unlike neural complexity, TE is asymmetric with respect to causal/temporal order in brain dynamics. Typically extracted from neuroimaging datasets to assess brain functional connectivity (Ito et al. 2011) and being equivalent to Granger statistical causality (Barnett et al. 2009), TE is often considered a candidate measure of consciousness (Seth et al. 2011; Mediano et al. 2019). Intuitively, however, TE can likely be a cognition-driven measure of the coherence of Bayesian learning which, indeed, results representatively in discrete conscious states. In effect, TE estimates the strength of causal rigidity of the structural–functional connectivity of \mathcal{N} over time at a hard level. Accordingly, brain processes with the maximum TE should generate identical dynamical patterns as if no new knowledge was acquired by the brain between those processes at a soft level. When imposed upon $(X, <)$, the patterns in turn should ignite the same conscious states that make no difference from the past at a psyche level. Thus, the difference between two successive conscious states in brain dynamics can be statistically defined by TE,

$$\Delta C_{i+1} = H(x_{i+1}) - TE \quad (17)$$

The increment ΔC_i reflects just the information gain in Bayesian learning, estimated now from the third-person perspective (Fig. 4b).

Now we make the general assumption on which CET is based. Intuitively, the brain’s way of knowing is the integration of knowledge freshly acquired and stored in memory networks. This must increase information capacities, imprinted on the finest structure of the brain that should somehow enrich and optimize the structural–functional connectivity of \mathcal{N} through rewiring neural networks. The goal of the assumption is to emphasize that (i) the brain evolves by accumulating new knowledge (difference), and (ii) consciousness is not a fundamental property of matter but rather a dynamical characteristic of the brain’s cognitive evolution.

Here “cognitive evolution” holds the dual hard-soft aspect of neural activity besides its more fundamental

³ Here the three successive states would be needed to capture *effective* causation by avoiding redundant statistical inferences; however, in brain dynamics it does not much matter. Yet, in neuroscientific practice, such a time resolution would be hard to perform. Instead, multivariate TE over slices $T = \sum \Delta t$ could be used there (Novelli et al. 2019; Ursino et al. 2020).

biological meaning (as discussed in “Introduction”). First, it is the cognitive evolution of the brain’s generative model by updating and memorizing information at a soft level. Second, it is cumulative causal dynamics that advances the brain’s computational power due to the connectome development via neurogenesis, cell migration, synaptogenesis, and Hebbian plasticity at a hard level (Kaiser 2017; Yuan et al. 2019). Both these contribute to an organism’s adaptive success over its lifespan. Ultimately, on the evolutionary timescale, they converge to the biological evolution of the brain over species. Recall, we have no access to how the cognitive evolution, secluded in the brain generative model, goes at a soft level. What neuroimaging datasets can tell us is how this is causally processed at a hard level then compared with a subjective report at a psyche level. CET does not specify how the *cumulative* cognitive evolution can be formalized in terms of critical dynamics which do not mathematically provide those cumulative features over time.⁴

Nevertheless, we can still obtain neuroimaging datasets and compute relevant statistical measures about how the evolution goes at the hard level. Its cumulative features are obviously manifested by the fact that the consciousness a subject had in youth is not the same in adulthood due to knowledge the brain had obtained over lifetime. In a timeless description over $(X, <)$, the difference depends on an increments ΔC_i summarized in working memory and updated over the stream $S(\tau)$ (Fig. 4c),

$$C(X) = \sum_{i=1}^{\infty} \Delta C_i \quad (18)$$

Now if some cognition quantity (CQ) can be adapted to a particular segment in the stream $S(\tau)$, the measure will reveal the dynamics of C_N in everyday activities. The aim is to estimate how the brain evolves causally while making its own subjective estimates to minimize prediction-error. In practice, the cognition quantity can then be experimentally defined by the increments averaged over time,

$$CQ = \langle \Delta C_i \rangle_t \quad (19)$$

Within the hard-soft duality of brain dynamics, CQ suggests an objective (causal) measure of subjective

information, which the brain has to store and exchange constantly in its generative model over all conscious states. For example, CQ should be lower in mind-wandering resulting from the computational ruminations running on autopilot (Christoff 2012; Maillet and Schacter 2016), but higher in cognitive performance accompanied with switches in functional connectivity of \mathcal{N} (Cabral et al. 2017). Thus, if $CQ = 0$, no knowledge has been acquired by the brain there.

Cognition quantity in disorders of consciousness and mental disorders

The aim of this section is to show how CET can contribute to our understanding of different brain disorders in terms of cognitive neurodynamics. To translate brain dynamics into the language of algorithmic information theory consider a segment $\{x_i, \dots, x_k\}$ of a chain $(X, <)$ like an individual string of letters. Let each letter stand for a particular x_i encoded by an N -sequence of 1 and 0 for active and inactive neurons at a given time. Clearly, the same letters will conform to the same conscious states with identical NCC. The segment is sufficient for the brain to have conscious experience that makes a difference for itself. Transitions between conscious states at critical points need the interval Δt for unconscious decision-making by the RCS. While C_N measures the level of consciousness at a given state, CQ can algorithmically quantify the dynamics of Bayesian leaning over that string of letters.

Intuitively, CQ has to vary in everyday activities, firstly, between sleep and waking states. Indeed, when quantified by TE between cortical and hippocampal neurons, CQ should be lower during NREM sleep than during wakefulness with a rapid shift in arousal (Olcese et al. 2018). One might then naively interpret CQ as “the more the better” like IQ. Nevertheless, it does not allow for the intrinsic dynamics of cognition depending neurophysiologically on informational coherency of $C(X)$ and resulting in the ability for consistent learning and logical reasoning. In entropy terms, $CQ = 0$ assumes (not necessarily) that many states are identical in $C(X)$ as if no useful work had been done by the brain over time. Conversely, the maximum CQ, conditioned on $TE = 0$, makes those states totally disconnected with empty MI (Fig. 4d). Thus, both low and high values of CQ have to be destructive in cognitive processing. They can be effectively interpreted only in respect to some ‘normal’ bandwidth $\delta = [\Delta C_{min}, \Delta C_{max}]$ extracted experimentally from dynamics of healthy controls.

CQ suggests a simple, albeit rough, marker that, nonetheless, may contribute to diagnostic tools in

⁴ In fact, the tools of mathematical neuroscience are borrowed from physics of complex dynamical systems which do not usually accumulate information. That is, a system that explores its state space does not evolve in a meaningful sense: it is the same system in time. In contrast, the brain learns, memorizes, and adapts to the environment. Moreover, a learning brain changes the neural landscape of its state space due to the connectome development. Neurologically, the astonishing ability of humans to actively change the world around them (via active inference in the FEP formalism (Friston 2010)) is a consequence of this cognitive evolution. Thus, cumulative brain dynamics require more sophisticated mathematical descriptions than those of complex dynamical systems.

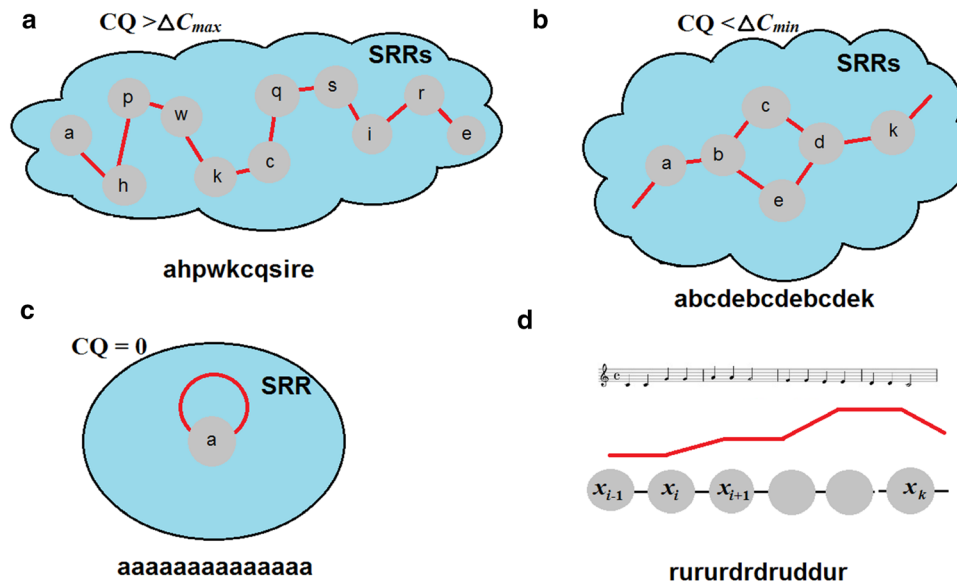


Fig. 5 Disorders in algorithmic coding. The stream of consciousness can be viewed as the brain's way of making a difference. The temporal coherency of the stream depends on how the difference is processed in Bayesian learning. If it is big, the brain captures too much information to learn something consistently. If it is small, learning fails. **a** Attention deficit disorder. Here $(X, <)$ is schematically depicted as a totally disconnected sequence of random letters, each standing for a particular state. While being all NFVM-initiated, the states are badly constrained by the RCS. The stream evolves with CQ extending the upper boundary of optimal cognitive processing. **b** Obsessive–compulsive disorder. The RCS is rigid by generating monotonically periodic sequences of states with cyclical loops in

