Consciousness and Cognition 82 (2020) 102937

Contents lists available at ScienceDirect

Consciousness and Cognition

journal homepage: www.elsevier.com/locate/concog

# Account of consciousness by Christof Koch: Review and questions

Talis Bachmann

*Department of Penal Law, School of Law, University of Tartu (Tallinn Branch), Kaarli Puiestee 3, 10119 Tallinn, Estonia*

ARTICLE INFO

*Keywords:*
Consciousness
Integrated information
Cognitive computation
Microgenesis
Phenomenal experience

ABSTRACT

This review is set to present the gist of the theoretical account of consciousness recently presented by Christof Koch and pose a couple of questions instigated by this account. The expected answers to these questions would hopefully help to advance our understanding of the basic nature of the conscious mind.

## 1. Theory of consciousness as seen by Christof Koch

Current theories of consciousness are notably influenced by works of Christof Koch. His views were developed in collaboration with Francis Crick (e.g., Crick & Koch, 1990; Crick & Koch, 2003), elaborted subsequently (e.g., Koch, 2004), and lately tuned in close collaboration with some colleagues such as Giulio Tononi and others (e.g., Koch, Massimini, Boly, & Tononi, 2016). Especially in recent years the influence of Tononi's integrated information theory (e.g., Tononi, 2012) on Koch's thinking is easy to spot. However, the integrated information theory (IIT) which was proposed by Giulio Tononi is not all that is described as a theory of consciousness by Koch. His recent conceptualization in its developed variety is quite comprehensively and in sufficient detail presented in (Koch, 2019) and could be seen as one of the most elaborated theoretical explanations of consciousness. As this recent work (2019) has brought together various ingredients of the Koch's theoretical account from his other publications, it stands as a pertinent source of the present review and will be further on used as a representative set for Koch's version of the integrated information theory.

The "pre-theoretical" axioms of Koch's theory of consciousness (KTC) are founded on how subjective experience feels from the first person perspective, which allows to put forward theoretical postulates for how one must describe a natural system capable of implementing or simulating the basic attributes of subjective experience characterised by the set of these axioms. KTC, substantially inspired by IIT, starts with experience and asks how matter must be organized to support the mental (Koch, 2019, p. 74, p. 190). Any experience is structured, but this structure is integrated within the irreducible wholistic felt present where nothing else is contained. Thus, any one experience is definite and at the same time informative – one out of the millennia of possibilities potentially achievable by the experiencing system. There are five phenomenological properties as axioms: *intrinsic* nature (conscious experience exists for himself, for the experiencing organism), *structuredness*, *informational* specificity (consciousness has specific contents), *integrated* oneness, *definiteness* (Koch, 2019, p. 9, 75). Consciousness properties are chracterized by mutual consistency and independence. The KTC deserves serious attention because it is bolstered by a large amount of empirical supportive material, is contextuated in the discussion of most of the current theories and research issues of consciousness science and, in its discourse, the logic of arguments in favor of it tends to prevail over some purely declarative statements. Importantly from the methodological point of view, the theory is bolstered by the fact that science deals with non-observable entities also, so subjective phenomenology is scientific subject matter. (The predicted matter of fact that other individuals are not zombies is extremely likely and this line of thinking is justified because abductive reasoning has a solid place in science.)

*E-mail address:* talis.bachmann@ut.ee.

https://doi.org/10.1016/j.concog.2020.102937
Received 7 January 2020; Received in revised form 16 April 2020; Accepted 18 April 2020

According to KTC, similarly to the IIT, consciousness is not determined by extrinsic causal powers, but intrinsic causal powers (ICP) determine consciousness as such. For the solution of the "hard problem" (Chalmers, 1996) of consciousness science the theory suggests a principally opposite strategy compared to most of the well known approaches: it goes from phenomenal experience to its implementation in the physical system. Neither environmental stimulation nor the afferent nervous system are the starting points for the theory. (I must add that while typical scientific information is a result of measurement by some measurement device, human subjects themselves are also a kind of measurement devices with reliability of the reports about experiences oftentimes higher than that of the many physical devices. Statistics of the deliberate human lies need not be "worse" than the "lies" of the physical device.) One must start with how the first-person introspective experiences are felt; this is what has to be explained. Indeed, a person in her awake state, but with eyes closed, ears shut with earplugs, laying motionless in a quiet room lacking any suprathreshold olfactory or other stimulation is nevertheless conscious, characterized by the feeling of being here and alive. Intrinsic nature of consciousness is characterized by Koch (2019, p. 153) by this: black empty wall is still experienced intrinsically, but has zero value extrinsically. Importantly, virtually all intelligent cognitive operations executed by the brain (especially as mediated by the frontal cortex) can be carried out in a zombie mode without the necessity for the accompanying subjective experience of these cognitive computations. Consciousness is not computed, it is an intrinsic property of the natural system capable of integrating information. The concept of information adopted in the KTC is not Shannonian (which is extrinsic, performed by an input–output device set to detect signals in noise), but comes from the fact that a sentient system (the whole) in one actual state of the specific pattern of the configuration of the states of its subsystems is different from the multitude of the potential different states of the same system. The concept of information here is intrinsic, contained in the signals' structure itself (*in-formare*, Aristotelian "give form or shape to") (Koch, 2019, p. 84-85). (The numerical value of the ratio of the number of different possible structurally specified state contents to the one actual structural state content is measurable by a virtually uncomputable number of variants, which therefore grants a huge informational capacity of the system by the repertoire of its states. What makes this possible is evident in the buildup of the fully functioning brains of the living organisms – this buildup consists of many billions of neurons and other types of cells and a multitude of biochemical substances necessary for the system's work, with neurons interconnected by a vast number of mutual connections. The key attribute of this networked interactive structure is *recurrence* of internal influences between the subsystems. The system has ICP as it is capable of causally changing its states "from within"; the present state causally grows out of the past state and causes the future state.)

