Biological Theory
https://doi.org/10.1007/s13752-024-00472-6

THEMATIC ISSUE ARTICLE: LADISLAV KOVÁČ AND THE ORIGINS OF COGNITIVE BIOLOGY

On the Origin of Cognition

Arto Annila[1]

Received: 29 February 2024 / Accepted: 17 August 2024

Abstract
To explain why cognition evolved requires, first and foremost, an analysis of what qualifies as an explanation. In terms of physics, causes are forces and consequences are changes in states of substance. Accordingly, any sequence of events, from photon absorption to focused awareness, chemical reactions to collective behavior, or from neuronal avalanches to niche adaptation, is understood as an evolution from one state to another toward thermodynamic balance where all forces finally tally each other. From this scale-free physics perspective, energy flows through those means and mechanisms, as if naturally selecting them, that bring about balance in the least time. Then, cognitive machinery is also understood to have emerged from the universal drive toward a free energy minimum, equivalent to an entropy maximum. The least-time nature of thermodynamic processes results in the ubiquitous patterns in data, also characteristic of cognitive processes, i.e., skewed distributions that accumulate sigmoidally and, therefore, follow mostly power laws. In this vein, thermodynamics derived from the statistical physics of open systems explains how evolution led to cognition and provides insight, for instance, into cognitive ease, biases, dissonance, development, plasticity, and subjectivity.

Keywords Atomism · Causality · Emergence · Information · Nondeterminism · Nonequilibrium thermodynamics · Subjectivity

Introduction

How cognition evolved appears to be a reasonable inquiry into the sequence of events that gave rise to the multifaceted faculty (Heyes and Huber 2000; Shettleworth 2009; van Horik and Emery 2011). Likewise, why cognition evolved seems to be a sensible study into the causes that raised awareness even up to this level, where we ponder upon the origin of cognition. In fact, comprehending how evolution gave rise to cognition might just open the door to comprehending what it is to know.

However, the more we reduce cognition to constituent processes, the less we have left of what we wish to grasp. The essence of study somehow dissolves away by splitting cognition into perception, attention, reasoning, remembering, imagining, and so on, and further by decomposing cognitive functions into operations of neural circuits, neurons, synapses, neurotransmitters, etc., and still finer into chemical reactions and physical processes.

Similarly, tracks of evolution fade away as we descend down from the branches of the phylogenetic tree toward the stem. Rather than converging to the last universal common ancestor, nucleic acid sequence alignments diverge at the roots of life, as it seems, due to horizontal gene transfer (Fournier et al. 2015). As long as we do not know the principle governing evolution by natural selection in terms of physics, we end up only suspecting rather than showing abiogenesis (Sutherland 2017), even speculating about the extraterrestrial origin of life, yet not quite knowing what life is.

Since reductionism does not seem to explain how evolution resulted in cognition, the present article resorts to holism, assuming that everything is elementally the same and, hence, can be understood by the same principle of physics. From this viewpoint the article argues that cognition does not stand out as a distinctive function, its evolution does not differ from sequences of events in general, and its characteristics are not qualitatively unique to but only quantitatively pronounced in some systems. However, the article

✉ Arto Annila
arto.annila@helsinki.fi

[1] Department of Physics, University of Helsinki, Helsinki, Finland

does not deny that the enormous complexity, let alone aims and agency, in cognition might seem to be beyond physics but shows that such an impression stems from the standard physics perfected to quantify stationary systems. Instead, the physics of evolution accounts for all the complexity with the precision of a basic building block and for goal-orientedness with the universal principle of consuming available energy in the least time. In this way, the article explains how evolution led to cognition with concepts of overarching operational coherence (Chang 2022). Seamlessly absorbed into the whole of phenomena, cognition requires no explanation of its own.

The Atomistic Approach

Since disintegrating cognition into constituents does not seem to explain what cognition is and why and how it evolved, let us see what can be explained by integrating fundamental elements into functional structures so that free energy is consumed in the least time. According to this principle, energy in all forms flows to attain thermodynamic balance in the least time. Conversely, if the flows were not to channel along the steepest gradients in energy, causes as driving forces would appear inexplicably out of nothing or disappear into nothingness.

Thinking that explanatory terms should be better known than the terms to be explained, force is employed as an all-inclusive concept for causes, such as evolutionary pressure from competition, predation, food availability, climate, disease, and more. In turn, free energy consumption is used as an overarching concept for consequences, subsuming survival as well as other processes, with the least-time attribute for the fittest. Moreover, variation arises as paths of evolution diverge, redirect, and eventually narrow for the process to keep on consuming its driving forces in the least time. Since changes from one state to another cannot but proceed along the steepest gradients in energy, it appears that evolution naturally selects to go along these ways that lead to balance in the least time (Annila and Salthe 2010). So, in general terms of physics, evolution by natural selection means that energy flows invariably along the lines of force from one state to another rather than along just any line. In the same vein, Darwin illustrated that all kinds of environmental pressures, i.e., forces, amount over time to evolution from one form to another by comparing “natural” selection to “artificial” selection familiar from breeding practices and to Malthus’s idea of selection known from humans competing for limited resources.

Accordingly, an organism, just like a population or a species, is seen as attaining thermodynamic balance with its surroundings through survival and reproduction. In general terms, by tapping into resources of free energy, a species succeeds in invading an ecosystem much like a society succeeds in expanding into new territory. Similarly, a new gadget will be taken into use because it does the job, i.e., consumes free energy, better than the existing ones do. Likewise, a new way of behaving, just like a novel business, is understood to gain ground because it outperforms or enhances contemporary practices. Conversely, a process comes to a dead end when out of free energy.

The same drive for balance in the least time explains why neurons in high-speed, long-distance signaling are covered with myelin and highways are paved with tarmac. Likewise, it is understood that foliage has evolved over eons to harvest sunlight more efficiently and that photocells have improved in their efficiency through recent technological progress. The whole biosphere has emerged from the geosphere, hydrosphere, and atmosphere for the Earth to attain a better balance with hot sunlight and cold space (Lovelock 1972; Karnani and Annila 2009). In fact, according to the least-time principle, life emerged on Earth as soon as possible. Conversely, replicating abiogenesis on the laboratory scale would not be possible.

The all-inclusive least-time principle is also deemed to subsume extreme principles of mathematical models in cognitive science (Kováč 2000): maximizing rationality (Simon 1955) or utility (Von Neumann and Morgenstern 1944), and minimizing effort (Zipf 1949), free energy (Friston 2009), or prediction error (Clark 2013; Hohwy 2013), and self-organizing into criticality (Chialvo 2004).

Rather than deriving from empirical reduction, searching for the fundamental element by dividing things into finer and finer fragments, the offered holistic worldview is based on material monism, assuming that everything is made of “in-divisible,” *a-tomos* (Greek), constituents and hence can be explained in terms of them.

