Review

Thoughts and thinkers: On the complementarity between objects and processes

Chris Fields [a,*], Michael Levin [a,b]

[a] *Allen Discovery Center at Tufts University, Medford, MA 02155, USA*
[b] *Wyss Institute for Biologically Inspired Engineering at Harvard University, Boston, MA 02115, USA*

ARTICLE INFO

ABSTRACT

We argue that "processes versus objects" is not a useful dichotomy. There is, instead, substantial theoretical utility in viewing "objects" and "processes" as complementary ways of describing persistence through time, and hence the possibility of observation and manipulation. This way of thinking highlights the role of memory as an essential resource for observation, and makes it clear that "memory" and "time" are also mutually inter-defined, complementary concepts. We formulate our approach in terms of the Free Energy Principle (FEP) of Friston and colleagues and the fundamental idea from quantum theory that physical interactions can be represented by linear operators. Following Levin (2024) [30], we emphasize that memory is, first and foremost, an interpretative function, from which the idea of memory as a record, at some level of accuracy, of past events is derivative. We conclude that the distinction between objects and processes is always contrived, and always misleading, and that science would be better served by abandoning it entirely.

> Thought is itself the thinker.
> [William James [1] p. 401]

1. Introduction

Is it advantageous to think of the world in terms of processes instead of objects, as suggested by Whitehead [2] among others? Is it, in particular, advantageous to think of organisms as processes instead of objects? Does doing so resolve, or avoid, the problem of identity over time, or does it merely reformulate this problem? Does adopting a process view of biology, for example, help us understand what it means to say that an individual human is an individual, something that maintains a single identity from birth to death, despite continuous change in properties at multiple levels? Does it help us to understand what it means to say that Homo sapiens sapiens has remained the same subspecies for the past ca. 250,000 years? Does it help us decide whether a clonal colony of E. coli, or a plate full of adult planaria that are all regenerated from the fragments of a single planarian cut up in a laboratory, comprises a single "individual organism" or many? And, reaching forward towards empirical utility in fields like regenerative medicine and bioengineering, does it facilitate interesting new discoveries, capabilities, and research programs?

The notations employed in much of mathematics suggest – though do not require – an interpretation in terms of processes: writing '$f(x)$', for example, suggests a process f acting on an object x, while the notation '$A \rightarrow B$' suggests a process, indicated by '$\rightarrow$' that

---

* Corresponding author.
   E-mail address: fieldsres@gmail.com (C. Fields).

converts A into B, or perhaps moves something from A to B. The current foundational language for mathematics is category theory. A (small) category comprises a set of O objects and a set M of morphisms, or arrows, such that:

1. for every object o in O, there is a morphism Id_o in M, the "identity morphism for o", such that $\text{Id}_o : o \to o$, and
2. for all objects o_1, o_2, o_3 and morphisms $m_{12} : o_1 \to o_2$ and $m_{23} : o_2 \to o_3$, there is a morphism $m_{13} : o_1 \to o_3$.

As Adámek, Herrlich and Strecker remark in their popular textbook [3], replacing every object with one of its identity morphisms (they are not in general unique) allows any category-theoretic statement to be made just in terms of morphisms. The Yoneda Lemma and the notion of object-free categories formalize this fact. Hence mathematics, in a deep sense, does not need objects. It just needs morphisms, some of which preserve an abstraction called "identity" and others of which do not.

Similar moves are made ubiquitously in physics. Quantum theory replaces "properties" of objects – position, momentum, spin, etc. – with operators that yield values of those properties when deployed by an observer [4]. Quantum field theory replaces "particles" with creation, destruction, and number operators acting on fields that permeate spacetime. Feynman's diagrams famously replace "particles" with sums of processes that yield equivalent results when interrupted by observations. Even "quantum systems" can be replaced with the operational procedures required to identify them [5]. An operational procedure is a process that may or may not preserve the identity of either the system that it acts on or the system – conventionally called an "observer" or "perceiver" – that implements it.

The idea that objects are constructed by perceivers is foundational in 2nd-order cybernetics [6,7]. This idea was introduced into evolutionary and developmental biology by Maturana and Varela [8], where it became the notion of the joint "enaction" of an object by an organism and its environment. It was introduced into perceptual psychology by Gibson [9] as the idea that objects and their properties are "affordances" of an environment for a specific, behaving organism. In all of these formulations, perceived "objects" are emergent from the joint activity of a perceiver and its environment, not fundamental "furniture of the world" as sought by traditional ontologists [10].

All of these process-oriented ways of thinking, however, employ object-like concepts in their formulation. One can think of a morphism $m_{12} : o_1 \to o_2$ simply as an ordered pair (o_1, o_2); the ordered-pair notation suggests an object, not a process. Quantum-theoretic measurement operators must be deployed by observers to yield results [11]; observers are themselves quantum systems, i.e. objects. When such systems are replaced by identification criteria, these criteria must be implemented, again by a system – an object – acting as an observer. Enaction requires two systems, or objects, an organism and its environment, which interact in some well-defined way as discussed below. Affordances are for something, an organism, hence an object, and are offered by something, the organism's environment, another object. Objects, in other words, seem to be required after all.

We will argue here that "processes v/s objects" is not a useful dichotomy in either physics or the life sciences. There is, instead, substantial theoretical utility in viewing "objects" and "processes" as complementary ways of describing persistence through time, and hence the possibility of observation and manipulation. This way of thinking highlights the roles of memory as an essential resource for observation and of time as an essential resource for action, and makes it clear that "memory" and "time" are also mutually inter-defined, complementary concepts that both rest on the underlying notion of an "identity" that persists. We formulate our approach in terms of the Free Energy Principle (FEP) of Friston and colleagues [12–15,17,18], noting the close relationship between the FEP and the Holographic Principle (HP) of physics [19–22] and that the fundamental idea of both of these principles – that information passing through a surface can be represented as being encoded on that surface – is implicit in the Divergence Theorem, one of the fundamental theorems of vector calculus, first formulated in the early 19th century [23]. The FEP characterizes all time-persistent physical systems as implementing a process, active inference [24,25], which is a combination of reacting to (learning from) and acting on (manipulating) their environments. As a generic description of persistent physical systems as information-processing systems, the FEP provides a uniform way of thinking about information and memory across biological scales, from molecular networks in morphogenesis to evolutionary dynamics.