Importantly for the KTC, quite complex information-processing operations at the level of intelligent behavior can be performed by systems without any capacity for conscious experience. Consciousness is not a result of cognitive computation or cannot be blown into some arbitrary system by emulating only the input-output functions of a system with known consciousness capability, no matter how sophisticated tasks these computational functions can solve. "Experience is different from thinking, being smart, or attending" (Koch, 2019, p. 31). Consciousness is not a clever algorithm, its essence is causal power upon itself, not computation (Koch, 2019, p. 149). Experience is present also in simple organisms, perhaps even in single cells. Genetic, behavioral and physiological *continuity* packed together in a system capable of integrating information is a premise for consciousness; experience without voice or button-pushing is still experience. In this capacity, the KTC supports the usefulness of the no-report experimental paradigms to tackle the classic issue of how direct first-person experience can be studied by the third person methodology (e.g., Block, 2019; Pitts, Metzler, & Hillyard, 2014; Tsuchiya, Wilke, Frässle, & Lamme, 2015). What is needed for consciousness is an intrinsic cause-effect structure capable of creating integrated information wholes subserving the qualitative feel from the inner perspective of the experiencing system. A system without consciousness can be more intelligent compared to a system possessing consciousness. (However, bear in mind that Koch (2019, p. 150) admits that neuromorphic electronic hardware could in principle allow computers that have experience. This admission is doubtful for those who attribute consciousness exclusively to living organisms formed along the millennia of years of biological evolution.) For specifying the state of the integrated configuration of subsystems (as elements of the whole system), a special integrated information measure termed Phi was deveoped (e.g., Albantakis, Hintze, Koch, Adami, & Tononi, 2014; Tononi, 2012) and is generically used by Koch. It quantifies to what extent the form that is generated by the whole is different from that generated by its parts. The system state is a conscious state insofar as Phi has a non-zero value; only a system with a non-zero maximum of integrated information is conscious (Koch, 2019, p. 76). The higher the value of Phi, the more conscious the system is. Among the different integrated constellations the one with highest Phi value exclusively defines the contents of present conscious experience; others can work in a zombie mode.

Most of the empirical support for the theory has come from the brain imaging and neurological data demonstrating that increased integration of multiple cortical areas by active connectivity correlates with consciousness (e.g., Casali et al., 2013; Massimini et al., 2005; Massimini & Tononi, 2018). Neural correlates of consciousness (NCC) for sensory experiences are marked by KTC in the so-called "hot zone" in the broadly specified temporo-parietal-occipital region of the posterior cerebral cortex (Koch, 2019, p. 61). However, prefrontal cortex seems not necessary as the constitutive part of NCC, which is supported by neurological and neurosurgery evidence, neuropsychological lesion data and stimulations. On the other hand, frontal part of the brain is necessary for intelligence and behavioral control. It is fair to add though that there is the distinction between NCC that mark unconscious prerequisite processes (NCCpr) necessarily preceding NCC and proper NCC that mark the ongoing conscious processes (Aru, Bachmann, Singer, & Melloni, 2012; Bachmann, 2009). Both are revealed as a result of contrastive analysis ([markers recorded when stimulus is consciously experienced] - [markers recorded when the stimulus is not consciously experienced] = NCC *OR* NCCpr). This allows to accept that frontal-cortex mechanisms may act as NCCpr in many circumstances such as hallucinations ignited by expectation or top-down inhibition of competitors for specific selective percept to occur. The latter possibility is consistent with Koch (2019, p. 116) notion that the *maximally irreducible cause-effect structure* associated with a whole (the equivalent of present subjective experience) depends not only on the interneuronal connectivity, but also on the current state of the neurons, including the inactive states. (Inactive units selectively constrain what is making up the actual active structure and what will be the future states of the system. Moreover, a

relatively inactive cortex may be constitutive of an experience. It should be added here, however, that while Koch mostly speaks about spiking and non-spiking neurons, the excitatory post-synaptic potential levels of the spike-wise "silent" neurons can be and perhaps must be also considered as the determinants of the wholistic irreducible structure of the integrated information.)

In what concerns the "hot zone" location for consciousness, the KTC stands in opposition to several influential consciousness theories where NCC are heavily loaded on frontal cortex (e.g., Brown, Lau, & Ledoux, 2019; Dehaene & Changeux, 2011; Michel & Morales, 2019). For the present reviewer, there is no convincing evidence that frontal activities correlating with conscious contents are the NCC proper and not NCCpr (Aru et al., 2012; Bachmann, 2009). On the other hand, given the current knowledge of the issue, one may want to agree with Michel and Morales (2019) sober comment that there might be more than one region or neuroanatomically mapped process for rendering consciousness. Moreover, the mechanisms marked by NCC recordings may be redundant, meaning also that what appears as NCC in one set of experimental and subject conditions may not be so in other conditions.