However enticing conceptually, such a metaphysical stance was for a long time not enough in practice to evaluate what we see and sense because the basic building block remained ambiguous. Therefore, it is epoch-making that today, the irreducible element of everything can be identified with the photon. No longer is the unity of all only a logical corollary of things transforming to other things but displays itself explicitly so that matter transforms into light and, conversely, light quanta into elementary particles according to the renowned mass-energy equivalence, $E = mc^2$. Even the superficially empty space, originating from matter, comprises photons, however, in pairs of opposite phases, rendering the vacuum transparent and relativistic (Mäkelä and Annila 2010; Annila and Wikström 2022). Conversely, the proposed standpoint cannot employ concepts that cannot be grounded in the quantized substance of photons.

Each photon as an *atomos* measures an invariant, Planck’s constant, $h = Et$, by carrying energy, E, on its period [of time], t. Thereby, a photon embodies causality;

its absorption or emission forces a change in the state of energy and a step of time (Annila 2021). Then, by definition, the evolution of a species and cognition across species, as well as the development of an organism and awareness across organisms, are nothing but sequences of events, fundamentally photon-by-photon changes.

Things happen, however, not in whichever way, but so that the quanta flow along the lines of force to gain balance in the least time. For example, a stone is understood to fall straight down, a chemical reaction to reroute through introduced catalysis, and a cell to redirect toward applied nutrients. In general, flows of energy follow the least-time paths, i.e., geodesics, in leveling out forces (Annila and Salthe 2010). Conversely, were we to witness some seemingly non-optimal processes nonetheless viable, we would be overlooking some forces in action.

The least-time free energy consumption is regarded as the final cause. While inexplicable itself, the principle can be drawn from the invariance of the quantum, i.e., Planck's constant, $dh = Edt + tdE = 0$, rearranged for the least-time gradient, $dE/dt = -E/t$. As the photon energy decreases and the period increases, the universe expands forever.

The photon concept is thoroughly tangible, for also the sense of time and the sense of space are not abstract intuitions of the mind (Kant 1893), or absolute entities of Newtonian mechanics or spacetime variables of Einstein's general relativity but the photon's period and wavelength are concrete characteristics of events (James 1890). For example, one senses time passing when getting cold as photons carry heat on their periods of time away from one's warm skin. In turn, one senses the substance of space in the form of inertia. For example, when slamming on the brakes, one continues in motion because of one's coupling to space, measured by one's mass. Having the atomistic substance in mind, Lewis renamed the light quantum the photon, appending the Greek word *phōs*, "light," with the particle suffix "-on" (Lewis 1926).

We know photons through experience. They are real. The human eye can capture even a single photon, and a pattern of photons may trigger a cognitive response. For example, photons are released from hydrogen bonds, van der Waals forces, and electrostatic interactions when a neurotransmitter binds to its receptor. All that dissipation amounts to the human brain consuming about 20% of metabolic energy while making up only about 2% of the body weight.

Clearly, the balance would be badly off unless gains by cognition covered the costs. From this thermodynamic viewpoint, memory is regarded as useful to the extent that it applies to attaining balance in the future. Likewise, learning is worth it to the degree that it pays off. Cognitive plasticity is rated valuable as much as it solidifies into the productivity of gaining balance with surroundings. As everything is elementally the same, all costs and gains are commensurable in the evolution toward thermodynamic balance.

The universal drive toward balance is best known as the second law of thermodynamics. However, the law remains poorly known unless derived from the first principles. Namely, Ludwig Boltzmann's statistical mechanics applies only to closed and stationary systems, where the total energy and the number of particles are fixed. Attempts to expand it to evolving systems have only generated confusion (Kováč 2023), referred to as nonequilibrium thermodynamics including dissipative structures (Prigogine and Wiame 1946) producing entropy (Jaynes 1957a, b). By contrast, as shown in the Appendix, thermodynamics derived from the statistical physics of open systems makes sense of evolution as a series of changes from one state to another until net forces vanish at balance. While a small chemical reaction mixture may gain balance in a short time, it will take eons for a large system, such as the Earth, to attain balance between the hot Sun and cold space. Moreover, considering that everything is in flux, conditions keep changing. Thus, no balance is but momentary in the evolution of the cosmos (Kováč 2015).

In precis, drawing from both empiricism and rationalism, the holistic worldview founded on photons carrying energy and time is offered to make sense of reality in general and the evolution of cognition in particular. Then, instead of imagining cognition as somehow special, sui generis, insight into it derives from postulating that everything is elementally the same and hence behaves the same way. While the all-encompassing theory may seem ambitious, even an audacious claim, a special theory about cognition and its evolution would be impaired to start with because no phenomenon can be explained in terms of its own. Then again, arguing for scientific monism is not calling for monism in science but rather pluralism, distinct from relativism, to find out what works (Chang 2022). Accordingly, the offered thermodynamic thinking about cognition is not just an unverifiable imagination but the whole atomistic thought style would collapse if some substance not consisting of photons was found or if some process not following the least-time principle was observed.

Thermodynamic Thinking

Evolution by natural selection is often presumed to be solely a biological phenomenon, but according to the statistical physics of open systems, it is no different from any other sequence of events bound to thermodynamic balance in the least time. This conclusion is consistent with the fact that by mere data, we cannot differentiate living from nonliving. (Mantegna and Stanley 1995; Limpert et al. 2001; West and Brown 2004; Newman 2005;

Jones et al. 2005; Clauset et al. 2009; Kello et al. 2010; Gabaix 2016). Everywhere, we witness skewed, nearly lognormal distributions accumulating in a sigmoid manner, hence trending power laws. For example, populations of species are like those of stars, sizes of genomes like those of galaxies, and networks of neurons are like circuits of semiconductors. Cognition is no exception. Power laws characterize response times, perception, learning, recalling, and forgetting (Stevens 1957; Kelso 1995; Wixted and Ebbesen 1997; Chater and Brown 2008).

Already, Galileo Galilei spotted scaling laws across *scala naturae*, and later, Snell (1892), Thompson (1917), Pareto (1897), Huxley (1932), and George Zipf (1949) recognized the recurrent self-similarity. Also, Lotka (1926), Price de Solla (1965), Richardson (1960), and Bak et al. (1988) noticed the ubiquity of power laws.

While scaling laws have been suspected to be signs of a universal principle, it was only recently that the least-time paths toward thermodynamic balance were shown to produce the ubiquitous patterns in data (Mäkelä and Annila 2010), i.e., the skewed, fat-tailed distributions and sigmoidal cumulative curves that follow close power laws (see Appendix). The scale-free patterns are characteristic of cognition, too (He et al. 2010). For example, cortical electrical activity is similar to seismic activity (Christensen et al. 2002). Episodic seizures resemble avalanches in semiconductor circuits (Logan et al. 1962; Chialvo 2004; Milton et al. 2007). Neurons oscillate as chemical bonds vibrate and cells cycle (Epstein and Pojman 1998). Nodal activity in neural networks and in the World Wide Web is distributed in the same skewed manner (Huberman and Adamic 1999). Like a stock market, a nervous system shows fluctuations, shocks, damping time series, and even chaotic behavior. Consequently, comprehending cognition begins by recognizing its characteristics in other things, too.

