bioRxiv preprint doi: https://doi.org/10.1101/060467; this version posted October 28, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

frontiers

Neurobiology as Information Physics

Sterling Street[1]

[1]Independent Researcher, Athens, GA, USA

*** Correspondence:**
Sterling Street
sterling.street@uga.edu

Keywords: information thermodynamics$_1$, Landauer limit$_2$, free energy principle$_3$, optimization$_4$, Bekenstein bound$_5$.

Abstract

This article reviews thermodynamic relationships in the brain in an attempt to consolidate current research in systems neuroscience. The present synthesis supports proposals that thermodynamic information in the brain can be quantified to an appreciable degree of objectivity, that many qualitative properties of information in systems of the brain can be inferred by observing changes in thermodynamic quantities, and that many features of the brain's anatomy and architecture illustrate relatively simple information-energy relationships. The brain may provide a unique window into the relationship between energy and information.

Introduction

That information is physical has been suggested by evidence since the founding of classical thermodynamics (J Gleick 2011; S Lloyd 2006). In recent years, Landauer's principle (CH Bennett 2003; R Landauer 1996), which relates information-theoretic entropy to thermodynamic information, has been confirmed (JMR Parrondo et al. 2015), and the experimental demonstration of a form of information-energy equivalence (A Alfonso-Faus 2013) has verified that Maxwell's demon cannot violate any known laws of thermodynamics (K Maruyama et al. 2009). The theoretical finding that entropy is conserved as event horizon area is leading to the resolution of the black hole information paradox (P Davies 2010; C Moskowitz 2015), and there is a fundamental relationship between information and the geometry of spacetime itself (R Bousso 2002; C Eling et al. 2006). Current formulations of quantum theory are revealing properties of physical information (Č Brukner and A Zeilinger 2003; S Lloyd 2006; V Vedral 2010; J Wheeler 1986), and information-interpretive attempts to show that gravity is quantized (JW Lee et al. 2013; L Smolin 2001) could even lead to the unification of quantum mechanics and the theories of relativity. Although similar approaches are increasingly influential in biology (JL England 2013; JC Flack 2014; ED Schneider and D Sagan 2005), "a formalization of the relationship between information and energy is currently lacking in neuroscience" (G Collell and J Fauquet 2015). The purpose of this article is to explore a few different sides of this relationship and, along the way, to suggest that many hypotheses and theories in neuroscience can be unified by the physics of information.

bioRxiv preprint doi: https://doi.org/10.1101/060467; this version posted October 28, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Information bounds

"How can the events in space and time which take place within the spatial boundary of a living organism be accounted for by physics and chemistry?" – (E Schrödinger 1944, from KJ Friston 2013)

As a fundamental physical entity (S Lloyd 2015), information is not fully understood, and there is currently a significant amount of disagreement over different definitions of information and entropy in the literature (A Ben-Naim 2015; B Poirier 2014). In thermodynamics, however, information can be defined as a negation of thermodynamic entropy (C Beck 2009):

$$I \equiv -S$$

A bit of thermodynamic entropy represents the distinction between two alternative states in a physical system (JV Stone 2015). As a result, the total thermodynamic entropy of a system is proportional to the total number of distinguishable states contained in the system (JD Bekenstein 2001; JD Bekenstein 2007). Because thermodynamic entropy is potential information relative to an observer (S Lloyd 2006), and an observer in a physical system is a component of the system itself, the total thermodynamic entropy of a system includes the portion of entropy that is accessible to the observer as relative thermodynamic information (G Collell and J Fauquet 2015; J Wheeler 1989):

$$I_{relative} = S_{total} - S_{relative}$$

Since entropy in any physical system is finite (S Lloyd 2006; C Rovelli 2015), the total thermodynamic entropy of any system of the brain can be quantified by applying the traditional form of the universal (JD Bekenstein 1981, 1984, 2001, 2004, 2007) information-entropy bound:

$$S_{sys} = \zeta \frac{AEk}{\hbar c}$$

where A is area, E is energy including matter, $\hbar$ is the reduced Planck constant, c is the speed of light, k is Boltzmann's constant, and ζ is a factor such that $0 \leq \zeta \leq 1$

Setting this factor to 1 in order to quantify the total thermodynamic entropy of a system at a certain level of structure now allows us to quantify thermodynamic information by partitioning the factor into a relative information component ($\zeta_I = 1 - \zeta_S$) and a relative entropy component ($\zeta_S = 1 - \zeta_I$),

$$I_{sys} = \zeta_I \frac{AEk}{\hbar c} = (1 - \zeta_S) \frac{AEk}{\hbar c}$$

Because a maximal level of energy corresponds to a maximal level of thermodynamic information, and a minimal level of energy corresponds to a minimal level of thermodynamic information (TL Duncan and JS Semura 2004), any transitions between energy levels occur as transitions between informational extrema. So, in the event that information enters a system of the brain,

$$\Delta I_{sys} = \frac{\Delta E_{sys}}{kT} = \Delta \zeta_I$$

where T is temperature

and, in the event that information exits a system,

This is a provisional file, not the final typeset article

bioRxiv preprint doi: https://doi.org/10.1101/060467; this version posted October 28, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

$$-\Delta I_{sys} = \frac{\Delta E_{surr}}{kT} = \Delta \zeta_S$$

Various forms of these relationships, including information-entropy bounds, have been applied in neuroscience (G Collell and J Fauquet 2015; B Sengupta et al. 2013; B Sengupta and KJ Friston 2015; P Sterling and S Laughlin 2015). The contribution of this review is simply to show that these relationships can be united into a common theoretical framework.

Neurobiology

"… classical thermodynamics… is the only physical theory of universal content which I am convinced, that within the framework of applicability of its basic concepts, will never be overthrown." – (A Einstein 1949, from JD Bekenstein 2001)

This section reviews thermodynamic relationships in systems neuroscience with a focus on information and energy. Beginning with neurons, moving to neural networks, and concluding at the level of the brain as a whole, I discuss the energetics of processes such as learning and memory, excitation and inhibition, and the production of noise in neurobiological systems.