Juxtaposing consciousness and computability is one of the most original features of the KTC (Koch, 2019, e.g., p. 129–137). There is a simple, but powerful distinction: replicating anything that consciousness-possessing humans *can do* can not be taken for proof that the replicating systems *can be* like humans and therefore conscious. Importantly and in support for non-computational, non-functionalist KTC, it must be acknowledged that phenomenal conscious experiences are often non-veridical from the third-person perspective (although they are always "veridical" from the first-person perspective, as they are definitely subjectively real). The main aim of computing artifacts is help (if not replace) humans by precise, errorless, correct problem solving and robotic actions. Phenomenology as experienced by humans features illusions and hallucinations (e.g., Aru, Tulver, & Bachmann, 2018; Powers, Mathys, & Corlett, 2017). Consciousness also goes hand in hand with autonoetic possibility to mentally travel in time – between past, present and imagined future and back to present. Objectively nonveridical, but subjectively real experiences can result from conditioned expectations (e.g., Aru et al., 2018; Powers et al., 2017) constructive expectational bias (e.g., Partos, Cropper, & Rawlings, 2016), cognitive beliefs (e.g., Schmack, Rothkirch, Priller, & Sterzer, 2017) and various other contextual influences (de Lange, Heilbron, & Kok, 2018). All these examples refer to internal actually emerging causal effects, which indirectly supports the KTC theoretical foundations. However, it is another matter that majority of cases of imaging future and hallucinating present are determined by learned priors with their footprints in memory.

At a more belletristic level, KTC as it is perused by its author does not exclude panpsychism (Koch, 2019, chapter 14): consciousness may be in many places, but this is on the premise that there is non-zero Phi Max integrated information by intrinsic causality. On the other hand, Koch dismisses the possibility that experiences could aggregate into larger, superordinate experiences such as consciousness of a human corporational collective or nation. The exclusion postulate for Wholes – perhaps the central concept of KTC, which associates it with neuro-Gestaltism – explains this postulate: aggregate collectives have no intrinsic causal power and thus no phenomenal reality. Koch also admits that insofar as highly intelligent artifacts will have no consciousness, principles of ethics and morale necessary for sentient beings need not be applied to the sentience-free machines.

Interestingly, some basic principles of KTC find some similarities in the microgenetic theory, described in essence by Jason Brown (2013, 2015, 2017). (How microgenetic approach deals with actual formation of the conscious perceptual experience as based on data from experiments is reviewed in Bachmann (2000) and contextuated in terms of the pending tasks for consciousness research by Aru and Bachmann (2017). I recommend these sources for more details on the microgenetic account.) First, in the prevailing theoretical models of the mental, mind is generated by experience of the external world, which becomes ingredient in mental process. However, externalist theoreticians must inevitably agree that knowledge and perception are mental processes realized entirely within the brain. For the adherents to microgenetic theory it is also obvious that mind is shaped and altered by experience, but experience is not brought in by the inflow of sense data into the mind from the world. Sense data influence the content and character of the *endogenous gradual process*, which is the basic source of mind (Brown, 2017). External conditions constrain and modify a process that originates and develops within the mind, allowing the endogenously determined mind become adaptively useful. By virtue of this internalism, the microgenetic view and KTC are mutually consistent. Second, microgenetic process theorists (e.g., Brown, 2017) emphasise contrast between constructivist models and holistic derivation models. These alternative positions about how perception is formed either accept that perception is analogous to building up an object by addition of elements or, alternatively, that in perception its object is revealed by the elimination of whatever is extraneous or irrelevant. The microgenetic approach accepts a wholistic stance and therefore KTC is similar to it. Notably for this aproach, wholes are not assembled from parts, but perceptual particulars arise from the modification of wholes.

In my opinion, the predictive coding conceptualization (e.g., Clark, 2015; Friston, 2005; Heeger, 2017; Hohwy, 2013; O'Callaghan, Kveraga, Shine, Adams, & Bar, 2017) is where a common ground for the KTC and the microgenetic approach can be found. We could accept that the initially intrinsic wholes, waiting to be unfolded along actual percept genesis (both along the phylogenetic and ontogenetix axes) act as priors, with prediction errors as the main mechanism for implementing the effect of external constraints. Adherents to KTC could accept that in the realiziation of ICP within the system, the whole structure with largest Phi Max computed value has the role of a continuously changing prior. This prior determines which zombie subsystems will be brought in to the succeeding instances of the conscious whole as its structural parts.

According to the KTC the content of consciousness is continually changing and the timescale of NCC mimics the rise and fall of percepts in the temporal dynamics of experience (Koch, 2019, p. 43, p. 196). It takes some 150–500 ms (ms) for the conscious percept to emerge (reviews: Bachmann, 2000; Rutiku & Bachmann, 2017). Relatively early ERP and MEG markers of conscious perception (with a 150–200 ms post-stimulus delay) as well as late ones (300–1000 ms delay) indicate that the hypothetical wholistic structure that stands for perception of one definite specific experience (with its unique content) must take some time to form (e.g., Derda et al., 2019; Eklund & Wiens, 2018; Förster, Koivisto, & Revonsuo, 2020; Haque, Lobier, Palva, & Palva, 2020; Jimenez, Grassini, Montoro, Luna, & Koivisto, 2018; Mai, Grootswagers, & Carlson, 2019; Rutiku, Aru, & Bachmann, 2016). Koch also appears to accept that

conscious percepts, i.e. irreducible wholes that stand for conscious phenomenal experience emerge rather early in the vicinity of about 200 ms.