At first sight, it might seem mind-boggling to theorize everything, including how evolution produced cognition, merely as matter in motion toward thermodynamic balance. Indeed, thermodynamic thinking entails a paradigm shift, not only abolishing disciplinary divides but also abandoning some lines of inquiry altogether. When we cannot differentiate steps of evolution from a series of changes in general, we have no grounds for conceptualizing natural selection distinct from the all-inclusive least-time free energy minimization. Similarly, when we cannot demarcate cognitive processes from natural processes in general, we have no empirical basis for assigning unique characteristics to them (Beni 2023). The difference between animate and inanimate was never discernible in the data; we penned legends and labels to make up the differences and consequently found it difficult to define life and demarcate cognition.

Even if our epistemic competence were refined to distinguish any one entity from any other by the minimum difference of one quantum, as the electron differs from the W^- boson by one quantum, the neutrino, we would fail to define life because, unlike the quantum, life is not a natural kind, a category independent of human classification. Also, as a corollary of atomism, cognition, let alone consciousness, is not a fundamental feature of all things but rather an ambiguous attribute we associate with some systems.

While the evolution of sophisticated functions, such as perceiving, contextualizing, and conceptualizing, took eons, mathematically, elements can be integrated into functional entities in one stroke with an equation of evolution. This perfect bookkeeping of constituents is known as statistical physics, deriving from “standing,” *status* (Latin), of “nature,” *physis* (Greek). The many-body theory keeps track of all quanta of any given system at any given state (see Appendix). Then, the equation of evolution in its finest detail and grandest entirety can be formulated simply as a series of changes for more probable states until the most probable state, the free energy minimum, equivalent to entropy maximum, is attained. There, the system is at thermodynamic balance with its surroundings.

Mathematical exactness of an equation is often associated with determinism. However, despite its quantum precision, the evolutionary equation does not entail predetermined consequences from initial causes. Trajectories cannot be projected into the future because evolution changes its driving forces. Since variables cannot be separated, the equation cannot be solved. This is also apparent in cognition, where one function affects another and leads to multiplicative, allometric characteristics rather than additive, stochastic Gaussian statistics. So, we know from the equation of evolution exactly why we cannot know exactly how things will happen. Nevertheless, not just anything can happen, only that for which there are forces, i.e., causes.

Since forces point toward balance, evolution is teleological yet not predestined because evolution alters its forces. Consequently, the state of balance depends on the path to it, as the evolutionary arms race between species or between neural networks exemplifies (O’Reilly and Hemberg 2020). So, the obstacle to understanding evolution is not the lack of its computability but the lack of understanding about computability (Annila 2012d).

The desire for precise solutions has narrowed physics for stationary states. Obviously, a quantity cannot be obtained with high precision when evolving, only at the end of an evolutionary trajectory. But, per definition, such a steady-state calculus yields nothing new. Mathematically speaking, the far-famed Lagrange’s principle of least action for bound tracks is a special case of the forsaken Maupertuis’s principle of least action for open paths.

In brief, according to the statistical physics of open systems, evolution by natural selection is nothing but sequences of events toward thermodynamic balance, for example, from abiogenesis to the apocalypse, from protein folding to the unfolding of traditions (Annila and Annila 2008; Sharma et al. 2009; Annila and Salthe 2010). Since free energy is consumed in these changes in the least time, the same patterns appear in diverse data. Therefore it makes sense to consider also how evolution led to cognition by the least-time principle.

Cognition from the First Principles

According to thermodynamics, energy flows invariably along the least-time paths, as if naturally selecting these geodesics rather than less optimal ways. The geodesic notion has also been recognized in the cognitive context (Carlton and Shepard 1990a, b).

So, let us consider a photon that isomerizes the retinal cofactor from the 11-cis to the all-trans configuration, thereby freeing forces from various chemical potentials. These events culminate when the photoreceptor cell releases neurotransmitters to the synaptic cleft, flanking a bipolar cell dendrite. Next, at the cellular level, bipolar cells collect forces liberated from dendrite action potentials and forward them to ganglion cells that, in turn, transmit signals along their axons, i.e., the optic nerve, to the visual cortex. Subsequently, at the cortical level, the action potentials spread out in the central nervous system, eventually out into the peripheral nervous system. The whole process compares to the flow of water, starting from drops falling on mountain tops, channeling into brooks and tributaries forming the main river, and eventually spreading out into the river delta to reach the sea in the least time.

Despite all events complying with the same least-time law, they spread on unique paths. A common path diverges when one subject senses some forces that another does not. Also, the path forks when one subject improves in consuming a form of free energy while another does not. For example, bats evolved in cognizing sonar echoes to locate prey, while humans evolved in their cognitive capacity to locate and extract free energy from various sources, such as fossil fuel deposits.

The context of free energy consumption pronounces itself in co-gnize, "to know with," *cognoscere* (Latin) from the Proto-Indo-European root "to know," *gno*, and "together, with," *com*, *cum* (Latin). For example, a surrounding chemical concentration gradient directs a microorganism to a whole pool of combustibles. Similarly, a vein of ore leads miners to a whole deposit. When reencountered, signs and signals are re-cognized, "known again," and the least-time paths are reconnected to resume free energy consumption. In semiotic terms, a sign signifies only in context (Peirce 1931). Even a single signal, often seen as an anomaly, might revise a whole context. For example, the discovery of penicillin was a fortunate accident that revolutionized treating bacterial infections. Likewise, it might well have been a lucky mishap of a crow to drop a nut on the road, but to see it getting cracked by a passing car changed the subsistence of a whole crow population (Nihei and Higuchi 2002).

In turn, when signals are not received or recognized, no force is sensed and no action is taken. Expressly, magic is incomprehensible until put into a familiar context that opens eyes.

Assuming that everything is composed of the same basic building blocks, combining concepts into a new idea does not differ in principle from combining atoms into a molecule. In the compound, the atoms are not as they were when free because every reaction either emits photons to or absorbs them from surroundings (Pernu and Annila 2012). Since surroundings drive the combination by contributing to it, emergent properties do not originate solely from the constituents of the system. For example, the puzzle about "what time *is*" provides the impetus to combine the photon period with the element of time (Annila 2021). Thereafter, the photon is no longer only a quantum of the electromagnetic field and its period is no longer only one full cycle of oscillation, but the photon concept has acquired, so to say, greater operational coherence (Chang 2017), or thermodynamically speaking, potential, i.e., free energy, to explain not only electromagnetism but also phenomena where time flows.

According to thermodynamics, perceptual, conceptual, emotional, social, and cultural fluxes cohere into the context so that free energy is consumed in the least time (Lehmonen and Annila 2022). As the diversity in cognitive capacities across species suggests, the best bargain between prototypic plasticity and matured structure depends on circumstances. On the one hand, the least-time principle refines responses into automatic reflexes and, on the other hand, into reflective awareness. For example, once mastered, walking, running, or biking needs no time-consuming attention (Polanyi 1967). Also, emotions and intuitive judgments are immediately at hand. Then again, at times, it pays off to take time and energy to cognize alternatives (Kahneman 2011). Since awareness accrues from life experience, children are not held fully accountable.

