bioRxiv preprint doi: https://doi.org/10.1101/060467; this version posted June 24, 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 The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity.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.

# Neurobiology as information physics

Sterling Street
Independent Researcher

---

## Abstract

This article presents the preliminary results of an analysis of the brain that may supplement our current understanding of its structure and function, even if by falsification. This analysis seems to reveal that the information in a system of the brain can be quantified to an appreciable degree of objective precision, that many qualitative properties of the information content of a system of the brain can be inferred by observing matter and energy, and that many general features of the brain's anatomy and architecture illustrate simple information-energy relationships. It is possible that the brain provides a unique window into the relationship between matter, energy, and physical information.

## Introduction

That information is physical has been suggested by evidence for over a century (S Lloyd 2006). In recent years, Landauer's principle (R Landauer 1961), which relates informational entropy to thermodynamic entropy, has been confirmed (A Bérut et al. 2012; Y Jun et al. 2014; AO Orlov et al. 2012; B Piechocinska 2000), and information-energy equivalence has been shown to prohibit the existence of Maxwell's Demon (S Toyabe et al. 2010). The finding that information is conserved on event horizons has led to the tentative resolution of major paradoxes in black hole thermodynamics (L Susskind and J Lindesay 2005; K Bradler and C Adami 2014), and there is evidence for a universal relationship between information and spacetime geometry (R Bousso 2002, T Asselmeyer-Maluga 2015). Many of the classic mysteries of quantum mechanics, including entanglement and double-slit interference, may simply demonstrate properties of physical information (Č Brukner and A Zeilinger 2003; A Zeilinger 2004), and information-interpretive approaches in quantum gravity research (A Alfonso-Faus 2011; JW Lee et al. 2013; E Verlinde 2011) could even lead to the unification of quantum mechanics and general relativity, a necessary prerequisite in the development of a unified physical theory of everything (SW Hawking and L Mlodinow 2011). Surprisingly, however, "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 competing models and theories in neuroscience can be unified through the physics of information.

bioRxiv preprint doi: https://doi.org/10.1101/060467; this version posted June 24, 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.

## I. 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)

Information in any physical system is finite, and its flow is inevitably tied to thermodynamics. As a result, the information content of any system in the brain at any level of structure can be quantified by applying a modified form of the universal (JD Bekenstein 1981, 1984, 2001, 2005; J Oppenheim 2015) information-entropy bound:

$$I_{sys} = \zeta \frac{\sqrt{A}\,E}{\hbar c \log 2}$$

*where $I_{sys}$ is the quantity of units of information in a physical system, A is area in $m^2$, E is energy in joules, ħ is the reduced Planck constant, c is* the *speed of light in meters per second, log 2 is a logarithm of arbitrary base, and $\zeta$ is a factor of information saturation such that $0 \leq \zeta \leq 1$*

In short, a system approaches its information bound as it approaches a state of maximum information saturation. In other words, the bound quantifies the absolute information present in a system, while the portion of this information available for inference (available for performing logical operations or receiving in an influence network node or observatorial apparatus) is necessarily observer-dependent and generally far lower than quantity given by the bound. It is critical to make this distinction; a number of researchers have used the bound to estimate the information capacity of the brain as a hypothetical upper limit rather than as an absolute quantity of information, and a central argument among their critics is that the bound overestimates the quantity of information that is functionally relevant in cognition. Further, in the case that quantum effects play a meaningful role in cognition, this bound includes both classical information in bits and quantum information in qubits. The bound quantifies units of information of any base or classification and varies only with the area and energy content of a system.

The factor $\zeta$ partitions absolute information from observer-dependent information by separating entropic information from relative information. In other words, while absolute saturation ($\zeta = 1$) specifies the absolute quantity of information present in a thermodynamic universe, relative saturation ($0 \leq \zeta \leq 1$) specifies the quantity of absolute information that is available for influence into a referential system. Since entropy is hidden information (C Adami 2011; L Susskind 2008; S Lloyd 2006) with respect to a referential system, a system transferring none of its information to a referential system (resulting in maximal KL entropy) carries an entropic saturation factor of 1 and an informational saturation factor of 0. On the other hand, a system transferring the entirety of its information to a referential system (resulting in minimal KL entropy) carries an entropic saturation factor of 0 and an informational saturation factor of 1. The flow of information from an index to a referential system would parallel the crossing of Markov boundaries (KJ Friston 2013) and Fisher information spaces (BR Frieden 2015).

bioRxiv preprint doi: https://doi.org/10.1101/060467; this version posted June 24, 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.

## II. Stratification

*"If information forms the most basic layer of reality, what does it refer to?"* – (A Aguirre et al. 2015)

Consider the information processed by the four distinct bases of DNA. The quantity of information contained in a single base molecule includes the information in its constituent atoms and subatomic particles, all the way down to the quantum information evaluated in their spin states, angular momenta, vibrational quanta, and so forth. How is it that we can ignore this information when analyzing base pair sequences? Although the absolute information in a single DNA base far exceeds the relative information it transfers to its referents in a process such as protein synthesis, DNA lies within an effectively base-four layer of influence – the alternative evaluation of a qubit that determines the spin of a proton in an adenine molecule is very unlikely to alter the effective structure of adenine, much less a resultant base pair sequence or the structure of a protein, cell, or organism. So, information is stratified between layers of influence: only the value of a nucleotide's macrostate is informative relative to a referential amino acid.

