ISSN 2976 – 730X
IPI Letters 2023, (1): 78-83
https://doi.org/10.59973/ipil.41

Received: 2023-12-01
Accepted: 2023-12-20
Published: 2023-12-22

IPI Letters

Opinions

Entropic Consciousness: A Proposal

Samuele Manco[1,*]

[1] University of Palermo, Viale delle Scienze, Ed. 15, Palermo, 90128, Italy
[*]Corresponding Author (Email: Saamux97@gmail.com)

Abstract – Tononi's *Integrated Information Theory* (IIT) established that consciousness is integrated information. Starting from IIT, the contribution has explored the nature of conscious experience from a new perspective. The Landauer limit, the MEI (Mass-Energy-Information) equivalence and Friston's "*Free Energy Principle*" were introduced, from which important implications were derived regarding the phenomenon of consciousness and its consequences, the most important of which was the one according to which consciousness could be a state that is in a stable level of energy on the border between order and disorder. The research undertaken aspires to contribute significantly to the field of consciousness science through a new conceptual framework formulation called "*Entropic Consciousness*" (EC). However, some open problems persist in the theoretical landscape of EC.

Keywords – ITT; EC; Consciousness; Entropy; Mass; Energy; Information; MEI principle.

1. Introduction

Starting from the dualism of the philosopher and mathematician René Descartes according to which "*res cogitans*" and "*res extensa*" are two substances ontologically irreducible to each other [1], up to the cognitive and computational neuroscience of the 21st century, numerous theories of consciousness have been developed, but none of them has so far been able to give a complete and exhaustive explanation of the basis of subjective experience [2]. Among these theories, the present work will consider that of the neuroscientist Giulio Tononi: the *Integrated Information Theory* (IIT) [3-7]. We start from the premise that consciousness is not a question of quantity, as the cortex has about 16 billion neurons compared to approximately 70 billion in the cerebellum, and, despite this, the removal of the latter via cerebellectomy doesn't affect consciousness. According to Tononi's theory, all conscious experiences are both informative and integrated, i.e. consciousness "*is*" integrated information, where:

- by "information" we mean "reduction of uncertainty", informatively understood, in that we live *this* experience, and not *that* or *that one*, in the sense that every conscious experience is different from any other "possible" experience, in the past, in the present, in the future: every moment we live *a* precise experience among the infinite possible ones;
- for "integrated", however, we refer to the unity of conscious experience as a coherent whole; in other words, it refers to the way in which experienced information is processed and connected in a unified way.

Tononi's merit was to start from phenomenology rather than neurobiology in asking what consciousness is, identifying its essential properties:

- *Existence*: consciousness exists, it is an undeniable fact of reality;
- *Information*: every experience is informative, it differs from every possible other experience;
- *Composition*: every conscious experience is structured in its elements, in the sense that it consists of multiple aspects in various combinations;
- *Integration*: every experience is unified in its constituents, i.e. the world is experienced in a unitary way;

- *Exclusion*: each experience, with defined boundaries, excludes all the others, having a specific spatio-temporal grain.