The KTC bears easily noticeable evolutionary flavor associated with the adherence to the principles of continuity and intrinsic causal power (e.g., Koch, 2019, pages 26–28, 122–124, 137, 202–203). For example, when asking why there are so many types of nervous system cells, Koch says that this is not for providing variety and efficiency of computational functions, but because of the evolutionary, developmental and metabolic constraints. Neurobiological constraints gradually evolving along the evolutionary ladder of brain development set the constraints for gradually unfolding stages of subjective perception. Brains evolved in evolution, providing by default the physico-chemical internal conditions that allow the same system state change from unconscious to conscious. (Let me suggest an overly simplified analogy: similarly to how the same volume of liquid, depending on conditions, changes from non-boiling state to boiling state and, concomitantly, the liquid state allows a surplus new state – gas, the unconscious state of the brain grows over to the conscious state of the brain.)

In the mental time there is an overlapping presence of past, present and expected future subjective contents (Brown, 2013, 2015, 2017; Koch, 2019). The mind/brain state of an alert person is not a discrete unit of isolated content, but a traversal from unconscious depth to conscious surface, characterized by gradual fading of the former final objectification, superimposed by newly emerging content. In successive states present at time-1 and time-2, the arising of time-2 content begins before the completion of time-1 content. This principle also makes it controversial to distinguish perception from memory. (See also Aru & Bachmann, 2017). Information integration takes place not only in space, but also in time.

## 2. Questions to be asked for further elaboration of the KTC

Even though the KTC is quite thoroughly elaborated and related to valid empirical research, there are many open questions and speculative standpoints in it. For the sake of conciseness, let me ask them, point by point.

(1) Acceptance of the possibility of contentless subjective experience (including the purported examples from meditation and mystics; see also Winter et al., 2019, for this doubtful stance) is questionable (e.g., Koch, p. 114). It is more likely that in most cases the criterion contents of report apply to some *seemingly* empty *aspect* of total experience in the context of other phenomenal feelings present in (the) other aspect(s) of the whole. Moreover, feeling of emptiness is also a feeling; blankness of visual experience comes at the background of bodily feelings or feelings of slight strain in the eyes; affective indifference to the (i.e., the “null affect”), and stillness of the “frozen” external world are nevertheless subjective qualities. “Cessation of all mental content” in the conscious state is extremely unlikely. If this emptiness applies to feelings of personal attitudes, me-ness, etc, it could be possible. However, if it applies to the whole of sensory experiences, it is not likely. Even if a meditator, due to an extremely well fixed and prolonged sensory adaptation reaches the state of fading of the content to which the perceptual system was adapted, the feeling that he is alive and present in this world – even having a fully indifferent attitude towards his existence or sensations – will not have disappeared. Maybe in terms of Block (1995) distinction the contentless subjective experience might be ascribed to nullification of access consciousness in some specific states, but not to the phenomenal consciousness.

(2) The “critical distinction” between states of consciousness and conscious states (Koch, 2019, p. 39, 43, etc) may be more nyanced than presented in the KTC. First, both parts of the distinction refer to subjective, phenomenal reality and therefore inevitably require description in terms of the KTC-axiomatic first-person perspective. Behavioral and physiological correlates of consciousness sates are just indirect markers lacking any phenomenal quality themselves. Therefore, consciousness states inevitably are phenomenal states. At the axiomatic subjective level they are expressed by psychophysically measurable clarity, level, distinctness, fragmentariness, fullness, vagueness, subjective stability etc subjective qualities (Bachmann, 2011, 2012). Conscious states are defined and measured by contents of states of consciousness. So there is some tautology between states and contents unless we say that by consciousness states we mean states as measured by the objective correlates of the hypothetical or postulated (or abducted) states of consciousness. State as estimated by objective correlates *characterizes* perceiving *subjects*, whereas state as experience is *felt* by perceiving subjects. Second, there are neurobiological reasons to believe that in terms of brain mechanisms state and content are mediated by the same system of subsystems (Aru, Suzuki, Rutiku, Larkum, & Bachmann, 2019; Bachmann & Hudetz, 2014; Bachmann, 1994; Phillips, Bachmann, & Storm, 2018). (As a side-note: Koch, 2019, seems to present KTC as a cortico-centric theory attributing the critical brain areas to the “hot zone” in the back of cortex. Massive evidence about the intact thalamocortical recurrent processing associated with reticular formation modulations as a necessary condition for consciousness speaks against corticocentrism. Moreover, non-zero Phi Max is likely with decorticated subjects. Consider also Björn Merker (2007, 2008) well known arguments about consciousness in hydrocephalous subjects. On page 55 Koch (2019) deprives reticular formation from its possible role as an ingredient of NCC. This position would be convincing if by microelectrode or optogenetic stimulation of some sensory cortical area one would evoke conscious sensation without any associated change in brainstem activity. No such experimental result is known.) Third, we cannot call states of being anesthetized, fainted or being in deep sleep conscious states because, axiomatically, there are no subjective feelings present. These states are unconscious states. For a conscious content of an experimental stimulus to be experienced, all necessary neurobiological conditions that altogether make up the sufficiency condition must be satisfied (Bachmann, 1984). Yet, subjectively, the consciousness state and state of consciousness are amalgamated.