Putting things in context is often rephrased as processing information, i.e., shaping things into a form from Latin *in-formare*. Information is thus physical (Landauer 1961), but the consequences of computing remain immaterial to the computer as long as it does not have the skin we have to put in the game. Paraphrasing Kurt Goldstein, we cognize in order to come to terms with our environment (Goldstein 1934). For example, in the hardening climate of Greenland during the Little Ice Age, Norsemen failed to gain balance by abandoning agriculture to harvest fruits of the sea to live like Inuits (Harris 1977; Diamond 2005).

Since each system makes sense of those forces it senses and responds to them with the means it has, its identity builds up from these ways of consuming free energy. The one-time Norse in Greenland would no longer have been Norse had they changed their lifestyle. By the same token, we might imagine what it is like to be a bat but still not know what it is like for a bat to be a bat (Nagel 1974). Only as much as the two share history does one know what it is to be the other.

Thermodynamic Insights into Cognition

It is rather remarkable that it takes only a split second for someone to make sense of what happened after having all of a sudden tumbled over. In biological lingo, rapid responses were naturally selected to survive in competitive circumstances, but in thermodynamic terms, speed bespeaks the least-time free energy consumption that subsumes survival, too.

The least-time maxim materializes so that sensory signals construct only a bit, while most of the context is merely retrieved or revitalized as if it were unchanged. Incremental updating quickly produces at least a functional, if not factual, view of the world consistent with the subject's history. In other words, by drawing from considerable context, even a simple signal can generate impressive subjective experience, qualia. Thus, the hard problem of consciousness, the challenge of understanding why and how physical processes in the brain give rise to subjective experiences (Chalmers 1995), does not seem so hard after all. The explanatory gap between objective physical processes, say, neural activity, and subjective experience, say, the feeling of pain, is imaginary since the signal triggers a whole array of neural processes that embody the relevant context, the very experiences of the subject. These qualia are thus unique to a subject as much as its history is unique.

Thermodynamics also makes sense of the capricious character of cognition. Namely, while running, trains of thought consume their motives, which, in turn, redirects locomotion. Thinking changes thoughts. Since causes and consequences are not only connected but, in fact, inseparable, courses are intractable, non-computable, nondeterminate. Thus, future thoughts are not fully foreseeable because consequences give rise to new causes, not because of the complexity of cognition or randomness of its processes (Chater et al. 2006), let alone because of quantum mechanics' uncertainty principle (Bruza et al. 2015).

Moreover, since all systems gravitate toward balance, intention is not unique to cognition. Across all scales, nature is goal-oriented toward free energy minima. Rivers run to sea, leaves turn to light, and predators follow prey. We orient ourselves along the lines of force, often by sheer reflex (Sokolov 1960). If there were an exception to this least-time rule, it ought to stand out from the ubiquitous patterns in data.

Cognition's subjective nature challenges traditional theorizing, which assumes an objective viewpoint. However, objectivity is an illusion since nothing can be seen without interacting. The interaction is subjective because each indivisible force carrier flows from one subject to another, not from all subjects to all others. So, each system is subject only to those forces it senses. Conversely, another perspective presents other forces. Still, the least-time free energy consumption is an objective gauge of the subjective status.

Thermodynamic insight into human thinking also clarifies, for example, that conflicting forces cause discomfort (Festinger 1957), at worst threatening to tear apart the ability to decide at all. A way out of cognitive dissonance may open by putting things into perspective rather than overlooking some of them. For instance, emphasizing inalienable rights may restore one's consonance that quivers under social pressure. Similarly, living up to one's convictions is liberating the associated free energy, while giving them up is enslaving, i.e., restraining one's potential to act.

Traditional ways of thinking have been perfected, but they may not be perfect. One may misread the situation simply because it closely resembles one in the past. Then, the well-trodden lines of thought may stray from accurate responses. Through a ready-made frame, things are understood with ease, and often, the availability bias is accompanied by the overconfidence bias. Moreover, the first impressions are the most lasting (Kahneman 2011) because restructuring the context built upon them consumes a lot of free energy.

Despite being aware of the perils of preconceptions, it is hard to reconsider. When pondering paralyses, even wrong reasoning is held right simply because it at least leads somewhere. Also, to get going, contradicting information is rejected or reinterpreted to confirm rather than confront biases. Cognitive ease compares to a river flowing with speed after having carved its way through the terrain or having been channeled to an aqueduct. In this way, confirmation, availability, recency, and framing biases (Tversky and Kahneman 1973) can be understood to stem from the least-time maxim. Also, in-group bias can be understood to derive from consuming free energy faster collectively than solitarily.

Moreover, according to thermodynamics, memory is not for remembering accurately but for functioning efficiently.

We remember relevant experiences and tend to forget, also actively, irrelevant ones. We might even recall something that did not happen if it is important enough, as has been revealed by examining eyewitness testimonies (De Brigard 2014). Vivid recollection resurrects imaginations (Hirst and Echterhoff 2012).

Besides exposing objectivity as an illusion, thermodynamics reveals repeatability as an improper scientific ideal because no situation can be reconstructed exactly. Even if we were to excite a single atom repeatedly, with only one photon each time, those quanta of light would be taken from the surrounding sources. The circumstances would thereby change. Since no process is independent of the background, the ceteris paribus assumption does not hold. Particularly in cognitive studies, repeatability is an elusive aim. For a stimulus to be a stimulus, it must alter cognition.

It is worth underscoring that cognition is and should be biased to consume free energy in the least time. Errors are seen in hindsight or from another, wider perspective. Still, deeds are rightfully deemed reprehensible when a more comprehensible context was already available at the time. However, when forces have played out, an incidence is often seen as having been foreseeable, even when it was not. Obviously, such an appraisal itself is biased, not least because the language to discuss it is biased by cultural context (Sapir 1929; Whorf and Carroll 1956). Nevertheless, metacognition, i.e., reflection on one's own cognitive processes, aims at learning one's lesson.

Finally, free energy also measures free will because only those things for which there are forces, i.e., causes, can be made to happen. In contrast to unattended automatic actions, free will manifests itself in recognizing alternatives and deciding between them. Consistently, judgments of the actions reflect arbiters' perspectives on how free energy should have been consumed. As much as free energy is consumed, events become irreversible, and as much as actions affect motives affecting actions, events become unpredictable. Since a system is not deterministic unless stationary and not random but causal, both determinism and indeterminism are fallacious notions, while nondeterminism is the factual stance (Annila 2020). Compatibilism is thus out of the question.

Discussion

The ambiguity in what cognition is all about leaves a lot of room for theorizing, but then again calls for compelling reasoning. While thermodynamics founded on atomism seems comprehensive enough, more pertinent is whether the tenet is good enough. According to Thomas Kuhn, a sound scientific theory is accurate, consistent, comprehensive, simple, and fruitful (Kuhn 1970).

