

# Christof Koch (2019): *The Feeling of Life Itself: Why consciousness is widespread but can't be computed*

Cambridge: The MIT Press.  
280 pp. ISBN 978-0262042819.

Acta Fakulty filozofické  
Západočeské univerzity v Plzni /  
West Bohemian Review of  
Social Sciences & Humanities  
2022, Vol. 14 (2), 26–30  
DOI: <https://doi.org/10.24132/actaff.2022.14.2.3>  
<https://actaff.zcu.cz/>  
ISSN 2336-6346



This paper is licensed under Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

Christof Koch is a Chief Scientist of the Mind-Scope Program at the Allen Institute for Brain Science and a Professor at the California Institute of Technology. His work includes the field of computational neuroscience and the neural basis of consciousness. His book, published in 2019, *The feeling of life itself: Why consciousness is widespread but can't be computed*, presents a novel argument on consciousness and the importance of a quantitative theory known as Integrated Information Theory. He argues that consciousness is a unique feature instilled in living beings and cannot be programmed in computers. He proves this statement with the help of prominent theories, i.e., Neural Correlates of Consciousness and Integrated Information theory. Koch's primary aim is to build a clinically valuable consciousness meter to help us interpret the true nature of being. The book engages us with a scientific and philosophical discussion of the mind and brain, which gives us a holistic view of consciousness. Moreover, he crafts his book by highlighting the features of IIT, which predicts the possibility that all cellular life feels like something.

In *The feeling of life itself*, Christof Koch presents interesting discussions on consciousness and his elaborate scientific practices. His primary aim in this book is to implement a quantitative theory that can infer and predict the existence of experience. Koch's work engages in the core discussion of Neural Correlates of Consciousness (NCC), Integrated Infor-

mation Theory (IIT), and consciousness. The focus of his argument is to prove that experiences and feelings cannot be computed. He attempts to bridge the gap between scientific and humanistic tendencies. One of his favoured theories, initially put forward by Giulio Tononi and subsequently supported by Koch, is Integrated Information Theory, which is the beating heart of this book. It outlines the theory as quantitative and empirically consistent, starting with experience and proceeding with neuronal mechanisms. The book comprises 14 chapters, each with a valuable discussion on consciousness, brain mechanism, and computation.

Chapters 1, 2, 5, 6, and 7 address the questions based on brain and consciousness; and, simultaneously, his effort to answer them with scientific evidence. Here, the focus is on understanding experience, what facilitates the feeling of conscious experience, and the limitations of this feeling in humans. Chapters 3, 4, and 8 introduce a mathematical understanding of the terms related to the brain and consciousness. Lastly, chapters 9–14 consist of rich scientific inquiry data, i.e., can we comprehend consciousness mathematically, or can it be computed? These chapters start with tools involved in understanding consciousness, the rich concept of IIT, and the global neuronal workspace. In the end, these concepts are narrowed down to understand the philosophical existence of consciousness in the natural environment.

Chapters 1 and 2 primarily aim to interpret how the brain gives rise to subjective experience. This aim eventually turns out to be the book's primary purpose, i.e., each chapter contributes to proving that consciousness is experience. This is solved with the help of scientific, mathematical, and neuroscientific resources. According to Koch, there are five fundamental axioms/properties to understand the concept – consciousness is experience. These five properties will be reflected and explained further in chapter 8, where he explains IIT. These fundamental axioms are the following: 1. Denying one's experience – This property focuses on the aspect of inner experience as believed by Koch but is denied by other Philosophers such as Paul and Patricia Churchland. Again, the idea of inner experience highlights his core belief in consciousness as an integral part of the experience. 2. The Challenge of Defining Consciousness as Experience – It is connected to the former property that consciousness is difficult to define or explain to another individual, making it specific to oneself. 3. Any experience is structured – Here, Koch explains that any experience has distinctions and is structured and composed of many internal phenomenal or subjective distinctions. 4. Any Experience is Informative, Integrated, and Definite – Here, it can be interpreted as an experience rich in information and cannot be generic. Each experience is unitary and integrated, including phenomenal distinctions and relations within that experience. 5. Any experience has a point of view and occurs in time – This property focuses on the fact that each experience occurs from a different perspective and in its own time. Therefore, these are the essential properties that conscious experiences exhibit.