On the other hand, while binary operations in an electronic device are nominally performed in a base-two layer of influence, a single bit of information registered by the presence or absence of electrons in a capacitor also contains the information within these electrons, and even in the atoms and subatomic particles of the capacitor itself. The absolute quantity of information processed by the entire device over a period of time would similarly include all of the information in any of its structural metals and plastics, in the photons that exit the screen, in the information lost as heat to the surroundings, and in the photons, air molecules, and external systems that interact with the device. The absolute information processed by the classical computing device clearly exceeds the quantity of information used in performing electronic computations. This stratification is arguably what allows the brain to contain a far greater quantity of relative information than an electronic device with equivalent physical dimensions.

Some other properties of information should be noted in the context of stratification. Each degree of freedom in a physical system is essentially an informational influence space, so that the realization of one state or elimination of one degree of freedom (i.e., the induction of many microstates into a single macrostate, the occurrence of a probabilistic event over its complement, the transition from a prior to a posterior probability, the collapse of a wavefunction, etc.) registers a single unit of physical information. It is for this reason that the finite number of quantum states imposed on any physical system ensures its containing and processing of finite information (S Lloyd 2006; C Rovelli 2015), that storing many degrees of freedom in one computational macrostate confers greater statistical stability (S Lloyd 2000), that the occurrence of an unlikely physical event registers more Shannon information, and that a relative quantity of information is measurable as entropy (both as $H = log_2 N$ and $S = k\ log_e W$). In any case, the information in any layer of influence can be quantified even if the content of its internal evaluation is unattainable by an external system.

bioRxiv preprint doi: https://doi.org/10.1101/060467; this version posted June 24, 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.

## III. Information-energy equivalence

*"It is important to realize that in physics today, we have no knowledge what energy* is.*"* – (RP Feynman 1964)

Because a state of maximal internal energy corresponds to a state of maximal information saturation, any transitions between energy levels occur as transitions between informational extrema. These transitions enclose a finite quantity of information, so the flow of information through a referential system is followed by an energy change:

$$\Delta I_{sys} = \sum_i k \log 2$$

*where k is Boltzmann's constant*

Reintroducing the universal information-entropy bound as

$$I_{univ} = \zeta_u \frac{\sqrt{A}\, Ek}{\hbar c}$$

*where A is the area of a thermodynamic universe in meters squared, E is energy in joules and* $\zeta_u = 1$

now allows us to partition the universal information saturation factor into a relative information component ($\zeta_i = 1 - \zeta_s$) and an entropic information component ($\zeta_s = 1 - \zeta_i$),

$$I_{sys} = \zeta_i \frac{\sqrt{A}\, Ek}{\hbar c} = (1 - \zeta_s) \frac{\sqrt{A}\, Ek}{\hbar c}$$

so that, in the case that information enters or exits a macroscopic physical system with an effectively constant mass and surface area, we can write

$$\Delta I_{sys} = W(kT \log 2) = \Delta E_{sys}$$

*where W is the number of microstates in a physical system*

Various forms of similar relationships have been considered in neuroscience (P Sterling and S Laughlin 2015). This article simply proposes their restatement as fundamental physical relationships in the brain requiring further investigation and critique. On a side note, these relationships also seem to be consistent with the possibility that the quantum of gravity is (roughly) the energy of one bit of entropy (A Alfonso-Faus 2011), the possibility that the Planck constant is the minimal unit of information (C Rovelli 2015), and the possibility that the asymmetry of time is an entropy gradient (L Mlodinow and TA Brun 2014).

bioRxiv preprint doi: https://doi.org/10.1101/060467; this version posted June 24, 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.

## IV. The optimization principle

*"To heat also are due the vast movements which take place on the earth. It causes the agitations of the atmosphere, the ascension of clouds, the fall of rain and of meteors, the currents of water which channel the globe… It is necessary to establish principles applicable not only to steam engines but to all imaginable heat engines, whatever the working substance and whatever the method by which it is operated."* — (S Carnot 1824)

The observation that biological systems use a steady influx of energy to maintain order and structural integrity seems to show that, since entropy can be considered a referent-dependent measure of inaccessible information, the continuous influence of energy through the information bound of a biological system necessarily decreases entropy and increases internal energy. Since the maximization of internal energy and the maximization of relative information occur as coupled, mutually-reinforcing processes – and both processes offset the transition into equilibrium – living systems can be seen as thermodynamic demons that locally maximize information and internal energy. It has indeed been suggested that there should be no distinction between formalized information and physical information in biology (C Adami 2004), that defining biological systems in terms of their management of information may condense a great deal of experimental evidence into a small number of theoretical models (S Brenner 2012, PM Binder et al. 2011), and that the development of a unified theory of biology may require the application of information thermodynamics (DR Brooks et al. 1989). It appears that, as a general trend, living systems conform to a principle of information-energy maximization, or optimization. Interestingly, because this maximization would lead to the minimization of informational free energy, this principle is describing one side of the same relationship represented by the variational free energy principle (KJ Friston 2010, B Sengupta et al. 2013). In other words, this is simply an alternative way of considering the free energy principle as an informational principle of least action.

bioRxiv preprint doi: https://doi.org/10.1101/060467; this version posted June 24, 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.