The central role of energy in determining the activity of neurons exposes the close connection between information and thermodynamics at the level of the cell. For instance, the process of depolarization, which occurs as a transition to E_{max} from a resting state E_{min}, clearly shows that cellular information content is correlated with energy levels. In this respect, the resemblance between ion concentration gradients in neurons and temperature gradients in thermodynamic demons (i.e., agents that use information from their surroundings to decrease their thermodynamic entropy) is not a coincidence – in order to acquire information, neurons must expend energy to establish proper membrane potentials. Recall that Landauer's principle (JMR Parrondo et al. 2015; MB Plenio and V Vitelli 2001) places a lower bound on the quantity of energy released into the surroundings with the removal of information from a system. Thus, reestablishing membrane potentials after depolarization – the neuronal equivalent of resetting a demon's memory – dissipates energy. Because Landauer's principle applies to all levels of structure, and cells process large quantities of information, neurons use energy efficiently despite operating at several orders of magnitude above the nominal limit. Parameters including membrane area, spiking frequency, and axon length have all been optimized over the course of evolution to allow neurons to process information efficiently (P Sterling and S Laughlin 2015). Examining the energetics of information processing in neurons reinforces the notion that, while it is often convenient to imagine the neuron to be a simple binary element, these cells are intricate computational structures that process more than one bit of information.

Relationships between information and energy can also be seen at the level of neural networks. Attractor networks naturally stabilize by seeking energy minima, and the relative positions of basins of attraction define the geometry of an energy landscape (DJ Amit 1992). As a result, the transition into an active attractor state occurs as a transition into an information-energy maximum. These transitions correspond to the generation of informational entities such as memories, decisions, and perceptual events (ET Rolls 2012). In this way, the energy basins of attractor networks may be analogous to lower-level cellular and molecular energy gradients; a transition between any number of distinguishable energy levels follows the passage of a finite quantity of information. Since processing information requires the expenditure of energy, competitive network features also underscore the need to minimize unnecessary information processing. Lateral inhibition at this level may optimize thermodynamic efficiency by reducing metabolic expenses associated with networks responding less robustly to entering signals. Another interesting thermodynamic property of networks concerns macrostates: the functional states of large-scale neural networks rest emergently on the states of neuronal assemblies (R Yuste 2015). As a result, new computational properties may arise with the

bioRxiv preprint doi: https://doi.org/10.1101/060467; this version posted October 28, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

addition of new layers of network structure. Finally, the energetic cost of information has influenced network connectivity by imposing selective pressures to save energy by minimizing path length between network nodes (E Bullmore and O Sporns 2009).

Again, in accordance with Landauer's principle, the displacement of information from any system releases energy into the surroundings (MB Plenio and V Vitelli 2001; TL Duncan and JS Semura 2004). This principle can be understood by imagining an idealized memory device, such as the brain of a thermodynamic demon. Since information is conserved (L Susskind and G Hrabovsky 2014), and clearing a memory erases information, the thermodynamic entropy of the surroundings must increase when a demon refreshes its memory to update information. This fundamental connection between information, entropy, and energy appears in many areas of the neurobiology of learning. For example, adjusting a firing threshold in order to change the probability that a system will respond to a conditioned stimulus (Y Choe 2015; T Takeuchi et al. 2014) optimizes engram fitness by minimizing the quantity of energy needed for its activation (S Still et al. 2012). Recurrent collateral connections further increase engram efficiency by enabling a minimal nodal stimulus to elicit its full energetic activation (ET Rolls 2012). Experimental evidence also shows that restricting synaptic energy supply impairs the formation of stable engrams (JJ Harris et al. 2012). Because the formation and disassembly of engrams during learning and forgetting optimizes the growth and pruning of networks in response to external conditions, the process of learning is itself a mechanism for minimizing entropy in the brain (KJ Friston 2003).

As another example of a multiscale process integrated across many levels by thermodynamics, consider the active balance between excitation and inhibition in neurobiological systems. Maintaining proper membrane potentials and adequate concentrations of signaling molecules requires the expenditure of energy, so it is advantageous for systems of the brain to minimize the processing of unnecessary information – to "send only what is needed" (P Sterling and S Laughlin 2015). Balancing excitation and inhibition is therefore a crucial mechanism for saving energy. Theoretical evidence that this balancing maximizes the thermodynamic efficiency of processing Shannon information (B Sengupta et al. 2013) is consistent with experimental findings in several areas of research on inhibition. For instance, constant inhibitory modulation is needed to stabilize internal states, and hyperexcitation (e.g., in epilepsy, intoxication syndromes, or trauma) can decrease relative information by reducing levels of consciousness (B Haider et al. 2006; K Lehmann et al. 2012). Likewise, selective attention is mediated by the activation of inhibitory interneurons (G Houghton and SP Tipper 1996), and sensory inhibition appears to sharpen internal perceptual states (JS Isaacson and M Scanziani 2011). The need to balance excitation and inhibition at all levels of structure highlights the energetic cost of information.

A final example worth discussing is the relationship between thermodynamics and the production of noise in neurobiological systems. Noise is present in every system of the brain, and influences all aspects of the organ's function (AA Faisal et al. 2008; ET Rolls and G Deco 2010; A Destexhe and M Rudolph-Lilith 2012). Even in the absence of any potential forms of classical stochastic resonance, the noise-driven exploration of different states may optimize thermodynamic efficiency by allowing a system to randomly sample different accessible configurations. Theoretical arguments suggest indeed that noise enables neural networks to respond more quickly to detected signals (ET Rolls 2012), and empirical evidence implicates noise as a beneficial means of optimizing the performance of diverse neurobiological processes (MD McDonnell and LM Ward 2011). For example, noise in the form of neuronal DNA breaking (JU Guo et al. 2011; K Herrup et al. 2013; P Tognini et al. 2015) could enhance plasticity, since any stochastically optimized configuration would be more likely to survive over time as, in this case, a strengthened connection in a modifiable network. Because noise is a form of relative entropy, optimizing the signal-to-noise ratio in any neurobiological system promotes the efficient use of energy.