From these "axioms", some "postulates" derive, according to which any physical substrate - be it the brain or a computer - which presents these fundamental characteristics manifests some form of consciousness, understood as an intrinsic property of the system (this has profound implications on the life-consciousness equivalence front: the main consequence is that consciousness has nothing to do with the vital functions of the conscious subject): hence the "accusation" raised against Tononi of panpsychism, i.e. of considering consciousness as a fundamental characteristic of the universe as well as mass, energy and/or electric charge. IIT therefore constitutes an "axiomatic" approach to consciousness (*phenomenology-first approach*), as it starts from theoretical principles rather than experimental data. According to Tononi's theory, therefore, the more integrated (read "*dependently interacting with each other*") the various parts of the thalamocortical system (the part of the brain where consciousness is believed to reside), the greater will be the degree of integrated information and, ultimately, of consciousness; this also offers an explanation to the "quantity problem": the cerebellum, despite the huge number of neurons and connections, is not integrated as it has highly differentiated modules that are independent of each other, so it doesn't integrate information and, therefore, doesn't show signals of consciousness. Looking at the following image (see figure 1), let's see better what it means for a mechanism to be informationally integrated. Consider the two extremes in the complexity-order graph in figure 1, on the left a random system, in which each element behaves stochastically and independently, while on the right a regular system, in which each element is completely determined by the state of the other elements, and therefore the degrees of freedom that the system can assume are limited. Now, the system on the left presents the maximum information, but the minimum integration, whereas the system on the right presents the maximum integration but the minimum information. In the middle of these two extremes, between order and disorder, there are complex systems like the brain, highly informative and at the same time highly integrated. If a brain behaved like the systems on the left (maximum information, minimum integration) all the neurons would be disconnected from each other, firing randomly and without coordination. Instead, if it behaves like the system on the right (maximum integration, minimum information) all the neurons would be interconnected with each other, firing in unison (as happens in epileptic seizures). In neither of the two opposite states would there be consciousness, but for the same reason: there is no "balance" between information and integration. This also suggests that the architecture of the brain could be a "small-world" network [8].

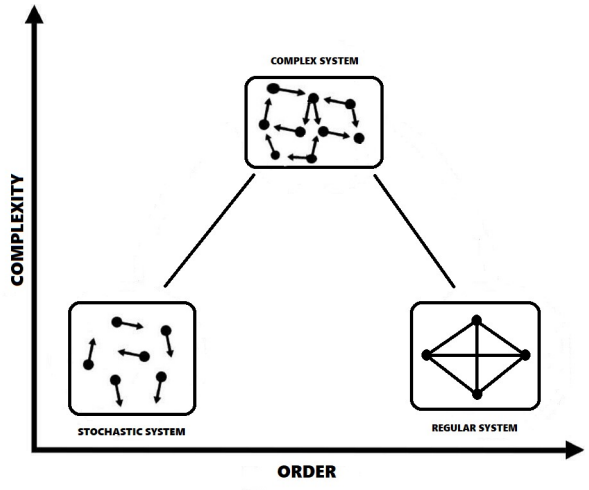

Figure 1. Relationship between Complexity and Order.

Tononi's theory is not limited only to theoretical conceptualizations, but is immersed in mathematical language through the formulation of a quantitative measurement called "*Phi*", Φ, which detects the "level", or rather, the "degree" of consciousness of a system. Φ is calculated by means of the difference between the total information entropy of the system as a whole and the sum of the information entropies of each possible configuration that the system can assume. This difference represents the amount of information integrated within the system compared to its fragmentation into independent partitions. In other words, "Phi" measures the amount of information that is unique and integrated into the system, compared to that which can be explained by the sum of the information of the individual subsystems: it can be considered the measure of how much a system is "*more than the sum*" of its parts in information terms, i.e. it determines how the system is an "integrated whole" beyond the amount of information that each of its sub-components generates independently. This underlies the main thesis of the theory, namely that *a*

system is conscious to the extent that its whole generates more information than its parts, so that there will be as much consciousness in proportion to how large Φ is. The greater Φ, the greater the degree of consciousness, conversely, the lower Φ, the lower the degree of consciousness, up to the limiting (but common) case of degree 0, in which there is no sense of consciousness in a system. This implies that one knows what the null state of consciousness is, but one doesn't know the maximum level that the degree of integrated information can reach. The hypothesis of the theory, therefore, is explained in a gradual model: consciousness is not an *all-nothing state*, but goes up the "great chain of Being", from inorganic materials (without consciousness) up to human beings (the highest degree of consciousness). Therefore, according to the IIT, a stone is not conscious like a chair is not, but an octopus has a minimal state of consciousness (even if we can never know what it means to "feel" for an octopus). A dog will have even more consciousness, up to the case of the human brain, in which the type of consciousness we know manifests itself. But, here another question arises, if the neural configurations of the human being are always the same, how can it happen that consciousness fails when "one loses consciousness", during anesthesia, or in the NREM dream (the deep sleep without dreams), or in pathological situations such as coma? And how can it happen that in the same state of sleep there are two phases with different states of consciousness, the NREM phase in which there is no hint of consciousness, and the REM phase, in which the mind is conscious in the elaboration of the most bizarre dreams? According to Tononi's theory, the loss of consciousness in the above situations is linked to the fact that integration in the neural architecture fails, and the system differentiates into independent modules, so consciousness disappears. Therefore, the neural circuitry itself gives rise to qualitatively different states of consciousness. Yet, another question is why in the case of a generalized epileptic seizure, in which there seems to be an excess of conscious integration, as the neurons fire unanimously, can there be no consciousness? To this question, there is no valid answer yet.