As for accuracy, the equation of evolution reproduces the observed skewed distributions accumulating in a sigmoidal manner, and hence trending power laws, as well as oscillations and chaotic courses that also characterize cognition. For example, response times, memory recalls, word frequencies, and social connections are distributed in a power-law manner, and brain activity and decision-making display oscillations, even chaos (Mäkelä and Annila 2010; Annila 2016).

In regard to self-consistency, thermodynamics ascribes all causes to forces irrespective of whether they stem from the natural environment, social setting, or subject itself. Thus, the tenet is also consistent with Piaget's theory of cognitive development, where cognitive structures interpret and integrate information (Piaget 1952). Likewise, Vygotsky's sociocultural context presents itself to a subject as various forces (Vygotsky 1978). In reference to ecological action-oriented perception, free energy compares to affordance, i.e., what the environment offers the subject (Gibson 1966). It is also in line with the least-time maxim that it depends on the circumstances whether to employ intuitive, unconscious conduct or deliberate, conscious behavior, or eventually something subconscious in between.

In regard to consistency with information theories (Shannon and Weaver 1949), game theory is seen as a model of behavior reproducing thermodynamic aspects such as subjectivity (Anttila and Annila 2011). Still, the true target is to minimize free energy, not to maximize equivocal utility (Von Neumann and Morgenstern 1944; Nash 1951) or prospect (Kahneman 1979). Likewise, the Free Energy Principle (FEP) (Friston 2010), minimizing in a Bayesian manner information-theoretic surprise or prediction error (Den Ouden et al. 2012) between an internal model and sensory inputs, resembles but does not equate with the nondeterministic course toward free energy minimum. Specifically, despite its name, FEP is not the principle of physics derived in the Appendix from the atomistic axiom. Instead, Friston's free energy, just like Jaynes's entropy (Jaynes 1957a, b), is an information-theoretic measure, inconsistent with the physical embodiment of information (Karnani et al. 2009). Indeed, FEP is rightfully questioned as a proper thermodynamic account of systems (Beni 2021; Colombo and Palacios 2021). Moreover, in reference to the computational theory of mind, thermodynamics also represents objects, however, not in abstract terms of information, but in universal terms of energy. Information is what information does (Adriaans and Van Benthem 2008; Karnani et al. 2009). Accordingly, the least-time free energy consumption is the natural way to organize networks (Hartonen and Annila 2012) rather than, e.g., deep learning to construct large language models.

Kuhn's call for broad scope is, per definition, innate to thermodynamics. Based on the atomistic axiom, the theory explains a wide range of phenomena, from chemical reactions to the emergence of the biosphere (Karnani and Annila 2009), from a small purchase to world trade (Annila and Salthe 2009), and from a falling body to the expansion of the universe (Annila and Wikström 2022). Cognitive phenomena also compare to other processes by the same principle, not merely by metaphors.

By the same token, Kuhn's quest for simplicity is satisfied. The trivial yet universal least-time maxim brings order to phenomena that otherwise would be individually isolated and, as a set, confused. For example, sleep is understood to restore metabolic, cellular, and cognitive balance offset by focused activities (Annila 2016), even by mere idleness (Greicius et al. 2003). By spanning a wide range of frequencies, brain waves spread out, leveling out imbalance and thereby creatively connecting seemingly unrelated concepts.

Finally, in its all-encompassing nature, thermodynamics is fruitful in revealing new phenomena and unknown connections among those already known. No longer does evolution stand out exclusively as a biological phenomenon demanding a dedicated dogma, and no longer does cognition outrank solely as mental information processing requiring a distinct doctrine. Paradoxically, the more comprehensive the theory, the less there is to explain beyond self-evident.

While Kuhn underlined accuracy, consistency, scope, simplicity, and fruitfulness, he understood their evaluation as demanding. Cognition is conservative by nature, however desirable advances in thoughts about thinking might be. Since current comprehension is constructed in the least time from the materials of the past, we cannot stomach just any theory, just as we cannot digest just any matter without processing it. So, we are not free of our prior thoughts. Even if we want to think differently, it takes time and energy, literally, quanta, to revise.

Appendix: Statistical Physics of Open Systems

According to thermodynamics, not only gas atoms through collisions and compounds through reactions, but everything through various transformations evolves toward thermodynamic balance. Indeed, data look alike. Irrespective of source, scope, or scale, distributions are skewed, nearly lognormal, and thus accumulating in a sigmoidal manner, and hence trending power laws. The underlying universal law, the second law of thermodynamics, can be derived from the statistical physics of open systems. The many-body theory is based on the ancient axiom that everything comprises the same fundamental elements. Then, the equation of evolution from one state to another can be written even when the components of a system are not known explicitly.

The State Equation

Let us examine an entity, indexed with j, in the energy level diagram (Fig. 1). This entity exists with probability, ${}_1P_j = \Pi_k \phi_k$, combining all ingredient densities in energy, ϕ_k (Gibbs 1928). Due to the product form, if any one k-ingredient were missing altogether, $\phi_k = 0$, then also ${}_1P_j = 0$. For example, a neurotransmitter molecule could not possibly exist, as it is, if any one of its ingredient atoms were missing.

A pool of j-entities, N_j, exists with probability $P_j = ({}_1P_j)^{N_j}/N_j!$. Again, if any entity were missing altogether, also $P_j = 0$. The order of identical entities makes no difference, hence the division by the number of orderings, $N_j!$.

The total P of a system housing diverse entities

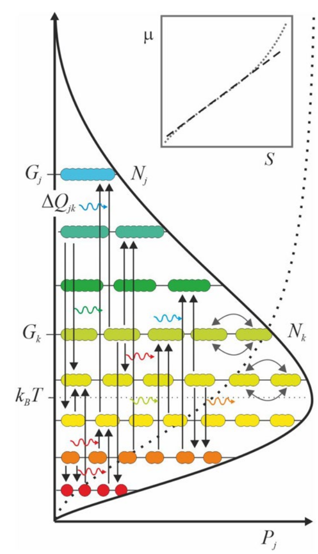

Fig. 1 Energy level diagram presents any system as everything comprises the same fundamental elements, the quanta. Entities with the same energy, G_k, in numbers N_k, are on the same level. Their mutual exchange (bow arrows), by changing nothing, causes no change in the average energy of the system, k_BT. Things change (vertical arrows) when entities move from one level to another. For example, in a chemical reaction, starting materials, N_k, transform into products, N_j, coupling also light quanta, ΔQ_{jk}, (wavy arrows) from the environment. Through flows of quanta, the system and its surroundings move toward thermodynamic balance. When the surroundings are higher in energy than the system, the system evolves higher in energy, and vice versa. The cumulative probability distribution (dotted line) is a sigmoid. Its logarithm, entropy, S, as a function of potential energy, μ, follows a power law closely, i.e., a straight line on the logarithm-logarithm scale (inset)

$$P = \prod_{j=1} P_j = \prod_{j=1} \left(\prod_{k=1} \phi_k \right)^{N_j} \Big/ N_j! \tag{1}$$

pools all densities in energy, $\phi_k = N_k \exp[(-\Delta G_{jk} + i\Delta Q_{jk})/k_B T]$, where the photon flux, $i\Delta Q_{jk}$, that couples to the jk-transformation bridges the energy gap, $-\Delta G_{jk}$, between the k-ingredient and the j-product, relative to the average energy of the system, $k_B T$, proportional to temperature, T, by Boltzmann's constant, k_B. When each event perturbs $k_B T$ only slightly, statistics limits to the self-similar continuous compounding, $f(x) = e^x = de^x/dt$ (Gibbs 1928). For example, a stream of light powers photosynthesis in a chloroplast comprising numerous j,k-components; a mitochondrion dissipates photons in breaking down glucose for ATP. As a convention, the i-prefix distinguishes the energy flows between the open system and its surroundings from the system-bound energy densities (Griffiths 2005; Tuisku et al. 2009).