(3) It is questionable that the results of fusiform face area (FFA) stimulation causing experience of faces allow limit the NCC for seeing faces only in this temporal area. FFA activity may not be minimally sufficient. Recurrence with thalamus and possibly with other subsystems such as cortical V2 and/or V1 circuits is a more likely NCC candidate.

(4) Koch accepts that visual, somatosensory and other subjective experiences have spatial structure, extension in space and they also show some granularity. Topography is what characterizes phenomenal experience (e.g., Koch, 2019, p. 64). On the other hand, Koch postulates that primary sensory cortices are not seats for NCC. There is some controversy in these views. First, primary sensory areas in the back of cortex have the best cortical spatial resolution, which makes these areas the best candidates for wide and at the same time visually fine-tuned subjective representation. Second, phenomenal perceptual experience includes not only generalized or superordinate formal categories, but also low-level features. Transcranial magnetic stimulation of higher areas capable of phosphenes induces top-down volleys to more primary areas (Marzi, Mancini, & Savazzi, 2009; Pascual-Leone & Walsh, 2001; Silvanto, Cowey, Lavie, & Walsh, 2005; Webster & Ro, 2017). We also know that, from the first-person perspective, the wholistic experience in space and time appears seamless, smooth, without any gaps in the sensed weft of experience. In the KTC conceptualization, the structured wholes have well expressed granularity, which means lack of seamlessness. Although it is not difficult to imagine how a formal multilevel structure where higher level nodes dissolve the granularity present in the lower level multiplicity of units by taking in the converging inputs at the "cardinal cell" level, Koch himself, and justifiably so, rejects the single-cell, gnostic units paradigm. Yet the way how granularity makes peace with seamlessness would need a more specific theoretical treatment.

(5) There is one tough question posed, of which Koch is well aware (2019, p. 76). *Expressis verbis*: "Even if everything about IIT is correct, why should it feel like anything to have a maximum of integrated information?". Koch's answer on the same page is not very convincing. He asserts that the five axiomatic properties of consciousness, by construction, fully delimit any experience. "Nothing else is left out." This begs a question: how do you *know*? Moreover, there may be other necessary axiomatic attributes or preconditions characteristic to conscious experience left out from the Koch's list. For example, for a putative sixths axiom one can suggest the microgenetic criterion: given experimental "litmus test" conditions for introspection, phenomenal perceptual experience of an object or an event has to develop through qualitatively different, maturing subjective-experience stages with systematically increasing structural-informational capacity. Otherwise there will be no conscious perception because conscious perception can not emerge only as its own end-state. Or that some sixth (or seventh) axiomatic characteristic would require that subjective experience is felt to be *projected* to inner or outer *environment* (the intentionality criterion). Or that a felt "me-ness" (subjective perspectivalness) would be another axiomatic characteristic. This list of additions could be even longer, especially if in addition to the subjective-experiential axiomatic characteristics some other ones revealed by future research will be added. For instance, this may happen if it turns out that only the neural units assembled according to the blueprint of genetic/evolutionary "instructions" (which are created in phylogenesis and ontogenesis) are the units capable of providing the capacity of subjective feelings to the structure of the whole.

(6) One of the critical experimental tests for KTC that has to be done, but is still lacking is this: (i) design about ten experiments for studying NCC according to the contrastive methodology which are based on the combination of psychophysical and brain-imaging techniques; (ii) run these experiments and seed out trials with target-stimulus consciousness versus trials with target stimulus consciousness absent; (iii) measure Phi Max for each of these two groups; (iv) find that in all of these about ten designs Phi Max value is non-zero for trials from conscious experience group, but zero for all trials in no-experience group. My intuition says that the possibility to falsify KTC by this approach is quite high. The main problem is that we do not know what exactly, in terms of the contents of consciousness, the Phi Max refers to. Maybe for the consciousness *state* measurement the Phi Max has merit, but for the measurement of whether *contents* of certain stimulation reach phenomenally conscious level the Phi Max measure may offer unreliable results. In several places in Koch (2019) the author himself speaks about many varying background conditions that contaminate signals to be recorded and measured.

(7) The examples on evolving integrated brains (Koch, 2019, p. 123) are related to the consciousness issue only indirectly, by a postulated analogy. Why these examples are not examples of evolution of clever zombies, we can not know.

(8) Perhaps there is one most interesting notion that comes with KTC – the notion about the quantity of consciousness. What does it mean to have *more or less consciousness* (with different non-zero levels of Phi Max)? Maybe it is just here where the most interesting potential of the discussion on KTC lies. I have already noted the need for more differentiated and specific use of the psychophysical/introspective scales for measurement of conscious experience (Bachmann, 2012). More or less consciousness – for what? Is it an underspecified, ambiguous scale for "vividness" or "clarity" (e.g., PAS scales)? Or would it be better to design a battery of more focused and contentfully more specific subjective scales such as "intensity", "contrast", "spatial extent", "variety", "granularity", "numerosity", "coarseness", "jaggedness", "smoothness/softness", "element numerosity", "fragmentariness/unity", "felt meaning", "intuitive confidence", "metacognitive confidence" etc etc. Koch's notion of "more distinctions and relations" (Koch, 2019, p. 126) seems insufficient. Importantly, it is quite likely that measurements by some different psychophysical subjective scales will not be mutually consistent or monotonically varying. According to one scale there is "more consciousness", but according to the other scale, for the same stimulation, there is "less consciousness". From the consciousness research point of view especially the scales that are universal for different or for all modalities must be valuable because consciousness capacity is super- and intermodal, essentially panmodal. Here we get back to the potential advantage of the Phi Max measures because they are abstract and, therefore, if not universal, then at least better generalizable across studies. But they must be linked better to the specific contents of conscious percepts. On the other hand, the need for more differentiated basic concepts for measurement of the phenomenal states of consciousness is echoed well in how the analogous calls for better conceptual distinctions have helped advance consciousness studies (e.g., Block, 1995), psychophysical measurement (e.g., Rahnev & Denison, 2018) or promise to help advance assessment of attention (e.g., Luo & Maunsell, 2019). Relating the Phi Max quantification measures as it is presented in KTC to measurements carried out with different content-wise focused and more specific scales might be a useful aim. Some first steps toward the contentful specification of the IIT-computational measures of