## V. Neurobiology

*"… classical thermodynamics… is the only physical theory of a universal content concerning 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)

### Neurons

Neuron activity is, to some extent, reducible to chemical and electrical activity, and such activity is necessarily accompanied by an exchange of energy that follows the transfer of information between systems. The occurrence of depolarization as a transition to $E_{max}$ and the maintenance of a resting state as $E_{min}$ may reflect coupled information-energy dynamics, and the maintenance of a negative resting membrane potential may show that depolarization requires the input of external information-energy. In addition, while the graded potential and analog influence is clearly important, a quintessential feature of neurons is their propagation of all-or-nothing action potentials. This feature seems to mirror the discrete, granular all-or-nothing mathematical structure of physical information. Although information is processed in the activity of the ions and neurotransmitters that generate an EPSP or IPSP, even at subthreshold levels, any transition into one state from two or more alternatives or degrees of freedom could be interpreted as carrying a single unit of information regardless of underlying dynamics. Further, that spatial and temporal summation of numerous dendritic spikes results in the propagation of a single axonal action potential – one macrostate – even suggests a compression algorithm. This operation may represent an analog of feature binding at the level of the neuron, particularly suitable for the layer III and IV pyramidal and stellate cells whose projections connect higher cortical areas. Additionally, the observation that the information in a sequence of action potentials is carried only by spike count and temporal spacing (E Kandel et. al 2000) may show that a frequency distribution code is partly an error-minimizing redundancy feature of neuronal information processing mechanisms. That signal strength increases linearly with frequency and is solely determined by frequency also seems to suggest that spikes represent discrete quantities of physical information. The possibility that bursting activity and repeated spiking responses may have evolved partly to allow error correction by repetition – perhaps the simplest and most rudimentary error correction algorithm, present everywhere from the organization of DNA (V Vedral 2010) to the structure of speech (J MacCormick 2012) – is supported by the observation that signal-to-noise ratios tend to increase with signal repetition (W Bialek and F Rieke 1992).

### Networks

The general features of the various network architectures in the brain also appear to show information-energy relationships. Networks naturally stabilize by seeking energy minima, and the various attractor states in one of these networks define the geometry of its energy landscape (JJ Hopfield 1982). As a result, transitions into spontaneous states follow transitions into information-energy maxima, and such transitions have been associated with the generation of an informational entity such as a memory or a decision (G Deco et al. 2009). In this way, the local energy basins of attractor networks may be closely analogous to the resting potentials of neurons and receptors; a transition between any two distinctive energy states follows the passage of a

bioRxiv preprint doi: https://doi.org/10.1101/060467; this version posted June 24, 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.

finite quantity of information. Even in the case of a massively integrated network managing a great deal of complex information, the observation that any energy landscape contains a finite number of basins at any moment may show that each energy transition in such a network carries a finite quantity of information. This architecture may even parallel the action potential as an mechanism for promoting stability (in the sense of minimizing prediction errors) simply by using many degrees of neuronal freedom to effect the transition into one macrostate from a large number of microstates. In other words, where the degrees of freedom governing the dynamics of a neuron rest on the degrees of freedom of its structural molecules, neurotransmitters, and ions, the degrees of freedom of a network rest (emergently) on the degrees of freedom in underlying neurons and their possible states and configurations. Of course, a network in the brain is still prone to chaos and nonlinear behavior, and the function of a single neuron could result in an altered network macrostate. For the very reason that this is true, a vastly connected network whose behavior is governed by the activity of large assemblies of integrated neurons is minimally likely to be altered by the macrostate of an individual subsystem (e.g., a single potassium ion) at a lower level of stratification. That a network's number of degrees of freedom is limited by connection density is also suggested by the observation that vocabulary size and memory capacity appear to correspond to a network's number of connections (ET Rolls 2012). If entities such as words and memories are considered to be units of information, there is clearly a correspondence between a network's number of connections and its number of effective degrees of freedom as possible units of information. Network competition may also restate the importance of increasing outcome stability by maximizing conditional network responses, since inhibiting more weakly activated surrounding networks would reduce the overall number of degrees of freedom of a network at a higher level of architecture, and for the brain as a whole.

## Inhibition

It is possible that inhibition only nominally reduces information processing in the brain. Since maintaining proper membrane potentials, energy gradients, neurotransmitter and signaling protein concentrations requires the expenditure of internal energy, it is advantageous for the brain to process only a minimal quantity of environmental information necessary for survival. For this reason, it is possible that inhibition in the brain still results in optimization, both by the brain itself and the organism as a whole. This possibility even seems to align with some very general observations at the level of behavior and cognition. For instance, the ability of a verbal command to rapidly inhibit the flow of information through non-salient networks (F Grabenhorst and ET Rolls 2008) may show that the affective nature of language parallels inhibitory neuromodulation: minimizing the expenditure of internal energy at any level of structure seems to maximize the internal information content of each influence node. In the case of social networks, integration enhances the efficiency of inhibition: the peripheral mediation of another's inner state is clearly advantageous for a function such as calling attention to a stimulus in order to coordinate a group response, and the rapid evaluation of another's inner state is inversely beneficial for the same reasons. The ability of disintegrated information flow to prevent the emergence of an optimized influence network may be further revealed by the observation that the maintenance of a stable internal state requires constant inhibitory modulation (K Lehmann et al. 2012), as well as the observation that large-scale network hyperexcitation can lead to the

bioRxiv preprint doi: https://doi.org/10.1101/060467; this version posted June 24, 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.

rapid loss of consciousness. The "attentional spotlight" of awareness noted by Global Workspace Theory (BJ Baars 2005) may even describe a form of inhibition mediated largely by frontal and parietal network competition. In the case of volitional attention, the willful selection of a salient perceptual stream (i.e., through a chosen communication channel) requires the activation of large assemblies of inhibitory interneurons.

