

# How Entropy Explains the Emergence of Consciousness: The Entropic Theory

## Abstract

**Aim:** Emergentism as an ontology of consciousness leaves unanswered the question as to its mechanism. I aim to solve the body–mind problem by explaining how conscious organisms emerged on an evolutionary basis at various times in accordance with an accepted scientific principle through a mechanism that cannot be understood, in principle. **Proposal:** The reason for this cloak of secrecy is found in a seeming contradiction in the behavior of information with respect to the first two laws of thermodynamics. Information, the microstate of particles within an isolated system's macrostate, can, like first-law energy, be neither created nor destroyed, yet the information in the system, like second-law entropy, will inevitably increase. To explain information increasing without being created, Laplace's demon is invoked, able to predict where each particle is destined. This doesn't work for emerging events like consciousness, which are unpredictable. This can be understood in terms of the derivation of entropy, and the emergence of classical physics, from the Relativistic Transactional Interpretation of Quantum Mechanics. I propose that the increased entropy in a time-irreversible, unpredictable (emergent) isolated system requires the simultaneous deletion of information concerning the steps, or calculations, involved. **Conclusion:** Thus, the steps leading to consciousness are immediately destroyed and must remain a mystery. Implications include that entropy, not panpsychism, is the universal principle generative of consciousness, that our being conscious proves that we are not predetermined, and that consciousness requires assuming an "entropy debt" that can only be repaid by living organisms, prohibiting the emergence of conscious machines.

**Keywords:** *Consciousness, emergence, entropy, evolution*

**Peter Cornelis Lugten**

*Independent Philosopher,  
Lindenhurst, NY, USA*

## The Ontology of Consciousness

### Introduction

The purpose of this paper is to solve the problem of emergentism within consciousness. There are two views as to how tissues inside the vaults of our craniums can create the wonders of this experience. Panpsychism holds that all the component subatomic particles of those tissues participate in a form of consciousness, and have done so since their origin at the beginning of the universe. Its proponents have acknowledged the theoretical "binding problem:" how the minute consciousnesses of so many components can be melded into the singular conscious entities of living organisms. In addition, I'll show that they face a sterner problem posed by an inability to repay the debt owed by

consciousness to entropy. The alternative to panpsychism is emergentism, which holds that consciousness arises exclusively within nervous systems of living animals (and possibly within homologous structures in plants, maybe even smaller organisms). Proponents of this theory are challenged to explain by what magic consciousness can emerge from nonconscious components within living tissues.

Consciousness can be loosely defined as what it is like to be oneself, including introspection, subjective sense perceptions (qualia), integrated sense perceptions, subjective awareness, self-awareness, emotions, feelings, wakefulness, thoughtfulness, memory, awareness of memory, inquisitiveness, communicativeness, purposefulness, a drive toward problem-solving, and, as I will maintain, causative agency.

In particular, Yunus Çengel described a hierarchy of causative agencies, with

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

For reprints contact: WKHLRPMedknow\_reprints@wolterskluwer.com

**Ethics committee approval:** There is no need for ethics committee approval

**Received :** 21-02-2024

**Accepted :** 01-03-2024

**Published :** 25-03-2024

### Orcid

Peter Cornelis Lugten  
{ORCID: 0009-0007-6744-8618}

### Address for correspondence:

Dr. Peter Cornelis Lugten,  
75 Feustal St., Lindenhurst,  
NY 11757, USA.  
E-mail: pclugten@gmail.com

### Access this article online

**Website:** <https://journals.lww.com/jons>

**DOI:** 10.4103/jnbs.jnbs\_6\_24

### Quick Response Code:



**How to cite this article:** Lugten PC. How entropy explains the emergence of consciousness: The entropic theory. *J Neurobehav Sci* 2024;11:10-8.

consciousness as a purposive agency, endowed with a directed causality.<sup>[1]</sup> This is opposed to the laws of physics, which are nonpurposive causative agencies, and simple emergent properties, which are noncausative. Emergent theories of consciousness are interesting because they can allow consciousness to interact with and control a material brain.

Karl Popper and John Eccles noted that monist theories of consciousness, such as radical materialism/behaviorism, and the dualisms of panpsychism, epiphenomenalism, parallelism, and identity theory deny that, or cannot explain how, consciousness can causally affect the material brain.<sup>[2]</sup> For example, they cannot account for the mental effort that can be needed for the mind to direct the brain to retrieve a specific memory, and then evaluate the accuracy of the brain's performance. Popper held that mental processes evolved under the pressure of natural selection, based on the need for purposeful behaviors to ensure survival and reproductive success. (Eccles noted that if consciousness hadn't made a difference, it wouldn't have evolved.) This is the starting point linking my proposed ontogeny of emergentism with a causally directive consciousness.