At the level of the brain as a whole, the connection between information and thermodynamics

This is a provisional file, not the final typeset article

bioRxiv preprint doi: https://doi.org/10.1101/060467; this version posted October 28, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

is readily apparent in the organ's functional reliance on energy (PJ Magistretti and I Allaman 2015), its seemingly disproportionate consumption of oxygen and energy substrates (e.g., ATP, glucose, ketones, etc.) (S Herculano-Houzel 2011; ME Raichle and DA Gusnard 2002), its vulnerability to hypoxic-ischemic damage (JP Dreier et al. 2013; PL Lutz et al. 2003) and in the reduction of consciousness often conferred by the onset of energy restrictions (RG Shulman et al. 2009; J Stender et al. 2016). All fMRI, PET, and EEG interpretation rests on the foundational assumption that changes in the information content of neurobiological systems can be inferred by observing energy changes (D Attwell and C Iadecola 2002; G Collell and J Fauquet 2015), and it is well known that the information processing capacities of neurobiological systems are limited by energy supply (C Howarth et al. 2012). Overall, these relationships are consistent with the form of information-energy equivalence predicted by Landauer's principle and information-entropy bounds. The living brain appears to maintain a state of thermodynamic optimization.

Consciousness and free will

"... science appears completely to lose from sight the large and general questions; but all the more splendid is the success when, groping in the thicket of special questions, we suddenly find a small opening that allows a hitherto undreamt of outlook on the whole." – (L Boltzmann 1886, from HC Von Baeyer 1999)

Although neuroscience has yet to explain consciousness or free will at any satisfactory level of detail, relationships between information and energy seem to be recognizable even at this level of analysis. This section reviews attempts to conceptualize major properties of consciousness (unity, continuity, complexity, and self-awareness) as features of information processing in the brain, and concludes with a discussion of free will.

At any given moment, awareness is experienced as a unified whole. Physical information is the substrate of consciousness (A Annila 2016), and the law of conservation of information requires any minimal unit of information to be transferred into a thermodynamic system as a temporally unitary quantity. As a result, it is possible that the passage of perceptual time itself occurs secondarily to the transfer of information, and that the information present in any integrated system of the brain at any observed time is necessarily cohesive and temporally unified. In this framework, the passage of time would vary in proportion to a system's rate of energy dissipation. Although it is possible that physical systems in general exchange information in temporally unitary quantities, it is likely that many of the familiar features of the perceptual unity of consciousness require the structure and activity of neural networks in the brain. The biological basis of this unity may be the active temporal consolidation of observed events by integrated higher-order networks (S Dehaene and JP Changeux 2011; SA Greenfield and TFT Collins 2005; A Revonsuo 1999; F Varela et al. 2001). An informational structure generated by the claustrum has been speculated to contribute to this experiential unity (FC Crick and C Koch 2005, MZ Koubeissi et al. 2014), but it has also been reported that complete unilateral resection of the system performed in patients with neoplastic lesions of the region produces no externally observable changes in subjective awareness (H Duffau et al. 2007). Overall, it appears unlikely that the presence of information in any isolated or compartmentalized network of the brain is responsible for generating the unified nature of conscious experience.

While perceptual time is likely the product of a collection of related informational processes rather than a single, globalized function mediated by any one specific system of the brain, some of the perceptual continuity of consciousness may result from the effectively continuous flow of thermodynamic information into and out of integrated systems of the brain. In this framework, the quantum (M Prokopenko et al. 2014) of perceptual time would be the minimal acquisition of

bioRxiv preprint doi: https://doi.org/10.1101/060467; this version posted October 28, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

information, and the entrance of information into neurobiological systems would occur alongside the entrance of energy. This relationship is implicit in the simple observation that the transition of a large-scale attractor network is progressively less discrete and smoother in time than the activation of a small-scale engram, the propagation of a cellular potential, the docking of a vesicle, the release of an ion, and so forth. Likewise, electroencephalography shows that the summation of a large number of discrete cellular potentials can accumulate into an effectively continuous wave as a network field potential (PL Nunez and R Srinivasan 2006), disruptions of which are often correlated with decreases in levels of consciousness (H Blumenfeld and J Taylor 2003). It is also well known that higher frequency network oscillations tend to indicate states of wakefulness and active awareness, while lower frequency oscillations tend to be associated with internal states of lesser passage of perceptual time, such as dreamless sleep or unconsciousness. The possibility that the experiential arrow of time and the thermodynamic arrow of time share a common origin in the flow of information is supported both by general models of time in neuroscience and the physical interpretation of time as an entropy gradient (L Mlodinow and TA Brun 2014; OC Stoica 2008).

The subjective complexity of consciousness may show that extensive network integration is needed for maximizing the mutual thermodynamic information and internal energy content of systems of the brain (JS Torday and WB Miller Jr 2016). An exemplary structure enabling such experience, likely one of many that together account for the subjective complexity of consciousness, is the thalamocortical complex (Y Hannawi et al. 2015; RS Calabrò et al. 2015). The functional architecture of such a network may show that, at any given moment in the internal model of a living brain, a wide range of integrated systems are sharing mutual sources of thermodynamic information. This pattern of structure may reveal that the perceptual depth and complexity of conscious experience is a direct product of recognizable features of the physical brain. However, it also seems that extensive local cortical processing of information is necessary for producing a refined and coherent sensorium within a system, and that both the thalamocortical complex and the brain stem are involved in generating the subjective complexity of consciousness (GM Edelman et al. 2011; LM Ward 2011). The dynamics of attractor networks at higher levels of network structure may show that quantities of complex internal information can be observed as changes in cortical energy landscapes (ET Rolls 2012), with a transition between attractor states following the transfer of information. The degree of subjective complexity of information enclosed by such a transition would be proportional to the degree of structural integration of underlying networks.