obsessive–compulsive periods where Bayesian updating fails. CQ is thus reduced below the lower boundary of optimal cognitive processing. **c** Unresponsive wakefulness syndrome. The RCS is disrupted while the stream is locked in a single SRR with no CQ. Note, a state, repeated many times like a freeze-frame, does not imply causally closed loops. **(d)** Music therapy. In Parsons (2008) code of melodic contours, a notation identifies movements of the pitches on each pair of consecutive notes as “u” (up) if the second note is higher than the first one, “d” (down) otherwise, and “r” (repeat) if the pitches are equal. While rhythm is completely omitted, a well-connected sequence of tones arises

classifications of mental disorders. Intuitively, the values $CQ > \delta$ should be predictable in hyperactive children with attention deficit disorder (ADD). Its behavioral symptoms such as impulsiveness, trouble with focusing on task, and low attention span are typically associated in numerous studies with weaker coherence of brain dynamics (for review see Cortese et al. 2012; Castellanos and Aoki 2016). Thus, ADD can be classified as the *cognition quantity excess*, resulting from the impaired RCS and accompanied with hyperactive, i.e., random and ill-adaptive behavior. This can be schematically presented with a string of random letters, each standing for a certain state (vector) $x_i = [n_1, \dots, n_N]$ (Fig. 5a). There could be found an interesting resemblance between mind-wandering and ADD. While the former occurs as a free-retrieval process well maintained by the NFVM + RCS complex within the bandwidth δ , the latter can be characterized as decoherent mind-wandering in desynchronized brain dynamics near a supercritical regime.

On the contrary, $CQ < \delta$ would lead the brain to structural–functional rigidity of \mathcal{N} in a subcritical regime. For example, obsessive–compulsive disorder (OCD) is commonly associated in psychology with cognitive impairments such as intrusive thoughts and ritualistic behaviors

(Oudheusden et al. 2018). In neuroscience, OCD is typically associated with overstability in brain dynamics at a hard level (Rolls et al. 2008) and concomitant deficits in Bayesian updating at a soft level (Seow and Gillan 2020). While having the NFVM intact, the brain could keep the RCS heavily cyclical in OCD by generating the stream $S(\tau)$ like a string of repeated letters encoding these states (Fig. 5b). Overall, CQ should be low in OCD, reflecting the functional inflexibility of \mathcal{N} , with a loss of cognitive abilities that would be subjectively experienced via disturbed self-monitoring in the *perpetum cogito* process.

One can then ask what the consciousness would be like if cognitive processes were fully disrupted in the brain. According to Eq. (18), if cognitive evolution $C(X)$ stops at some moment of time, then $CQ = 0$, and a corresponding state x_i at that moment will be preserved like a freeze-frame on a screen for all future states within a segment $\{x_i, \dots, x_k\}$,

$$(\forall k \geq i)x_k = x_i \quad (20)$$

This can explain what can occur to patients with UWS. Instead of losing consciousness, typically manifested in coma, a certain state can persist in those patients, however, with no cognitive evolution. This is just a snapshot

triggered by the NFVM and locked in the same stimulus-reaction repertoire when the RCS is severely disrupted (Fig. 5c). Because of the circularity, both the perpetum cogito and memory contents could not be updated as if time perception was suspended too. Indeed, after recovery patients with UWS have usually vague or no recollections of their time in rehabilitation (Gosseries et al. 2014).

A growing body of evidence shows that the spectrum of all psychological (normal and abnormal) states observed in humans depends neurophysiologically on balanced synchronization/desynchronization patterns in brain dynamics, varying from the subcritical to supercritical regime (for review see Zimmern 2020; Heiney et al. 2021). It is known that epilepsy, schizophrenia, dementia and Parkinson's disease come with pathological synchronization phenomena in brain dynamics (Uhlhaas and Singer 2006; Broyd et al. 2009; Yu et al. 2013; Cabral et al. 2013; Northoff and Gomez-Pilar 2021), often accompanied with synaptic disruption, whereas consciousness in normal general state exhibits more rich dynamical patterns of functional connectivity of \mathcal{N} (Bartfeld et al. 2015; Cavanna et al. 2018; Golkowski et al. 2019; Demertzi et al. 2019). It is also shown that the stream $S(\tau)$ relies on temporal circuitry between default mode network and dorsal attention network, which alternate their activity in an anticorrelated manner (Huang et al. 2020). Thus, functional disturbances in brain dynamics are crucial to causing different brain disorders, which are characterized in medical coding systems such as ICD or DSM mainly with the help of diagnostic tools applied to phenomenally grouped symptoms (Allsopp et al. 2019).

CET proposes a systematic approach to brain disorders (of any etiology) that focuses on cognitive impairments which are present in all mental disorders (Ganguli et al. 2011), and on the fact that all diagnostic tools for assessing those rely on the representative nature of consciousness. Consciousness is like a river buoy, fluctuating and drifting always on the surface of water regardless of its depth. The behavior of such a float can be very indicative of the underwater landscape and invisible flows, for example, in fishing or navigation. Accordingly, in clinical approaches, the stream $S(\tau)$ of consciousness is taken to be a valid indicator of symptoms to which diagnostic tools apply.

CET adopts the inverted perspective, according to which consciousness is not guilty there. By Eq. (10), consciousness is a passive phenomenon which manifestations depend entirely on how well the NFVM + RCS complex is able to maintain self-organized criticality to provide the NCC of a particular conscious state with a corresponding mental force. Thus, damage to any part of the complex directly affects the stream. If the RCS is impaired, the brain's cognitive evolution $C(X)$ becomes suppressed with a consequent reduction in both the level and conscious

contents measured by $C_{\mathcal{N}}$ and CQ respectively. For example, lesions of particular thalamocortical networks in the RCS can selectively disrupt conscious experience, causing numerous deficits such as blindsight, agnosia, or akinetic mutism. If the RCS does not function at all, while the NFVM is still at work, it can still provide the minimal correlates of consciousness (Owen et al. 2006) with $CQ = 0$ (Fig. 5c).

$$NCC^{\min} = NC(\text{NFVM} + \emptyset) \stackrel{\text{def}}{=} \text{UWS} \quad (21)$$

Typically, the brainstem is relatively spared in UWS whereas both cerebral hemispheres are widely and severely damaged. Recovery of consciousness depends then upon the functional reemergence of the ARAS, which must provide sufficient input via the thalamic projections to the anterior forebrain mesocircuit and frontoparietal network (Schiff et al. 2014; Giacino et al. 2014). Indeed, it is known that full recovery from UWS can be accompanied by restoration of activity solely in frontoparietal areas (Lau-reys 2005).

On the contrary, brainstem lesions cause immediate coma by damaging the ARAS and its associated neuro-modulatory systems (Parvizi and Damasio 2001). Thus, if the NFVM was severely injured even if the RCS remained entirely or partially intact, no conscious state could be initiated in $C(X)$.

$$NCC^{\text{null}} = NC(\emptyset + \text{RCS}) \stackrel{\text{def}}{=} \text{coma} \quad (22)$$

The inverted perspective implies that mental disorders (MDs) and disorders of consciousness (DoC) are to be intrinsically coupled in aberrant neural activities over long-term brain dynamics (Breakspear 2017). The divide between DoC and MDs stems from the common agreement to distinguish levels of consciousness, attested by arousal criteria, from conscious contents, related to cognitive function. Accordingly, ICD classifies DoC as pathologies per se, not related to MDs. On the other hand, DSM-5 does not use the term 'consciousness' at all, and operationalizes it as 'changes in attention' that should be just related to conscious contents. In clinical practice, attention and arousal are explicitly linked: the level of arousal must be sufficient before attention can be reasonably tested (Fig. 6a). However, as the conscious contents depend entirely on cognitive function, DSM-5 already implicitly maintains the inverted perspective: conscious states emerge after they are cognitively (unconsciously) processed. The level of consciousness (arousal) should thus be neuro-physiologically affected at a psyche level if cognitive function (attention) was depressed at a soft level. These both in turn depend on brain dynamics which show the reduced heterogeneity at a hard level (López-González et al. 2021).