## Learning

Lowering the firing threshold in order to increase the likelihood of individual cellular response to a stimulus (DO Hebb 1952) may optimize the activation of an engram by minimizing the quantity of energy needed to elicit maximal influence from its connections. Interestingly, the influence of a minimal energetic stimulus in an optimally integrated recurrent network has even been reported to enable a full energetic activation and an effectively perfect recall of a memory (ET Rolls 2012). For this reason, it is possible that a learned engram is not the storage of information as a dormant memory or response per se but rather a mechanism allowing the temporal convergence of otherwise separate units of information upon the engram's future activation. It is also possible that neurons tend to form composite systems only if the internal energy and information management capacity of the composite system is likely to exceed that of the individual neurons in their prior state of separation. This possibility seems to be supported by the general molecular and cellular energetics of learning and plasticity. The minimization of free energy achieved through learning (KJ Friston 2010) would then maximize internal information, and the continual formation and disassembly of engrams during learning and forgetting would optimize the growth and pruning of a network in response to external states.

## Noise

Any physical system whose state values are coupled to the exchange of energy and the transfer of information is inherently noisy, even if effectively non-stochastic and roughly predictable. Noise does appear to be an innate and unavoidable feature of information processing in the brain (ET Rolls 2012), scaling in proportion to influence network size (K Josic et al. 2009) and varying inversely with the coupling of network components (DJ Mar et al. 1999; B Kia et al. 2015). Even in the absence of any potential forms of stochastic resonance (F Moss et al. 2004), the noise-driven exploration of various state spaces could enhance the ability of a system in the brain to reach an optimal information-energy extremum. Interestingly, noise has been found to enhance learning and serve beneficial functions (MD McDonnell and LM Ward 2011); the finding that double-stranded DNA breaking (E Suberbielle et al. 2013) and reversible methylation (P Tognini et al. 2015) both occur as normal, everyday processes in the living brain could perhaps be explained by considering that any stochastically optimized structure would be more likely to survive over time as a strengthened connection or increased integration in a modifiable network. So, noise could enhance brain function in much the same way that noise at the level of the genome (i.e., recombination) enhances adaptation and evolution of species (JV Stone 2015; S Lloyd 2006). In other words, it possible that noise in the brain can be interpreted as informational entropy, which necessarily decreases when any systems integrate, couple their influence with mutual information, and jointly reduce KL and thermodynamic entropy. The signal-to-noise ratio could then be seen as a measurement of relative information.

bioRxiv preprint doi: https://doi.org/10.1101/060467; this version posted June 24, 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.

## Brain

The brain's dependence on energy (P Sterling and S Laughlin 2015) is readily apparent in its disproportionate consumption of energy substrates (e.g., oxygen, glucose, and ATP), in its vulnerability to hypoxic-ischemic damage, and in the rapid loss of consciousness conferred by the onset of an energy restriction. It is also interesting to consider that the information content of a large-scale system in the brain can change while mass and surface area remain effectively constant, and that the brain consumes a greater quantity of energy expected from its mass alone (ME Raichle and DA Gusnard 2002). These findings may even relate to the tradeoff of these parameters predicted by the information bound: a change in the information content of a system with constant mass and surface area must produce a change in internal energy. All fMRI and PET interpretation rests on the assumption that changes in the information content of a structure can be inferred by observing internal energy changes, and it is well known that the information processing capacities of systems in the brain are limited by energy supply. The living brain appears to dynamically maintain a state of optimization, while brain death could be said to occur as a rapid transition into a local information-energy minimum.

bioRxiv preprint doi: https://doi.org/10.1101/060467; this version posted June 24, 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.

## VI. 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)

### Unity

At any given moment, awareness is experienced as a unified whole. It is possible that physical information is the substrate of consciousness, and it appears that any unit of information – the bit, trit, nat, qubit, and so forth – is necessarily transferred into a referential system as a discrete and temporally unitary quantity. In other words, it is possible that the passage of time itself is influence-dependent, occurring secondarily to the transfer of information, and that the information present in any system at any time is always cohesive and temporally unified. From this perspective, the dilation or contraction of time would vary as a rate of information influence and entropy generation – the flow of absolute information into a relative information space would increase the internal energy of a referential demon by $kT \ln 2$ J for every macrostate observed, while the opposite would occur with the flow of relative information into an absolute information space (i.e., the generation of entropy). The possible requirement for any information to be unified at its time of influence into a referential system of the brain may even originate in fundamental conservation laws; any quantity of information is fully conserved during its course of influence, and the influence of unified quantities of information necessarily occurs alongside the influence of conserved quantities of energy. Although it is possible that all physical systems exchange information-energy in temporally unitary quantities, it is likely that many of the familiar features of unity in phenomenological human consciousness require the structure and function of neural networks in the brain. It appears that this unity, as an emergent informational structure (F Dretske 1981) requires an integration of thalamocortical and localized neocortical influence (E Zillmer et al. 2007; DF Swaab 2014; C Koch and S Greenfield 2007; S Greenfield and TFT Collins 2005). It is also possible that an emergent informational structure in the claustrum generates some of this experiential unity (FC Crick and C Koch 2005, MZ Koubeissi et al. 2014). However, it has been reported that complete unilateral resection of the claustrum performed in patients with neoplastic lesions of the region appears not produce any significant change in subjective awareness or observable functionality (Duffau et al. 2007). It appears unlikely that the flow of information into any isolated or compartmentalized network of the brain is responsible for producing all aspects of the unified nature of conscious experience.