Self-awareness likely arose as a survival necessity rather than as an accident of evolution (F Fabbro et al. 2015), and rudimentary forms of self-awareness likely began to appear early in the course of brain evolution as various forms of perceptual self-environment separation. As a simple example, consider the tickle response (DJ Linden 2007), which requires the ability to differentiate self-produced tactile sensations from those produced by external systems. The early need to distinguish between self-produced tactile states and those produced by more threatening non-self sources may be reflected by the observation that this recognition process is mediated to a great extent by the cerebellum (SJ Blakemore et al. 2000). While it is possible that other similar developments began occurring very early on, the evolutionary acquisition of the refined syntactical and conceptual self present in the modern brain likely required the merging of pre-existing self networks with higher-level cortical systems. The eventual integration of language and self-awareness would have been advantageous for coordinating social groups (MS Graziano 2013), since experiencing self-referential thought as inner speech facilitates verbal communication. Likewise, the coupling of self-awareness to internal sensory, cognitive, and motor states (T Metzinger 2004; G Northoff et al. 2006) may be advantageous for maximizing information between systems within an individual brain. Neuropsychological conditions involving different forms of agnosia, neglect, and self-awareness deficits do show that a reduced awareness of self-ownership of motor skills, body parts, or perceptual states can result in significant disability (F Fabbro et al. 2015; S Chokron et al. 2016; MD Orfei et al.

This is a provisional file, not the final typeset article

bioRxiv preprint doi: https://doi.org/10.1101/060467; this version posted October 28, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

2007; M Overgaard 2011; A Parton et al. 2004; M Tsakiris 2010; A Morin 2006; GP Prigatano 2009). Since experiencing self-awareness optimizes levels of mutual information between the external world and the brain's internal model (MA Apps and M Tsakiris 2014), and this activity decreases thermodynamic entropy (JS Torday and WB Miller Jr 2016), self-awareness may be a mechanism for optimizing the brain's consumption of energy.

Thermodynamic information is also interesting to consider in the context of free will. The brain is predictable within reason, and the performance of an action can be predicted before a decision is reported to have been made (P Haggard 2008). Entities such as ideas, feelings, and beliefs seem to exist as effectively deterministic evaluations of information processed in the brain. Whether or not the flow of information is subject to the brain's volitional alteration, neuroscience also shows that information can be internally real to a system of the brain, even if this information is inconsistent with an external reality. That the brain can generate an externally inconsistent internal reality is demonstrated by phenomena such as confabulation, agnosia, blindsight, neglect, commissurotomy and hemispherectomy effects, placebo and nocebo effects, reality monitoring deficits, hallucinations, prediction errors, the suspension of disbelief during dreaming, the function of communication in minimizing divergence between internal realities, the quality of many kinds of realistic drug-induced experiences, and the effects of many neuropsychological conditions. The apparent fact that subjective reality is an active construction of the physical brain has even led to the proposal of model-dependent realism (SW Hawking and L Mlodinow 2011) as a philosophical paradigm in the search for a unified theory of physics. In any case, it is likely that beliefs, including those in free will, exist as information, and that their internal reality is a restatement of its frequently observer-dependent nature.

Empirical outlook

Before concluding, it is worth reviewing a few notable experiments in greater detail. While considerable advances have been made in discovering how neurobiological systems operate according to principles of thermodynamic efficiency (S Laughlin and P Sterling 2015), relationships between information and energy in the brain are only beginning to be understood. The following studies are examples of elegant and insightful experiments that should inspire future research.

Several recent brain imaging studies support the proposal (A Annila 2016) that thermodynamics is able to explain a number of mysteries involving consciousness. For example, J Stender et al. 2016 used PET to measure global resting state energy consumption in 131 brain injury patients with impairments of consciousness as defined by the revised Coma Recovery Scale (CRS-R). The preservation of consciousness was found to require a minimal global metabolic rate of ≈ 40% of the average rate of controls; global energy consumption above this level was reported to predict the presence or recovery of consciousness with over 90% sensitivity. These results must be replicated and studied in closer detail before their specific theoretical implications are clear, but it is now established that levels of consciousness are correlated with energetic metrics of brain activity. To what extent there exists a well-defined "minimal energetic requirement for the presence of conscious awareness" (J Stender et al. 2016) remains an open question. However, the empirical confirmation of a connection between consciousness and thermodynamics introduces the possibility of developing new experimental methods in consciousness research.

Neurobiological systems, and biological systems in general (HC Von Baeyer 1999; ED Schneider and D Sagan 2005), can be considered thermodynamic demons in the sense that they are agents using information to decrease their thermodynamic entropy. Landauer's principle requires that, in order not to violate any known laws of thermodynamics, such agents dissipate heat when erasing information from their memory storage devices. In an experimental test of this principle, reviewed along with similar experiments in JMR Parrondo et al. 2015, A Bérut et al. 2012 studied

bioRxiv preprint doi: https://doi.org/10.1101/060467; this version posted October 28, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

heat dissipation in a simple memory device created by placing a glass bead in an optical double-well potential. Intuitively, this memory stored a bit of information by retaining the bead on one side of the potential rather than on the alternative. By manipulating the height of the optical barrier between wells, researchers moved the bead to one side of the memory without determining its previous location in the potential. This process was therefore logically irreversible, requiring the erasure of prior information from the memory device. Landauer's principle predicts that, since information is conserved, the entropy of the memory's surroundings must increase when this occurs. A Bérut et al. 2012 have verified that energy is emitted when a memory is cleared. As noted by the authors, "this limit is independent of the actual device, circuit or material used to implement the irreversible operation." It would be interesting to study the erasure principle in the context of neuroscience.