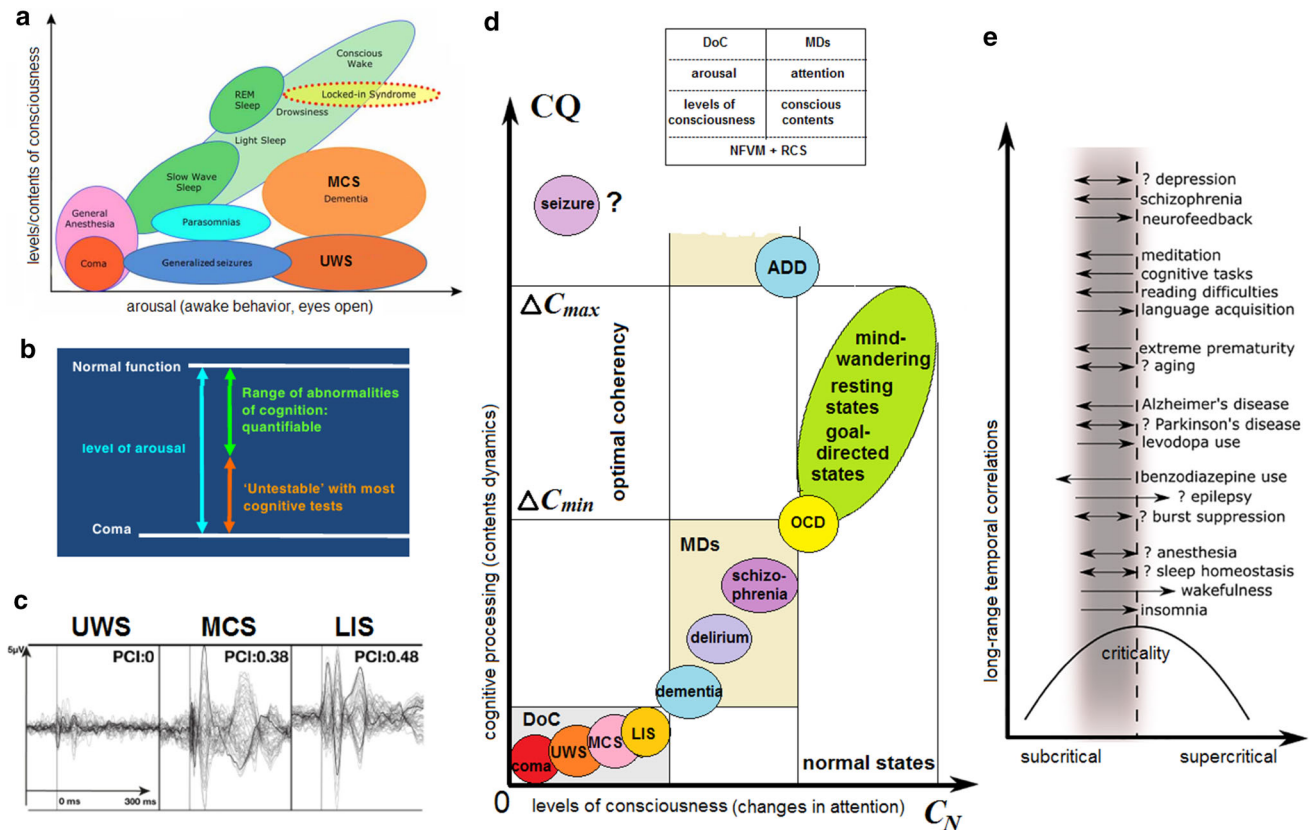


Fig. 6 A conceptual diagram for a quantitative classification of DoC and MDs. **a** Conscious states are traditionally compared to levels of arousal ranged from coma to full alertness over UWS, MCS, and LIS. Typically, high conscious levels are associated with an increased range of conscious contents (adapted from Boly et al. 2013). **b** DSM-5 states that changes in cognition must not occur in states of severely reduced level of consciousness such as coma. Taking into account that there is a continuum of levels of consciousness (arousal), it is more accurate to recognize that it is not possible to determine a threshold for cognitive processing between coma and normal states (European Delirium Association 2014). It makes also impossible to separate DoC from MDs. **c** DoC are estimated by PCI, based on the analysis of EEG-responses to transcranial magnetic stimulation to distinguish altered states of consciousness (adapted from Bodart et al. 2017). **d**

CET argues that DoC and MDs both result from impairments (of any etiology) in the NFVM + RCS complex. This predicts that MDs could also affect the level of consciousness to some extent, and DoC should arise when the RCS was completely or partially impaired while preserving the NFVM intact (Eq. (21)). Thus, the divide between DoC and MDs relies on the mere fact that cognitive function (conscious contents) under MDs criteria cannot be tested at all in non-communicative patients with DoC (Fig. 6b). From the inverted perspective, these both would continuously converge to coma (Fig. 6d) where the whole complex or, at least, the NFVM, was disrupted after severe brain damage (Eq. (22)).

Thus, a classification quantitatively expressed in values of CQ may shed light on the nature of both DoC and MDs

The coherency of cognitive processing can be quantified by CQ over the stream of conscious states. Coma is placed at the bottom of the diagram with both C_N and CQ tending to zero. Yet, ascribing higher values of CQ to epileptic seizures, accompanied with loss of consciousness, is controversial as if the brain might gain too much information there. The controversy arises due to chaotic brain dynamics in a supercritical regime with which seizures are associated (Meisel et al. 2012; Jirsa et al. 2014). **e** Tendencies in long-range temporal correlations between synchronization (integration) and chaos (segregation) balanced near criticality are recorded in different psychological states (adapted from Zimmern 2020). These findings, although not related directly to complexity measures, are indicative of CQ

by algorithmically estimating the cognitive processes on which consciousness depends. Today many studies explore this approach—though separately—by comparing different complexity measures and related scale-freeness signatures across MDs (for review see Fernandez et al. 2013; Hager et al. 2017; Douw et al. 2019; Zimmern 2020; Rolls et al. 2021) and DoC (for review see Chennu et al. 2014; Pal et al. 2020; Sarasso et al. 2021). To illustrate how CQ can estimate the coherency of cognitive processing over $C(X)$, consider music therapy.

Music therapy is commonly acknowledged in treatment of patients with MDs such as ADD, autism, schizophrenia, Alzheimer's disease (Jackson 2003; Talwar et al. 2006; Trimble and Hesdorffer 2017). A general explanation of such effects comes from the fact that music melodies with

their easily discernible frequency patterns can be consistently laid upon the stream (Sanyal et al. 2019). Music therapy can almost hypnotically constrain the brain to follow a melody's patterns and, thus, actually benefit from cognitive training of the RCS a subject would spontaneously be involved in by normalizing CQ to the bandwidth δ (Fig. 5d). In other words, a melody, perceived by patients as a temporal sequence of acoustic scenes with multiple regularities, can motivate a corresponding segment $\{x_i, \dots, x_k\}$ of their stream to hold more optimized values of CQ in $C(X)$, thereby advancing its causal/cognitive coherency in time.

How could these various kinds of regularities be statistically estimated over $C(X)$? Recently, much attention was drawn to the so-called “perturbational complexity index” (PCI) based on the Lempel–Ziv complexity (LZC), a modified version of Kolmogorov algorithmic complexity, which is equal to the length of the shortest computer program that can reproduce an n -bit string. The program is then considered to be a compressed description of the string by eliminating redundant information about any regularity there. LZC can estimate the bandwidth of non-synchronized processes and the harmonic variability in quasi-periodic signals. This also allows to examine the rate of new patterns along brain dynamics with the EEG time-series data (Aboy et al. 2006). A design was to use transcranial magnetic stimulation to obtain data for LZC normalized by PCI (Casali et al. 2013). The results were well documented on patients with different DoC such as UWS, MCS, LIS, and epileptic seizures (Bodart et al. 2017; Mateos et al. 2018). PCI also reliably discriminates the level of consciousness during wakefulness, sleep, and different rates of anesthesia (Sarasso et al. 2015; Hudetz et al. 2016).

The techniques and computational tools of PCI are relevant to measuring CQ because the level of consciousness quantified by PCI depends entirely on the coherency of cognitive processes, which are based on the underlying brain dynamics, observable over much longer time spans than PCI. For example, it has shown PCI = 0–0.2 (Bodart et al. 2018) for patients with UWS (Fig. 6c) in accordance with our prediction that $C(X)$ should stop in brains of those patients like a freeze-frame on a screen. In terms of algorithmic complexity, with $CQ \rightarrow 0$, the variability of functional patterns in brain dynamics would be reduced in UWS (López-González et al. 2021) so that the stream of consciousness could be ‘compressed’ into a single state by Eq. (20).

CET suggests that with an appropriate well-elaborated methodology, a unified quantitative classification of both DoC and MDs, based on complexity measures and critical dynamics, and coupled with insights from clinical network neuroscience (Douw et al. 2019) and genomics (Torres

2020), could improve our understanding of their neuropathology (Fig. 6e).