Following a brief review of the FEP and its conceptual underpinnings in §2, we apply this view of objects and processes as complementary to evolution and development in §3, asking "what evolves?" and "what develops?" We show how thinking of evolution and development as processes acting on processes enables a fully scale-free biology, i.e. a biology in which processes at every scale can be described using a single set of concepts and methods. We then consider the observer as a developmental system in §4, asking what is required for an observer to model itself as a persistent system, and hence to model its observations as a historical sequence. We re-examine in §5 the distinction between stigmergic memories – particularly those encoded as permanent, public records, such as this paper – and implicit or procedural memories such as those stored by machine learning (ML) systems, cellular networks [26,27], or biochemical pathways [28,29]. Following Levin [30], we emphasize that memory is, first and foremost, an interpretative function, from which the idea of memory as a record, at some level of accuracy, of past events is derivative. We conclude in §6 that the distinction between objects and processes is always contrived, and always misleading, and that science would be better served by abandoning it entirely. We then look forward to how doing so enables a bioscience of agentive materials that integrates biomedicine with bioengineering, a bioscience that is difficult to conceptualize from the traditional viewpoint that considers organisms – with the possible exception of humans – as biological machines implementing programs that are largely, if not completely, fixed by their genomes.

2. Systems, distinguishability, and persistence

The FEP is a mathematical framework for talking about persistence through time; indeed, the FEP can be stated, in obviously tautological form, as the claim that any system that persists through time must behave in a way that allows its persistence through time [14]. From a formal perspective, the FEP is an interpretation of dynamical systems that can be decomposed into components
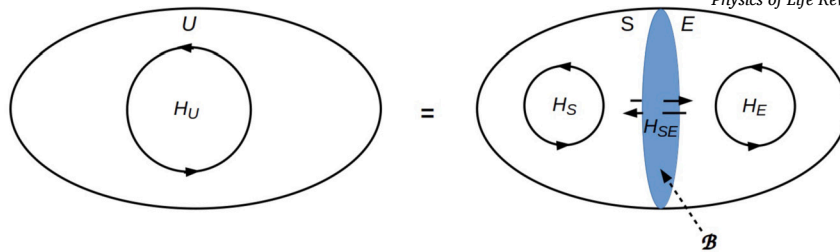
Fig. 1. Decomposing a system U into components S and E decomposes the self-interaction H_U of U into the sum $H_U = H_S + H_E + H_{SE}$. In the weak interaction limit, in which the dimension $\dim(H_{SE}) \ll \dim(H_S), \dim(H_E)$, the systems S and E will have "internal" states not directly involved in the interaction, as required for systems compliant with the FEP. In this case, the interaction H_{SE} can be represented as defined at a notional boundary $\mathcal{B}$ separating S from E. This boundary $\mathcal{B}$ is not a component of the joint state space $\mathcal{H}_U$, but is rather an ancillary construct that provides a convenient representation for the interaction, as described in [22].

that maintain conditionally statistically independent states over some period of time. Note that the FEP characterizes a system only while it persists – maintains a state that is conditionally statistically independent from that of its environment – through time, and says nothing about how long any system will be successful in behaving in a way that allows its continued persistence. Understanding the FEP as a statement of principle requires understanding what "persistence", "time", and "behavior" are. The formal theory of the FEP is, effectively, an answer to these three questions. This formal theory has evolved significantly since it was first proposed; see [14,17,18] for current statements of the theory using classical statistical physics, [15] for a restatement in the language of quantum information theory, and [16] for a detailed comparison of the classical and quantum formulations. We will not replicate these formal presentations here, but will rather examine, in reverse order, how the theory addresses the fundamental questions of persistence, time, and behavior.

2.1. Behavior and time in the FEP

First, what is behavior? The FEP follows physics in general in identifying behavior with physical interaction, and then follows statistical physics – equivalently, information theory, cybernetics, or even computer science – in identifying physical interaction with information exchange. The "space" in which this behavior takes place is the joint state space of the system of interest S and its environment E, i.e. the state space of U in Fig. 1. The roots of the identification of behavior with information exchange again go back to the 19th century, to Clausius' original definition of entropy, $\Delta \mathbf{S} = \Delta \mathbf{E}/T$, where $\mathbf{S}$ is entropy, $\mathbf{E}$ is energy, and T is temperature, and Boltzmann's interpretation of entropy in terms of uncertainty, $\mathbf{S} = k_B \ln\Omega$, where k_B is Boltzmann's constant and Ω is the number of observationally indistinguishable states of the system of interest. Combining these two in the case of a binary system ($\Omega = 2$) yields $\Delta \mathbf{E} = \ln 2 k_B T$, which is Landauer's Principle [31,32], now recognized as the fundamental connection between energy and information [33,34].

Physical interactions can be of two types: interactions within a system (internal or "self" interactions) and interactions between systems. In order to describe these interactions, we make a fundamental assumption: that "standard" or "textbook" quantum theory, with the Dirac-von Neumann axioms as given in, e.g. [35], is correct. With this assumption, the state spaces of physical systems can be represented as Hilbert spaces, which for convenience we take to be finite dimensional, and interactions between physical systems can be represented as (again finite dimensional) linear operators, specifically Hamiltonians or total energy operators, operating on Hilbert spaces. This assumption rules out, in particular, any nonlinear interactions between quantum systems, and hence any objective, observer-independent "collapse of the wavefunction." While such nonlinear interactions have never been observed, they have not yet been conclusively ruled out by experiments [36,37].

Given the assumption that standard quantum theory is correct, we can write the total energy of any isolated system U as an operator, the Hamiltonian $H_U = H_S + H_E + H_{SE}$, where H_U, H_S, and H_E are the internal interactions of U, S, and E respectively and H_{SE} is the interaction between S and E. Fig. 1 illustrates this equation. The distinction between interactions within or between systems is, therefore, dependent on, and relative to, a decomposition of whole systems such as U into components such as S and E. Decomposing U in a different way, e.g. as $U = AB$ with $A \neq S$, would define different interactions: $H_U = H_A + H_B + H_{AB}$. This dependence of the distinction between self-interactions and other-interactions – interactions between systems – on the decomposition of U tells us that the distinction cannot be fundamental. Nothing about the physics of U favors any one decomposition – any one placement of an inter-component boundary – over any other.

Assuming standard quantum theory also provides a strict criterion for determining whether the components S and E specified by a decomposition of a system U have independently-specifiable, and hence conditionally statistically independent states, as is required if the FEP is to apply. The components S and E have independently-specifiable states if and only if their joint state factors, i.e. if and only if $|SE\rangle = |S\rangle|E\rangle$ in Dirac's notation. If the joint state does not factor, it is, by definition, entangled. Interactions between physical systems generically induce entanglement; an assumption that two systems are unentangled is, therefore, an assumption that they interact only weakly, and that even this weak interaction has not been observed for an asymptotically long time. Determining experimentally whether two systems are entangled is very difficult; the 2022 Nobel Prize in Physics was awarded to Aspect, Clauser, and Zeilinger for showing how this could be done. In practical settings, therefore, interacting systems that appear to be mutually distinguishable are assumed to be unentangled, i.e. assumed to have joint states that factor. The FEP applies only to systems that can

each be assumed to have its own, independently-specifiable, conditionally statistically independent internal states. We will refer to systems as "distinct" or "distinguishable" if they can, in practice, be assumed to satisfy this condition.