Experimental applications of information theory in cell biology have already led to the discovery of general principles of brain organization related to thermodynamics (S Laughlin and P Sterling 2015). In one particularly interesting study, JE Niven et al. 2007 measured the energetic efficiency of information coding in retinal neurons. Intracellular recordings of membrane potential and input resistance were used to calculate rates of ATP consumption in response to different background light intensities. These rates of energy consumption were then compared with rates of Shannon information transmission in order to determine metabolic performance. It was found that metabolic demands increase nonlinearly with respect to increases in information processing rate: thermodynamics appears to impose a "law of diminishing returns" on systems of the brain. The authors interpret these results as evidence that nature has selected for neurons that minimize unnecessary information processing. Studying how thermodynamics has influenced cellular parameters over the course of evolution is likely to raise many new empirically addressable questions.

Conclusion

This article has reviewed information-energy relationships in the hope that they may eventually provide a general framework for uniting theory and experiment in neuroscience. The physical nature of information and its status as a finite, measurable resource are emphasized to connect neurobiology and thermodynamics. As a scientific paradigm, the information movement currently underway in physics promises profound advances in our understanding of the relationship between energy, information, and the physical brain.

Conflict of Interest

The author confirms that this research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgements

I am grateful to Baroness Susan Greenfield, Dr. Francesco Fermani, Dr. Karl Friston, Dr. Biswa Sengupta, Dr. Roy Frieden, Dr. Bernard Baars, Dr. Brett Clementz, Dr. Cristi Stoica, Dr. Satoru Suzuki, Dr. Paul King, Guillem Collell Talleda, Dr. Jordi Fauquet, and others who have helped me improve these ideas. I am also grateful to Dr. Shanta Dhar and her team for introducing me to academic research, to Jim Reid for introducing me to biology, to Alex Tisch for introducing me to physics, and to those affiliated with the Department of Neurosurgery at the University of Virginia Medical Center for introducing me to neuroscience.

References

This is a provisional file, not the final typeset article

bioRxiv preprint doi: https://doi.org/10.1101/060467; this version posted October 28, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Alfonso-Faus, A. (2013). Fundamental Principle of Information-to-Energy Conversion. *Proceedings of the 7th European Computing Conference.* WSEAS Press.

Amit, DJ. (1992). *Modeling Brain Function.* Cambridge: Cambridge University Press.

Annila, A. (2016). On the character of consciousness. *Frontiers in Systems Neuroscience, 10*. doi: 10.3389/fnsys.2016.00027.

Apps, M. A., and Tsakiris, M. (2014). The free-energy self: a predictive coding account of self-recognition. *Neuroscience & Biobehavioral Reviews, 41*, 85-97. doi: 10.1016/j.neubiorev.2013.01.029.

Attwell, D., and Iadecola, C. (2002). The neural basis of functional brain imaging signals. *Trends in Neurosciences, 25*(12), 621-625. doi:10.1016/S0166-2236(02)02264-6

Beck, C. (2009). Generalised information and entropy measures in physics. *Contemporary Physics, 50*(4), 495-510. doi: 10.1080/00107510902823517

Bekenstein, J. D. (1981). Universal upper bound on the entropy-to-energy ratio for bounded systems. *Physical Review D, 23*(2), 287. doi:10.1103/PhysRevD.23.287

Bekenstein, J. D. (1984). Entropy content and information flow in systems with limited energy. *Physical Review D, 30*(8), 1669. doi:10.1103/PhysRevD.30.1669

Bekenstein, J. D. (2001). The limits of information. *Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics, 32*(4), 511-524. doi: 10.1016/S1355-2198(01)00020-X

Bekenstein, J. D. (2004). Black holes and information theory. *Contemporary Physics, 45*(1). PACS: 89.70.+c,03.67.-a,04.70.-s,04.70.Dy,65.40.Gr

Bekenstein, J. D. (2007). Information in the holographic universe. *Scientific American, 17,* 66-73.

Ben-Naim, A. (2015). *Information, Entropy, Life and the Universe.* Singapore: World Scientific. 4-5.

Bennett, C. H. (2003). Notes on Landauer's principle, reversible computation, and Maxwell's Demon. *Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics, 34*(3), 501-510. doi:10.1016/S1355-2198(03)00039-X

Bérut, A., Arakelyan, A., Petrosyan, A., Ciliberto, S., Dillenschneider, R., and Lutz, E. (2012). Experimental verification of Landauer/'s principle linking information and thermodynamics. *Nature, 483*(7388), 187-189. doi: 10.1038/nature10872

Blakemore, S. J., Wolpert, D., and Frith, C. (2000). Why can't you tickle yourself? *Neuroreport, 11*(11), R11-R16. doi: 10.1097/00001756-200008030-00002

Blumenfeld, H., and Taylor, J. (2003). Why do seizures cause loss of consciousness? *The Neuroscientist, 9*(5), 301-310. doi: 10.1177/1073858403255624

Bousso, R. (2002). The holographic principle. *Reviews of Modern Physics, 74*(3), 825. doi: 10.1103/RevModPhys.74.825

bioRxiv preprint doi: https://doi.org/10.1101/060467; this version posted October 28, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Brukner, Č., and Zeilinger, A. (2003). Information and fundamental elements of the structure of quantum theory. In *Time, Quantum and Information*, ed. L. Castell and O. Ischebeck. Heidelberg: Springer. 323-354.