Discussion

There are now dozens of various theories of consciousness. They are fragmentary in explaining how the brain integrates consciousness, volition, and cognition seamlessly across three hierarchical levels—causal, computational, and phenomenal at a micro-, a meso- and a macroscale respectively. In general, most of the theories are silent about brain dynamics from which conscious states have to emerge. At the same time, all the theories ascribe a special, active role to consciousness while being either indifferent to the free will problem or superdeterministic, i.e., controlled by hidden deterministic variables λ despite Eq. (9). It is not surprisingly therefore that they suggest no principled obstacle to creating machine consciousness (Dehaene et al. 2017; VanRullen and Kanai 2021). In contrast to the classical theories, quantum-inspired models (Hameroff and Penrose 2014; Fisher 2015; Georgiev 2020) take the free will problem seriously by involving quantum entanglement across the whole brain to account for active consciousness that could not be (classically) machine-generated.

It is commonly acknowledged that it is unsatisfactory to have a plethora of very different theories, each suggesting its own meaning, function, and neural account of consciousness. Many authors try to converge the theories to a unified framework (Shea and Frith 2019; Hohwy and Seth 2020; Mashour et al. 2020; Chang et al. 2020; Safron 2020; Cofré et al. 2020). However, such a framework must operate exclusively on the stream of consciousness (Fingelkurts et al. 2010; Northoff and Huang 2017), not on “consciousness” with its philosophical baggage. Accordingly, static theories have also missed out another important psychological aspect of brain dynamics, namely, its cognitive evolution with aging. Hence, they are too unable to explain the origins of consciousness in biological terms.

CET starts from the claim that brains should have primarily evolved as volitional (quantum in origin) subsystems of organisms, not as deterministic prediction machines (Knill and Pouget 2004; Friston 2008; Clark 2013). Contrary to quantum-inspired models, CET requires only a minimal use of quantum indeterminism for rudimentary volitional mechanisms. These biophysical mechanisms should be initially involved in reflexive (rapid and random) reactions of simplest organisms, thereby laying a foundation for the psyche-matter divide between non-living systems, restricted to cause-effect interactions, and those organisms, exploiting their SRRs freely (Fig. 1). CET postulates the NFVM to be inherited from those

reflective mechanisms as a key mechanism for the emergence of consciousness from subcortical arousal systems. In this sense, the search for artificial consciousness based on deep machine learning can diverge crucially from the way Nature had chosen to evolve biological brains. More exactly, CET predicts:

If a system possesses the NFVM + RCS complex maintaining self-organized criticality, the system can be conscious

What follows is that whatever superior performance AI systems might reach by using statistical learning over a large amount of data, they should still have no mental force. On this (perhaps, counterintuitive) condition, the machines that cannot be Bell-certified have to be treated as unconscious even without resorting to a Turing test.

However, it does not rescue conscious will in biological brains at all. Are humans quantum computers, or merely clever robots (Fisher 2017), in any case consciousness cannot have causal power over its own physical substrate, unless Cartesian dualism applies. Only unconscious brain dynamics can be free of predetermination by the NFVM. It refers to most reliable mechanisms like the Beck–Eccles (1992, 1998) quantum trigger of exocytosis in a synaptic cleft. The NFVM only initiates neuronal firing, amplified by spontaneous scale-free avalanches inherent to critical dynamics (Beggs and Plenz 2003; Shew et al. 2011). All brain processes relevant to consciousness remain apparently classical, causally unclosed, and time-irreversible.

Unlike the above theories, CET emphasizes the dynamical nature of consciousness (psyche), then coupling it with volition (hard) and cognition (soft) seamlessly over micro-, meso-, and macroscales in brain dynamics. In this unified framework, conscious states merely enter the global workspace as representations of working memory after cognitive processing (Aly and Yonelinas 2012; Shea and Frith 2019). According to Eq. (3), conscious experience (way of being) is initially NFVM-driven, and can be derived—in any meaningful sense—only from the brain's cognitive evolution, going in unconscious ways over Bayesian learning (way of knowing). The stream $S(\tau)$ cannot go on at all, if the cognitive evolution $C(X)$ stops and working memory has nothing to update as it occurs in patients with UWS. CET also proposes a quantitative classification of both DoC and MDs with the unified framework, based on empirical measures of neural complexity and cognition quantity.

While the emergence of consciousness from brain dynamics is commonly accepted, many authors still prescribe a special biological function to consciousness. In CET, consciousness is representative which—like a river buoy on the surface of water—characterizes underlying brain dynamics, however, with no influence on it. It also

explains why consciousness can lightly be manipulated by physiologically or mentally affecting different brain systems, for example, through anesthetic and psychedelic drugs (at a hard level) or in music listening and hypnosis (at a soft level). The illusion of conscious control emerges due to the *perpetum cogito* process that follows the recurrent unconscious cognitive processing implemented by the RCS.

Conclusion

CET is a physicalist theory. However, it adopts a quantum kind of physicalism not a classical (superdeterministic) one. Together with two other prerequisites, dynamism and contextuality, CET entails five consequences about the nature of consciousness: discreteness, passivity, uniqueness, integrity (unitarity), and graduation. Consciousness unfolds as a chain of separate states, each ignited at critical points of continuous brain dynamics: *it is discrete*. As being phenomenal, consciousness has no causal power over its physical substrate: *it is passive*. The stream of states, each triggered by the NFVM, cannot be copyable by a perfect (omniscient) predictor: *it is unique*. Conscious states result from cognitive (unconscious) processing as ultimate decisions the brain has just computed at that moment of time: *it is unitary*. The level of consciousness depends on the NFVM + RCS complex, and fluctuates continuously in brain dynamics: *it is gradual*.

In CET, the only biological function assigned to consciousness is self-awareness, i.e., the ability not only to be but also to have *the sense of being*. Consciousness is merely necessary to feel the biological value of life. Evolution could not have succeeded on the earth if organisms did not appreciate the sense of being. But what is life for a cell, for an AI system, or even for a human in coma? Leaving aside the biological definition of life, consciousness is genuine life, which sense and value vanish together with loss of consciousness. Organisms are not machines programmed to survive; they struggle for existence to enjoy consciousness, 'something it is like to be'.

Funding The authors did not receive support from any organization for the submitted work.

Declarations

Conflict of interest The authors declare they have no financial interests.