The above discussion refers to the "system" U and its components, the distinct "systems" S and E. A system is naturally thought of as an object: a collection of "degrees of freedom" each of which can be in some state or other. As seen in Fig. 1, however, each system is also characterized by its self-interaction. This self-interaction is a process; in the quantum formalism, the propagator $P_U(t) = exp((-\iota/\hbar)H_U(t))$, where $\hbar$ is the reduced Planck's constant, is the unitary operator that evolves the state of U forward in an assumed "background" [38] time t. Each degree of freedom x of U can, moreover, be replaced with a collection of binary degrees of freedom, one for each of the discernable values of x. We can, therefore, think of U simply as a collection of physically-implemented bits, or in the quantum case, qubits. We can, in other words, think of any state of U as an encoding of data, and think of the action of P_U on any such state as a computation on those data; as shown in [39], doing so only requires finding a mapping – a semantic interpretation – from states of U to some data structure, and from finite samples of the action of P_U to the action of some function on that data structure. Many such mappings are, moreover, always possible [40]. Provided that their joint state remains separable, the same goes for S and E.

If H_U, H_S, and H_E can be interpreted as implementing computations, how are we to interpret the interaction H_{SE}? Let us suppose that information in some measurable form is carried by discrete, detectable physical systems, such as individual photons, with each photon encoding one bit of information. The Divergence Theorem tells us that the number of photons, and hence the amount, in bits, of information emitted by any source within a system can be computed by computing the number of photons crossing any closed boundary around the system. The HP extends this result, postulating that the only information that can be obtained about a system by an external observer is the information that is carried across its boundary by physical systems capable of traversing the boundary. It states, moreover, that for any finite system, the maximal number of physical information carriers – again, e.g. photons – that can traverse its boundary is finite, and indeed proportional to the area of the boundary (formally, the maximum classical entropy $\mathbf{S}(\mathscr{B})$ of the boundary $\mathscr{B}$ is $A_{\mathscr{B}}/4$, where $A_{\mathscr{B}}$ is the area of $\mathscr{B}$ in Planck units). The boundary $\mathscr{B}$ separating S from E can, therefore, also be thought of as a collection of discrete, physically-implemented bits, or in the quantum case, physically-implemented qubits (see [16] or [22] for mathematical details and [41] for a less formal discussion). The interaction H_{SE} can, therefore, be regarded as finite bit exchange between S and E, i.e. as communication between S and E.

With this understanding of behavior, we are in a position to address time, the background time t employed above to construct the propagator $P_U(t)$. The FEP is about persistence through time, but where does "time" come from, and what role is it playing in the theory? To answer this question, it is useful to ask another: what happens to the information that flows, via the action of $P_U(t)$, into S from E? If H_S is viewed as implementing a computation, the answer is clear: information from E is transferred across $\mathscr{B}$, after which it serves as an input to whatever computation H_S implements. In FEP language, information arriving across $\mathscr{B}$ is "sensation." Information flow from S back to E is output from H_S – "action" in FEP language. Sending E a bit string involves writing these bits on $\mathscr{B}$; this takes energy, at least $ln2k_BT$ per bit. It also takes time: formally, action is the product, $\Delta E \Delta t$, of energy and time. Hence formally, time is built into the concept of interaction: it is what separates input from output. Any system S able to distinguish or record successive inputs from, or outputs to, another system E must implement, as one function of its internal interaction H_S, a time counter or clock that "ticks" between each input and the following output, and thus provides an available timestamp for recording inputs or outputs [15]. The internal time t_S counted by S is discrete, as it counts particular events, and is by definition coarse-grained with respect to the background time t. This internal time is strictly system-relative, as stressed by Di Biagio, Donà and Rovelli [42] among others.

Let us now describe S in the language of the FEP, considering S to both observe and manipulate E by reading from and writing on $\mathscr{B}$; this alternation between reading and writing is implemented by H_S and is, in FEP language, the process of active inference. Observation can be informative only if most of what is observed is not changing (or changing only slowly) in the internal time t_S; if everything is observed to be changing, all that is "seen" is noise. Manipulation similarly makes sense only if there is something to manipulate, and an effectively fixed background against which to manipulate it. Formally, these are conditions of "sparse coupling" or "weak interaction" required by the FEP [16]. Processes in E that act on $\mathscr{B}$, in other words, can be observed and manipulated only against a background of stasis. Time enables stasis: time provides a redundancy resource for – a "place to put" – the multiple "copies" of the observed background that enable it to be seen as unchanging, and hence allow other things to be seen as changing. Time is not just Nature's way of assuring that everything does not happen at once [43]; it is also Nature's way of allowing what happens to be noticed by observers.

Time has, therefore, two roles in the FEP: it is something that agents count by acting, and it is a resource made available to observers by their weak interactions with their environments. We can view these two roles at the level of fundamental physics, where weak interaction is required for distinctness from – absence of entanglement with – the environment as noted above. Hence weak interaction is required for H_S to be locally well-defined, i.e. defined on S alone, which is in turn required for H_S to implement an S-specific clock with which to count time relative to S. We can also view these roles at the level of problem solving, where distinctness from the environment is required for active inference, and active inference requires a distinction between input and output – sensation and action – that is provided by counting time. Active inference agents can manipulate their coupling to their environments by coarse-graining their inputs, or by employing selective attention – manipulating Bayesian precision – to ignore some inputs altogether. Change blindness [44], from this perspective, is not a cognitive deficit, but rather a way of maintaining the fixed background that renders attended-to changes salient. Maintaining a conversation in a crowded, noisy room, for example, would be impossible without the ability to not notice – be blind to – other conversations in the background. Manipulating coupling to the environment can be achieved by manipulating (agent-relative) time – processing fewer inputs – in order to minimize the energy
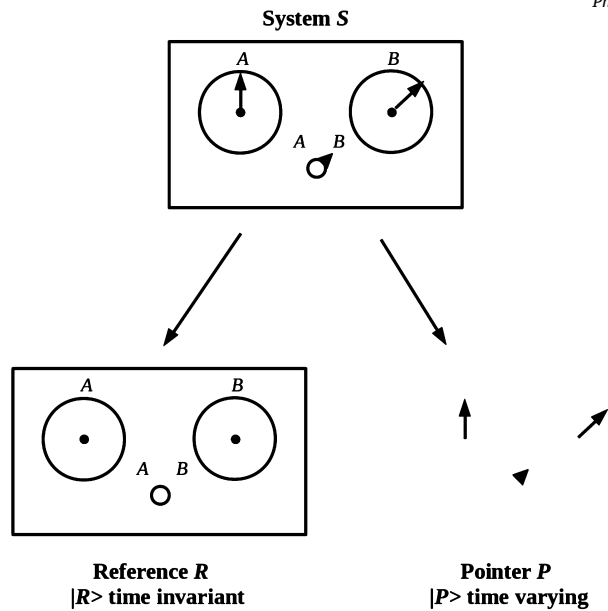
Fig. 2. Observing the state of any system requires first identifying a time-invariant "reference" component $|R\rangle$ of its current state, and then, having distinguished the object from its perceptible background, measuring any time-varying "pointer" components $|P\rangle$ of its state that are of interest. From [48], CC-BY license.

expended to keep track of what is happening. Agents accomplish this by using specific reference frames to selectively observe and act on their environments.