Chapters 5, 6, and 7 builds a base for understanding the coordinates of the brain, emergence, and the need to understand consciousness theory. Chapter 5 aims to acknowledge that the brain is intimately tied to the sensation and movement of the human body. This is done with the help of supporting claims from historical literature by Socrates and Aristotle. Simultaneously after introducing the base of history,

Koch shifts his focus to the modern discussion of the brain and experience. He admits that the discussion about consciousness is complex and must first be approached from technical points. Primarily, he explains the distinction between *conscious states* and *states of consciousness*. The former one, i.e., *conscious states*, happens when we are awake from sleep and ready in the physiological and psychological sense to respond to the sound, sight, and touch.

On the other hand, the latter, i.e., *states of consciousness*, can be understood as the process between the conscious, unconscious, and dream levels. When we sleep, consciousness fades. According to Koch, these three distinct states are reflected differently when the electrical brain activities are recorded in an electroencephalogram (EEG). While recording a human brain through EEG, these are patterns observed: Awake, drowsy (alpha waves), REM, and deep sleep (delta waves) stages. Thus, capturing the stages and technical differences between each stage of sleep or consciousness. In addition, he highlights the importance of the topic – Neural correlates of consciousness (NCC). NCC has been defined as a minimal neural mechanism that is necessary and sufficient for experiencing any conscious percept (Tononi, Koch 2015: 2). “*The brain as a whole can be considered an NCC; the brain generates experience, day in and day out. But Crick and I were after the synapses, neurons, and circuits constitutive of experience*” (p. 49). In this context, one of the essential questions we encounter is the significance of NCC. NCC primarily works as a process necessary for capturing experience, found in the broad regions of the temporo-parietal-occipital region of the posterior and prefrontal cortex, and cerebellum. The evidence of these recordings can be found again in EEG and fMRI. “*Activating the NCC by such means should trigger the associated percept while suppressing the NCC should prevent the experience*” (p. 50). Both these possibilities can alter the character of experience. In essence, I believe NCC acknowledges techniques of understanding experience and provides further information about consciousness in patients with

severe brain damage. This is certainly a step forward to open the door for research expectations and success in new findings related to consciousness.

Chapter 6 aims to identify other brain regions that are closely tied to understanding consciousness, i.e., the role played by the central nervous system. Neuroscience hypothesizes that the brainstem, cerebellum, and cerebral cortex are the prime areas engaged in enabling consciousness. Each section is supported by studies anticipating what would happen if specific nerves or parts of the brain were operated on to activate consciousness. Furthermore, if specific processes in the brain are suppressed, will the patient lose the sense of experience? The possibilities are endless, but it is only with the help of case studies we have come to realize that the inactivity or damage of one part cannot destroy a patient's consciousness.

Chapter 7 answers the 'why' – i.e., why we need a theory of consciousness. Koch implements IIT as a link to understanding mental and physical states. "*IIT is a fundamental theory that links ontology and phenomenology to the realm of physics and biology. In short, it starts with experience and asks how matter must be organized to support the mental*" (p. 74). This is a concise statement of the problem, i.e., how can we explain consciousness in ourselves and animals? The theory takes five phenomenological postulates of experience to prove that consciousness does exist. They are – intrinsic existence, composition, information, integration, and exclusion. Koch states confidently that despite IIT covering in-depth mathematical equations, it correctly describes aspects of systems that support consciousness. He adds, "*by construction, these five properties fully delimit any experience. What people mean by subjective feelings is precisely described by these five axioms*" (p. 76).

Fundamentally, the axioms mentioned in chapters 1 and 2 are intended to capture the essential aspects of every conscious experience. On the other hand, the properties required of a conscious physical substrate are called the postulates. Therefore, we can understand that ax-

ioms entail the role of experience while postulates entail the role of proving the physical existence of consciousness. "*Each axiom has an associated postulate, a bridging principle that the system under consideration has to obey*" (p. 75).