the experimental results have been taken (e.g., Haun et al., 2017).

(9) On page 67 (Koch, 2019) we read that we "… might even have to hunt for the NCC down at the subcellular level, seeking mechanisms operating inside cells rather than across large neural coalitions, as is widely assumed". This is clearly a view consistent with the non-computational stance and a theoretical position to be eagerly agreed with by the adherents to the subcellular consciousness mechanisms perspective (e.g., Larkum, 2013; Meyer, 2015; Bachmann, 2015; Phillips et al., 2018; Takahashi, Oertner, Hegemann, & Larkum, 2016). Yet this view calls for at least two comments. First, while Koch appears to link foundational NCC processes to spiking neurons, it is equally possible that membrane potentials of the neurons that do not send out axonal action potentials contribute no less. The same applies to the dendritic spikes. Note that conscious experience is characterised by spatiotemporal smoothness, but axonal spiking manifests quantally, by the all-or-none rule whereas membrane potentials involve gradual changes. Second, while computations carried out by networks of discretely firing neurons can be simulated and implemented by traditional artificial neural nets, single-cell level processes with their more intimate association with genomic and proteomic processes and with their indiscrete activity dynamics can be also modelled by certain pertinent computational mechanisms (e.g., biocomputing with analog signal capacity). Whether this is a standard computing anymore is a question. Of course, the other adherents to IIT besides Koch may have a bit more specific view on the levels where IIT mechanisms are implicated (e.g., Honjoh et al., 2018). Moreover, in recent views on multilevel consciousness mechanisms spiking of neurons is by no means considered as the only or decisive mechanism (e.g., Boly et al., 2013).

(10) Now, consider a very simple organism with a non-zero Phi Max (i.e., a conscious animat) and compare it (her?) to a sleeping human who is in an unconscious state. My intuition says that non-zero Phi Max computed value must be higher for the sleeping human brain than for that small system. If I am wrong, my other intuition says that for KTC to be correct there must be a dedicated mechanism in the brain that from time-to time counteracts meaningful information integration and actively dumps down the integrated whole. (Some precisely tuned counterphase mechanism?) This also means that the standard state of the consciousness-possessing organisms is alert wakefulness with accompanying consciousness with its content and that sleeping is the way to deprive us of that wonderful gift of nature, the feeling of life itself.

## Acknowledgments

Development of the ideas expressed in this paper were partly supported by the scholarly conditions made available through the Estonian Science Agency institutional support grant IUT20-40/TSVPH14140I.

## References

Albantakis, L., Hintze, A., Koch, C., Adami, C., & Tononi, G. (2014). Evolution of integrated causal structures in animats exposed to environments of increasing complexity. *PLoS Computational Biology, 10*(12), e1003966.

Aru, J., & Bachmann, T. (2017). In and out of consciousness: How does conscious processing (d)evolve over time? *Frontiers in Psychology, 8*, 128. https://doi.org/10.3389/fpsyg.2017.00128.

Aru, J., Bachmann, T., Singer, W., & Melloni, L. (2012). Distilling the neural correlates of consciousness. *Neuroscience and Biobehavioral Reviews, 36*(2), 737–746.

Aru, J., Suzuki, M., Rutiku, R., Larkum, M. E., & Bachmann, T. (2019). Coupling the state and contents of consciousness. *Frontiers in Systems Neuroscience, 13*, 43. https://doi.org/10.3389/fnsys.2019.00043 (1-9).

Aru, J., Tulver, K., & Bachmann, T. (2018). It's all in your head: Expectations create illusory perception in a dual-task setup. *Consciousness and Cognition, 65*, 197–208.

Bachmann, T. (1984). The process of perceptual retouch: Nonspecific afferent activation dynamics in explaining visual masking. *Perception and Psychophysics, 35*(1), 69–84.

Bachmann, T. (1994). *Psychophysiology of visual masking. The fine structure of conscious experience.* Commack, NY: Nova Science Publishers298.

Bachmann, T. (2000). *Microgenetic approach to the conscious mind.* Amsterdam/Philadelphia: John Benjamins.

Bachmann, T. (2009). Finding ERP-signatures of target awareness: Puzzle persists because of experimental co-variation of the objective and subjective variables. *Consciousness and Cognition, 18*(3), 804–808.

Bachmann, T. (2011). Attention as a process of selection, perception as a process of representation, and phenomenal experience as the resulting process of perception being modulated by a dedicated consciousness mechanism. *Frontiers in Psychology, 2*, 387. https://doi.org/10.3389/fpsyg.2011.00387.