2.2. Examples: detecting changes in signals and objects

Let us first consider the detection of a signal in some medium. Analog radios that employ either amplitude modulation (AM) or frequency modulation (FM) provide a particularly clean example, and one with historical importance for theories of perception. Analog radio transmitters encode signals in the ambient photon field, the same medium that encodes visually-accessible information. We access these signals with appropriate detectors, i.e. analog radios.

Suppose now that you are listening to a song on the radio. How does the joint system comprising you and your radio extract the information of interest – the song – from the ambient photon field? Extracting the song is a two-step process that mirrors the two-step process of encoding it. Your radio is a tunable resonator that detects the "carrier wave" (CW) of the transmitter. This CW is a constant signal at some designated frequency. It encodes, effectively, the name of the transmitter. Once the radio has found this signal, it amplifies small, very slow (tens of Hz to kHz as opposed to GHz) AM or FM perturbations of the CW, using filters and feedback (i.e. measures of the differences between current and expected states) to reduce noise. It is these small, slow perturbations that encode the song.

The ability of radios to extract songs from the ambient photon field is of interest historically because Gibson [45] used it as a model of visual perception, and used it in particular to argue that perception by "resonance" required no memory for prior information on the part of the perceiver. This argument has since been used extensively to argue that organisms neither store nor process information [46,47]. The brief description above highlights the stored information that this argument misses: radios store information about CWs that, for example, rocks of similar size and weight do not, and also store extensive information about how CWs are modulated by radio transmitters. This latter information allows them to process perturbations of the CW of the right kind, and at the right timescale, to extract the encoded information, e.g. a song.

A human using a radio to find songs is faced with the same problem facing the radio, and solves this problem is a similar way. In order to use the radio to find songs, a human must first find the radio. This is usually done by interacting with the ambient photon field, i.e. by visual perception. The human must scan the available components of this field to locate an expected, constant signal, what we ordinarily call the "visual appearance" of the radio. It is only after doing this that the human can detect or manipulate time-varying properties of the radio, such as the station (i.e. CW) to which it is tuned. Fig. 2 illustrates this difference between constant (or "reference") and variable (or "pointer") properties. The specificity of the reference state $|R\rangle$ (using Dirac's notation) used to identify a system – a radio, or even a CW in the radio-frequency spectrum – determines the specificity with which a particular system can be detected. If one wants a specific radio station, or a specific radio, one needs to detect an $|R\rangle$ specific to, and known to be specific to, that individual system and no other.

The above characterization of detecting changes in signals and objects is completely general, and in fact characterizes any physical system capable of responding in some specific way to some specific property of its environment [5]. In the language of physics, detectors of specific kinds of signals or systems are quantum (because physically implemented) reference frames (QRFs [49,50]). A rhodopsin molecule responsive to photons in a specific frequency range, for example, is, in virtue of its physical implementation, a

QRF for any signals carried by photons in that frequency range. Such molecules are not isolated or passive, but rather exist within larger systems on which they depend for biochemical stability, and on which they act when excited by photons. In the language of the FEP, QRFs are components of a system's generative model that specify what kinds of signals or objects are present in its environment, what kinds of variation they are capable of, what kinds of actions can be taken with respect to them, and what the anticipated consequences of such actions are [15,16]. Implementing a QRF confers observational and action capabilities on a system; hence characterizing a system in terms of its implemented QRFs is a functional, not structural, characterization. Any such characterization is, moreover, a hypothetical model; how a system implements its QRFs, and even which QRFs it implements, cannot be determined by finite observations [51]; see [52] for a review in more contemporary language.

As an organism develops, we can expect what it can detect, and what it does with the information it obtains, to change. Hence we can expect the QRFs it implements to change. Such changes may be gradual, for example as a child increases her native-language vocabulary, or radical, for example when metamorphosis constructs an entirely new body, with new sensors and effectors, for an insect. Such changes naturally raise questions about the identity of the system undergoing change, and whether this identity itself is changing over time.

2.3. What is persistence?

We can now ask what it means to say that an object S, or its internal self-interaction H_S, persists through time. Friston [14] introduces this topic with: "we start by asking how something can be distinguished from everything else." We adopt this tactic here, starting by considering persistence relative to some specified observer. We then show why persistence can only be defined in this observer-relative way.

The examples above highlight what is required to distinguish something from everything else: a QRF of sufficient specificity. The vertebrate visual system, for example, distinguishes "objects" – with no further characterization – by their motion with respect to an unmoving (or more slowly moving) background [53]. This motion, or visual trajectory, detection process is highly sensitive to gaps in observation, and infers object persistence only following very specific kinds of occlusion, and only for timeframes of seconds. It is implemented in mammals primarily by the dorsal visual stream, and relies on spatial and temporal coordinates implemented in medial-temporal areas [54]; these networks together can be considered the mammalian short-term motion QRF. Longer gaps between observations require additional inferences about object properties and history, and involve not only the ventral visual stream and hippocampus, but also more frontal areas. What is inferred in these cases is a hypothetical process: a causal chain that preserves object identity [55]. Multiple spatial, temporal, and contextual points of reference are required to, for example, re-identify an old friend not seen for decades; hence in these cases, much of the brain is functioning as an object-recognition QRF.

Persistence is thus a combination of observed distinctions – "detector clicks" from a QRF – and a hypothetical process. The process, as mentioned above, is what happens between observations. We cannot talk about objects as persistent without the processes that extend them through time, and we cannot talk about processes, except as pure abstractions, without talking about observations that interrupt them with "snapshots" that produce some particular observational outcomes. We cannot, in particular, talk about time without some means of experiencing it – some reference points that distinguish "now" from "then". Hence we cannot talk about persistence with talking about memory. And as [30] makes clear, we cannot talk about memory without talking about coarse-graining, interpretation, and the drive to make sense of things [56]. The FEP has, in other words, taken us from physics back to psychology: by formulating a criterion for persistence, it has formulated an implicit definition of sense-making.