### Why consciousness is necessary

Consciousness is a property of a neurological analytic facility that, to survive, must blend all our experienced perceptions with the signals of internal homeostasis. It must match these with expectations, apprehend the nature of problems, rank their urgency, and solve them based on an analysis that has to have a representation of the self in the world. For this, the brain creates a metaphorical surround-sound movie screen that Daniel Dennett called the "Cartesian theater."<sup>[3]</sup> This must transpire without the benefit of a "little man," or homunculus, to view the scene from within, since the homunculus, in turn, would need a homunculus, regressing infinitely. As Antonio Damasio put it: "The sense of the self in the act of knowing emerges within the movie. Self awareness is actually part of the movie, thus it creates the "seen" and the "seer," the "thought" and the "thinker," with no separate spectator for the movie-in-the-brain."<sup>[4]</sup> Our conscious experience is the homunculus. It is rapidly supplied with pertinent information about our internal and external milieus in a format that's readily understood, including the ability to feel changes in our bodies that are provoked in our minds by its directing our own actions. So, in effect, we are left with the question, not "Why do we have consciousness?", but "What possible alternative way of experiencing the world could there be?"

But why did the need for conscious experience evolve? Imagine a world in which we had no more consciousness than a machine, but more intelligence than a rock. In this world, we could computationally reason that it was time to seek food or shelter, and that we must earn money in order to pay for these, but we could not reason or imagine, in

the absence of the sensations of pleasure, love, reward, or enjoyment, that it is time to seek romance, or to perform an act that will lead to the birth of children, or to care for them afterward. Consciousness, therefore, is more than a result of Darwinian selection for self-preservation. It is also the result of sexual selection for the sights, smells, and sounds of beauty and the emotional package leading to reproduction. The environment of mating rituals and the efforts required to make ourselves (and then our children) fit, educated and attractive, is the primary driver of much conscious programming.

Consciousness is a very neat trick, but how does the brain pull it off? To answer this, it is necessary to understand the universal drive toward diversity and disorder.

## The Interaction between Consciousness and Entropy

### Thermodynamic entropy: Microstates within macrostates

Living organisms are islands of reduced thermodynamic entropy within their environments, far from thermal equilibrium. Consciousness further reduces this entropy multifocally around the world, as is seen by an examination of the structures conscious minds have built. Compared to the thermodynamic entropy of the brain's sensory inputs, the thermodynamic entropy of the output of consciousness is lower, i.e. it is more organized, which is what makes the brain interesting.

Thermodynamic entropy, as described by the second law of thermodynamics, is the tendency for disorder in a closed system of particles, i.e. gas molecules, always to increase. In the 19<sup>th</sup> century, Ludwig Boltzmann linked this concept mathematically to the distribution of molecules in space, thereby determining the foundational equation of statistical mechanics:  $S$  (the entropy) =  $k \cdot \log W$ , where  $W$  is the number of real microstates corresponding to a gas's macrostate. The entropy of a macroscopic state is proportional to the number of configurations of microscopic states of a system where all microscopic states are equiprobable.<sup>[5]</sup> Later, it was realized that there was an analogy between the distribution of molecules in a gas and the amount of information in a message.

### Thermodynamic entropy and information

According to John Wheeler, information is fundamental to the physics of the universe.<sup>[6]</sup> He suggested the emergence of the physical from enough information, saying "more is different" and "The rich complexity of the whole does not in any way preclude an extremely simple element such as a bit of information from being what the universe is made of."<sup>[7]</sup> David Chalmers built on this in his analysis of consciousness, writing "If so, then information is a natural candidate to also play a role in a fundamental theory of consciousness. We are led to a conception of the world on

which information is truly fundamental, and on which it has two basic aspects, corresponding to the physical and the phenomenological features of the world.”<sup>[8]</sup> There is a connection between information and entropy, microstates and macrostates, which needs exploration to understand consciousness.

The amount of information in a message is, in most contexts, proportional to its length in characters or digits. Likewise, Entropy, per Boltzmann’s equation, is the number of digits of probability in a system, and represents the possible combinations of activity that we’re ignorant of. The more certain an event is, the less surprising it will be and the less information it will contain, and therefore, a gain in information (by which, I don’t imply known information, which has zero entropy) is an increase in uncertainty or entropy. An increasing entropy implies an increasing uncertainty, or number of possible outcomes, being associated with an increased number of microstates within a macrostate. (Microstates are subunits of a system, or macrostate, which can be imperceptibly rearranged within it.) If we are about to toss a coin, or roll a die, there isn’t yet any information about the outcome, and zero entropy. Having tossed the coin, with two equally likely outcomes, the information gained, or surprise upon learning the result will be less than the information gained, or surprise on learning the result, of the die roll, with six possible outcomes. An increase in the number of possible outcomes in “information space” is equivalent to an increase in disorder in the world. The information space can refer to the possible arrangements of sand grains on a beach, or atoms in a jar, impossible to apprehend. Entropy is the amount of “missing information” needed to determine what specific microstate your system (or information space) is in. While the thermodynamic entropy of a physical system is measured in physical units (Joules of energy divided by the absolute temperature), the informational entropy is measured in abstract mathematical units - bits, short for “binary digits.”