Bullmore, E., and Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience, 10*(3), 186-198. doi: 10.1038/nrn2575

Calabrò, R. S., Cacciola, A., Bramanti, P., and Milardi, D. (2015). Neural correlates of consciousness: what we know and what we have to learn! *Neurological Sciences, 36*(4), 505-513. doi: 10.1007/s10072-015-2072-x

Choe, Y. (2015). Hebbian Learning. In *Encyclopedia of Computational Neuroscience*, ed. D Jaeger and R Jung. New York: Springer. 1305-1309. doi: 10.1007/978-1-4614-7320-6_672-1

Chokron, S., Perez, C., and Peyrin, C. (2016). Behavioral Consequences and Cortical Reorganization in Homonymous Hemianopia. *Frontiers in Systems Neuroscience, 10*. doi: 10.3389/fnsys.2016.00057

Collell, G., and Fauquet, J. (2015). Brain activity and cognition: a connection from thermodynamics and information theory. *Frontiers in Psychology, 6*. doi: 10.3389/fpsyg.2015.00818

Crick, F. C., and Koch, C. (2005). What is the function of the claustrum? *Philosophical Transactions of the Royal Society B: Biological Sciences, 360*(1458), 1271-1279. doi: 10.1098/rstb.2005.1661

Davies, P. (2010). Universe from bit. In *Information and the Nature of Reality*, ed. P Davies and N. H. Gregersen. Cambridge: Cambridge University Press. 83-117.

Dehaene, S., and Changeux, J. P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron, 70*(2), 200-227. doi: 10.1016/j.neuron.2011.03.018

Destexhe, A., and Rudolph-Lilith, M. (2012). *Neuronal Noise*. New York: Springer.

Dreier, J. P., Isele, T., Reiffurth, C., Offenhauser, N., Kirov, S. A., Dahlem, M. A., and Herreras, O. (2013). Is spreading depolarization characterized by an abrupt, massive release of Gibbs free energy from the human brain cortex? *The Neuroscientist, 19*(1), 25-42. doi: 10.1177/1073858412453340.

Duffau, H., Mandonnet, E., Gatignol, P., and Capelle, L. (2007). Functional compensation of the claustrum: lessons from low-grade glioma surgery. *Journal of Neuro-oncology, 81*(3), 327-329. doi: 10.1007/s11060-006-9236-8

Duncan, T. L., and Semura, J. S. (2004). The deep physics behind the second law: information and energy as independent forms of bookkeeping. *Entropy, 6*(1), 21-29. doi:10.3390/e6010021

Edelman, G. M., Gally, J. A., and Baars, B. J. (2011). Biology of consciousness. *Frontiers in Psychology, 2*, 4.

Eling, C., Guedens, R., and Jacobson, T. (2006). Nonequilibrium thermodynamics of spacetime. *Physical Review Letters, 96*(12), 121301. doi: 10.1103/PhysRevLett.96.121301.

England, J. L. (2013). Statistical physics of self-replication. *The Journal of Chemical Physics, 139*(12), 121923. doi: 10.1063/1.4818538

This is a provisional file, not the final typeset article

bioRxiv preprint doi: https://doi.org/10.1101/060467; this version posted October 28, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Fabbro, F., Aglioti, S. M., Bergamasco, M., Clarici, A., and Panksepp, J. (2015). Evolutionary aspects of self-and world consciousness in vertebrates. *Frontiers in Human Neuroscience, 9*, 157. doi: 10.3389/fnhum.2015.00157

Faisal, A. A., Selen, L. P., and Wolpert, D. M. (2008). Noise in the nervous system. *Nature Reviews Neuroscience, 9*(4), 292-303. doi: doi:10.1038/nrn2258

Flack, J. C. (2014). Life's information hierarchy. *Santa Fe Institute Bulletin, 28*, 13-24.

Fox, D. (2015). The limits of intelligence. *Scientific American, 305*(1), 36-43.

Friston, K. J. (2003). Learning and inference in the brain. *Neural Networks, 16*(9), 1325-1352. doi:10.1016/j.neunet.2003.06.005

Friston, K. J. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience, 11*(2), 127-138. doi: 10.1038/nrn2787

Friston, K. J. (2013). Life as we know it. *Journal of The Royal Society Interface, 10*(86), 20130475. doi: 10.1098/rsif.2013.0475

Gleick, J., (2011). *The Information*. New York: Random House. 269-270; 355-372.

Graziano, M. S. (2013). *Consciousness and the Social Brain.* Oxford: Oxford University Press.

Greenfield, S. A., and Collins, T. F. T. (2005). A neuroscientific approach to consciousness. *Progress in Brain Research, 150*, 11-23. doi:10.1016/S0079-6123(05)50002-5

Guo, J. U., Ma, D. K., Mo, H., Ball, M. P., Jang, M. H., Bonaguidi, M. A., et al. (2011). Neuronal activity modifies the DNA methylation landscape in the adult brain. *Nature Neuroscience, 14*(10), 1345-1351. doi: 10.1038/nn.2900

Haggard, P. (2008). Human volition: towards a neuroscience of will. *Nature Reviews Neuroscience, 9*(12), 934-946. doi:10.1038/nrn2497

Haider, B., Duque, A., Hasenstaub, A. R., and McCormick, D. A. (2006). Neocortical network activity in vivo is generated through a dynamic balance of excitation and inhibition. *The Journal of Neuroscience, 26*(17), 4535-4545. doi: 10.1523/JNEUROSCI.5297-05.2006

Hannawi, Y., Lindquist, M. A., Caffo, B. S., Sair, H. I., and Stevens, R. D. (2015). Resting brain activity in disorders of consciousness. *Neurology, 84*(12), 1272-1280. doi: 10.1212/WNL.0000000000001404

Harris, J. J., Jolivet, R., and Attwell, D. (2012). Synaptic energy use and supply. *Neuron, 75*(5), 762-777. doi: 10.1016/j.neuron.2012.08.019

Hawking, S. W. and Mlodinow, L. (2011). *The Grand Design*. New York: Random House. 7, 46.