References

- Aaronson S (2016) The ghost in the quantum turing machine. In: Cooper SB, Hodges A (eds) *The once and future turing: computing the world*. Cambridge University Press, Cambridge, pp 193–294
- Aboy M, Abasolo D, Hornero R, Álvarez D (2006) Interpretation of the Lempel-Ziv complexity measure in the context of biomedical signal analysis. *IEEE Trans Biomed Eng* 53(11):2281–2288
- Aguilera M (2019) Scaling behaviour and critical phase transitions in integrated information theory. *Entropy* 21:1198
- Allsopp K, Read J, Corcoran R, Kinderman P (2019) Heterogeneity in psychiatric diagnostic classification. *Psychiatry Res* 279:15–22
- Aly M, Yonelinas AP (2012) Bridging consciousness and cognition in memory and perception: evidence for both state and strength processes. *PLoS ONE* 7(1):e30231
- Aru J, Suzuki M, Rutiku R, Larkum ME, Bachmann T (2019) Coupling the state and contents of consciousness. *Front Syst Neurosci* 13:43
- Atmanspacher A, Rotter S (2008) Interpreting neurodynamics: concepts and facts. *Cogn Neurodyn* 2:297–318
- Baars BJ (2003) The conscious access hypothesis: origins and recent evidence. *Trends Cogn Sci* 6(1):47–51
- Baars BJ, Franklin S, Ramsoy TZ (2013) Global workspace dynamics: cortical “binding and propagation” enables conscious contents. *Front Psychol* 4:200
- Bachmann T, Hudetz AG (2014) It is time to combine the two main traditions in the research on the neural correlates of consciousness: $C = L \times D$. *Front Psychol* 5:940
- Bak P, Tang C, Wiesenfeld K (1987) Self-organized criticality: an explanation of the $1/f$ noise. *Phys Rev Lett* 59:381–384
- Baluška F, Miller WB, Reber AS (2021) Biomolecular basis of cellular consciousness via subcellular nano-brains. *Inter J Mol Sci* 22:2545
- Barnett L, Barrett AB, Seth AK (2009) Granger causality and transfer entropy are equivalent for Gaussian variables. *Phys Rev Lett* 103:238701
- Barttfeld P, Uhrig L, Sitt JD, Sigman M, Jarraya B, Dehaene S (2015) Signature of consciousness in the dynamics of resting-state brain activity. *PNAS USA* 112(3):887–892
- Beck F, Eccles J (1992) Quantum aspects of brain activity and the role of consciousness. *PNAS USA* 89:11357–11361
- Beck F, Eccles J (1998) Quantum processes in the brain: a scientific basis of consciousness. *Cogn Stud* 5(2):95–109
- Beggs JM, Plenz D (2003) Neuronal avalanches in neocortical circuits. *J Neurosci* 23(35):11167–11177
- Beggs JM, Timme N (2012) Being critical of criticality in the brain. *Front Physiol* 3:163
- Bell J (1993) *Speakable and unspeakable in quantum mechanics*. Cambridge University Press, Cambridge
- Blanchard P, Cessac B, Krueger T (2000) What can one learn about self-organized criticality from dynamical system theory? *J Stat Phys* 98:375–404
- Bodart O, Gosseries O, Wannez S, Thibaut A, Annen J, Boly M et al (2017) Measures of metabolism and complexity in the brain of patients with disorders of consciousness. *NeuroImage Clin* 14:354–362
- Bodart O, Amico E, Gomez F, Casali AG, Wannez S, Heine L et al (2018) Global structural integrity and effective connectivity in patients with disorders of consciousness. *Brain Stimul* 11:358–365
- Boly M, Seth AK, Wilke M, Ingmundson P, Baars B, Laureys S et al (2013) Consciousness in humans and non-human animals: recent advances and future directions. *Front Psychol* 4:625
- Boly M, Massimini M, Tsuchiya N, Postle BR, Koch C, Tononi G (2017) Are the neural correlates of consciousness in the front or in the back of the cerebral cortex? Clinical and neuroimaging evidence. *J Neurosci* 37:9603–9961
- Brascamp JW, van Ee R, Noest AJ, Jacobs RH, van den Berg AV (2006) The time course of binocular rivalry reveals a fundamental role of noise. *J vis* 6(11):1244–1256
- Breakspear M (2017) Dynamic models of large-scale brain activity. *Nat Neurosci* 20(3):340–352
- Brembs B (2011) Towards a scientific concept of free will as a biological trait: spontaneous actions and decision-making in invertebrates. *Proc R Soc B* 278:930–939
- Bressler SL, Kelso JAS (2001) Cortical coordination dynamics and cognition. *Trends Cogn Sci* 5:26–36
- Broyd SJ, Demanuele C, Debener S, Helps SK, James CJ, Sonuga-Barke EJ (2009) Default-mode brain dysfunction in mental disorders: a systematic review. *Neurosci Biobehav Rev* 33(3):279–296
- Cabral J, Fernandes HM, Van Hartevelt TJ, James AC, Kringelbach ML, Deco G (2013) Structural connectivity in schizophrenia and its impact on the dynamics of spontaneous functional networks. *Chaos* 23:046111
- Cabral J, Vidaurre D, Marques P, Magalhães R, Moreira PS, Soares JM et al (2017) Cognitive performance in healthy older adults relates to spontaneous switching between states of functional connectivity during rest. *Sci Rep* 7:5135
- Casali AG, Gosseries O, Rosanova M, Boly M, Sarasso S, Casali KR et al (2013) A theoretically based index of consciousness independent of sensory processing and behavior. *Sci Transl Med* 5:1–10
- Castellanos FX, Aoki Y (2016) Intrinsic functional connectivity in attention-deficit/hyperactivity disorder: a science in development. *Biol Psychiatry: Cogn Neurosci Neuroimaging* 1:253–261
- Cavanna F, Vilas MG, Palmucci M, Tagliazucchi E (2018) Dynamic functional connectivity and brain metastability during altered states of consciousness. *Neuroimage* 180:383–395
- Chang AYC, Biehl M, Yu Y, Kanai R (2020) Information closure theory of consciousness. *Front Psychol* 11:1504
- Cheng-Yu TL, Poo MM, Dan Y (2009) Burst spiking of a single cortical neuron modifies global brain state. *Science* 324:643–646
- Chennu S, Finoia P, Kamau E, Allanson J, Williams GB, Monti MM et al (2014) Spectral signatures of reorganised brain networks in disorders of consciousness. *PLoS Comput Biol* 10(10):e1003887
- Chenu A, Scholes GD (2015) Coherence in energy transfer and photosynthesis. *Ann Rev Phys Chem* 66:69–96
- Chialvo DR (2010) Emergent complex neural dynamics: the brain at the edge. *Nat Phys* 6:744–750
- Christoff K (2012) Undirected thought: neural determinants and correlates. *Brain Res* 1428:51–59
- Clark A (2013) Whatever next? Predictive brains situated agents and the future of cognitive science. *Behav Brain Sci* 36(3):181–204
- Cleeremans A (2011) The radical plasticity thesis: how the brain learns to be conscious. *Front Psychol* 2:86
- Cocchi L, Gollo LL, Zalesky A, Breakspear M (2017) Criticality in the brain: a synthesis of neurobiology models and cognition. *Prog Neurobiol* 158:132–152
- Cofré R, Herzog R, Mediano PAM, Piccinini J, Rosas FE, Sanz Perl Y et al (2020) Whole-brain models to explore altered states of consciousness from the bottom up. *Brain Sci* 10:626
- Conway JH, Kochen S (2008) The strong free will theorem. *Not Am Math Soc* 56:226–232
- Cortese S, Kelly C, Chabernaud C, Proal E, Di Martino A, Milham MP, Castellanos FX (2012) Toward systems neuroscience of ADHD: a meta-analysis of 55 fMRI studies. *Am J Psychiatry* 169:1038–1055