In fact, once we learn the result of the coin toss, or the roll of the die, the information, its uncertainty and its entropy, drops to zero. Information here is defined as the opposite of knowledge. The loss of entropy associated with gaining knowledge is compensated for by the increase in entropy associated with our brain’s activity, especially that associated with maintaining consciousness, as well as the activities it directs. Additionally, the more we learn, the more we realize what we still don’t know, increasing our uncertainty. Or, according to a quote generally attributed to John Wheeler, “We live on an island surrounded by a sea of ignorance. As our knowledge grows, so does the shore of our ignorance.”<sup>[9]</sup> One could reasonably describe the brain as an instrument for expanding entropy through the conversion of information into knowledge.

### The entropic brain theory and Shannon entropy

To avoid possible confusion, I’ll mention here that the Entropic Theory of the Emergence of Consciousness is not

related to, and contradicts aspects of, the Entropic Brain Hypothesis of Carhart-Harris *et al.*<sup>[10]</sup> The latter considers the human brain to have a higher entropy, or disorder, than the brains of other animals, i.e. a greater repertoire of potential mental states, which expanded greatly relatively recently in our evolutionary history. This occurs with an increased entropy suppression during normal consciousness, relative to the degree of suppression characteristic of archaic humans and also of infants. This hypothesis is couched in terms of thermodynamic entropy, but also described in terms of Shannon entropy. Shannon entropy is a measure of the amount of information, or surprise, contained in a variable, or unit of storage/transmission that can take different values following some process, such as in a message. It is eponymic for Claude Shannon, who realized that the equation for representing this concept resembled the Boltzmann equation for thermodynamic entropy, except with a minus sign in front.<sup>[11]</sup> The minus sign in this context means that information is the opposite of entropy.<sup>[12]</sup> It means that a message with zero surprise has no Shannon entropy, whereas a room with almost zero thermodynamic entropy would have all the air molecules concentrated in one corner: a very surprising state of affairs indeed! The Entropic Brain Hypothesis is also consistent with Karl Friston’s “Free-Energy Principle” theory of consciousness, which asserts that conscious organisms seek to minimize their Free Energy.<sup>[13]</sup> For clarity’s sake, I’ll continue the discussion in terms of thermodynamic entropy.

The Entropic Brain Hypothesis posits that the entropy reduction manifest by consciousness is associated with highly organized activity in the Default Mode Network of the brain, as described by Raichle *et al.*, and with related neural connections.<sup>[14]</sup> Carhart-Harris considers this to be secondary consciousness. Under conditions such as sleep, general anesthesia, seizures, and depression, entropy is further suppressed, focused inwards, and unable to access a healthy diversity of thoughtful states. On the other hand, infant consciousness, dreaming, magic thinking, near-death experiences, and psychedelic drug trips increase entropy, allowing access to mental states no longer tethered to reality. These “hyperconscious” states, which Carhart-Harris considers to be primary consciousness, can approach “self-organized criticality,” the transition point from order to disorder, where so-called “power-law scaling” can result in avalanches of neural cascades. This is clinically manifest in the “dissolution of the self” reportedly experienced during psychedelic drug use.

The Entropic Theory of the Emergence of Consciousness, while in agreement with the idea of entropy reduction during normal consciousness, disagrees with Entropic Brain Theory on significant points. In particular, the idea that sedation, anesthesia, depression, etc., represent further decreases in entropy beyond the reduction associated with normal consciousness is a good description in terms of Shannon entropy, because in these states, one is less



apt to produce surprising behavior. However, in terms of thermodynamic entropy, in these states, one is much less likely to do anything constructive that decreases the entropy of the immediate outside world. Because of this, subconscious states are equivalent to an increase in thermodynamic entropy relative to normal consciousness. I'd argue that during evolution, organisms of relatively low, poorly organized levels of consciousness, having high thermodynamic, but low and subcritical Shannon entropy, developed higher levels of more organized consciousness. This reduced their thermodynamic entropy (but increased their Shannon entropy to a point little below criticality). It seems intuitively sensible that the state of consciousness would arise from a state of unconsciousness rather than from a state of hyperconsciousness: otherwise, one is left to wonder how the hyperconscious state first appeared. I submit that consciousness emerged from unconsciousness, and the uncertainty associated with this, the seeming impossibility of our being able to understand it, guarantees that thermodynamic entropy will, perhaps counterintuitively, favor this process of emergence.