Finally, one would wonder whether artificial intelligence could it be conscious? According to Tononi, the configurational architecture at the basis of today's computers – built by networks of transistors – precludes the necessary level of integration of information that is necessary to have consciousness, so the answer, at the moment, is negative.

What about a society? Can it give life only to collective-connective intelligence or also to a global consciousness? Recently, integrated information theory was tested on groups of people, with the demonstration that estimates of integrated information shared by team members could predict group performance on various tasks [9]. Furthermore, in the context of studies on the effects of drugs [10] seems to support the theses of IIT. In this work it was discovered that the intake of psychedelic drugs involves evident alterations of brain dynamics, according to which some brain regions whose activity was usually coordinated became uncoupled, while other brain regions whose activities were usually independent connected coherently with each other. As reported in [11], "in general, the image was that of a rupture of the connectivity configurations that characterized the brain under normal conditions". In relation to these results, the author of the study hypothesized that "this rupture could provide an explanation of some characteristic aspects of the psychedelic state, such as the dissolution of the boundaries between self and the world, as well as the mixing of the senses (...). Brain activity in the psychedelic state became more stochastic over time, consistent with the loose reorganization of perceptual experience that people frequently report during a psychedelic trip". In relation to this, the conclusions we reach is that the integration of information in the brain, following the intake of this type of drug, disintegrates and reassembles in unusual ways, giving life to an "other" experience of consciousness.

2. Discussion

Tononi has paved the way in the problem of consciousness. His work has opened a gap in the dark mystery of subjective experience, and now it is time to put on glasses to look more clearly at the discoveries of the IIT from another perspective. What will be proposed below are a set of intuitions and considerations that relate phenomenological states with the quantities that underlie the physical world as we know it.
We start from Landauer's principle [12], formulated by the physicist Rolf Landauer in 1961:

$$E = kT \log 2$$

Where E is the amount of energy; k is Boltzmann's constant, equal to approximately $1{,}380649\times10^{-23}$ J/K (Joule/Kelvin); T is the absolute temperature of the circuit in degrees Kelvin; *log 2* is the natural logarithm of 2, which is approximately 0.69315.

It establishes that there is a quantity of energy below which it is not possible to go in the process of deleting a bit of information, so the equivalence *cancellation-information = use-energy* can be established. In other words, the change

in information involves a dissipation of energy in the form of heat. More specifically, the energy dissipated is proportional to the temperature of the system and to the logarithm of the number of possible states in which the information can be represented (which in the case of information is binary, and therefore is 2). The deleted information results in an increase in entropy (which can therefore be considered a measure of erased information) in the system. This principle provides a bridge between information theory [13] and thermodynamic physics, according to which information, understood as the reduction of uncertainty, has a physical nature [14]. On this bridge stands the discipline of thermodynamics of information, which links energy, information and entropy together. It should be noted, however, that the Landauer limit doesn't apply to all information processing processes, but only to the deletion of information. Today, however, the scientific community is unanimous in accepting the Landauer's equation, after its experimental confirmations [15], considering information as a form of energy. Indeed, in [16], Vopson gives an elegant formulation to the so-called MEI equivalence principle (*Mass-Energy-Information equivalence principle*), on the basis of which, just as mass is a form of energy, according to Einstein's famous equation ($E=mc^2$), information is also a form of energy and mass. Figure 2 shows the famous equivalence:

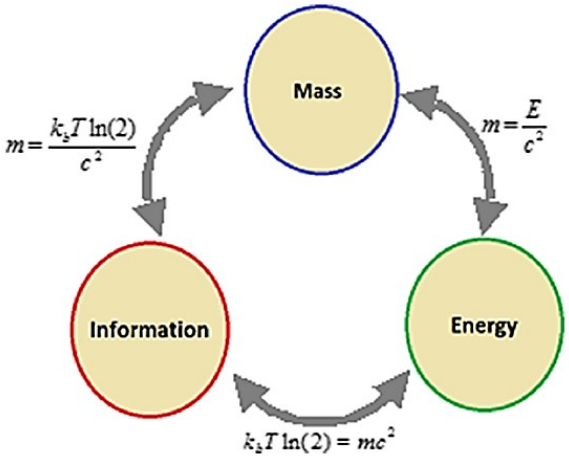

Figure 2. Representation of the mass – energy – information equivalence. Image taken from [16].

The implications of this principle are extraordinary. Firstly, Vopson suggests that, playing with equivalences, information has a non-zero mass, predicting that in about 350 years, given unlimited digital growth, the weight of digital bits could be so great as to exceed that of all the atoms present on the Earth, a phenomenon known as "the information catastrophe" [17]. Secondly, Vopson goes so far as to hypothesize that information could be the *fifth state of matter*, alongside solid, liquid, gas and plasma [18]. Having thus established that information is another form of energy, let us turn again to the study of IIT, considering consciousness, which is integrated information, under this new perspective. The thesis here is divided into two hypotheses, as following:

1. Starting from the MEI principle, we can consider integrated information, which is a reduction in uncertainty and is measured by entropy, as the energy generated by the interconnection of the brain regions of the thalamocortical system;
2. Consciousness is an emergent property of the interconnected regions of the thalamocortical system, which is found in a stable energy (read "integrated information") interval on the border between order and disorder, i.e. on the edge of chaos.

The "Phi" formulated by Tononi could be the measure of entropy, which is the degree of disorder or uncertainty of a system. In other words, *entropy "is"* (the quantity of) *information* (contained in a system) [19]. The greater the disorder, or uncertainty, or lack of information of a system, the higher the level of entropy. Conversely, the greater the degree of order, certainty, precision of the information of a system, the lower the level of entropy. From what has been said, it follows that "human" consciousness is correlated to a "balanced" range of entropy in the complex brain system. I have emphasized "*human*" as I believe that consciousness follows both a *graduated* model and a *threshold* model, based on the point of view we adopt. To explain this, just like Tononi, we believe that consciousness is not an *all-nothing* phenomenon, but we favor the idea that there are various degrees of consciousness depending on the "object" being examined. A stone will not have consciousness, an octopus will have a little consciousness, a dog will have more consciousness up to the human being, in which conscious experience reaches its peak. However, if we focus on the human being, whose subjective experience we have defined as an "*emergent property*", by definition the

emergency needs a threshold beyond which there is conscious experience as well as all we experience, and below which there is no consciousness as we are used to (attention: there is no "consciousness as…", not that there is no consciousness!), so from this point of view the two models can coexist together. It is also plausible that NREM sleep, the anesthetic state, the infantile state, the "animal" state, the pathological states, are conditions in which there is more or less consciousness, perhaps dormant, but there is, to a certain extent, nevertheless, not enough to exceed the threshold that makes consciousness a "human emergent state". And yet, when a child grows up and becomes self-conscious, or when one wakes up from a deep sleep (whether normal, anesthetic or pathological) the entropy (read "energy", or "integrated information") increases, exceeding the critical threshold and giving new awareness of oneself as a human being experiencing conscious states. Now, we wonder if this threshold is the one indicated in [20], equal to 0.31, measured with a mix of neuroscientific techniques through the PCI (Perturbational Complexity Index) developed by the neuroscientist Marcello Massimini: the threshold of 0.31 could identify the point at which neural dynamics reach a sufficient "complexity" to be considered related to consciousness.