Einstein famously asked whether the moon was there when no one looked at it [57], a question that quantum theory has answered, as Berkeley answered a similar question posed by Hume, by asking what "no one" means. Let us return to Fig. 1, and ask whether S can tell whether E persists through time. This is, of course, also the question of whether S can tell whether it itself persists through time, or equivalently, whether S can tell whether the boundary $\mathscr{B}$ maintains its integrity through time. Since S being separate from E is a matter of mutual conditional independence, or in quantum-theoretic terms, absence of entanglement, we can formulate this question as whether S can tell whether it is itself conditionally independent of E, or in quantum terms, whether it is entangled with E.

The answer to this question is a categorical "no" [52]. No system can determine whether it is entangled with its environment. No system can determine, in other words, whether it has an effective boundary. The reason is simple, and follows immediately from the HP, or in FEP terms, from the definition of a Markov blanket [58]: systems interact with their environments via their boundaries, and cannot "reach outside their boundaries" to more directly measure or manipulate their environments. In particular, they cannot measure the entanglement entropies across their boundaries, so they cannot determine whether they are separable from, and hence conditionally independent of, their environments. This is independent of how the boundary is defined, leading immediately to the consequence that a system can neither determine nor specify "where its boundary is" or even measure the size (i.e. Hilbert-space dimension) of its boundary (again see [52] for formal proofs). This gives the "extended mind hypothesis" [59] – a notion usually thought of as involving cognitive aids, analogous to the extended body schema induced by tool use [60] – considerable teeth: no system can demonstrate that any perceived object is outside of its boundary, so no system can demonstrate that any perceived object is not part of its mind. Experiments in which cultured cells interact with non-biological materials – i.e. with "hybrots" – demonstrate effects analogous to body-schema extension outside of humans or even evolved organisms, suggesting that "extended minds" may be ubiquitous [61]. We can also access this result from a fully external, abstract perspective. Whether two systems are entangled – whether their joint state space factors – is not an observer-independent fact, but rather depends on how their joint state space is described [62,63]. Hence there is no such thing as an observer-independent boundary.

2.4. Persistence involves risk and requires tradeoffs

Mathematical models such as the FEP show their greatest value as thinking tools in situations in which decisive experiments are impossible but circumstantial evidence and intuitions are strong. The FEP describes agents as minimizing a variational free energy (VFE) that is, effectively, a measure of prediction error. They do this by active inference, a combination of learning environmental regularities and acting on the environment – in ways ranging from moving around to doing experiments – in order to obtain both thermodynamic resources and new information. The latter activity is essential for any systems not living in an unrealistically ideal environment [64] and its motivation – whether termed epistemic hunger, infotaxis, or curiosity – is a fundamental drive [65]. Whether for food or knowledge, foraging involves risk; indeed [14] shows how risk and (predicted) information gain are just two ways of describing the same statistical measure. Hence active inference is intrinsically risky, and systems that engage in it can suffer irreversible damage or destruction, i.e. boundary collapse [66].

As discussed above, to persist as an entity is to maintain statistical separability from one's environment by maintaining a collection of internal states that are causally "far from one's boundary" and hence available for implementing computations under one's own control. How many such states there are depends on the size of one's boundary – effectively, on one's surface to volume ratio. The size of the boundary, in turn, depends via the HP on the strength of the system-environment interaction. Hence persistent systems must maintain small boundaries to avoid being penetrated or incinerated by their environments. Boundary size, however, also determines both the amount of information and the amount of thermodynamic resources that can be obtained from the environment [67,68]. Both are needed, and they trade off against each other when boundary size is fixed. Hence all active inference agents face coupled tradeoffs related to boundary size or, equivalently, total strength of interaction with their environments [41]. For any active inference agent, different choices of environmental exposure and different tradeoffs between feeding, learning, and exploration can be expected to yield different lifestyles, and hence pre-adaption to different niches.

3. What evolves? What develops?

3.1. Life as a process of boundary expansion

The NeoDarwinian idea that only genes evolve, while all other aspects of biology are just along for the ride [69–72], has been widely criticized, and to some extent supplanted, by theoretical approaches that countenance multi-scale evolution [73–76]. Even in a single cell, far more heritable information is encoded by the cytoplasm, cytoskeleton, and membrane than by the genome; models that focus exclusively on genes miss the inheritance of this information altogether [77]. The idea that both organismal variation and the environmental variation that drive selection are random, and hence essentially thermal processes has also been challenged by models of organisms as active inference agents that modify their environments, and thereby modify the selective landscapes faced by both themselves and others [78–82]; see [83–85] for discussions of this point from a more philosophical perspective.

These and other, similar considerations motivated our previous proposal that the processes we distinguish by the terms "evolutionary" and "developmental" are manifestations of a single mechanism that differ only in spatial and temporal scale [86]. This perspective suggests, in turn, that Life as we know it, the lineage of the Last Universal Common Ancestor (LUCA) of extant organisms, can be viewed as a single, temporally-extended, morphogenetic process [87]. Life as we currently observe it, in other words, has a spatially-extended, multicellular body, all of the cells of which bear the same lineal relationship to LUCA that cells in a conventional multicellular body – setting aside commensal bacteria and other organisms – bear to a zygote. With the recognition that all living multicellular organisms are holobionts [88–90], it indeed becomes clear that conventional multicellulars are composites of multiple lineages from LUCA.

Adopting the notation of Fig. 1, we can represent the developmental process of Life very simply: as shown in Fig. 3a, Life starts out as a small system with very few degrees of freedom and hence a very small share of the total state space of U. Most degrees of freedom, and hence most of the state space, compose the abiotic environment with which Life interacts. As Life develops, however, its identity as a system – a collection of degrees of freedom – changes. As the "background" time t becomes larger, degrees of freedom that were formally part of the abiotic environment, e.g. available water, carbon, nitrogen, and other molecules, become part of Life, which employs them to synthesize new macromolecules. This process makes Life a larger system, with both a larger share of the state space and a larger boundary. The larger state space enables more complex information processing, while the larger boundary affords more thermodynamic resources, more useful information, and more possibilities for action. The morphogenetic process of Life, from this perspective, can be described as the elaboration of a complex computational process, implemented by a time (t) dependent self-interaction H_{Life}, at the expense of its abiotic environment. This expansion of Life as a system is a phenomenon that the overall self-interaction H_U implements as it drives the overall state $|U\rangle$ around in its very large state space.