## The Paradox of Emergence without Creation

### The emergence and destruction of information in a determined universe

By now, it should be clear that the term “information” has meanings at different levels which could be confused, especially with regard to the (imperfect) analogy between information and entropy. For instance, the first law of thermodynamics states that in an isolated system, energy can be neither created nor destroyed; it is believed that the same applies to information, with two caveats. What's known as the “Black Hole paradox” suggests that information may be destroyed by a black hole that subsequently radiates away; this paradox may have been solved at the time of writing,<sup>[15]</sup> but doesn't concern this discussion. Confusion is added by the second law of thermodynamics, which states that the amount of entropy in an isolated (closed) system cannot decrease – it tends inevitably to increase until the system achieves equilibrium. This is a property which seems to “emerge” from quantum physics, in which all interactions are perfectly reversible. According to Ruth Kastner's Relativistic Transactional Interpretation of Quantum Mechanics, a “direct action theory” of objective reduction, this occurs with the loss of any “phase coherence” in quantum states, with the resultant “throwing out of information” to create “blurring” at the classical level.<sup>[16]</sup> In particular, it emerges when we are dealing with large numbers of particles in a statistical fashion. Also, while information is conserved, entropy is not. Therefore, if the amount of information is proportional to the entropy, then this means that the amount of information, too, must increase until the universe reaches the equilibrium of heat death far off in the future. Sean Carroll explained that the information

which is conserved is that of the microstate, made up of the positions and momenta of particles, unknown to us, and not the information in the system's macrostate, of which we might or might not have knowledge.<sup>[17]</sup> The embedded, classical, or macroscopic information, of which we can seek knowledge, is not conserved, and can be copied or deleted perfectly. Therefore, a book, full of classical-level information, can be destroyed in a fire, which will increase the entropy of the macrostate, as well as (one might think) that of the scattered atoms in the microstate. However, the radiation and the atoms in the smoke and the ash could theoretically be traced back to their original positions in the book, so this conserves information at the level of the microstate (it is not destroyed), even as we know this, being equivalent to the reassembly of a broken egg, would decrease the entropy of the scattered particles and will never happen. The law of conservation of information seems to assume that the second law of thermodynamics does not apply at the microstate level. Likewise, the question of what happens to all the information in our brains, if not our minds, at death is analogous to what happens to the information in the incinerated book. Someone who knew all the trajectories of all the particles after classical information is destroyed could reconstruct all the information laid out in neural pathways.

This explains how information in the book escapes destruction, but how is information not created as the book's entropy is increased, and its atoms scattered? Anyone who knew the trajectories of all the relevant particles before the book was burnt would know the information concerning each particle that was about to change as the printed paper burned. Since they would be able to predict what would happen next, the information of this system is thus changed without being “created,” and doesn't actually increase. Again, the conservation of information requires sacrificing the second law of thermodynamics at the level of the microstate. It simply denies that the scattering of all that particulate information is chaotic, or able to increase entropy. This needs to be rethought.

Consider the process of my writing this paper, and thus increasing the amount of information in the world, about a new theory. The creation of information by Laplace's Demon would mean a rearrangement of known information without any unpredictable ideas being able to emerge. Since he could predict the trajectories of all the atoms in my brain as they directed my fingers on the keyboard, he would explain the increased information in this paper as being a process of reshuffling the ideas of other people. But this would not produce a surprising, emergent, unpredictable creative idea.

The key feature of this paper, an unpredictable solution to the problem of emergence, is an emergent result of brain evolution as an organ of problem-solving. Its biological drive to ensure the survival and success of its owner was

facilitated by the emergence of consciousness, so that the nature of the problems and the implementation of their solutions could be understood. Creative new ideas are thus a feature of an evolved computational process resulting in the emergent unpredictability of consciousness in our brains.

The first and second laws can be reconciled by asserting that information in a single system can be increased by a process of rearrangement, without being created. However, this comes with a serious problem. It requires that the future course of the atoms and particles be predetermined by knowledge of the system at present, and denies the existence of spontaneous creativity, or of emergent phenomena: those which are, by definition, unpredictable, even given a complete understanding of their underlying constituents. If the universe is not deterministic, then the paradox of the conservation of information and its increase with increasing entropy has not been solved at all.

### **The emergence of information in an undetermined universe**

A nondeterministic theory suggests that information is related to but not equivalent to entropy. What is conserved is some combination of them, with one increasing as the other decreases. For instance, as information is converted into knowledge, the information decreases in amount, and, locally, so does entropy, but, overall, as knowledge increases, like the shore of Wheeler's island, entropy does also. A deterministic knowledge of the future is no longer required. Given the quantum necessity of chaotic indeterminism in the universe, for instance, as described by Ilya Prigogine,<sup>[18]</sup> I believe this is key to resolving the paradox.

Yunus Çengal has suggested that "the notion of conservation of information should be limited to the physical universe governed by the laws and forces of physics, and it should be referred to as *physical information*...to clearly distinguish it from other forms of information or knowledge."<sup>[19]</sup> There are limits to the reach of the theory of the Conservation of Information, such as occurs when it encounters an unpredictable, irreversible transaction such as the leap into consciousness. This leap is a computational process, which decreases our personal entropy, but must balance this with an increase in entropy. It turns out that the act of performing calculations causes this increase in entropy.