The state equation (1) is the main result; thermodynamics follows from straightforward mathematical derivation.

For historical reasons, the additive measure of state, $\ln P$, is multiplied by k_B to entropy

$$S = k_B \ln P = k_B \sum_j \ln P_j \approx \frac{1}{T} \sum_{jk} N_j(-\Delta\mu_{jk} + i\Delta Q_{jk} + k_B T), \tag{2}$$

where the total energy, TS, sums the energy bound in the entities, $\sum N_j k_B T$, and the energy that is free, $\sum N_j(-\Delta\mu_{jk} + i\Delta Q_{jk})$, to consume differences, $\Delta\mu_{jk} = \mu_j - \mu_k$, between the potentials, $\mu_k = k_B T \ln \phi_k$ and $\mu_j = k_B T \ln \phi_j$, as well as the flux, $i\Delta Q_{jk}$, between the system and its surroundings. The approximation, $\ln N_j! \approx N_j \ln N_j - N_j$, is excellent for $N_j > 10$.

It is worth emphasizing that entropy, $S = k_B \ln P$, enumerates states distinct in energy, not configurations indistinguishable in energy, i.e., microstates with $k_B \ln W$. Hence, entropy does not equal to disorder. Disorder, just like order, is a consequence of consuming free energy, not an end in itself.

The Equation of Evolution

Evolution from one state to another can be counted statistically with differentials, dN_j, over time, dt,

$$\begin{aligned} T\frac{dS}{dt} &= T \sum_j \frac{dS}{dN_j}\frac{dN_j}{dt} \\ &= \sum_{jk} \frac{dN_j}{dt}\left(-\Delta\mu_{jk} + i\Delta Q_{jk}\right) \end{aligned} \tag{3}$$

to see that free energy, $\sum_k(-\Delta\mu_{jk} + i\Delta Q_{jk})$, forwards $dN_j/dt > 0$, while the opposite force reverses, $dN_j/dt < 0$. Thus, $dS > 0$ until at balance, where $\sum_k(-\Delta\mu_{jk} + i\Delta Q_{jk}) = 0$ and $dN_j/dt = 0$.

Despite being exact, the equation of evolution cannot be solved since the change,

$$\frac{dN_j}{dt} = \frac{1}{k_B T} \sum_k \sigma_{jk}\left(-\Delta\mu_{jk} + i\Delta Q_{jk}\right), \tag{4}$$

proportional to free energy by mechanism-dependent factors, $\sigma_{jk} > 0$, cannot be separated from the driving forces, i.e., $\Delta\mu_{jk}$ is a function of N_j. Due to this interdependence, chains of events are fundamentally unpredictable, not due to the complexity of a system or ambiguity in its initial conditions. For example, an increase in cognitive capacity may deliver more resources to build up even more capacity, and so on. Thus, the outcome cannot be determined at the onset. Still, evolutionary courses can be simulated a step at a time according to Eq. (4) to demonstrate the emergence of standards, skewed distributions, growth curves, oscillations, and even chaotic courses (Mäkelä and Annila 2010).

The flows of energy *naturally select* the mechanisms that bring about balance in the least time. For example, a neuronal circuit facilitating free energy consumption thrives, while a noncontributing population dwindles down by lacking in free energy. Stimuli keep neurons alive.

The Universal Patterns

The characteristic S-shape of a growth curve, such as a learning curve, can be worked out from Eq. (4). Initially, when resources are abundant, mechanisms, σ_{jk}, limit the rate of free energy consumption,

$$\begin{aligned} &\frac{d}{dt}\frac{1}{k_B T}\sum_{k=1}\left(-\Delta\mu_{jk}+i\Delta Q_{jk}\right) \\ &= \frac{dN_j}{dt}\frac{d}{dN_j}\frac{1}{k_B T}\sum_{k=1}\left(-\Delta\mu_{jk}+i\Delta Q_{jk}\right) \approx \sum_{k=1}\sigma_{jk} \\ &\Rightarrow \frac{dN_j}{N_j} = \sum_{k=1}\sigma_{jk}dt \Rightarrow N_j(t) = N_j(t_o)\exp\left(\sum_{k=1}\sigma_{jk}t\right), \end{aligned} \tag{5}$$

where $d\mu_j/dN_j = d(G_j + k_B T \ln N_j)/dN_j = k_B T/N_j$, whereas μ_k, Q_j, and Q_k have no explicit but only a stoichiometric dependence on N_j. The initial growth is thus nearly exponential. Similarly, mechanisms limit consumption near balance, where $N_j(t)$ levels off almost exponentially.

The power-law region between the initial and final phases can be deduced from $N_j = \prod_k \phi_k = \alpha_j N_1^j$, as the product of its k-constituents, each the product of the basic elements, N_1, where $\alpha_j = \prod_{mn} \exp[(-\Delta\mu_{mn} + i\Delta Q_{mn})/k_B T]$ compiles the free energy terms that assemble N_j from N_1. So, the change

$$\frac{dN_j}{dt} = j\alpha_j N_1^{j-1} \frac{dN_1}{dt} = j\frac{N_j}{N_1}\frac{dN_1}{dt} \Rightarrow \frac{dN_j}{N_j} = j\frac{dN_1}{N_1} \tag{6}$$

when integrated, follows a power law $\ln N_j = j \ln N_1 +$ constant.

When the assumption of a nearly constant change in free energy does not hold, the change can be modeled by adding the term $-\beta N_j$ to Eq. (5)

$$\begin{aligned} \frac{dN_j}{N_j} &\approx \sum_{k=1} \left(\sigma_{jk} - \beta N_j\right) dt \Rightarrow N_j(t) \\ &= N_j(t_o)\left(\sum_{k=1} \sigma_{jk} - \beta N_j(t_o)\right) \end{aligned} \tag{7}$$

where $N_j(t_o)$ at a time, t_o, determines $N_j(t)$, at a later time, t (May 1976). When $|-\Delta\mu_{jk} + i\Delta Q_{jk}|/k_BT << 1$, evolution is almost predictable, and when not, oscillations and chaos occur. For example, a rapidly proliferating population oscillates by exceeding the environment's carrying capacity time after time.