- Crick F, Koch C (2003) A framework for consciousness. *Nat Neurosci* 6(2):119–126
- Dahmen D, Grün S, Diesmann M, Helias M (2019) Second type of criticality in the brain uncovers rich multiple-neuron dynamics. *PNAS USA* 116(26):13051–13060
- de Graaf TA, Hsieh PJ, Sack AT (2012) The ‘correlates’ in neural correlates of consciousness. *Neurosci Biobehav Rev* 36:191–197
- Deco G, Jirsa V (2012) Ongoing cortical activity at rest: criticality multistability and ghost attractors. *J Neurosci* 32:3366–3375
- Deco G, Tononi G, Boly M, Kringelbach ML (2015) Rethinking segregation and integration: contributions of whole-brain modeling. *Nat Rev Neurosci* 16(7):430–439
- Deco G, Cruzat J, Kringelbach ML (2019) Brain songs framework used for discovering the relevant timescale of the human brain. *Nat Commun* 10:583
- Dehaene S, Changeux JP (2011) Experimental and theoretical approaches to conscious processing. *Neuron* 70:200–227
- Dehaene S, Naccache L (2001) Towards a cognitive permanence of consciousness: basic evidence and a workspace framework. *Cognition* 79:1–37
- Dehaene S, Lau H, Kouider S (2017) What is consciousness and could machines have it? *Science* 358:486–492
- Del Papa B, Priesemann V, Triesch J (2017) Criticality meets learning: criticality signatures in a self-organizing recurrent neural network. *PLoS ONE* 12(5):e0178683
- Del Pin SH, Skóra Z, Sandberg K, Overgaard M, Wierzchoń M (2021) Comparing theories of consciousness: why it matters and how to do it. *Neurosci Conscious* 7(2):1–8
- Demertzi A, Tagliazucchi E, Dehaene S, Deco G, Bartfeld P, Raimondo F et al (2019) Human consciousness is supported by dynamic complex patterns of brain signal coordination. *Sci Adv* 5(2):eaat7603
- Doerig A, Schurger A, Herzog MH (2020) Hard criteria for empirical theories of consciousness. *Cogn Neurosci* 12:41–61
- Douw L, van Dellen E, Gouw AA, Griffa A, de Haan W, van den Heuvel M et al (2019) The road ahead in clinical network neuroscience. *Netw Neurosci* 3(4):969–993
- Drissi-Daoudi L, Doerig A, Herzog MH (2019) Feature integration within discrete time windows. *Nat Commun* 10(1):4901
- Edelman GM (1989) *The remembered present: a biological theory of consciousness*. Basic Books, New York
- Edelman GM (2003) Naturalizing consciousness: a theoretical framework. *PNAS U.S.A.* 5520–5524
- Edelman GM, Gally JA, Baars BJ (2011) Biology of consciousness. *Front Psychol* 2:4
- Elamrani A, Yampolskiy RV (2019) Reviewing tests for machine consciousness. *J Conscious Stud* 26(5–6):35–64
- European Delirium Association (2014) The DSM-5 criteria level of arousal and delirium diagnosis: inclusiveness is safer. *BMC Med* 12:141
- Ezaki T, Fonseca dos Reis E, Watanabe T, Sakaki M, Masuda N (2020) Closer to critical resting-state neural dynamics in individuals with higher fluid intelligence. *Commun Biol* 3(1):52
- Fekete T, Van de Cruys S, Ekroll V, van Leeuwen C (2018) In the interest of saving time: a critique of discrete perception. *Neurosci Conscious* 4(1):niy003
- Fernandez A, Gomez C, Hornero R, Lopez-Ibor JJ (2013) Complexity and schizophrenia. *Prog Neuropsychopharmacol Biol Psychiatry* 45:267–276
- Fields C, Glazebrook JF, Levin M (2021) Minimal physicalism as a scale-free substrate for cognition and consciousness. *Neurosci Cons* 7(2):niab013
- Fingelkurts AA, Fingelkurts AA, Neves CF (2010) Natural world physical, brain operational, and mind phenomenal space-time. *Phys Life Rev* 7(2):195–249
- Fingelkurts AA, Fingelkurts AA, Neves CF (2013) Consciousness as a phenomenon in the operational architectonics of brain organization: criticality and self-organization considerations. *Chaos Solitons Fractals* 55:13–31
- Fisher MP (2015) Quantum cognition: the possibility of processing with nuclear spins in the brain. *Ann Phys* 362:593–602
- Fisher MP (2017) Are we quantum computers or merely clever robots? *Int J Mod Phys B* 31:1743001
- Freeman WJ (2007) Proposed cortical ‘shutter’ mechanism in cinematographic perception. In: Perlovsky L, Kozma R (eds) *Neurodynamics of cognition and consciousness*. Springer, Heidelberg, pp 11–38
- Fried I, Mukamel R, Kreiman G (2011) Internally generated preactivation of single neurons in human medial frontal cortex predicts volition. *Neuron* 69:548–562
- Friston K (2008) Hierarchical models in the brain. *PLoS Comput Biol* 4(11):e1000211
- Friston KJ (2010) The free-energy principle: a unified brain theory? *Nat Rev Neurosci* 11:127–138
- Friston K (2018) Am I self-conscious? (Or does self-organization entail self-consciousness?). *Front Psychol* 9:579
- Friston KJ, Breakspear M, Deco G (2012) Perception and self-organized instability. *Front Comput Neurosci* 6:44
- Friston K, Schwartenbeck P, Fitz GT, Moutoussis M, Behrens T, Dolan RJ (2013) The anatomy of choice: active inference and agency. *Front Hum Neurosci* 7:598
- Fujisawa S, Matsuki N, Ikegaya Y (2006) Single neurons can induce phase transitions of cortical recurrent networks with multiple internal states. *Cereb Cort* 16:639–654
- Gallicchio JS, Friedman AS, Kaiser DI (2014) Testing Bell’s inequality with cosmic photons: closing the setting-independence loophole. *Physics Rev Lett* 112:110405
- Ganguli M, Blacker D, Blazer DG, Grant I, Jeste DV, Paulsen JS et al (2011) Classification of neurocognitive disorders in DSM-5: a work in progress. *Am J Geriatr Psychiatry* 19(3):205–210
- Georgiev DD (2020) Quantum information theoretic approach to the mind–brain problem. *Prog Biophys Mol Biol* 158:16–32
- Giacino JT, Fins JJ, Laureys S, Schiff ND (2014) Disorders of consciousness after acquired brain injury: the state of the science. *Nat Rev Neurol* 10:99–114
- Gisin N (2013) Are there quantum effects coming from outside space–time? Nonlocality, free will and no many-worlds. In: Suarez A, Adams P (eds) *Is science compatible with free will?* Springer, New York, pp 23–39
- Golesorkhi M, Gomez-Pilar J, Zilio F, Berberian N, Wolff A, Mustapha CE et al (2021) The brain and its time: intrinsic neural timescales are key for input processing. *Commun Biol* 4:970
- Golkowski D, Larroque SK, Vanhauzenhuysse A, Plenevaux A, Boly M, Di Perri C et al (2019) Changes in whole brain dynamics and connectivity patterns during sevoflurane- and propofol-induced unconsciousness identified by functional magnetic resonance imaging. *Anesthesiology* 130:898–911
- Gosseries O, Dim H, Laureys S, Boly M (2014) Measuring consciousness in severely damaged brains. *Annu Rev Neurosci* 37:457–478
- Guggisberg AG, Mottaz A (2013) Timing and awareness of movement decisions: does consciousness really come too late? *Front Hum Neurosci* 7:385
- Hager B, Yang AC, Brady R, Meda S, Clementz B, Pearson GD et al (2017) Neural complexity as a potential translational biomarker for psychosis. *J Affect Disord* 216:89–99
- Hahn G, Petermann T, Havenith MN, Shan Y, Singer W, Plenz D et al (2010) Neuronal avalanches in spontaneous activity in vivo. *J Neurophysiol* 104:3312–3322

- Haimovici A, Tagliazucchi E, Balenzuela P, Chialvo DR (2013) Brain organization into resting state networks emerges at criticality on a model of the human connectome. *Phys Rev Lett* 110:178101
- Haldeman C, Beggs JM (2005) Critical branching captures activity in living neural networks and maximizes the number of metastable states. *Phys Rev Lett* 94:058101
- Hameroff S (2012) How quantum brain biology can rescue conscious free will. *Front Integr Neurosci* 6:93
- Hameroff S, Penrose R (2014) Consciousness in the universe: a review of the ‘Orch OR’ theory. *Phys Life Rev* 11:39–78
- Haynes JD, Sakai K, Rees G, Gilbert S, Frith C, Passingham RE (2007) Reading hidden intentions in the human brain. *Current Bio* 17:323–328
- He BJ (2014) Scale-free brain activity: past, present, and future. *Trends Cogn Sci* 18:480–487
- He BJ, Zempel JM, Snyder AZ, Raichle ME (2010) The temporal structures and functional significance of scale-free brain activity. *Neuron* 66:353–369
- Heiney K, Huse Ramstad O, Fiskum V, Christiansen N, Sandvig A, Nichele S, Sandvig I (2021) Criticality, connectivity and neural disorder: a multifaceted approach to neural computation. *Front Comput Neurosci* 15:611183
- Herzog MH, Kammer T, Scharnowski F (2016) Time slices: what is the duration of a percept? *PLoS Biol* 14(4):e1002433
- Herzog MH, Drissi-Daoudi L, Doerig A (2020) All in good time: long-lasting postdictive effects reveal discrete perception. *Trends Cogn Sci* 24(10):826–837
- Hesse J, Gross T (2014) Self-organized criticality as a fundamental property of neural systems. *Front Syst Neurosci* 8:166
- Hiscock HG, Worster S, Kattinig DR, Steers C, Jin Y, Manolopoulos DE et al (2016) The quantum needle of the avian magnetic compass. *PNAS USA* 113:4634–4639
- Hohwy J (2009) The neural correlates of consciousness: new experimental approaches needed? *Conscious Cogn* 18:428–443
- Hohwy J, Seth A (2020) Predictive processing as a systematic basis for identifying the neural correlates of consciousness. *Philos Mind Sci* 1(II):3
- Hohwy J, Roepstorff A, Friston K (2008) Predictive coding explains binocular rivalry: an epistemological review. *Cognition* 108:687–701
- Hossenfelder S, Palmer T (2020) Rethinking superdeterminism. *Front Phys* 8:139
- Houweling AR, Doron G, Voigt BC, Herfst LJ, Brecht M (2010) Nanostimulation: manipulation of single neuron activity by juxtacellular current injection. *J Neurophysiol* 103:1696–1704
- Huang Z, Zhang J, Wu J, Qin P, Wu X, Wang Z et al (2016) Decoupled temporal variability and signal synchronization of spontaneous brain activity in loss of consciousness: an fMRI study in anesthesia. *Neuroimage* 124:693–703
- Huang Z, Zhang J, Wu J, Mashour GA, Hudetz AG (2020) Temporal circuit of macroscale dynamic brain activity supports human consciousness. *Sci Advan* 6:eaz0087
- Hudetz AG, Humphries CJ, Binder JR (2014) Spin-glass model predicts metastable brain states that diminish in anesthesia. *Front Syst Neurosci* 8:234
- Hudetz AG, Liu X, Pillay S, Boly M, Tononi G (2016) Propofol anesthesia reduces Lempel-Ziv complexity of spontaneous brain activity in rats. *Neurosci Lett* 628:132–135
- Hunt T, Schooler JW (2019) The easy part of the hard problem: a resonance theory of consciousness. *Front Hum Neurosci* 13:378
- Ito S, Hansen ME, Heiland R, Lumsdaine A, Litke AM, Beggs JM (2011) Extending transfer entropy improves identification of effective connectivity in a spiking cortical network model. *PLoS ONE* 6:e27431
- Jackson NA (2003) A survey of music therapy methods and their role in the treatment of early elementary school children with ADHD. *J Music Ther* 40(4):302–323
- Jedlicka P (2017) Revisiting the quantum brain hypothesis: toward quantum (neuro)biology? *Front Mol Neurosci* 10:366
- Jirsa VK, Stacey WC, Quilichini PP, Ivanov AI, Bernard C (2014) On the nature of seizure dynamics. *Brain* 137:2210–2230
- Kaiser M (2017) Mechanisms of connectome development. *Trends Cogn Sci* 21:9
- Kelso JS (1995) *Dynamic patterns: the self-organization of brain and behavior*. MIT Press, Cambridge
- Khalighinejad N, Schurger A, Desantis A, Zmigrod L, Haggard P (2018) Precursor processes of human self-initiated action. *Neuroimage* 165:35–47
- Kim H, Lee U (2019) Criticality as a determinant of integrated information ϕ in human brain networks. *Entropy* 21:981
- Kleiner J, Hoel E (2021) Falsification and consciousness. *Neurosci Conscious* 7(1):niab001
- Knauer B, Stüttgen MC (2019) Assessing the impact of single-cell stimulation on local networks in rat barrel cortex—a feasibility study. *Int J Mol Sci* 20:2604
- Knill DC, Pouget A (2004) The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends Neurosci* 27(12):712–719
- Koch C (2009) Free will physics biology and the brain. In: Murphy N et al (eds) *Downward causation and the neurobiology of free will*. Springer, Berlin, pp 31–52
- Koch C, Massimini M, Boly M, Tononi G (2016) Neural correlates of consciousness: progress and problems. *Nat Rev Neurosci* 17:307–321
- Kozma R, Freeman WJ (2008) Intermittent spatio-temporal desynchronization and sequenced synchrony in ECoG signals. *Chaos* 18:037131
- Kozma R, Freeman WJ (2017) Cinematic operation of the cerebral cortex interpreted via critical transitions in self-organized dynamic systems. *Front Syst Neurosci* 11:10
- Lake BM, Ullman TD, Tenenbaum JB, Gershman SJ (2017) Building machines that learn and think like people. *Behav Brain Sci* 40:e253
- Lau H, Rosenthal D (2011) Empirical support for higher-order theories of conscious awareness. *Trends Cogn Sci* 15(8):365–373
- Laureys S (2005) The neural correlate of (un)awareness: lessons from the vegetative state. *Trends Cogn Sci* 9(12):556–559
- Lavazza A (2016) Free will and neuroscience: from explaining freedom away to new ways of operationalizing and measuring it. *Front Hum Neurosci* 10:262
- Lee U, Oh G, Kim S, Noh G, Choi B, Mashour GA (2010) Brain networks maintain a scale-free organization across consciousness anesthesia and recovery: evidence for adaptive reconfiguration. *Anesthesiology* 113:1081–1091
- Lee H, Golkowski D, Jordan D, Berger S, Ilg R, Lee J et al (2019) Relationship of critical dynamics functional connectivity and states of consciousness in large-scale human brain networks. *Neuroimage* 188:228–238
- Libet B (1985) Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behav Brain Sci* 8:529–566
- Liu X, Ward BD, Binder JR, Li SJ, Hudetz AG (2014) Scale-free functional connectivity of the brain is maintained in anesthetized healthy participants but not in patients with unresponsive wakefulness syndrome. *PLoS ONE* 9:e92182
- London M, Roth A, Beeren L, Häusser M, Latham PE (2010) Sensitivity to perturbations in vivo implies high noise and suggests rate coding in cortex. *Nat Neurosci* 13:123–127
- López-González A, Panda R, Ponce-Alvarez A, Zamora-López G, Escrichs A, Martia S et al (2021) Loss of consciousness reduces the stability of brain hubs and the heterogeneity of brain dynamics. *Commun Bio* 4:1037