This view of Life as a process of state-space and boundary expansion is purely topological, and neglects the geometric process of cell division that produces the branching pattern of lineages that we observe, as in Fig. 3b, as phylogenetic diversity. This branching process renders the boundary between Life and its environment geometrically complex, with different components of Life accessing thermodynamic resources and information from, and acting on, different components of the abiotic environment as well as on each other. Representing Life as an abstract system S, we can describe this geometric complexity in computational terms by viewing S as deploying a set of QRFs $S_1, S_2, \ldots S_n$ that access mutually disjoint "slices" of the state space $\mathcal{H}_{\mathcal{B}}$ of the boundary $\mathcal{B}$ as shown in Fig. 3c; see [91] for details. We observe this kind of boundary segmentation at all scales, from the asymmetric placement of membrane receptors and effectors that makes chemotaxis possible to the independently mobile, exploratory pseudopodia of amoeboid cells to the branching patterns of neurons, higher plants, and animal appendages. Indeed, we showed in [91] that such "neuromorphic"
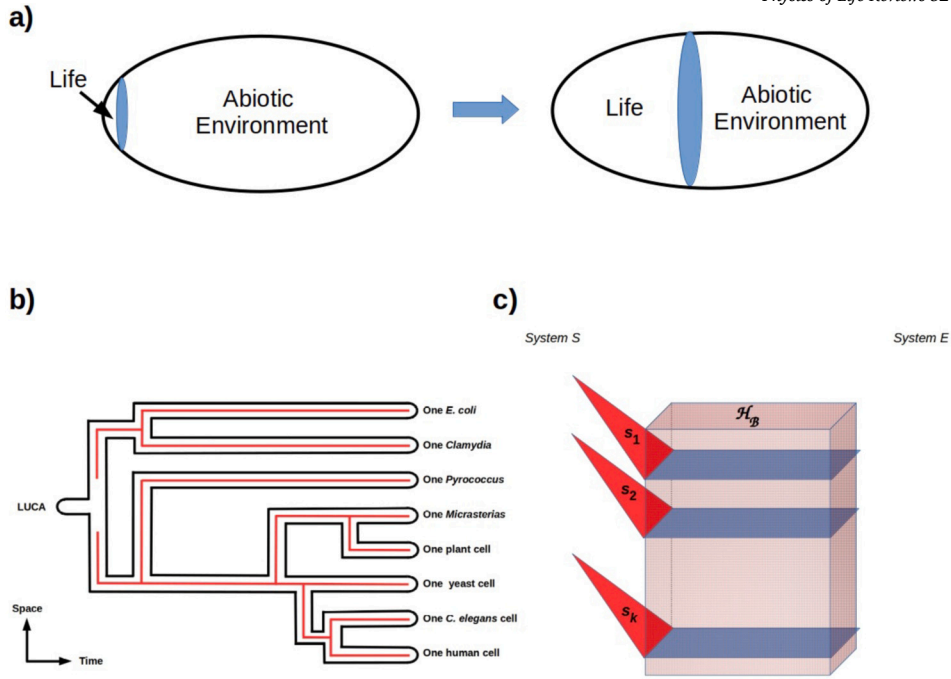
Fig. 3. Life as boundary expansion. a) Life begins with only a few degrees of freedom, and hence a small share of the overall state space and a small boundary. As Life develops, it incorporates degrees of freedom from the abiotic environment, expanding both its state space and its boundary. b) Viewing Life as a lineage from LUCA reveals phylogenetic branching not visible in the purely topological representation of panel a). From [87]; used with permission. c) Any system S with spatial degrees of freedom will be driven by the FEP to probe its boundary in distinct "slices" that yield different information and afford different kinds of actions. From [91]; used with permission.

branching patterns are induced in any active inference agent that has a boundary with spatial degrees of freedom. From an FEP perspective, they increase both predictive power and the effectiveness of actions by distributing computational engagement with the boundary evenly across its surface.

3.2. Life as a multiscale competency architecture

Maturana and Varela [8], Polanyi [92], Rosen [93], Noble [94], and many others have noted the bidirectional flows of both constraints and their complements, enabling conditions, that characterize biological systems. Organic chemistry, for example, both constrains and enables macromolecular biochemistry, while macromolecular biochemistry enormously expands, within the constraints imposed by enzymatic structures and activities, the diversity and hence the reaction possibility space of organic chemistry. Animal bodies similarly enable and constrain the built environments that animals can construct, while these environments both enable and constrain the behaviors of both their builders and other organisms occupying overlapping niches. Both the upward and downward flows of constraints and enabling conditions and the absence of local determinism by either higher or lower levels can be expressed by characterizing living systems at any scale as "agential materials" that both act on and are influenced by the behavior of systems at larger and smaller scales but retain their own autonomy. Multiscale systems comprising agential materials at every scale implement multiscale competency architectures (MCAs), i.e. structural and functional organizations in which, at every scale, components are competent to go about their business with guidance, but not explicit instructions, from other scales [27,61,95]. The idea that components of MCAs are autonomous – free from local determinism, even by all events in their causally-relevant pasts taken together – rests on a fundamental theorem of physics showing that such autonomy is required for consistency with both special relativity and quantum theory [96].

Viewed as a computational architecture, an MCA is a hierarchy of active-inference agents, each complying with the FEP with respect to its own environment. Computation, however, is a process. Viewed in the notation of Fig. 1, an MCA is a decomposition of the internal interaction H_S of the system S implementing the MCA. In particular, we can write:

$$H_S = \sum_i (H_i + \sum_{j \neq i} H_{ij}),$$
(1)

where the H_i are component self-interactions and the H_{ij} are interactions between components. Equation (1) follows from our assumption of standard quantum theory, which requires all operators representing observables, including H_S as a representation of total energy of S, to be linear. The only constraints on this linear decomposition of H_S into components labeled by i and j are that the H_i and the H_{ij} are all well-defined, which in turn requires that the states $|i\rangle$ of the components are all well-defined. Any number and size of components can be chosen provided that these constraints are satisfied. As described in connection with Fig. 1, systems
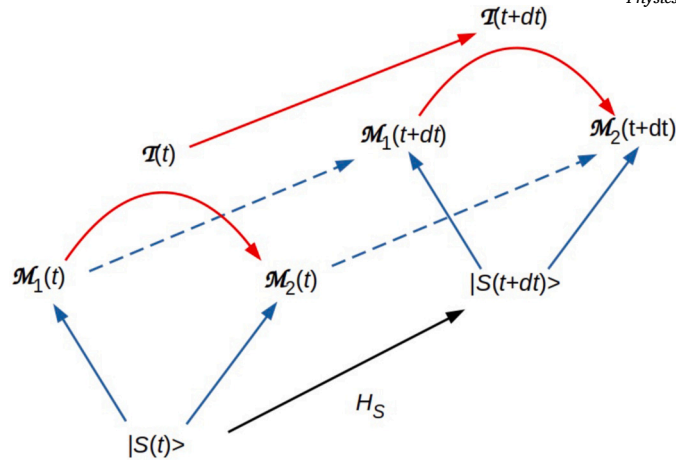
Fig. 4. Mutual coherency condition for MCA models of a system S. Provided that both $\mathscr{M}_1$ and $\mathscr{M}_2$ are coherent models of S [39], and that the map $\mathscr{T}$ between models remains constant over time, the diagram commutes and mutual coherence between the models $\mathscr{M}_1$ and $\mathscr{M}_2$ is guaranteed.

and their internal dynamics defined within standard quantum theory do not impose any mereological descriptions a priori; all such descriptions are model-dependent, and hence dependent on how the system and its behavior are measured [62,63]. No particular decomposition of any system into "parts" can be assumed to be given, a priori, by the "physics" of the system. Hence no description of a system as an MCA can be assumed to be given, or implied, by its physical dynamics.