### **Computation, entropy, and the deletion of information in emergence**

Computation requires a temporary storage of information, upon which the mind (or calculator) acts in order to perform the calculation. It cannot be stored indefinitely, and must be erased in order to proceed to the next calculation. Rolf Landauer proposed that any logically irreversible computation, i.e. erasing a bit of information, requires work, expels heat, and increases entropy;

information is physical.<sup>[20]</sup> Ruth Kastner and Andreas Schlatter have resolved earlier controversies concerning this by asserting that it applies not to epistemic information loss (our ignorance of the microstate) during resets, but to ontological uncertainty about the quantum nature of the microstate itself.<sup>[21]</sup>

It is clear there is a great decrease in entropy resulting from the creation of conscious knowledge, which is mirrored in the organization we have imposed on the world around us. Considering the activity of each brain cell involved with the generation of consciousness as an informational transaction, the decrease in our personal entropy must be at least balanced by the simultaneous increase in entropy associated with the possible microstates involved with consciousness. Physically, these computations must involve, at a minimum, interacting electrons and their exchange of photons which occur during exothermic chemical reactions. The aromatic amino acid tryptophan, ubiquitous in microtubules, cell membranes, and other important biological and neural mega-networks, strongly absorbs ultraviolet photons and fluoresces in response.<sup>[22]</sup> The entropy of the mysterious computations that engender consciousness and the uncertainty surrounding them being significantly high, the likelihood of our ever understanding these transactions becomes correspondingly small. I propose that emergent systems, such as the emergence of classical from quantum physics, the emergence of life from chemistry, and the emergence of consciousness, involve the irretrievable destruction of microscopic information and that the uncertainty about them is the result of entropy.

Kastner has discussed how entropy itself may arise together with classical physics from a loss of quantum-level information when photons, emitted by excited electrons, are absorbed by other electrons.<sup>[16]</sup> This collapse of the photon's wave function, she wrote: "can be understood as leading to a generalized form of spontaneous symmetry breaking," a natural emergent phenomenon. Entropy thus arises from the spontaneous breaking of the symmetry of the unitary time evolution of the quantum state according to Schrödinger's wave function equation, which otherwise would result in the possible outcomes of particle interactions always being deterministic. The relevant context for the Relativistic Transactional Theory of Quantum Mechanics is that the nonunitary evolution is nondeterministic, even though the probabilities involved sum to a unitary 100%. This occurs with the destruction of any "computation" involved in the symmetry breaking, in the form of the loss of phase coherence in quantum states which, as we've seen, "throws out the information" to create the emergent, but "blurred," classical level. The more the microscopic information about a process is erased, the less we can predict about that process macroscopically, and uncertainty increases. An emergent increase in macroscopic (conscious) information in this setting is simply not predictable. I propose that to avoid the "creation" of information during emergence,

confounding the putative first law, the increase in entropy “requires,” or occurs with, the simultaneous destruction of the computational pathways involved in the emergence. This, then, destroys an equivalent amount of information. Only in this way can the first and second laws be reconciled during the phenomenon of emergence. The destruction of the information at the level of microstates required by entropy means that information is related to but not equivalent to entropy, and that only some combination of the two is being conserved. (As a corollary to this, I think we can say that physical determinism is incompatible with the emergent phenomenon of consciousness, and that therefore, since we are conscious, we are not predetermined.)

### Quantum effects and negative entropy

David Layzer and Robert O. Doyle have shown that the creation and embodiment of information occurs with a local decrease in entropy, or a pocket of negative entropy.<sup>[23]</sup> Entropy greater than the information increase must be radiated away as heat or as pure information. In quantum mechanics, information is governed by a conservation law, which prohibits the exchange of heat for negative entropy. Doyle notes that quantum mechanics combines a deterministic wave aspect with an indeterministic particle aspect. An electron can end up randomly in any one of the physically possible states of a measuring apparatus plus the electron, with the probabilities of each state given by the wave function. This “collapse of the wave function,” reducing multiple probabilities into one actuality, drops local entropy of the measuring device commensurate with the increased information and there is a discharge of heat to carry away the positive entropy. This irreversibly creates information at a purposive level (the deliberate measurement) and negative entropy newly embodied in the apparatus. Adequate but imperfect determination occurs, said Doyle, through averaging huge numbers of quantum interactions over large objects.

I propose the following solution for the enigma of how quantum information may increase while still being conserved.

### A proposed new thermodynamic principle solves the paradox of emergence

This reconciliation between the first and second laws required by the quantum conservation of information can be stated as a new principle within thermodynamics: the increase in entropy in a time-irreversible, unpredictable (emergent) isolated system requires the simultaneous permanent deletion of information concerning the steps, or computations, involved. The local increase in negative entropy is balanced by positive entropy radiated as heat.

It says, in effect, that to increase information without creating it, the process of creating the information must be deleted simultaneously with creating it.