- López-Ruiz R, Mancini HL, Calbet X (1995) A statistical measure of complexity. *Phys Lett A* 209:321–326
- Maillet D, Schacter DL (2016) From mind wandering to involuntary retrieval: age-related differences in spontaneous cognitive processes. *Neuropsychology* 80:142–156
- Marr D (2010) *Vision: a computational investigation into the human representation and processing of visual information*. MIT Press, Cambridge
- Mashour GA, Hudetz AG (2018) Neural correlates of unconsciousness in large-scale brain networks. *Trends Neurosci* 41(3):150–160
- Mashour GA, Roelfsema P, Changeux JP, Dehaene S (2020) Conscious processing and the global neuronal workspace hypothesis. *Neuron* 105:776–798
- Mateos DM, Guevara Erra R, Wennberg R, Perez Velazquez JL (2018) Measures of entropy and complexity in altered states of consciousness. *Cogn Neurodyn* 12:73–84
- Mediano PAM, Seth AK, Barrett AB (2019) Measuring integrated information: comparison of candidate measures in theory and simulation. *Entropy* 21:17
- Mediano PAM, Farah JC, Shanahan MP (2016) Integrated information and metastability in systems of coupled oscillators. [arXiv:1606.08313](https://arxiv.org/abs/1606.08313)
- Mehler DMA, Kording KP (2018) The lure of misleading causal statements in functional connectivity research. [arXiv 2018, arXiv:1812.03363](https://arxiv.org/abs/1812.03363)
- Meisel C, Storch A, Hallmeyer-Elgner S, Bullmore E, Gross T (2012) Failure of adaptive self-organized criticality during epileptic seizure attacks. *PLoS Comput Biol* 8(1):e1002312
- Merker B (2007) Consciousness without a cerebral cortex: a challenge for neuroscience and medicine. *Behav Brain Sci* 30:63–134
- Miller EK, Cohen JD (2001) An integrative theory of prefrontal cortex function. *Annu Rev Neurosci* 24:167–202
- Naccache L (2018) Why and how access consciousness can account for phenomenal consciousness. *Philos Trans R Soc B* 373:20170357
- Noë A, Thompson E (2004) Are there neural correlates of consciousness? *J Conscious Stud* 11:3–28
- Northoff G, Gomez-Pilar J (2021) Overcoming rest–task divide—abnormal temporospatial dynamics and its cognition in schizophrenia. *Schizophr Bull* 47:751–765
- Northoff G, Huang Z (2017) How do the brain's time and space mediate consciousness and its different dimensions? Temporo-spatial theory of consciousness (TTC). *Neurosci Biobehav Rev* 80:630–645
- Northoff G, Lamme V (2020) Neural signs and mechanisms of consciousness: is there a potential convergence of theories of consciousness in sight? *Neurosci Biobehav Rev* 118:568–587
- Northoff G, Zilio F (2022) Temporo-spatial theory of consciousness (TTC)—bridging the gap of neuronal activity and phenomenal states. *Behav Brain Res* 424:113788
- Novelli L, Wollstadt P, Mediano P, Wibral M, Lizier JT (2019) Large-scale directed network inference with multivariate transfer entropy and hierarchical statistical testing. *Netw Neurosci* 3(3):827–847
- O'Reilly EJ, Olaya-Castro A (2014) Non-classicality of the molecular vibrations assisting exciton energy transfer at room temperature. *Nat Commun* 5:3012
- Oizumi M, Albantakis L, Tononi G (2014) From the phenomenology to the mechanisms of consciousness: integrated information theory 3.0. *PLoS Comput Biol* 10:e1003588
- Olcese U, Bos JJ, Vinck M, Pennartz CMA (2018) Functional determinants of enhanced and depressed inter-areal information flow in NREM sleep between neuronal ensembles in rat cortex and hippocampus. *Sleep* 41(11):zsy167
- Oudheusden LJB, Draisma S, van der Salm S, Cath D, van Oppen P, van Balkom AJL et al (2018) Perceptions of free will in obsessive-compulsive disorder: a quantitative analysis. *BMC Psychiatry* 18:400
- Owen AM, Coleman MR, Boly M, Davis MH, Laureys S, Pickard JD (2006) Detecting awareness in the vegetative state. *Science* 313:1402
- Pal D, Li D, Dean JG, Brito MA, Liu T, Fryzel AM et al (2020) Level of consciousness is dissociable from electroencephalographic measures of cortical connectivity slow oscillations and complexity. *J Neurosci* 40(3):605–618
- Parsons D (2008) *Directory of classical themes*, Spencer Brown
- Parvizi J, Damasio A (2001) Consciousness and the brainstem. *Cognition* 79:135–160
- Pennartz CMA, Farisco M, Evers K (2019) Indicators and criteria of consciousness in animals and intelligent machines: an inside-out approach. *Front Syst Neurosci* 13:25
- Pezzulo G, Rigoli F, Friston K (2018) Hierarchical active inference: a theory of motivated control. *Trends Cogn Sci* 22:4
- Pironio S, Acín A, Massar S, de la Giroday AB, Matuskevich DN, Maunz P et al (2010) Random numbers certified by Bell's theorem. *Nat Phys* 6:1021–1024
- Rees G, Kreiman G, Koch C (2002) Neural correlates of consciousness in humans. *Nat Rev Neurosci* 3:261–270
- Reid AT, Headley DW, Mill RD, Sanchez-Romero R, Uddin LQ, Marinazzo D et al (2019) Advancing functional connectivity research from association to causation. *Nat Neurosci* 22(11):1751–1760
- Ritz T (2011) Quantum effects in biology: bird navigation. *Procedia Chem* 3:262–275
- Rolls ET (2020) Neural computations underlying phenomenal consciousness: a higher order syntactic thought theory. *Front Psychol* 11:65
- Rolls ET, Loh M, Deco G (2008) An attractor hypothesis of obsessive-compulsive disorder. *Eur J Neurosci* 28(4):782–793
- Rolls ET, Cheng W, Feng J (2021) Brain dynamics: the temporal variability of connectivity, and differences in schizophrenia and ADHD. *Transl Psychiatry* 11:70
- Sabbadini SA, Vitiello G (2019) Entanglement and phase-mediated correlations in quantum field theory. Application to brain-mind states. *Appl Sci* 9:3203
- Safron A (2020) An integrated world modeling theory (IWMT) of consciousness: combining integrated information and global neuronal workspace theories with the free energy principle and active inference framework; toward solving the hard problem and characterizing agentic causation. *Front Artif Intell* 3:30
- Sanyal S, Nag S, Banerjee A, Sengupta R, Ghosh D (2019) Music of brain and music on brain: a novel EEG sonification approach. *Cogn Neurodyn* 13:13–31
- Sarasso S, Boly M, Napolitani M, Gosseries O, Charland-Verville V, Casarotto S et al (2015) Consciousness and complexity during unresponsiveness induced by propofol xenon and ketamine. *Curr Biol* 25:1–7
- Sarasso S, Casali AG, Casarotto S, Rosanova M, Sinigaglia C, Massimini M (2021) Consciousness and complexity: a concision of evidence. *Neurosci Conscious* 7(2):1–24
- Sattin D, Magnani FG, Bartesaghi L, Caputo M, Fittipaldo AV, Cacciatore M et al (2021) Theoretical models of consciousness: a scoping review. *Brain Sci* 11:535
- Schiff ND, Nauvel T, Victor JD (2014) Large-scale brain dynamics in disorders of consciousness. *Curr Opin Neurobiol* 25:7–14
- Schreiber T (2000) Measuring information transfer. *Phys Rev Lett* 85:461
- Schultze-Kraft M, Birman D, Rusconi M, Allefeld C, Görden K, Dähne S et al (2016) The point of no return in vetoing self-initiated movements. *PNAS USA* 113:1080–1085