### Complexity

The subjective complexity of conscious experience may reveal that extensive influence network integration is needed for maximizing mutual information and internal energy. An exemplary structure enabling such integration, likely one of many that account for the sensory complexity of consciousness, is the cluster of projections that form the internal capsule. A structure such as this may show that, at any given moment in the influence space of a living brain, a wide range of integrated referential networks are sharing mutual sources of sensory information. This pattern of structure may reveal that the perceptual depth and complexity of consciousness is an

bioRxiv preprint doi: https://doi.org/10.1101/060467; this version posted June 24, 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.

inevitable emergent effect of influence rather than a deliberate construction of the brain per se. It also seems that extensive localized regional cortical consolidation of sensory information is necessary for producing a refined and coherent complex sensorium within a referential network, which could then pass into other higher-order networks through commissural and association fibers. In addition, the dynamics of two-state attractor networks at these levels of network structure may also show that quantities of complex sensory-perceptual information can be observed as changes in cortical energy landscapes, with a single transition between states following the transfer of a finite quantity of information between any two systems. The complexity of the information encapsulated by such a transition would reflect the degree of consolidation of information-energy from underlying influence networks. In short, it is possible that the sensory and perceptual complexity of consciousness is, to at least some extent, apparent in the properties of observable structures that are fairly well understood. The structural analogs of the complexity of conscious experience may be observable as integration between large-scale influence networks that receive and share surroundings-sourced information.

## Self-awareness

It is possible that the informational structure of self-awareness arose as a necessity for survival rather than as an accident of evolution, and rudimentary forms of self-awareness likely began to appear early in the course of brain evolution as various forms of sensory self-environment separation. The extent of this perceived separation would likely vary as a function of Markov boundary density and internal information content (enclosed by the information bound). An example of this division is the tickle response, which clearly (DJ Linden 2007) requires the ability to partition self-produced (i.e., non-threatening) tactile sensory states from those produced by external influence (i.e., potentially threatening). The early need for the ability to cancel self-produced tickles may be further elucidated by the observation that the cancellation process is largely mediated by the cerebellum (SJ Blakemore et al. 1998). The perception of a well-defined self is clearly advantageous (PS Churchland 2002, MS Graziano 2013), and conditions involving neglect do show that a reduced awareness of self-ownership of motor skills or perceptions can result in significant impairment even if potential sensory or motor function remains intact (A Parton et al. 2004). While it is possible that the perceptual self arose from early subcortical origins, the complex syntactical and conceptual self structure present in the modern brain likely requires the neocortical influence networks that began integrating with (i.e., essentially grew over) pre-existing rudimentary subcortical self networks. The innate capacity for developing sophisticated self-modeling informational structures may have evolved, along with language, as a requirement for coordinating group behavior as social dynamics began to grow increasingly less predictable. Further, in the same way that the specific state values of informational language components (e.g., graphemes, phonemes, words, constructions, etc.) are the products of external informational structures (i.e., a learning environment) it is likely that states of self-awareness are learned. This possibility is supported by the finding that self-touch seems to enhance body ownership in stroke patients with acquired body awareness disorders (HE Van Stralen 2011) and that mirror therapy in such groups can often alleviate self-awareness deficits (A Fotopoulou et al. 2009; AS Rothgangel et al. 2011). It is also notable that self structures and language structures in the brain both traverse a wide range of cortical and subcortical regions (G Northoff 2006, W Penfield and L Roberts 2014), and that self-referential

bioRxiv preprint doi: https://doi.org/10.1101/060467; this version posted June 24, 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.

thought informs the content of inner speech (A Morin and J Michaud 2007). The continual experience of a first-person syntactical self thus enables the rapid, thorough evaluation and influence of another's inner state, which could be said to optimize social networks.

## Continuity

While perceptual time likely results from a group of related processes rather than a single, global function localized to any specific influence network in the brain, it is possible that some of the perceptual continuity of conscious experience results from the effectively continuous flow of discrete quantities of physical information into and out of systems of the brain. In such a framework, the "quantum" (M Prokopenko et al. 2014) of perceptual time would be the minimal flux of information, and an effectively continuous influence of relative information would appear as as an effectively continuous influence of energy into an internal space. This possibility seeems to be supported by the observation that the transition of a large-scale attractor network is regressively more continuous than the activation of a small-scale engram, the propagation of an action potential, the release of a vesicle, or the passage of an ion through a membrane. Likewise, electroencephalography shows that the ongoing summation of a large number of neuronal potentials can converge into an effectively continuous wave as a network field potential (SJM Smith 2005) whose disruption may cause discontinuity of experience and induction of unconsciousness (H Blumenfeld and J Taylor 2003, B Haider et al. 2006). In addition, higher frequency network oscillations are generally indicative of wakeful states and active awareness (M Teplan 2002), while lower frequency oscillations tend to be associated with states of lesser passage of perceptual time, such as sleep and unconsciousness. The possibility that the perceptual arrow of time and the external arrow of time share a common origin in the flow of physical information is supported by general models of time in cognitive neuroscience and the interpretation of time in physics as a statistically asymmetrical increase in relative entropy (L Maccone 2009; L Mlodinow and TA Brun 2014; OC Stoica 2008; T Asselmeyer-Maluga 2015).

## Free will

Evidence suggests that the brain is often predictable within reason, and the performance of an action can be predicted before a decision is reported to have been made (B Libet et al. 1983). Evidence suggests that entities such as ideas, feelings, and beliefs exist as largely deterministic, even if stochastically mutable, evaluations of information processed in the brain. Whether this flow is subject to its volitional alteration or retroactively assigned as an "illusory" feeling or belief, evidence also shows that information in the brain is subjectively real to its influence space, even if this information is inconsistent with an external reality or the experienced reality of an external system. That an influence space of the brain can contain a subjectively true and independently real but objectively inaccurate and externally inconsistent reality is exemplified by phenomena such as confabulation, neglect, commissurotomy effects, placebo and nocebo effects, hallucinations, prediction errors, the suspension of disbelief during dreaming, the requirement for channels in the form of synapses and white matter tracts, the function of social communication in minimizing divergence between individual realities, the quality of many kinds of realistic drug-induced experiences, and the observable effects of many cognitive conditions. The fact that subjective reality is an active construction of the physical brain has even led to the

bioRxiv preprint doi: https://doi.org/10.1101/060467; this version posted June 24, 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.