## The Emergence of Consciousness: A Tale of Two Demons

### Emergence and convergence exchange information for negative entropy

This new law seems to me to be necessary to explain emergent phenomena. At the quantum level, information is binary bits related to the microstates of particles. This is believed to be conserved in a manner analogous to energy. It cannot be created because it represents the arrangement of energy in the universe, and since that cannot be created, there is a limit on its possible arrangements. Its destruction would be equivalent to the destruction of the “missing” information needed to determine what microstate your system is in. This would be equivalent to the destruction of entropy, and is, therefore, at least in nonemergent systems, impossible. But at the classical level, information can be in the form of ideas, and can be copied or deleted perfectly – it can be scrambled without being lost, and Laplace’s Demon keeps track of it all. The emergence of the classical level, like the emergence of consciousness, requires the creation of an emergent level of new, unpredictable information. To increase the level of order in the world gained by our becoming conscious, entropy must simultaneously increase through the loss of certainty associated with the unpredictable process. This occurs through the destruction of a portion of the information space that could become known to us, specifically, that portion involved in the process of emergence. When the “Entropy Demon” responsible for this opens the gate between our subconscious and our consciousness, the information that is erased is that describing the pathway of how consciousness emerges. This might most economically happen immediately after all the qualia associated with the perception, say, a face, a voice, and an emotion, have been assembled together by transcortical communications such as axonal connections, brain waves, or electromagnetic waves. However, the Orchestrated Objective Reduction Theory of Roger Penrose and Stuart Hameroff, places these calculations at the level of microtubules within neurons.<sup>[24]</sup> One level up, nerve cell membranes have unique ion channel proteins and G-protein coupled receptors of different masses and energy levels. Mostyn Jones described how their electrical activity generates nerve impulses and fields in different sensory detector cells.<sup>[25]</sup> These, together with limbic hormone receptors, in a noncomputational manner, could be responsible for qualia and the differences between them.<sup>[26]</sup> Oscillating electromagnetic fields can cause oscillations in brain circuits, which, in turn, help guide cognition. In “The Computational Brain,” Clem *et al.* give a brief overview of how research is expected to elucidate the circuits responsible for emotional valences, cognitive decision-making, and memory.<sup>[27]</sup> So, these are all ways that our subconscious brain cells and their circuits can prepare communications to send past the Entropy Demon. A recent study by Panda *et al.* as part of the Human Brain



Project suggests that this may require two neural circuits.<sup>[28]</sup> The posterior cortical regions and Default Mode Network hubs are needed to convey information while thalamic, frontal, and temporal-parietal regions, responsible for various cognitive processes, must be correctly integrated to broadcast information, with appropriate glucose consumption, or else consciousness is lost or disturbed. According to my proposal, in all these activities, to the extent that they help generate consciousness, something is going on that must be considered irreversible computations in the context of the Entropy Demon. Furthermore, whatever is going on that allows our subconscious to become conscious must be accompanied by a complementary irreversible, unpredictable process that converts our thoughts back again into subconscious instructions to move muscles, for neurons to follow. Call it a process of “convergence.” During convergence as thus defined, information is deleted and the entropy of the subconscious is temporarily reduced during the performance of the corresponding activity.

#### **Entropy debt: Implications for consciousness artificial intelligence and the reality of our existence**

There is a startling implication for general artificial intelligence. Since we can never learn for ourselves the computations that lead to consciousness, we’ll never be able to program them into a computer. Furthermore, a computer will never need to be conscious to perform its computations. The computer selects from its database to create an emergent answer, but unlike the emergence of consciousness, this is a noncausal emergence. Although the steps involved aren’t retained, they don’t need to be destroyed. The actual causation of the emergence of the answer lies in the algorithm, which originated with a conscious person. The reduction in entropy necessary to become conscious does not come for free, and conscious organisms must continually repay their entropy debt. A computer will not be able to do this unless being conscious enables it to perform computations that it couldn’t otherwise perform, or need to perform. Even then, it would still need to be able to print them out or otherwise act on them somehow, reintroducing Descartes’ mind–“body” problem. I believe that the process of becoming conscious must be evolutionary, perhaps associated with, if not fully originating from, the massive increase in complexity brought about when asexually reproducing organisms began to reproduce sexually. With this development arose a need for a primitive understanding of why and how to perform reproduction successfully, without which consciousness would have remained superfluous. It follows that ethicists need not worry about fair treatment of conscious computers until two computers should fall in love, and are prepared to die in order to defend their baby. At the moment, the development of “kinematic self-replication in reconfigurable organisms,” or “xenobots,” may open a pathway leading to self-motivated, programmable robots, but kinematic self-replication doesn’t involve genetic reproduction, and

the replicants are therefore not “babies.”<sup>[29]</sup> What type of “intelligence,” or consciousness they might be capable of remains to be seen. They show no evidence of emotions whatsoever. Furthermore, if consciousness can’t be programmed into a computer, we cannot be simulations in some future computer algorithm, contrary to ideas proposed by Bostrom<sup>[30]</sup> and Chalmers.<sup>[31]</sup> This ensures that our reality is really “real,” and should be especially reassuring to those who lie awake at night, fearful that they may exist as a Boltzmann’s brain that could disintegrate before morning. This is impossible because there is no way to permit the requisite reduction in thermodynamic entropy, especially in a condition of thermodynamic equilibrium, and no way to repay the consequent entropy debt, necessary for such a structure to become conscious.