This approach, as argued, doesn't envisage panpsychism, unlike the IIT, according to which a physical substrate that presents the properties listed by the theory, whether biological or not, "deterministically" presents consciousness.

Our explanation is that according to the assumptions made here, panpsychism applies only to biological (organic) organisms, which could, to a greater or lesser degree, integrate information as IIT wants. Regarding inanimate matter, such as a stone, we are inclined to say that it doesn't integrate information, since the main form of energy in this case is mass ($E=mc^2$) and not information. In line with Tononi, and based on the previous conclusion, we believe that an AI (Artificial Intelligence) may be able to reach, equal and perhaps surpass human beings in intelligence, but will never show signs of consciousness. There can in fact be intelligence without consciousness, so it is important to distinguish between these two aspects of human nature. At this point, a further step could be taken, as discussed below. The "*free energy principle*", formulated by the neuroscientist Karl Friston [21-23], establishes that natural systems tend to reach a state of minimum free energy or maximum stability. That is, they tend to decrease the uncertainty or entropy in their environment by interacting with it in a way that maximize their adaptability. Free energy, in this context, represents a measure of the reduction of predictive error in the system through the process of "active inference" and updating of predictions. According to this theory, the brain tries to minimize the free energy or the uncertainty to reach a state of equilibrium, reducing the error or the discrepancy between its predictions and the observations coming from the outside world through appropriate decisions and actions. Following the hypotheses of this conceptual framework, the reduction of uncertainty through the difference between prediction and observation increases learning, and therefore *information is acquired*, and therefore free energy is minimized (*entropy is reduced*) stabilizing around a point (or better, interval) of equilibrium. Therefore, systems such as the brain that follow the principle of free energy are based on *meanness* ("*katà métron*", or "according to the right measure"). In conclusion, the idea is that the processing of free energy in the brain could be involved in the generation of states of consciousness and, therefore, in the integration of information. Here, finally, that even the disappearance of consciousness following epileptic attacks could be explained by these conjectures. In these situations, there could be so much entropy that the range of entropic equilibrium is exceeded, and therefore the states of consciousness fail or are altered.

3. Conclusion

We have therefore seen the cornerstones of this new proposal of "entropic consciousness" (EC): it is based on a stable interval of energy, on the margin between entropic order and disorder, within which the emergence of human consciousness takes place as we all experience it. We have also seen that, through the MEI equivalence principle and the Landauer limit, information is a form of energy, that entropy can be the "unit of measurement" of this energy and that the "free energy principle" is consistent with this theorization. Finally, we have seen the conjecture on panpsychism under another lens, through the coexistence of the threshold model with the graduated model, so all that remains for us to do is deal with the open problems that stand out on the conceptual horizon of EC. Here are the main issues highlighted:

- Currently there is no direct method to measure the entropy of the human brain: could PCI and PHI be reliable suspects to fill this role? Will PCI and PHI be able to measure brain entropy in relation to the emergence of consciousness?
- Could neuropsychiatric diseases be the result of an excess or a lack of cerebral entropy, which alters the functioning of neural networks, even if it leaves consciousness unchanged?
- And what state of consciousness is experienced when reaching "nirvana" through meditation? Can one "consciously" calm the "chatter" of conscience? And what effect does psychotherapy have on consciousness?

And hypnosis? And what relationship does conscious experience have with the *Default Mode Network* (DMN), in which attention fluctuates? And with the state of "flow", in which attention is focused? Are these states a "subtype" of consciousness or an "other" form of consciousness?

- Is a nation state or every emerging complex system in general conscious? Does human society, a colony of ants, a school of fish, a flock of birds, a pack of wolves integrate information? Are they higher forms of consciousness? And does "*Gaia*", the Earth-system, truly represent a global consciousness?