proposal of model-dependent realism (SW Hawking and L Mlodinow 2011) as a theoretical approach in developing a unified physical "theory of everything" (ignoring unsolved problems in other fields). Even the scientific method itself seems to hint some elements of information processing (ET Jaynes 2003), such as error correction by experimental replication, updating of priors with the falsification of predictions, and the information-energy maximization that could be said to result in the integrated social networks that benefit from its presence and application. In any case, it is likely that beliefs, including those in free will, exist as physical information, and that their internal reality is a restatement of its inherently observer-dependent nature. It is likely that a scientific debate over the existence of objective, external free will independent of reference is analogous to a debate over the existence of a similar informational entity such as a perception, experience, or quale. It is likely that the internal reality of informational entities in the brain is further revealed by their existence as physical information, their inevitably stochastic nature, their effectively deterministic evaluation, their requirement for physical media, and their ability to determine the states of physical systems. It is likely that a simple inquiry currently offers the most accurate, valid, and reliable test for the presence of free will.

## Conclusion

This article has presented information-energy relationships in the hope that they may eventually provide a general framework for uniting theory and experiment in neuroscience. This analysis has been very superficial and preliminary, and largely limited to speculation and suggestion. Nevertheless, it is possible that these or similar relationships are present, to at least some extent, in all brain structure and function. Refinement or falsification is needed.

## Acknowledgements

I am grateful for the helpful discussions, friendly criticism, and generous insight and assistance provided by Baroness Susan Greenfield, Dr. Francesco Fermani, Dr. Karl Friston, Dr. Sengupta Biswa, Dr. Chris Adami, Dr. Roy Frieden, Dr. Bernard Baars, Dr. Brett Clementz, Dr. Cristi Stoica, Dr. Satoru Suzuki, Dr. Paul King, Guillem Collell Talleda, and Dr. Jordi Fauqet.

## References

Adami, C. (2004). Information theory in molecular biology. *Physics of Life Reviews.*

Adami, C. (2011). Toward a fully relativistic theory of quantum information. *arXiv:1112.1941.*

Alfonso-Faus, A. (2011). Quantum gravity and information theories linked by the physical properties of the bit. *arXiv:1105.3143*

Alfonso-Faus, A. (2013). Fundamental Principle of Information-to-Energy Conversion. *arXiv:1401.6052.*

Ash, Robert B. (1990). *Information Theory*.

Asselmeyer-Maluga, T. (2015). Spacetime Weave—Bit as the Connection Between Its or the Informational Content of Spacetime. *It From Bit or Bit From It?*

Attwell, D., & Laughlin, S. B. (2001). An energy budget for signaling in the grey matter of the brain. *Journal of Cerebral Blood Flow & Metabolism.*

Baars, B. J. (2005). Global workspace theory of consciousness: toward a cognitive neuroscience of human experience. *Progress in Brain Research.*

bioRxiv preprint doi: https://doi.org/10.1101/060467; this version posted June 24, 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.

Bérut, A., et al. (2012). Experimental verification of Landauer's principle linking information and thermodynamics. *Nature.*

Bekenstein, J. D. (1981). Universal upper bound on the entropy-to-energy ratio for bounded systems. *Physical Review D.*

Bekenstein, J. D. (1984). Entropy content and information flow in systems with limited energy. *Physical Review D.*

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*.

Bekenstein, J. D. (2004). Black holes and information theory. *Contemporary Physics.*

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.*

Bennett, C. H., & Landauer, R. (1985). The fundamental physical limits of computation. *Scientific American.*

Bialek, W., & Rieke, F. (1992). Reliability and information transmission in spiking neurons. *Trends in Neurosciences.*

Binder, P. M., & Danchin, A. (2011). Life's demons: information and order in biology. *EMBO.*

Blakemore, S. J., et al. (1998). Central cancellation of self-produced tickle sensation. *Nature.*

Blumenfeld, H., & Taylor, J. (2003). Why do seizures cause loss of consciousness? *The Neuroscientist.*

Boltzmann, L. (2012). *Theoretical Physics and Philosophical Problems: Selected Writings*.

Borchers, S., et al. (2012). Direct electrical stimulation of human cortex—the gold standard for mapping brain functions? *Nature Reviews Neuroscience.*

Bousso, R. (2002). The holographic principle. *Reviews of Modern Physics.*

Bradler K. & Adami C. (2014). The capacity of black holes to transmit quantum information *Journal of High Energy Physics.*

Bremermann, H. J. (1982). Minimum energy requirements of information transfer and computing. *International Journal of Theoretical Physics.*

Brenner, S. (2012). Turing centenary: Life's code script. *Nature.*

Brooks, D. R., & Wiley, E. O. (1988). *Evolution as entropy*.

Brooks, D. R., et al. (1989). Entropy and information in evolving biological systems. *Biology and Philosophy.*

Brukner, Č., & Zeilinger, A. (2003). *Information and fundamental elements of the structure of quantum theory.*

Churchland, P. S. (2002). *Brain-Wise: Studies in Neurophilosophy.*

Collell, G., & Fauquet, J. (2015). Brain activity and cognition: a connection from thermodynamics and information theory. *Frontiers in Psychology.*

Collier, J. (1986). Entropy in evolution. *Biology and Philosophy.*

Crick, F., & Koch, C. (2002). The problem of consciousness. *Scientific American.*

bioRxiv preprint doi: https://doi.org/10.1101/060467; this version posted June 24, 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.