- Schurger A, Mylopoulos M, Rosenthal D (2016) Neural antecedents of spontaneous voluntary movement: a new perspective. *Trends Cogn Sci* 20:77–79
- Seow TXF, Gillan CM (2020) Transdiagnostic phenotyping reveals a host of metacognitive deficits implicated in compulsivity. *Sci Rep* 10(1):2883
- Seth AK (2013) Interoceptive inference emotion and the embodied self. *Trends Cogn Sci* 17(11):565–573
- Seth AK (2014) A predictive processing theory of sensorimotor contingencies: explaining the puzzle of perceptual presence and its absence in synesthesia. *Cogn Neurosci* 5(2):97–118
- Seth AK, Barrett AB, Barnett L (2011) Causal density and integrated information as measures of conscious level. *Philos Trans R Soc A* 369:3748–3767
- Shapiro KL, Raymond JE, Arnell KM (1997) The attentional blink. *Trends Cogn Sci* 1(8):291–296
- Shea N, Frith CD (2019) The global workspace needs metacognition. *Trends Cogn Sci* 23(7):560–571
- Shew WL, Yang H, Yu S, Roy R, Plenz D (2011) Information capacity and transmission are maximized in balanced cortical networks with neuronal avalanches. *J Neurosci* 31:55–63
- Signorelli CM, Meling D (2021) Towards new concepts for a biological neuroscience of consciousness. *Cogn Neurodyn* 15:783–804
- Signorelli CM, Szczotka J, Prentner R (2021) Explanatory profiles of models of consciousness—towards a systematic classification. *Neurosci Conscious* 7(2):1–13
- Soon CS, He AH, Bode S, Haynes JD (2013) Predicting free choices for abstract intentions. *PNAS USA* 110:5733–5734
- Stam CJ, de Bruin EA (2004) Scale-free dynamics of global functional connectivity in the human brain. *Hum Brain Mapp* 22:97–109
- 't Hooft G (2016) The cellular automaton interpretation of quantum mechanics. Springer, Berlin
- Tagliazucchi E (2017) The signatures of conscious access and its phenomenology are consistent with large-scale brain communication at criticality. *Conscious Cogn* 55:136–147
- Tagliazucchi E, Chialvo DR, Siniatchkin M, Amico E, Brichant JF, Bonhomme V et al (2016) Large-scale signatures of unconsciousness are consistent with a departure from critical dynamics. *J R Soc Interface* 13:20151027
- Talwar N, Crawford MJ, Maratos A, Nur U, McDermott O, Procter S (2006) Music therapy for patients with schizophrenia. Exploratory randomized controlled trial. *Br J Psychiatry* 189:405–409
- Tanke N, Borst JGG, Houweling AR (2018) Single-cell stimulation in barrel cortex influences psychophysical detection performance. *J Neurosci* 38:2057–2068
- Tegmark M (2015) Consciousness as a state of matter. *Chaos Solit Fract* 76:238–270
- Tognoli E, Kelso JA (2014) The metastable brain. *Neuron* 81:35–48
- Tononi G (2008) Consciousness as integrated information: a provisional manifesto. *Biol Bull* 215:216–242
- Tononi G, Koch C (2015) Consciousness: here there and everywhere? *Philos Trans R Soc B* 370:20140167
- Tononi G, Sporns O, Edelman GM (1994) A measure for brain complexity: relating functional segregation and integration in the nervous system. *PNAS USA* 91:5033–5037
- Torday JS, Miller WB (2016) On the evolution of the mammalian brain. *Front Syst Neurosci* 10:31
- Torres EB (2020) Reframing psychiatry for precision medicine. *J Pers Med* 10:144
- Trimble M, Hesdorffer D (2017) Music and the brain: the neuroscience of music and musical appreciation. *Br J Psychiatry* 142(2):28–31
- Uhlhaas PJ, Singer W (2006) Neural synchrony in brain disorders: relevance for cognitive dysfunctions and pathophysiology. *Neuron* 52:155–168
- Ursino M, Ricci G, Magosso E (2020) Transfer entropy as a measure of brain connectivity: a critical analysis with the help of neural mass models. *Front Comput Neurosci* 14:45
- VanRullen R, Kanai R (2021) Deep learning and the global workspace theory. *Trends Neurosci* 10:1016
- VanRullen R, Koch C (2003) Is perception discrete or continuous? *Trends Cogn Sci* 7:207–213
- VanRullen R, Zoefel B, Ilhan B (2014) On the cyclic nature of perception in vision versus audition. *Philos Trans R Soc B* 369:20130214
- Varela FJ et al (1999) The specious present: a neurophenomenology of time consciousness. In: Petitot J et al (ed) *Naturalizing phenomenology*. Stanford University Press, Redwood City, pp 266–314
- Wang R, Liu M, Cheng X, Wu Y, Hildebrandt A, Zhou C (2021) Segregation, integration, and balance of large-scale resting brain networks configure different cognitive abilities. *PNAS USA* 118(23):e2022288118
- Ward LW (2011) The thalamic dynamic core theory of conscious experience. *Conscious Cogn* 20(2):464–486
- Watanabe T, Rees G, Masuda N (2019) Atypical intrinsic neural timescale in autism. *Elife* 8:e42256
- Weichwald S, Peters J (2021) Causality in cognitive neuroscience: concepts, challenges, and distributional robustness. *J Cogn Neurosci* 33(2):226–247
- Werner G (2009) Consciousness related neural events viewed as brain state space transitions. *Cogn Neurodyn* 3:83–95
- Wooters WK, Zurek WH (2008) The no-cloning theorem. *Phys Today* 62:2
- Xu L, Feng J, Yu L (2022) Avalanche criticality in individuals, fluid intelligence, and working memory. *Hum Brain Mapp* 43(8):2534–2553
- Yu Y, Shen H, Zeng LL, Ma Q, Hu D (2013) Convergent and divergent functional connectivity patterns in schizophrenia and depression. *PLoS ONE* 8:e68250
- Yuan Y, Liu J, Zhao P, Xing F, Huo H, Fang T (2019) Structural insights into the dynamic evolution of neuronal networks as synaptic density decreases. *Front Neurosci* 13:892
- Yurchenko SB (2017) Can “theory of everything” be global theory of consciousness? Ontology and psychodynamics of I-observer. *Neuroquantology* 15(2):118–131
- Yurchenko SB (2021) The importance of randomness in the universe: superdeterminism and free will. *Axiomathes* 31(4):453–478
- Zeki S (2003) The disunity of consciousness. *Trends Cogn Sci* 7:214–218
- Zhang J, Huang Z, Chen Y, Zhang J, Ghinda D, Nikolova Y et al (2018) Breakdown in the temporal and spatial organization of spontaneous brain activity during general anesthesia. *Hum Brain Mapp* 39:2035–2046
- Zimmerman V (2020) Why brain criticality is clinically relevant: a scoping review. *Front Neural Circuit* 14:54

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.