Bachmann, T. (2012). How to begin to overcome the ambiguity present in differentiation between contents and levels of consciousness? *Frontiers in Psychology, 3*, 82. https://doi.org/10.3389/fpsyg.2012.00082.

Bachmann, T. (2015). How a (sub)cellular coincidence detection mechanism featuring layer-5 pyramidal cells may help produce various visual phenomena. *Frontiers in Psychology, 6*, 1947. https://doi.org/10.3389/fpsyg.2015.01947.

Bachmann, T., & Hudetz, A. G. (2014). It is time to combine the two main traditions in the research on the neural correlates of consciousness: C = L × D. *Frontiers in Psychology, 5*, 940. https://doi.org/10.3389/fpsyg.2014.00940.

Block, N. (1995). On a confusion about a function of consciousness. *Behavioral and Brain Sciences, 18*(2), 227–247.

Block, N. (2019). What is wrong with the no-report paradigm and how to fix it. *Trends in Cognitive Sciences, 23*(12), 1003–1013.

Boly, M., Seth, A. K., Wilke, M., Ingmundson, P., Baars, B., Laureys, S., … Tsuchiya, N. (2013). Consciousness in humans and non-human animals: Recent advances and future directions. *Frontiers in Psychology, 4*, 625. https://doi.org/10.3389/fpsyg.2013.00625.

Brown, J. W. (2013). *The life of the mind.* Psychology Press.

Brown, J. W. (2015). *Microgenetic theory and process thought.* Imprint Academic.

Brown, J. W. (2017). Microgenetic theory of perception, memory, and the mental state: a brief review. *Journal of Consciousness Studies, 24*(11–12), 51–70.

Brown, R., Lau, H., & Ledoux, J. (2019). Understanding the higher-order approach to consciousness. *Trends in Cognitive Sciences, 23*(9), 754–768.

Casali, A. G., Gosseries, O., Rosanova, M., Boly, M., Sarasso, S., Casali, K. R., … Massimini, M. (2013). A theoretically based index of consciousness independent of sensory processing and behavior. *Science Translational Medicine, 5*(198) 198ra105-198ra105.

Chalmers, D. J. (1996). *The conscious mind: In search of a fundamental theory.* Oxford University Press.

Clark, A. (2015). *Surfing uncertainty: Prediction, action, and the embodied mind.* Oxford University Press.

Crick, F., & Koch, C. (1990). Towards a neurobiological theory of consciousness. *Seminars in Neuroscience, 2*, 263–275.

Crick, F., & Koch, C. (2003). A framework for consciousness. *Nature Neuroscience, 6*, 119–126.

de Lange, F. P., Heilbron, M., & Kok, P. (2018). How do expectations shape perception? *Trends in Cognitive Sciences, 22*(9), 764–779.

Dehaene, S., & Changeux, J. P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron, 70*(2), 200–227.

Derda, M., Koculak, M., Windey, B., Gociewicz, K., Wierzchoń, M., Cleeremans, A., & Binder, M. (2019). The role of levels of processing in disentangling the ERP