Crick, F. C., & Koch, C. (2005). What is the function of the claustrum?. *Philosophical Transactions of the Royal Society B: Biological Sciences*.

Deco, G., Rolls, E. T., & Romo, R. (2009). Stochastic dynamics as a principle of brain function. *Progress in Neurobiology.*

Demetrius, L. (2000). Thermodynamics and evolution. *Journal of Theoretical Biology.*

Dretske, F. (1981). *Knowledge and the Flow of Information.*

Duffau, H., et al. (2007). Functional compensation of the claustrum: lessons from low-grade glioma surgery. *Journal of Neuro-oncology.*

Duncan, T. L., & Semura, J. S. (2004). The deep physics behind the second law: information and energy as independent forms of bookkeeping. *Entropy.*

Feynman, R. P., & Zee, A. (2006). *QED: The strange theory of light and matter*.

Fotopoulou, A., et al. (2009). Self-observation reinstates motor awareness in anosognosia for hemiplegia. *Neuropsychologia.*

Frieden, B. R. (2015). Estimating a Repeatable Statistical Law by Requiring Its Stability During Observation. *Entropy.*

Frieden, B. R., & Gatenby, R. A. (2013). Principle of maximum Fisher information from Hardy's axioms applied to statistical systems. *Physical Review E.*

Frieden, B. R., & Gatenby, R. A. (2011). Information dynamics in living systems: prokaryotes, eukaryotes, and cancer. *PloS one.*

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature.*

Friston, K. (2013). Life as we know it. *Journal of The Royal Society Interface.*

Greenfield, S. (2002). Mind, brain and consciousness. *The British Journal of Psychiatry.*

Greenfield, S. A., & Collins, T. F. (2005). A neuroscientific approach to consciousness. *Progress in Brain Research.*

Haider, B., Duque, A., Hasenstaub, A. R., & McCormick, D. A. (2006). Neocortical network activity in vivo is generated through a dynamic balance of excitation and inhibition. *The Journal of Neuroscience.*

Hawking, S. W. (2005). Information loss in black holes. *Physical Review D.*

Hawking, S. W. and Mlodinow, L. (2011). *The Grand Design.*

Hebb, D. O. (1952). *The Organization of Behavior: A Neuropsychological Theory*.

Herculano-Houzel, S. (2009). The human brain in numbers: a linearly scaled-up primate brain. *Frontiers in Human Neuroscience.*

Hobson, J. A., & Friston, K. J. (2012). Waking and dreaming consciousness: neurobiological and functional considerations. *Progress in Neurobiology.*

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences.*

Jaynes, E. T. (2003). *Probability Theory: the Logic of Science.*

Josic, K., Rubin, J., Matias, M., Romo, R. (2009). *Coherent Behavior in Neuronal Networks*.

bioRxiv preprint doi: https://doi.org/10.1101/060467; this version posted June 24, 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.

Jun, Y., Gavrilov, M., & Bechhoefer, J. (2014). High-precision test of Landauer's principle in a feedback trap. *Physical Review Letters.*

Kandel, E. R., et al. (2000). *Principles of Neural Science.*

Katsikis, G., et al. (2015). Synchronous universal droplet logic and control. *Nature Physics.*

Kia, B., Lindner, J. F., & Ditto, W. L. (2015). Nonlinear dynamics based digital logic and circuits. *Frontiers in Computational Neuroscience.*

Koubeissi, M. Z., Bartolomei, F., Beltagy, A., & Picard, F. (2014). Electrical stimulation of a small brain area reversibly disrupts consciousness. *Epilepsy & Behavior.*

Knuth, K. H. (2015). Information-based physics and the influence network. In *It From Bit or Bit From It?*

Landauer, R. (1961). Irreversibility and heat generation in the computing process. *IBM Journal of Research and Development.*

Landauer, R. (1996). The physical nature of information. *Physics Letters A.*

Laughlin, S. B. (2001). Energy as a constraint on the coding and processing of sensory information. *Current Opinion in Neurobiology.*

Lavis, D. A., & Streater, R. F. (2002). Physics from Fisher information. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*.

Lehmann, K., et al. (2012). GABA through the ages: regulation of cortical function and plasticity by inhibitory interneurons. *Neural Plasticity*.

Lee, J.W., et al. (2013). Gravity from quantum information. *J Korean Physical Society.*

Leifer, M. S. (2015). "It from bit" and the quantum probability rule. *It From Bit or Bit From It?*

Libet, B., et al. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). *Brain.*

Linden, D. J. (2007). *The Accidental Mind.*

Liu, H. (2005). New pathway for glutamate release from brain glial cells upon ischemic stress. *Journal of the Physiological Society of Japan.*

Liu, K. C., Oztaskin, M., & Burchiel, K. J. (2012). Basics of neurosurgical techniques and procedures. *Essentials of Neurosurgical Anesthesia & Critical Care.*

Lu, Z., Mandal, D., & Jarzynski, C. (2014). Engineering Maxwell's demon. *Physics Today.*

Lloyd, S. (2000). Ultimate physical limits to computation. *Nature.*

Lloyd, S. (2006). *Programming the Universe.*

Lodish, H., et al. (2000). Neurotransmitters, Synapses, and Impulse Transmission. *MCBio.*

Maccone, L. (2009). Quantum solution to the arrow-of-time dilemma. *Physical Review Letters.*

MacCormick, J. (2012). *Nine Algorithms that Changed the Future.*

Mar, D. J., et al. (1999). Noise shaping in populations of coupled model neurons. *PNAS.*

Martyushev, L. M. (2013). Entropy and entropy production: Old misconceptions and new breakthroughs. *Entropy.*

Martyushev, L. M., & Seleznev, V. D. (2006). Maximum entropy production principle in physics, chemistry and biology. *Physics Reports.*

bioRxiv preprint doi: https://doi.org/10.1101/060467; this version posted June 24, 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.