#### **Can this New Thermodynamic Principle be Falsified?**

In principle, the Entropic Theory of the Emergence of Consciousness should be falsifiable, in keeping with Popperian conjectures and refutations, but probably not with current technology. One would have to be able to calculate the amount of heat emitted by the known activities of all the components of a network of neurons that are active during consciousness, both during consciousness and during states of unconsciousness. Then, one would need to be able to measure the actual heat emitted by those same neurons, say, during wakefulness and during deep sleep. My theory predicts that there would be a close match between the two results during deep sleep, but that during conscious activity, the measured amount of heat production would exceed the amount predicted. Furthermore, the possibility that the entropy debt may be repaid through the violation of a conserved quantity, not heat, cannot be excluded, as suggested by Vaccaro and Barnett.<sup>[32]</sup> Popper would have probably considered a proposal such as mine to be a “metaphysical research programme,” which can be evaluated by its logical consistency, explanatory power, problem-solving capacity, and successfulness empirically.<sup>[33]</sup> In this regard, it is worth remembering that the only theory to compete with emergentism in consciousness is the nonemergent “theory” of panpsychism. Panpsychism regards consciousness as a trait which has been possessed by all physical particles since the beginning of the universe, with no explanation as to how or why. It has the same explanatory power as the phrase “God wills it.” There is no conceivable test to falsify panpsychism at all! My proposal posits consciousness, and emergent phenomena in general, to be the result of an accepted scientific principle, entropy, that has existed since the beginning of the universe, and yet can account for the emergence of discrete, causally conscious living entities on an evolutionary basis at appropriate time points distant from the beginning of the universe. Unlike panpsychism, entropy requires ongoing

increases in complexity, and it is paid for. Also, it explains why consciousness must emerge, bind into functional units and evolve. It provides the best complement, in my opinion, to the various nonpanpsychist theories of mind, such as Global Neuronal Workspace Theory, Higher Order Theory, Conscious Electromagnetic Information Field Theory, and Orchestrated Objective Reduction, as they try to explain the cerebral cellular goings on, or neural correlates of consciousness.

## Conclusion

The emergence and convergence of consciousness are hidden deep within the computational workings of the brain by the inescapable tyranny of thermodynamic entropy and its irreversible tendency toward increasing diversity and disorder. As Colin McGinn noted, the problem is “cognitively closed” to us, but not because it’s too difficult for our limited minds.<sup>[34]</sup> Instead, it is impossible to understand, in principle. However, we can say that the situation is inconsistent with simple monism, but is an emergent dualism, specifically, a causally interactive dualism that is, at a hidden level, monist. There’s no need for the explanatory gap to be bridged by panpsychism. We must return to the question: “What possible alternative way of experiencing the world could there be?” as being the best answer to the body–mind question. I doubt it will ever be possible to characterize consciousness more descriptively than that. Indeed, I must conclude that the body–mind problem will never be solved until entropy can be defied, sometime after all the world’s broken eggs have reassembled, and all the world’s toothpaste has squeezed back into the tubes.

## Acknowledgments

I am grateful to Dr. Ruth E. Kastner for her correspondence concerning entropy and the relativistic transactional interpretation. Any deficiencies in this paper, nonetheless, remain my own responsibility.

## Patient informed consent

There is no need for patient informed consent.

## Ethics committee approval

There is no need for ethics committee approval.

## Financial support and sponsorship

No funding was received.

## Conflict of interest

There is no conflict of interest to declare.

## Author contribution subject and rate

- Peter Lugten (100%): Literature search, manuscript writing and editing.

## References

1. Çengel YA. The mind-brain problem from the perspective of agency. *J Neurobehav Sci* 2022;9:7-16. [doi: 10.4103/jnbs.

jnbs\_46\_21].