We can, therefore, describe any MCA in multiple ways, and ask how these descriptions relate [40]. Let $\mathscr{M}_1$ and $\mathscr{M}_2$ be descriptions of S as an MCA. The descriptions $\mathscr{M}_1$ and $\mathscr{M}_2$ can be stated in any language(s), formal or informal, and can divide S into "parts" and characterize the states of the parts at any scale and in any way. To "make sense" or be coherent as models of the behavior of S, each of $\mathscr{M}_1$ and $\mathscr{M}_2$ must, individually, satisfy the simple criterion given by [39]: the map from states $|S(t)\rangle$ to their descriptions in the model must commute with both the physical dynamics H_S and the postulated model dynamics. Describing some state $|S_{t_i}\rangle$ in the model and then running, or inferring, the model dynamics on the model's description of $|S_{t_i}\rangle$ for a simulated time dt must, in other words, yield the same result as describing the state $|S_{t_i+dt}\rangle$ in the model. How precise the criterion of "sameness" is depends on the scale, precision, and typically level of formality of the model. If the model is stated in a programming language, the notion of "the same result" means: as computed using an implementation of that language. If the model is stated in ordinary English, e.g. with diagrams such as typically used in biochemistry, what "the same result" means may be considerably more vague.

Assuming that $\mathscr{M}_1$ and $\mathscr{M}_2$ are each individually coherent models of the behavior of S, we can ask what is required for them to be mutually coherent, or inter-translatable. Suppose that the descriptions $\mathscr{M}_1(t)$ and $\mathscr{M}_2(t)$ of $|S(t)\rangle$ can be related by some translation $\mathscr{T}(t)$. If $\mathscr{M}_1$ and $\mathscr{M}_2$ are formal models, this $\mathscr{T}$ will be a mapping from the symbols used by $\mathscr{M}_1$ to the symbols used by $\mathscr{M}_2$. If $\mathscr{M}_1$ and $\mathscr{M}_2$ are informal models, $\mathscr{T}$ may just be an explanation, in some natural language, of how the models relate; nearly all models of genetic interactions, cellular biochemistry, or organismal physiology or behavior are of this informal kind. We can now ask: does the same translation between the descriptions of S given by $\mathscr{M}_1$ and $\mathscr{M}_2$ work at subsequent times, e.g. for the descriptions $\mathscr{M}_1(t + dt)$ and $\mathscr{M}_2(t + dt)$ of $|S(t + dt)\rangle$, as illustrated in Fig. 4? If so, the translation $\mathscr{T}$ "works" or "makes sense," at least for the limited collection of descriptions generated by observations of S between t and $t + dt$. If $\mathscr{T}$ works for descriptions of large numbers of systems, over long observation times, that are generated in the languages of $\mathscr{M}_1$ and $\mathscr{M}_2$, we can say, at least tentatively, that $\mathscr{T}$ is a coherent translation between these two languages.

The assumption of linearity that allows us to write Eq. (1) provides an alternative approach to checking whether two models of a system S are mutually coherent. Suppose that $\mathscr{M}_1$ describes S in terms of a mereological decomposition into subsystems S_i, while $\mathscr{M}_2$ describes S in terms of a mereological decomposition into subsystems S_k. If the two models can each fully reconstruct a representation of H_S, i.e. if each model can fully describe both the self-interactions, H_{S_i} and H_{S_k} respectively, and the pairwise interactions, $H_{S_{ij}}$ and $H_{S_{kl}}$ respectively, of the subsystems that they postulate, then the translation $\mathscr{T} : \mathscr{M}_1 \to \mathscr{M}_2$ between the two models is just a change in decomposition, a redrawing of component boundaries within S that, by satisfying Eq. (1), leaves the dynamics H_S and hence the state transition $|S(t)\rangle \to |S(t + dt)\rangle$ unchanged. The map $\mathscr{T}(t) \to \mathscr{T}(t + dt)$ is, in this case, the Identity, which guarantees that the diagram shown in Fig. 4 commutes, i.e. that the models $\mathscr{M}_1$ and $\mathscr{M}_2$ are mutually coherent as MCA models of S.

It is important to emphasize that the coherence of models of physical systems is an empirical question, and that whether any model of any system S is coherent, or accurate, in the above sense cannot be proven a priori. The Conway-Kochen theorem [96], which rules out local determinism for any physical system if quantum theory is true, guarantees that even models that are highly accurate for ensembles cannot precisely predict outcomes of experiments on individual systems. In biological systems, where even clonal populations cannot be expected to be uniform in all characteristics potentially relevant to their dynamics or interactions, determining the accuracy of the decompositions of H_S proposed by competing models is correspondingly more difficult.

If two models $\mathscr{M}_1$ and $\mathscr{M}_2$ are decompositions of S into active inference agents, and of H_S into interactions between such agents, at two different scales – that is, if $\mathscr{M}_1$ and $\mathscr{M}_2$ are both MCA models – then $\mathscr{T}$ describes how the active inference agents

defined at the smaller scale embed into, or collectively implement, the active inference agents defined at the larger scale. The theory $\mathscr{T}$ tells us, therefore, how the larger-scale components and their interactions "emerge" from the smaller-scale components and their interactions. Because $\mathscr{T}$ is, in this case, a mapping between the labels assigned to degrees of freedom of S by $\mathscr{M}_1$ and $\mathscr{M}_2$ respectively, this "emergence" process is also just a change in labeling. What S does – the dynamics H_S – does not change, the only change between small- and large-scale models is in how H_S is being described. As relabeling the components of S has no effect on E, the interaction H_{SE} also remains unchanged. What appears as new in the larger-scale description are the interactions H_{ij} between larger-scale components i and j, which in general cannot be predicted from any smaller-scale description of H_S and may be completely unexpected. Any larger scale component k that is bounded in part by some sector of $\mathscr{B}$ – i.e. that is "next to" the boundary $\mathscr{B}$ and hence directly interacts with E – will exhibit a local interaction H_{kE} that is similarly unpredictable, in general, from any smaller-scale description. Changing the modeling scale does not, therefore, generate any new physics – any change in either H_S or H_{SE} – but it does generate new knowledge and new possibilities for intervention and control [97].