According to IIT, consciousness is determined by the fundamental property of any mechanism that has cause-effect power upon itself. The theory unfolds the intrinsic causal powers of any system which obeys all five postulates. The causal power and whole in this section play an essential role. The whole can be understood as any physical system with intrinsic causal powers acting on itself and being irreducible and conscious. IIT's core claim is that a system is conscious if it possesses a property called Phi, which measures the system's integrated information. Phi can be a property of an entity, biological or non-biological. Koch explains this concept as a mathematical measure called integrated information. This conceptual structure specifies the quantity and quality of experience. It is measured by phi max value, i.e., if a system has larger phi max, it can lay claim to existing in a fuller sense than systems with lower phi max. "*The experience is identical to the cause-effect structure – not to its physical substrate, it's whole just as my experience of feeling kind of blue is not identical to the gray goo in my head that is the physical substrate of this experience*" (p. 87).

This is supported by a physical explanation of the brain. "*The cerebellum, which seems to have less internal connectivity, and hence lower Phi, than other neural regions, can be damaged without significantly affecting consciousness. On the other hand, brain scans of paralyzed or uncommunicative patients reveal higher Phi showing signs of being conscious*" (Horgan 2015). Here, it can be interpreted with the help of an example. For instance, if I am watching a movie or a drama featuring a famous actress, the central complex in my brain is made up of neurons within certain parts of the cerebral cortex. Thereby, every neuron shapes or predicts its probability of past states, i.e., the cause and effects of the complex, depending on how it is connected to the other neurons. The neu-

rons firing in the visual area correspond with past events, thus recognizing the actress's face. Even though Phi has worked towards the relation between conscious and unconscious factors, we could understand it with the help of such examples.

The explanation of Phi, Whole, and IIT has been intermittently covered in chapters 7, 8 and 10. It is vital to understand the technical section, i.e., discussion on the mathematical measure of consciousness, answering whether consciousness holds a function, and, if so, how IIT aids in proving the same.

The above-explained chapters explicitly work towards interpreting the core analogy of IIT and consciousness. Fundamentally, chapter 9 and 11–14 engages with explanation, case studies, and intense research on IIT, coordination of varied concepts with consciousness, questioning machine consciousness, and future progress with consciousness studies. With this understanding, I want to further shift my focus on Koch's discussion of IIT and consciousness, which is crucial in this book.

Starting with the explanation and focus on IIT, intelligence, and consciousness, Koch emphasizes that consciousness can be interpreted only in humans since this aspect makes a human recognized as intelligent. He argues that consciousness is proportional to the degree of integrated information in a system and that humans are more conscious than simpler mammals, bugs, or single-celled organisms because they have more integrated brains. "*Consciousness is not a clever algorithm. Its beating heart is causal power upon itself, not computation. The causal power, the ability to influence oneself or others, cannot be simulated. Not now nor in the future*" (p. 149). Here, he points out the ability and restrictions of humans and computers if induced with the idea of consciousness and experience. Computers or machines are minimally conscious since they can be reduced to individual components without losing their functionality. However, machines do not have the ability or the algorithm to experience or have feelings, and he states that they have no access to integrated information. Primarily, "*consciousness is*

*filled with highly structured percepts and memories of sometimes unbearable intensity*" (p. 121). Koch further explains the function of consciousness and intelligence, highlighting the feature present in humans and that they cannot be comparable to machines. These chapters evaluate an understanding based on the concept of experience, which can be captured with consciousness as a by-product. Similarly, it holds significance when Koch questions the function of consciousness and its ability to be computerized. I believe the answer and content lies in the feeling of experience; and that we have many years open to research and development on the claim whether machines can be generated with this quality.

To conclude, Koch's book contains valuable insights for the future developments of brain science. He attempts to dismiss the narcissistic belief prevalent in our society that animals and plants exist for our benefit. He urges us to embrace the principle of the whole, the moral status of any subject rooted in consciousness and not specifically humanness. This helps us in observing consciousness from the human-centric perspective.