McDonnell, M. D., & Ward, L. M. (2011). The benefits of noise in neural systems: bridging theory and experiment. *Nature Reviews Neuroscience.*

Morin, A., & Michaud, J. (2007). Self-awareness and the left inferior frontal gyrus: inner speech use during self-related processing. *Brain Research Bulletin.*

Mlodinow, L., & Brun, T. A. (2014). Relation between the psychological and thermodynamic arrows of time. *Physical Review E.*

Colloquium: The physics of Maxwell's demon and information. *Reviews of Modern Physics.*

Northoff, G., et al. (2006). Self-referential processing in our brain—a meta-analysis of imaging studies on the self. *Neuroimage.*

Orlov, A. O., et al. (2012). Experimental test of Landauer's Principle at the sub-$k_BT$ level. *Japanese Journal of Applied Physics.*

Penfield, W., & Roberts, L. (2014). *Speech and Brain Mechanisms*.

Piechocinska, B. (2000). Information erasure. *Physical Review A.*

Raichle, M. E., & Gusnard, D. A. (2002). Appraising the brain's energy budget. *PNAS.*

Ramachandran, V. S. (1998). The neurology and evolution of humor, laughter, and smiling: the false alarm theory. *Medical Hypotheses.*

Rolls, E. T. (2012). *Neuroculture.*

Rolls, E. T., et al. (2008). Selective attention to affective value alters how the brain processes olfactory stimuli. *Journal of Cognitive Neuroscience.*

Rothgangel, A. S., Braun, S. M., Beurskens, A. J., Seitz, R. J., & Wade, D. T. (2011). The clinical aspects of mirror therapy in rehabilitation: a systematic review of the literature. *International Journal of Rehabilitation Research.*

Rovelli, C. (2015). Relative information at the foundation of physics. *It From Bit or Bit From It?*

Schneider, E. D., & Kay, J. J. (1994). Life as a manifestation of the second law of thermodynamics. *Mathematical and computer modelling.*

Schrödinger, E. (1992). *What is life?: With mind and matter and autobiographical sketches.*

Sengupta, B., et al. (2013). Information and efficiency in the nervous system – a synthesis. *PLoS Computational Biology.*

Shannon, C. E., & Weaver, W. (2015). *The Mathematical Theory of Communication.*

Smith, S. J. M. (2005). EEG in the diagnosis, classification, and management of patients with epilepsy. *Journal of Neurology, Neurosurgery & Psychiatry.*

Smolin, L. (2002). *Three Roads to Quantum Gravity.*

Sparavigna, A. C. (2015). Dimensional equations of entropy. *arXiv:1511.07822.*

Sperry, R. W., et al. (1969). Interhemispheric relationships: the neocortical commissures; syndromes of hemisphere disconnection. *Handbook of Clinical Neurology.*

Stoica, O. C. (2008). "Flowing with a frozen river." *FQXi.*

Stoica, O. C. (2015). The Tao of It and Bit. *It From Bit or Bit From It?*

Stone, J. V. (2015). *Information Theory: A Tutorial Introduction*.

Styer, D. F. (2008). Entropy and evolution. *American Journal of Physics.*

bioRxiv preprint doi: https://doi.org/10.1101/060467; this version posted June 24, 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.

Suberbielle, E., et al. (2013). Physiologic brain activity causes DNA double-strand breaks in neurons, with exacerbation by amyloid-β. *Nature Neuroscience.*

Susskind, L. (2008). *The Black Hole War.*

Susskind, L., & Lindesay, J. (2005). *An Introduction to Black Holes, Information and the String Theory Revolution.*

Swaab, D. F. (2014). *We are Our Brains.*

Teplan, M. (2002). Fundamentals of EEG measurement. *Measurement Science Review.*

Tognini, P., et al. (2015). Dynamic DNA methylation in the brain: a new epigenetic mark for experience-dependent plasticity. *Frontiers in Cellular Neuroscience.*

Toyabe, S., et al. (2010). Information heat engine: converting information to energy by feedback control. *Nature Physics.*

Umpleby, S. A. (2007). Physical relationships among matter, energy and information. *Systems Research and Behavioral Science.*

Van Stralen, H. E., et al. (2011). The role of self-touch in somatosensory and body representation disorders after stroke. *Philosophical Transactions of the Royal Society B: Biological Sciences.*

Van Tol, M. J., et al. (2010). Regional brain volume in depression and anxiety disorders. *Archives of general psychiatry.*

Vedral, V. (2010). *Decoding Reality.*

Verlinde, E. (2011). On the Origin of Gravity and the Laws of Newton. *J High Energy Physics.*

Weber, B. H., et al. (1989). *Evolution in thermodynamic perspective: an ecological approach. Biology and Philosophy.*

Wheeler, J. A. (1989). *Proceedings of the 3rd International Symposium on the Foundations of Quantum Mechanics.* Physical Society of Japan.

Wicken, J. S. (1987). *Evolution, thermodynamics, and information.*

Willinsky, J. (2006). The access principle: The case for open access to research and scholarship.

Ye, B., et al. (2006). The secretory pathway and neuron polarization. *J Neuroscience.*

Zeilinger, A. (2004). Why the quantum? It from bit? A participatory universe? *Science and Ultimate Reality: Quantum Theory, Cosmology, and Complexity*.

Zillmer, E., et al. (2007). *Principles of Neuropsychology.*