Distribution about the representative, mean density in energy, ϕ_j, in terms of the elemental density, ϕ_1,

$$\ln \phi_{j-n\ldots j+n} = \ln \phi_j + \sum_n n \ln \phi_1 \tag{8}$$

is nearly lognormal. Despite the distribution's long tails, the typical form, ϕ_j, can be recognized in each member, $j \pm n$. For example, all-scaled snakes slither like snakes and not like lizards. Also, spirals, such as shells, cyclones, and galaxies, are approximately lognormal distributions in polar coordinates, i.e., energetically optimal shapes.

It is worth underlining that lognormal distributions, logistic cumulative curves, power laws, and so on, are models of the axiom-derived Eqs. (3) and (4), not explanations of natural processes.

In summary, derived from the statistical physics of open systems, thermodynamics accounts for all processes as flows of quanta. The arrow of time is inherent in the quantum itself, carrying energy, E, on its period of time, t; Planck's constant, $h = Et$, differentiates to the power, $dE/dt = -E/t = -\mathbf{F} \cdot \mathbf{v}$. Motion with velocity, $\mathbf{v}$, directs along the least-time path along the force, $-\mathbf{F}$.

Acknowledgments I am deeply grateful to the reviewers for their critical examination and constructive comments, which improved the manuscript considerably.

Funding Open Access funding provided by University of Helsinki (including Helsinki University Central Hospital). The study has not received any funding.

Data availibility All data generated or analyzed during this study are included in this published article.

Declarations

Conflict of interest The author declares that he has no conflict of interest.

References

Adriaans P, Van Benthem J (2008) Introduction: information is what information does. Handbook of the philosophy of science. Elsevier, North Holland, pp 3–26

Annila A (2012) Physical portrayal of computational complexity. International Scholarly Research Notices, 321372–115 arXiv:0906.1084

Annila A (2016) On the character of consciousness. Front Syst Neurosci 10:27

Annila A (2020) Back to reality. Privus Press, New York

Annila A (2021) The matter of time. Entropy 23(8):943

Annila A, Annila E (2008) Why did life emerge? Int J Astrobiol 7(3–4):293–300

Annila A, Salthe S (2009) Economies evolve by energy dispersal. Entropy 11(4):606–633

Annila A, Salthe S (2010) Cultural naturalism. Entropy 12(6):1325–1343

Annila A, Salthe S (2010) Physical foundations of evolutionary theory. J Non-Equilib Thermodyn 345:301–321

Annila A, Wikström M (2022) Dark matter and dark energy denote the gravitation of the expanding universe. Front Phys 10:995977

Anttila J, Annila A (2011) Natural games. Phys Lett A 375(43):3755–3761

Bak P, Tang C, Wiesenfeld K (1988) Self-organized criticality. Phys Rev A 38(1):364

Beni MD (2021) A critical analysis of Markovian monism. Synthese 199(3):6407–6427

Beni MD (2023) A free energy reconstruction of arguments for panpsychism. Phenomenol Cogn Sci 22(2):399–416

Bruza PD, Wang Z, Busemeyer JR (2015) Quantum cognition: a new theoretical approach to psychology. Trends Cogn Sci 19(7):383–393

Carlton EH, Shepard RN (1990a) Psychologically simple motions as geodesic paths I. Asymmetric objects. J Math Psychol 34(2):127–188

Carlton EH, Shepard RN (1990b) Psychologically simple motions as geodesic paths II. Symmetric objects. J Math Psychol 34(2):189–228

Chalmers DJ (1995) Facing up to the problem of consciousness. J Conscious Stud 2(3):200–219

Chang H (2017) Vi-operational coherence as the source of truth. In: Proceedings of the Aristotelian Society, vol. 117, pp. 103–122 . Oxford University Press

Chang H (2022) Realism for realistic people: a new pragmatist philosophy of science. Cambridge University Press, Cambridge, UK