signatures of conscious visual processing. *Consciousness and Cognition, 73*, 102767.
Eklund, R., & Wiens, S. (2018). Visual awareness negativity is an early neural correlate of awareness: A preregistered study with two Gabor sizes. *Cognitive, Affective, & Behavioral Neuroscience, 18*(1), 176–188.
Förster, J., Koivisto, M., & Revonsuo, A. (2020). ERP and MEG correlates of visual consciousness: The second decade. *Consciousness and Cognition, 80*(102917), 1–24.
Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences, 360*(1456), 815–836.
Haque, H., Lobier, M., Palva, J. M., & Palva, S. (2020). Neuronal correlates of full and partial visual conscious perception. *Consciousness and Cognition, 78*, 102863.
Haun, A. M., Oizumi, M., Kovach, C. K., Kawasaki, H., Oya, H., Howard, M. A., ... Tsuchiya, N. (2017). Conscious perception as integrated information patterns in human electrocorticography. *eNeuro, 4*(5), https://doi.org/10.1523/ENEURO.0085-17.2017 ENEURO.0085-17.2017.
Heeger, D. J. (2017). Theory of cortical function. *Proceedings of the National Academy of Science U S A. 114*(8), 1773–1782.
Hohwy, J. (2013). *The predictive mind.* Oxford University Press.
Honjoh, S., Sasai, S., Schiereck, S. S., Nagai, H., Tononi, G., & Cirelli, C. (2018). Regulation of cortical activity and arousal by the matrix cells of the ventromedial thalamic nucleus. *Nature Communications, 9*, 2100. https://doi.org/10.1038/s41467-018-04497-x.
Jimenez, M., Grassini, S., Montoro, P. R., Luna, D., & Koivisto, M. (2018). Neural correlates of visual awareness at stimulus low vs. high-levels of processing. *Neuropsychologia, 121*, 144–152.
Koch, C. (2004). *The quest for consciousness.* Englewood: Roberts & Co.
Koch, C. (2019). *The feeling of life itself. Why consciousness is widespread but can't be computed.* Cambridge, MA: The MIT Press.
Koch, C., Massimini, M., Boly, M., & Tononi, G. (2016). Neural correlates of consciousness: Progress and problems. *Nature Reviews Neuroscience, 17*(5), 307–321.
Larkum, M. (2013). A cellular mechanism for cortical associations: An organizing principle for the cerebral cortex. *Trends in Neurosciences, 36*(3), 141–151.
Luo, T. Z., & Maunsell, J. H. (2019). Attention can be subdivided into neurobiological components corresponding to distinct behavioral effects. *Proceedings of the National Academy of Sciences, 116*(52), 26187–26194.
Mai, A. T., Grootswagers, T., & Carlson, T. A. (2019). In search of consciousness: Examining the temporal dynamics of conscious visual perception using MEG time-series data. *Neuropsychologia, 129*, 310–317.
Marzi, C. A., Mancini, F., & Savazzi, S. (2009). Interhemispheric transfer of phosphenes generated by occipital versus parietal transcranial magnetic stimulation. *Experimental Brain Research, 192*(3), 431–441.
Massimini, M., Ferrarelli, F., Huber, R., Esser, S. K., Singh, H., & Tononi, G. (2005). Breakdown of cortical effective connectivity during sleep. *Science, 309*(5744), 2228–2232.
Massimini, M., & Tononi, G. (2018). *Sizing up Consciousness: Towards an objective measure of the capacity for experience.* Oxford University Press.
Merker, B. (2007). Consciousness without a cerebral cortex. Behavioral and Brain Sciences, 30, 63-81. Reprinted with permission in H. Liljenström, P. Århem (Eds.) (2008), Consciousness Transitions. Phylogenetic, Ontogenetic, and Physiological Aspects, Chapter 9 (pp. 193–230), Amsterdam: Elsevier. ())
Meyer, K. (2015). The role of dendritic signaling in the anesthetic suppression of consciousness. *Anesthesiology, 122*, 1415–1431. https://doi.org/10.1097/ALN.0000000000000673.
Michel, M., & Morales, J. (2019). Minority reports: Consciousness and the prefrontal cortex. *Mind & Language.* 1–21.
O'Callaghan, C., Kveraga, K., Shine, J. M., Adams, R. B., Jr, & Bar, M. (2017). Predictions penetrate perception: Converging insights from brain, behaviour and disorder. *Consciousness and Cognition, 47*, 63–74.
Partos, T. R., Cropper, S. J., & Rawlings, D. (2016). You don't see what I See: Individual differences in the perception of meaning from visual stimuli. *PLoS ONE, 11*(3), e0150615. https://doi.org/10.1371/journal.pone.0150615.
Pascual-Leone, A., & Walsh, V. (2001). Fast backprojections from the motion to the primary visual area necessary for visual awareness. *Science, 292*, 510–512.
Phillips, W. A., Bachmann, T., & Storm, J. F. (2018). Apical function in neocortical pyramidal cells: A common pathway by which general anesthetics can affect mental state. *Frontiers in Neural Circuits, 12*(50), 1–15. https://doi.org/10.3389/fncir.2018.00050.
Pitts, M. A., Metzler, S., & Hillyard, S. A. (2014). Isolating neural correlates of conscious perception from neural correlates of reporting one's perception. *Frontiers in Psychology, 5*, 1078. https://doi.org/10.3389/fpsyg.2014.01078.
Powers, A. R., Mathys, C., & Corlett, P. R. (2017). Pavlovian conditioning–induced hallucinations result from overweighting of perceptual priors. *Science, 357*(6351), 596–600.
Rahnev, D., & Denison, R. N. (2018). Suboptimality in perceptual decision making. *Behavioral and Brain Sciences, 41*(e223), 1–107.
Rutiku, R., Aru, J., & Bachmann, T. (2016). General markers of conscious visual perception and their timing. *Frontiers in Human Neuroscience, 10*, 23. https://doi.org/10.3389/fnhum.2016.00023 (1-15).
Rutiku, R., & Bachmann, T. (2017). Juxtaposing the real-time unfolding of subjective experience and ERP neuromarker dynamics. *Consciousness and Cognition, 54*, 3–19.
Schmack, K., Rothkirch, M., Priller, J., & Sterzer, P. (2017). Enhanced predictive signalling in schizophrenia. *Human Brain Mapping, 38*(4), 1767–1779.
Silvanto, J., Cowey, A., Lavie, N., & Walsh, V. (2005). Striate cortex (V1) activity gates awareness of motion. *Nature Neuroscience, 8*, 143–144.
Takahashi, N., Oertner, T. G., Hegemann, P., & Larkum, M. E. (2016). Active cortical dendrites modulate perception. *Science, 354*, 1587–1590. https://doi.org/10.1126/science.aah6066.
Tononi, G. (2012). The integrated information theory of consciousness: An updated account. *Archives italiennes de biologie, 150*(2/3), 56–90.
Tsuchiya, N., Wilke, M., Frässle, S., & Lamme, V. A. F. (2015). No-report paradigms: Extracting the true neural correlates of consciousness. *Trends in Cognitive Sciences, 19*, 757–770. https://doi.org/10.1016/j.tics.2015.10.002.
Webster, K., & Ro, T. (2017). Retinal and visual cortex distance from transcranial magnetic stimulation of the vertex affects phosphene perception. *Experimental Brain Research, 235*(9), 2857–2866.
Winter, U., LeVan, P., Borghardt, T. L., Akin, B., Wittmann, M., Leyens, Y., & Schmidt, S. (2019). Content-free awareness: EEG-fcMRI correlates of consciousness as such in an expert meditator. *Frontiers in Psychology, 10*.