2. Popper KR, Eccles JC. *The Self and Its Brain: An Argument for Interactionism*. New York: Routledge; 1977.
3. Dennet DC. *Consciousness Explained*. Boston: Little, Brown and Company; 1991.
4. Damasio A. *The Feeling of What Happens. Body and Emotion in the Making of Consciousness*. New York: Harcourt Brace; 1999.
5. Demirel Y, Gerbaud V. *Nonequilibrium Thermodynamics*. 4<sup>th</sup> ed. Amsterdam: Elsevier; 2018.
6. Wheeler JA. *At Home in the Universe*. Melville, NY: The American Institute of Physics; 1994.
7. Wheeler JA, Ford KW. *Geons, Black Holes and Quantum Foam: A Life in Physics*. New York: W.W. Norton and Company, Inc.; 1998. p. 341.
8. Chalmers DJ. The problem of consciousness. *Discuss Philos* 2011;12:29-59.
9. Gleiser M. *The Island of Knowledge: The Limits of Science and the Search for Meaning*. New York: Public Affairs; 2015.
10. Carhart-Harris RL, Leech R, Hellyer PJ, Shanahan M, Feilding A, Tagliazucchi E, *et al*. The entropic brain: A theory of conscious states informed by neuroimaging research with psychedelic drugs. *Front Hum Neurosci* 2014;8:20. [doi: 10.3389/fnhum.2014.00020].
11. Shannon CE, Weaver W. *The Mathematical Theory of Communication*. Urbana: University of Illinois Press; 1998.
12. Davies PC. *The Demon in the Machine: How Hidden Webs of Information Are Solving the Mystery of Life*. Chicago: University of Chicago Press; 2019. p. 76.
13. Friston K. The free-energy principle: A unified brain theory? *Nat Rev Neurosci* 2010;11:127-38. [doi: 10.1038/nm2787].
14. Raichle ME, MacLeod AM, Snyder AZ, Powers WJ, Gusnard DA, Shulman GL. A default mode of brain function. *Proc Natl Acad Sci U S A* 2001;98:676-82. [doi: 10.1073/pnas.98.2.676].
15. Almheiri A. Black holes, wormholes and entanglement. *Sci Am* 2022;327:34-41. [doi: 10.1038/scientificamerican0922-28].
16. Kastner RE. On quantum non-unitarity as a basis for the second law of thermodynamics. *Entropy* 2017;19:106. [doi: 10.48550/arXiv.1612.08734].
17. Carroll S. *The Big Picture. On the Origins of Life, Meaning and the Universe Itself*. New York: Dutton, Random House; 2016.
18. Prigogine IR. *The End of Certainty: Time, Chaos and the New Laws of Physics*. New York: Free Press, Simon and Schuster; 1997.
19. Çengel YA. On entropy, information, and conservation of information. *Entropy (Basel)* 2021;23:779. [doi: 10.3390/e23060779].
20. Landauer R. Irreversibility and heat generation in the computing process. *IBM J* 1961;5:183-91.
21. Kastner RE, Schlatter A. The entropy cost of “Erasure” in physically irreversible processes. *Mathematics* 2024;12:206 [doi: 10.3390/math12020206].
22. Babcock NS, Montes-Cabrera G, Oberhofer KE, Chergui M, Celardo GL, Kurian P. Ultraviolet superradiance from meganetworks of tryptophan in biological architectures. 2023. [doi: 10.48550/arXiv.2302.01469].
23. Doyle RO. *The Information Philosopher: How is Information Created?* Available from: <https://www.informationphilosopher.com>. [Last accessed on 2021 Jun 03].
24. Penrose R. *Shadows of the Mind: A Search for the Missing Science of Consciousness*. New York: Oxford University Press; 1994.
25. Jones MW. Growing evidence that perceptual qualia are



- neuroelectrical not computational. *J Conscious Stud* 2019;26:89-116.
26. Jones MW, LaRock E. Are qualia computations or substances? *Mind Matter* 2023;21:89-126.
27. Clem RL, Cai DJ, Rajan K, Shuman T, Wu X, Wu Z. The Computational Brian. Washington D. C: The Frontiers of Medical Research: Brain Science (Science/AAAS); 2023. p. 19-20.
28. Panda R, López-González A, Gilson M, Gosseries O, Thibaut A, Frasso G, *et al.* Whole-brain analyses indicate the impairment of posterior integration and thalamo-frontotemporal broadcasting in disorders of consciousness. *Hum Brain Mapp* 2023;44:4352-71. [doi: 10.1002/hbm.26386].
29. Kriegman S, Blackiston D, Levin M, Bongard J. Kinematic self-replication in reconfigurable organisms. *Proc Natl Acad Sci U S A* 2021;118:e211267211810. [doi: 10.1073/pnas.2112672118].
30. Bostrom N. Are you living in a computer simulation? *Philos Q* 2003;53:243-55. [doi: 10.1111/1467-9213.00309].
31. Chalmers DJ. *Reality+: Virtual Worlds and the Problems of Philosophy*. New York: W.W. Norton and Company; 2022.
32. Vaccaro J, Barnett S. Information erasure without an energy cost. *Proc R Soc A* 2011;467:1770-8. [doi: 10.1098/rspa.2010.0577].
33. Popper KR. *The Unending Quest: An Intellectual Autobiography*. London: Routledge; 1976.
34. McGinn C. Can we solve the mind-body problem? *Mind* 1989;98:349-66. [doi: 10.1093/mind/XCVIII.391.349].