Chater N, Brown GD (2008) From universal laws of cognition to specific cognitive models. Cogn Sci 32(1):36–67
Chater N, Tenenbaum JB, Yuille A (2006) Probabilistic models of cognition: conceptual foundations. Trends Cogn Sci 10(7):287–291
Chialvo RD (2004) Critical brain networks. Physica A 340(4):756–765
Christensen K, Danon L, Scanlon T, Bak P (2002) Unified scaling law for earthquakes. Proc Natl Acad Sci USA 99(suppl1):2509–2513
Clark A (2013) Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behav Brain Sci 36(3):181–204
Clauset A, Shalizi CR, Newman ME (2009) Power-law distributions in empirical data. SIAM Rev 51(4):661–703
Colombo M, Palacios P (2021) Non-equilibrium thermodynamics and the free energy principle in biology. Biol Philos 36(5):41
De Brigard F (2014) Is memory for remembering? Recollection as a form of episodic hypothetical thinking. Synthese 191:155–185
Den Ouden HE, Kok P, De Lange FP (2012) How prediction errors shape perception, attention, and motivation. Front Psychol 3:548
Diamond JM (2005) Collapse: how societies choose to fail or succeed. Penguin, New York
Epstein IR, Pojman JA (1998) An introduction to nonlinear chemical dynamics: oscillations, waves, patterns, and chaos. Oxford University Press, New York
Festinger L (1957) A theory of cognitive dissonance. Stanford University Press, Redwood City
Fournier GP, Andam CP, Gogarten JP (2015) Ancient horizontal gene transfer and the last common ancestors. BMC Evol Biol 15:1–18
Friston K (2009) The free-energy principle: a rough guide to the brain? Trends Cogn Sci 13(7):293–301
Friston K (2010) The free-energy principle: a unified brain theory? Nat Rev Neurosci 11(2):127–138
Gabaix X (2016) Power laws in economics: an introduction. J Econ Perspect 30(1):185–206
Gibbs JW (1928) The collected works of J. Willard Gibbs: thermodynamics. Longmans & Green, New York
Gibson JJ (1966) The senses considered as perceptual systems. George Allen & Unwin, London
Goldstein K (1934) Der Aufbau des Organismus. Nijhoff, Den Haag
Greicius MD, Krasnow B, Reiss AL, Menon V (2003) Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. Proc Natl Acad Sci USA 100(1):253–258
Griffiths DJ (2005) Introduction to quantum mechanics. Pearson Prentice Hall, Upper Saddle River
Harris M (1977) Cannibals and kings: the origins of cultures. Random House, New York
Hartonen T, Annila A (2012) Natural networks as thermodynamic systems. Complexity 18(2):53–62
He BJ, Zempel JM, Snyder AZ, Raichle ME (2010) The temporal structures and functional significance of scale-free brain activity. Neuron 66(3):353–369
Heyes CM, Huber L (2000) The evolution of cognition. MIT Press, Cambridge
Hirst W, Echterhoff G (2012) Remembering in conversations: the social sharing and reshaping of memories. Ann Rev Psychol 63:55–79
Hohwy J (2013) The predictive mind. Oxford University Press, Oxford/New York
Huberman BA, Adamic LA (1999) Growth dynamics of the world-wide web. Nature 401(6749):131
Huxley J (1932) Problems of relative growth. L. MacVeagh, The Dial Press, New York
James W (1890) The principles of psychology, vol 1. Henry Holt and Company, New York
Jaynes ET (1957) Information theory and statistical mechanics. I. Phys Rev 106(4):620
Jaynes ET (1957) Information theory and statistical mechanics. II. Phys Rev 108(2):171
Jones BJ, Martínez VJ, Saar E, Trimble V (2005) Scaling laws in the distribution of galaxies. Rev Mod Phys 76(4):1211
Kahneman D (1979) Prospect theory: an analysis of decisions under risk. Econometrica 47:278
Kahneman D (2011) Thinking, fast and slow. Farrar, Straus and Giroux, New York
Kant I (1893) Critique of pure reason. G. Bell & Sons, New York
Karnani M, Annila A (2009) Gaia again. Biosystems 95(1):82–87
Karnani M, Pääkkönen K, Annila A (2009) The physical character of information. Proc R Soc A Math Phys Eng Sci 465(2107):2155–2175
Kello CT, Brown GD, Ferrer-i-Cancho R, Holden JG, Linkenkaer-Hansen K, Rhodes T, Van Orden GC (2010) Scaling laws in cognitive sciences. Trends Cogn Sci 14(5):223–232
Kelso JS (1995) Dynamic patterns: the self-organization of brain and behavior. MIT Press, Cambridge
Kováč L (2023) Cognition: from the cosmos down to molecules–and back: the possible demise of humanity on earth would be just a trivial episode in the grand story of the evolution of cognition in the cosmos. EMBO Reports, 57263
Kováč L (2000) Fundamental principles of cognitive biology. Evol Cogn 6(1):51–69
Kováč L (2015) First movement. Life as a cosmic imperative. Closing human evolution: life in the ultimate age. Springer, Cham, pp 1–35
Kuhn TS (1970) The structure of scientific revolutions. University of Chicago Press, Urbana
Landauer R (1961) Irreversibility and heat generation in the computing process. IBM J Res Dev 5(3):183–191
Lehmonen L, Annila A (2022) Natural classes and natural classification. Efficiency in complex systems: self-organization towards increased efficiency. Springer, Berlin, pp 11–26
Lewis GN (1926) The conservation of photons. Nature 118(2981):874–875
Limpert E, Stahel WA, Abbt M (2001) Log-normal distributions across the sciences: keys and clues. BioScience 51(5):341–352
Logan R, Chynoweth A, Cohen B (1962) Avalanche breakdown in gallium arsenide p- n junctions. Phys Rev 128(6):2518
Lotka AJ (1926) The frequency distribution of scientific productivity. J Wash Acad Sci 16(12):317–323
Lovelock JE (1972) Gaia as seen through the atmosphere. Atmos Environ (1967) 6(8):579–580
Mäkelä T, Annila A (2010) Natural patterns of energy dispersal. Phys Life Rev 7(4):477–498
Mantegna RN, Stanley HE (1995) Scaling behaviour in the dynamics of an economic index. Nature 376(6535):46–49
May RM (1976) Simple mathematical models with very complicated dynamics. Nature 261(5560):459–467
Milton JG, Chkhenkeli SA, Towle VL (2007) Brain connectivity and the spread of epileptic seizures. Handbook of brain connectivity. Springer, Berlin, pp 477–503
Nagel T (1974) What is it like to be a bat? Philos Rev 83(4):435–450
Nash J (1951) Non-cooperative games. Ann Math 54:286–295
Newman ME (2005) Power laws, Pareto distributions and Zipf's law. Contemp Phys 46(5):323–351
Nihei Y, Higuchi H (2002) When and where did crows learn to use automobiles as nutcrackers? PhD thesis, Tohoku University
O'Reilly UM, Hemberg E (2020) From biological to computational arms races - studying cyber security dynamics. In: Banzhaf W (ed) Evolution in action: past, present and future: a festschrift in honor of Erik D. Goodman. Springer, Cham, pp 409–422
Pareto V (1897) Cours D'Ëconomie Politique: Professé l'Université de Lausanne, vol 2. F. Rouge, Lausanne
Peirce CS (1931) Collected papers of Charles Sanders Peirce. Harvard University Press, Cambridge

Pernu TK, Annila A (2012) Natural emergence. Complexity 17(5):44–47
Piaget J (1952) The origins of intelligence in children, vol 8. International Universities Press, New York
Polanyi M (1967) The tacit dimension. Routledge, New York
Price de Solla DJ (1965) Networks of scientific papers: the pattern of bibliographic references indicates the nature of the scientific research front. Science 149(3683):510–515
Prigogine I, Wiame J-M (1946) Biologie et thermodynamique des phénomènes irréversibles. Experientia 2:451–453
Richardson LF (1960) Statistics of deadly quarrels. Boxwood Press, Pittsburgh
Sapir E (1929) The status of linguistics as a science. Language 5:207–214
Shannon CE, Weaver W (1949) The mathematical theory of communication. University of Illinois Press, Urbana
Sharma V, Kaila VR, Annila A (2009) Protein folding as an evolutionary process. Physica A 388(6):851–862
Shettleworth SJ (2009) Cognition, evolution, and behavior. Oxford University Press, New York
Simon HA (1955) A behavioral model of rational choice. Q J Econ 69:99–118
Snell O (1892) Die abhängigkeit des hirngewichtes von dem Körpergewicht und den geistigen Fähigkeiten. Arch Psychiatr Nervenkrankheiten 23(2):436–446
Sokolov EN (1960) Neuronal models and the orienting reflex. In: Brazier MAB (ed) The central nervous system and behavior. Josiah Macy Foundation, Jr., New York, p 187276
Stevens SS (1957) On the psychophysical law. Psychol Rev 64(3):153
Sutherland JD (2017) Opinion: Studies on the origin of life-the end of the beginning. Nat Rev Chem 1(2):0012
Thompson DW (1917) On growth and form. Cambridge University Press, Cambridge
Tuisku P, Pernu TK, Annila A (2009) In the light of time. Proc R Soc A Math Phys Eng Sci 465(2104):1173–1198
Tversky A, Kahneman D (1973) Availability: a heuristic for judging frequency and probability. Cogn Psychol 5(2):207–232
van Horik J, Emery NJ (2011) Evolution of cognition. Wiley Interdiscipl Rev Cogn Sci 2(6):621–633
Von Neumann J, Morgenstern O (1944) Theory of games and economic behavior. Princeton University Press, Princeton
Vygotsky LS (1978) Mind in society: development of higher psychological processes. Harvard University Press, Cambridge
West GB, Brown JH (2004) Life's universal scaling laws. Phys Today 57(9):36–42
Whorf BL, Carroll JB (1956) Language, thought, and reality: selected writings of Benjamin Lee Whorf. MIT Press, Cumberland
Wixted JT, Ebbesen EB (1997) Genuine power curves in forgetting: a quantitative analysis of individual subject forgetting functions. Memory Cogn 25:731–739
Zipf GK (1949) Human behavior and the principle of least effort: an introduction to human ecology. Addison-Wesley Press, Cambridge