Herculano-Houzel, S. (2011). Scaling of brain metabolism with a fixed energy budget per neuron: implications for neuronal activity, plasticity and evolution. *PLoS One, 6*(3), e17514. doi: 10.1371/journal.pone.0017514

bioRxiv preprint doi: https://doi.org/10.1101/060467; this version posted October 28, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Herrup, K., Chen, J., and Li, J. (2013). Breaking news: thinking may be bad for DNA. *Nature Neuroscience, 16*(5), 518-519. doi: 10.1038/nn.3384

Houghton, G., and Tipper, S. P. (1996). Inhibitory mechanisms of neural and cognitive control: Applications to selective attention and sequential action. *Brain and Cognition, 30*(1), 20-43. doi: 10.1006/brcg.1996.0003

Howarth, C., Gleeson, P., and Attwell, D. (2012). Updated energy budgets for neural computation in the neocortex and cerebellum. *Journal of Cerebral Blood Flow and Metabolism, 32*(7), 1222-1232. doi: 10.1038/jcbfm.2012.35

Isaacson, J. S., and Scanziani, M. (2011). How inhibition shapes cortical activity. *Neuron, 72*(2), 231-243. doi: 10.1016/j.neuron.2011.09.027

Koubeissi, M. Z., Bartolomei, F., Beltagy, A., and Picard, F. (2014). Electrical stimulation of a small brain area reversibly disrupts consciousness. *Epilepsy & Behavior, 37*, 32-35. doi: 10.1016/j.yebeh.2014.05.027

Landauer, R. (1996). The physical nature of information. *Physics Letters A*, 217(4), 188-193. doi: 10.1016/0375-9601(96)00453-7

Lee, J. W., Kim, H. C., and Lee, J. (2013). Gravity from quantum information. *Journal of the Korean Physical Society, 63*(5), 1094-1098. doi: 10.3938/jkps.63.1094

Lehmann, K., Steinecke, A., and Bolz, J. (2012). GABA through the ages: regulation of cortical function and plasticity by inhibitory interneurons. *Neural Plasticity*, 892784. doi: 10.1155/2012/892784

Linden, D. J. (2007). *The Accidental Mind.* Cambridge: Harvard University Press. 9-12.

Liu, K. C., Oztaskin, M., and Burchiel, K. J. (2012). Basics of neurosurgical techniques and procedures. In *Essentials of Neurosurgical Anesthesia and Critical Care*. New York: Springer. 145. doi: 10.1007/978-0-387-09562-2_14

Lloyd, S. (2006). *Programming the Universe.* New York: Random House.

Lloyd, S. (2015). Interview in *Closer to Truth: Is Information Fundamental?* Retrieved from https://www.closertotruth.com/series/information-fundamental#video-2621

Lutz, P. L., Nilsson, G. E., and Prentice, H. M. (2003). *The Brain Without Oxygen.* New York: Kluwer.

Magistretti, P. J., and Allaman, I. (2015). A cellular perspective on brain energy metabolism and functional imaging. *Neuron, 86*(4), 883-901. doi: 10.1016/j.neuron.2015.03.035

Maruyama, K., Nori, F., and Vedral, V. (2009). Colloquium: The physics of Maxwell's demon and information. *Reviews of Modern Physics, 81*(1), 1. doi: 10.1103/RevModPhys.81.1

McDonnell, M. D., and Ward, L. M. (2011). The benefits of noise in neural systems: bridging theory and experiment. *Nature Reviews Neuroscience, 12*(7), 415-426. doi: 10.1038/nrn3061

This is a provisional file, not the final typeset article

bioRxiv preprint doi: https://doi.org/10.1101/060467; this version posted October 28, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Metzinger, T. (2004). *Being No One*. Cambridge: MIT Press.

Mlodinow, L., and Brun, T. A. (2014). Relation between the psychological and thermodynamic arrows of time. *Physical Review E, 89*(5), 052102. doi: 10.1103/PhysRevE.89.052102

Morin, A. (2006). Levels of consciousness and self-awareness: A comparison and integration of various neurocognitive views. *Consciousness and Cognition, 15*(2), 358-371. doi: 10.1016/j.concog.2005.09.006

Moskowitz, C. (2015). Stephen Hawking Hasn't Solved the Black Hole Paradox Just Yet. *Scientific American, 27*.

Niven, J. E., Anderson, J. C., and Laughlin, S. B. (2007). Fly photoreceptors demonstrate energy-information trade-offs in neural coding. *PLoS Biology, 5*(4), e116. doi: 10.1371/journal.pbio.0050116

Northoff, G., Heinzel, A., De Greck, M., Bermpohl, F., Dobrowolny, H., and Panksepp, J. (2006). Self-referential processing in our brain—a meta-analysis of imaging studies on the self. *Neuroimage, 31*(1), 440-457. doi: 10.1016/j.neuroimage.2005.12.002

Nunez, P. L., and Srinivasan, R. (2006). *Electric Fields of the Brain*. New York: Oxford University Press.

Orfei, M. D., Robinson, R. G., Prigatano, G. P., Starkstein, S., Rüsch, N., Bria, P., et al. (2007). Anosognosia for hemiplegia after stroke is a multifaceted phenomenon: a systematic review of the literature. *Brain, 130*(12), 3075-3090. doi: 10.1093/brain/awm106

Overgaard, M. (2011). Visual experience and blindsight: a methodological review. *Experimental Brain Research, 209*(4), 473-479. doi: 10.1007/s00221-011-2578-2

Parrondo, J. M. R., Horowitz, J. M., and Sagawa, T. (2015). Thermodynamics of information. *Nature Physics, 11*(2), 131-139. doi: 10.1038/nphys3230

Parton, A., Malhotra, P., and Husain, M. (2004). Hemispatial neglect. *Journal of Neurology, Neurosurgery & Psychiatry, 75*(1), 13-21.

Plenio, M. B., and Vitelli, V. (2001). The physics of forgetting: Landauer's erasure principle and information theory. *Contemporary Physics, 42*(1), 25-60. doi: 10.1080/00107510010018916

Poirier, B. (2014). *A Conceptual Guide to Thermodynamics*. Chichester: Wiley. 77-78.