There are a few instances in the book that could have been clearer regarding the definition and articulation. For instance, it was difficult to apprehend mathematical terminologies, figures, and the correlation of the same with consciousness context. It is the technical and formal language required to understand digital computationalism. In chapter 13, it was challenging to understand the term functionalism with context to cause-effect power. Koch states, "*applying IIT to programmable digital computers, we come to an even more startling conclusion that violates strongly held intuition about functionalism. The maximally irreducible cause-effect power of real physical computers is tiny and independent of the software running on the computer*" (Koch 2019: 144). Here, it was complex to understand what he meant by the cause-effect power of the physical computer. Koch could have facilitated better clarity about this statement with the help of a philosophical explanation of cause-effect phenomena so that interpreting it would have been simpler.

In chapter 7, he discusses IIT's five phenomenological postulates. To be precise, he explains axioms in Chapter 2 and postulates in Chapter 7. However, it is challenging to analyze the difference between postulates and axioms. There are terms such as abductions and inference, which he uses while explaining postulates. As a reader, it is not easy to comprehend its base and, more importantly, the need to differentiate them. In my opinion, the axioms and postulates feature similar principles rendering to consciousness.

On the other hand, I liked the setup of chapter 14 concerning ethical and moral consciousness. It was intriguing to see how Koch introduced this book with a particular agenda and desired to end it with an ethical and philosophical discussion based on panpsychism and consciousness in the universe. He urges that evolution explains the makeup of our bodies and the constitution of our minds. This can be modeled with an understanding that experience is found in the most unexpected places, which includes all species. Moreover, the field of brain science is exploding with new experiments and data analysis, which need to be interpreted in the light of various extant theories of consciousness. It eventually gets more interesting to engage with such theoretical conversation and thrive for a scientific outbreak.

Koch vouches for his opinion that IIT offers a principled, coherent, and significant account of the relationship between our physical and mental state. This is grounded in extrinsic and intrinsic causal powers. I found it interesting when he explained that IIT and consciousness relate to causal powers. "*IIT can be thought of as an extension of physics. Textbook physics deals with the interaction of objects with each other, dictated by extrinsic causal powers. My and your experiences are how brains with irreducible intrinsic causal powers feel from the inside*" (p.166).

IIT holds a mathematical feature of identifying a complex system, stating that conscious-

ness has an intrinsic causal power. Any system that has this potential is conscious. The larger the system's integrated information, i.e.,  $\Phi$ , the more conscious system is. IIT, moreover, has encouraged the ideas of pondering and working further on where in the brain does consciousness arise? What about animal consciousness? Etc. This theory has practical consequences, encouraging clinicians or researchers to build a consciousness meter. This device determines speech activity and eye movements for those who are anesthetized or suffered a brain injury.

I highly recommend this book to anyone interested in learning more and engaging in a different perspective of IIT and consciousness. The book certainly gives a rich output on Neuroscience and Philosophy, which is an exciting feature for those who desire to pursue their interest and career in Neuro-Philosophy. I believe our experience of the world, our thoughts, memories, dreams, desires, and fear arise from the brain. The theories and extensive research discussed in this book will help us understand the brain's functioning and how unique the trait of experience and consciousness is to humans. This can assist us in harnessing our brain power in service of a less stressful and more mindful life.

### Bibliography

- Horgan, J. 2015. "Can Integrated Information Theory Explain Consciousness?" [online] *Scientific American* [retrieved 5. 12. 2022]. <https://blogs.scientificamerican.com/cross-check/can-integrated-information-theory-explain-consciousness/>.
- Tononi, G., Ch. Koch. 2015. "Consciousness: Here, there and everywhere." *Philosophical Transactions of the Royal Society B* 370: 1–18, <https://doi.org/10.1098/rstb.2014.0167>.

**Akanksha Nambiar**  
University of West Bohemia  
akanksha@kfi.zcu.cz