Summing up the discussion, it must be noted that human consciousness, even if it takes place in exceptional conditions of emergency and self-organization, is still an *"accident" of natural selection.* Consciousness is not given to us (only) to be able to write poems or paint pictures, but because it is useful in the adaptation of the human species to the Earth's environment. So, one must deconstruct the narrative of human exceptionalism, and understand that we are not separate from the rest of nature, but are an integral part of it, as consciousness could be found in various degrees of development in other living forms. From this it follows that - the last point to understand - it is not the energy itself (whatever it is) that creates consciousness, but that this derives from a "particular form" of energy, *information*, which furthermore must be "integrated" as a *conditio sine qua non*. The "science of the mind" therefore, after that of Copernicus and Darwin, inflicts the third humiliation on human pride, which began over a hundred years ago with the theories on the *unconscious*. It is hoped that this work will stimulate further theoretical and experimental research, bringing the scientific community closer to understanding the nature of matter, energy, information and consciousness in the Universe.

References

[1] Descartes, R. (1641), Meditazioni metafisiche.
[2] Seth, A. K., & Bayne, T. (2022). Theories of consciousness. Nature Reviews Neuroscience, 23(7), 439-452.
[3] Tononi, G. (2003). Galileo e il fotodiodo: cervello, complessità e coscienza. Laterza.
[4] Tononi, G. (2008). Consciousness as integrated information: a provisional manifesto. The Biological Bulletin, 215(3), 216-242.
[5] Tononi, G. (2012). Phi: A Voyage from the Brain to the Soul. Pantheon.
[6] Tononi, G., & Massimini, M. (2013). Nulla di più grande. Baldini & Castoldi.
[7] Oizumi, M., Albantakis, L., & Tononi, G. (2014). From the phenomenology to the mechanisms of consciousness: integrated information theory 3.0. PLoS computational biology, 10(5), e1003588.
[8] Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of "small-world" networks. Nature, 393 (6684), 440-442
[9] Engel, D., & Malone, T. W. (2018). Integrated information as a metric for group interaction. PLoS One, 13(10), e0205335
[10] Carhart-Harris, R. L., Erritzoe, D., Williams, T., Stone, J. M., Reed, L. J., Colasanti, A., ... & Nutt, D. J. (2012). Neural correlates of the psychedelic state as determined by fMRI studies with psilocybin. Proceedings of the National Academy of Sciences, 109(6), 2138-2143.
[11] Seth, A. (2021). Being you: A new science of consciousness. Penguin.
[12] Landauer, R. (1961). Irreversibility and heat generation in the computing process. IBM journal of research and development, 5(3), 183-191.
[13] Shannon, C. E. (1948). A mathematical theory of communication. The Bell system technical journal, 27(3), 379-423.
[14] Landauer, R. (1991). Information is physical. Physics Today, 44(5), 23-29.
[15] Bérut, A., Arakelyan, A., Petrosyan, A., Ciliberto, S., Dillenschneider, R., & Lutz, E. (2012). Experimental verification of Landauer's principle linking information and thermodynamics. Nature, 483(7388), 187-189.
[16] Vopson, M.M., The mass-energy-information equivalence principle. AIP Advances, 9, 095206 (2019). https://doi.org/10.1063/1.5123794
[17] Vopson, M.M., The information catastrophe, AIP Advances, 10, 085014 (2020). https://doi.org/10.1063/5.0019941
[18] Vopson, M.M., Experimental protocol for testing the mass–energy–information equivalence principle. AIP Advances, 12, 035311 (2022). https://doi.org/10.1063/5.0087175
[19] Ben-Naim, A. (2008). Entropy Demystified: The Second Law Reduced to Plain Common Sense (Revised Edition). World Scientific.
[20] Casarotto, S., Comanducci, A., Rosanova, M., Sarasso, S., Fecchio, M., Napolitani, M., ... & Massimini, M. (2016). Stratification of unresponsive patients by an independently validated index of brain complexity. Annals of neurology, 80(5), 718-729.
[21] Friston, K., Kilner, J., & Harrison, L. (2006). A free energy principle for the brain. Journal of physiology-Paris, 100(1-3), 70-87.
[22] Friston, K. (2009). The free-energy principle: a rough guide to the brain? Trends in cognitive sciences, 13(7), 293-301.
[23] Friston, K. (2010). The free-energy principle: a unified brain theory? Nature reviews neuroscience, 11(2), 127-138.