Prigatano, G. P. (2009). Anosognosia: clinical and ethical considerations. *Current Opinion in Neurology, 22*(6), 606-611. doi: 10.1097/WCO.0b013e328332a1e7

Prokopenko, M., and Lizier, J. T. (2014). Transfer entropy and transient limits of computation. *Scientific Reports, 4*. doi: 10.1038/srep05394

Raichle, M. E., and Gusnard, D. A. (2002). Appraising the brain's energy budget. *Proceedings of the National Academy of Sciences, 99*(16), 10237-10239. doi: 10.1073/pnas.172399499

Revonsuo, A. (1999). Binding and the phenomenal unity of consciousness. *Consciousness and Cognition, 8*(2), 173-185. doi: 10.1006/ccog.1999.0384

bioRxiv preprint doi: https://doi.org/10.1101/060467; this version posted October 28, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Rolls, E. T. (2012). *Neuroculture.* Oxford: Oxford University Press.

Rolls, E. T., and Deco, G. (2010). *The Noisy Brain.* Oxford: Oxford University Press.

Rovelli, C. (2015). Relative information at the foundation of physics. In *It From Bit or Bit From It?* ed. A Aguirre, B. Foster, and Z. Merali. Cham: Springer. doi: 10.1007/978-3-319-12946-4_7

Schneider, E. D., and Sagan, D. (2005). *Into the Cool.* Chicago: University of Chicago Press.

Sengupta, B., and Friston, K. J. Personal communication, October 7th, 2015.

Sengupta, B., Faisal, A. A., Laughlin, S. B., and Niven, J. E. (2013). The effect of cell size and channel density on neuronal information encoding and energy efficiency. *Journal of Cerebral Blood Flow & Metabolism, 33*(9), 1465-1473. doi: 10.1038/jcbfm.2013.103

Sengupta, B., Laughlin, S. B., and Niven, J. E. (2013). Balanced excitatory and inhibitory synaptic currents promote efficient coding and metabolic efficiency. *PLoS Comput Biology, 9*(10), e1003263. doi: 10.1371/journal.pcbi.1003263

Sengupta, B., Stemmler, M. B., and Friston, K. J. (2013). Information and efficiency in the nervous system—a synthesis. *PLoS Computational Biology, 9*(7), e1003157. doi: 10.1371/journal.pcbi.1003157

Sengupta, B., Tozzi, A., Cooray, G. K., Douglas, P. K., and Friston, K. J. (2016). Towards a neuronal gauge theory. *PLoS Biology, 14*(3), e1002400. doi: 10.1371/journal.pbio.1002400

Shulman, R. G., Hyder, F., and Rothman, D. L. (2009). Baseline brain energy supports the state of consciousness. *Proceedings of the National Academy of Sciences, 106*(27), 11096-11101. doi: 10.1073/pnas.0903941106

Smolin, L. (2002). *Three Roads to Quantum Gravity.* New York: Basic Books. 103, 169-178.

Stender, J., Mortensen, K. N., Thibaut, A., Darkner, S., Laureys, S., Gjedde, A., and Kupers, R. (2016). The minimal energetic requirement of sustained awareness after brain injury. *Current Biology, 26*(11), 1494-1499. doi: 10.1016/j.cub.2016.04.024

Sterling, P., and Laughlin, S. (2015). *Principles of Neural Design.* Cambridge: MIT Press.

Still, S., Sivak, D. A., Bell, A. J., and Crooks, G. E. (2012). Thermodynamics of prediction. *Physical Review Letters, 109*(12), 120604. doi: 10.1103/PhysRevLett.109.120604

Stoica, O. C. (2008). "Flowing with a frozen river." FQXi, *The Nature of Time* essay contest.

Stone, J. V. (2015). *Information Theory*. Sheffield: Sebtel Press. 171.

Susskind, L., and G Hrabovsky (2014). *The Theoretical Minimum.* New York: Basic Books. 9; 170.

Takeuchi, T., Duszkiewicz, A. J., and Morris, R. G. (2014). The synaptic plasticity and memory hypothesis: encoding, storage and persistence. *Philosophical Transactions of the Royal Society B, 369*(1633), 20130288. doi: 10.1098/rstb.2013.0288

This is a provisional file, not the final typeset article

bioRxiv preprint doi: https://doi.org/10.1101/060467; this version posted October 28, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

555 Tognini, P., et al. (2015). Dynamic DNA methylation in the brain: a new epigenetic mark for
556 experience-dependent plasticity. *Frontiers in Cellular Neuroscience, 9*. doi:
557 10.3389/fncel.2015.00331

558 Torday, J. S., and Miller Jr, W. B. (2016). On the evolution of the mammalian brain. *Frontiers in*
559 *Systems Neuroscience, 10*. doi: 10.3389/fnsys.2016.00031

560 Tsakiris, M. (2010). My body in the brain: a neurocognitive model of body-ownership.
561 *Neuropsychologia, 48*(3), 703-712. doi: 10.1016/j.neuropsychologia.2009.09.034

562 Varela, F., Lachaux, J. P., Rodriguez, E., and Martinerie, J. (2001). The brainweb: phase
563 synchronization and large-scale integration. *Nature Reviews Neuroscience, 2*(4), 229-239.
564 doi:10.1038/35067550

565 Vedral, V. (2010). *Decoding Reality*. Oxford: Oxford University Press.

566 Von Baeyer, H. C. (1999). *Maxwell's Demon*. New York: Random House. 100-101; 165.

567 Ward, L. M. (2011). The thalamic dynamic core theory of conscious experience. *Consciousness and*
568 *Cognition, 20*(2), 464-486. doi: 10.1016/j.concog.2011.01.007

569 Wheeler, J. (1986). In *The Ghost in the Atom*, ed. P Davies and J Brown. Cambridge: Cambridge
570 University Press. 62.

571 Wheeler, J. A. (1989). Information, physics, quantum: the search for links. *Proceedings of the Third*
572 *International Symposium on Foundations of Quantum Mechanics*, Tokyo, 354-368.

573 Yuste, R. (2015). From the neuron doctrine to neural networks. *Nature Reviews Neuroscience, 16*(8),
574 487-497. doi: 10.1038/nrn3962