This notion of "emergence" as a consequence of a change in decomposition is consistent with the idea of "weak emergence" defined by Bedau [100]. A large-scale behavior of a system S is weakly emergent if it cannot be predicted from a smaller-scale description, but can be observed when the smaller-scale dynamics are simulated. Implicit in this definition, clearly, is that the simulation results are computed at the larger scale. As Bedau points out, S can have weakly-emergent behaviors only if it is not isolated, a condition that any system that interacts with an environment, and *ipso facto* any system that is observable, satisfies. Emergence being a consequence of decompositional change is, moreover, also consistent with ideas of "strong emergence" [101] that understand strong emergence in terms of symmetry breaking [102]. Inserting new decompositional boundaries into a system does not break a physical symmetry, but it does break a descriptive symmetry by introducing the new interactions H_{ij} and H_{kE}. Introducing new boundaries is adding new information to the description of the system, and this new information is manifested in the unpredictability of these new interactions from smaller-scale descriptions. Hence the claim that these interactions are "not determined by the underlying physical laws" [102], p. 6, is clearly correct; these interactions can only be predicted given the smaller-scale physics and the larger-scale boundaries. This is because the larger-scale boundaries are themselves not determined by the underlying physical laws. Mereology is imposed by observers, not generated by physics, as shown explicitly by Zanardi [62] and by many others subsequently.

From a practical point of view, Eq. (1) and the commutativity of Fig. 4 allow descriptions of S at multiple scales to serve as an error-correction mechanism for theorists: since all models must yield the same overall dynamics H_S, modeling errors at one scale can be detected by examining models at other scales. The utility of requiring consistency between scales – i.e. between "levels" in a virtual-machine hierarchy – for debugging computer software is obvious. In a setting more relevant to the biosciences, Friston [14] has shown that all random dynamical systems that exhibit sparse enough coupling to be decomposable into conditionally-independent components are renormalizable. In such systems, scale transitions can be represented by actions of the Renormalization Group on the generative models that the components can be described as implementing [103]. Precise mappings between algorithmically-specified models, e.g. dynamical systems or generative models specified using the FEP, and "hardware-level" descriptions of the physical implementation of such models in particular systems at particular scales are needed to apply such multi-scale representations in practice. The availability of such model-to-hardware maps can be expected to improve as new measurement techniques and formal languages are developed.

3.3. Life as a multiscale memory

Machine learning has revolutionized artificial intelligence because training an artificial neural network on a large set of examples effectively constructs an implicit statistical model of what those examples have in common. Running such a network "backwards" generates new instances of these commonalities, as the behavior of generative systems such as ChatGPT or Midjourney demonstrates. All that is needed to trigger the generation of a new instance is an input from the environment, a "prompt" of some specified type.

Active inference agents are generative systems in precisely this sense: any agent S encodes a statistical model of its environment E's observed behaviors, including E's responses to S's previous prompts. Active inference agents are, therefore, systems capable of memory, i.e. systems in which at least some state changes induced by external perturbations – actions by the environment - persist long enough to affect at least some future behavior [104]. Inputs from E also prompt S to take some action, even if it is a "non-action" such as waiting for more input. As discussed above, this model is implemented by H_S, which can be decomposed at multiple scales to yield multiple, distinct descriptions of S as an information-processing system.

We can now apply these ideas to Life as an evolving/developing system. We can describe Life as a single system with an internal interaction H_{Life} that interacts with its abiotic environment with some interaction $H_{Life,Env}$. Alternatively, we can decompose Life into a collection of lineages of various lengths and at various scales. A lineage is a sequence of cell divisions, with some branches being effectively immortal and others ending in cell death; if the cells in a lineage interact mostly with each other, we can also describe the lineage at the scale of multicellular organisms that reproduce in some fashion and sometimes die. We can distinguish, at this organismal scale, between lineages of organisms that are effectively immortal and lineages of organisms that die at the end of some finite, individually variable but statistically definable lifespan, noting that the latter correspond exactly to the organisms with obligate sexual reproduction [105]. For any given lineage L, we can write an internal interaction H_L and an interaction with L's environment $H_{L,E(L)}$, noting that $E(L)$ contains many lineages of Life that are disjoint from L. When these inter-lineage interactions occur inside an individual multicellular body, e.g. between neurons and non-neurons, we describe them in terms of organismal physiology; when they occur between organisms in the open environment, we describe them in terms of ecology. Such inter-lineage interactions can, however, be represented by the same formal structure at any scale, and we can treat them, at any scale, as interactions between

active inference agents, i.e. as interactions between physical systems with memory that each persist for some period of time while exchanging information bidirectionally with their environments.

Lineages imply reproduction, which can be described from multiple perspectives. Consider symmetric cell division as a canonical example. From the perspective of Fig. 3a, cell division co-opts environmental degrees of freedom into an expanded biological system. From the perspective of Fig. 1, cell division creates a new boundary, and approximately duplicates both the previous internal interaction H_C and the previous environmental interaction $H_{C,E(C)}$. This duplication is inexact, as exact duplication of an unknown quantum state – and hence of an unknown quantum process – is forbidden by the no-cloning theorem [106]. Ashby's [107] Law of Requisite Variety prevents any physical system from knowing its own state, a result that can also be proved from quantum theory alone [52]; therefore, the no-cloning theorem applies. From an information-theoretic perspective, even inexact duplication of any system – or of components within a system – creates redundancy, i.e. high mutual information between (sub)states, that can serve as a basis for error detection and correction. Redundancy of components is ubiquitous in biology, from the molecular scale to that of organismal populations. From the perspective of either of the two daughter cells produced by cell division, the environment is more predictable than it was for the parent cell, as the environment now contains something with which it shares mutual information [108]. The same is true for reproduction at any scale.

Predicting what E will do next requires memory for what E did previously. Redundancy increases predictive power by providing additional states on which to encode memories. We can, therefore, view reproduction as memory construction, and a lineage as a tree-structured memory. This is consistent with the common view of DNA as a record of past selection; it simply extends this picture to the lineage as a whole. As reproduced cells are "written" on the environment, they can be considered stigmergic memories. Cells can "write on" and "read from" other cells by biochemical, bioelastic, or bioelectric processes; hence other cells can be viewed as explicit as well as external memories, as opposed to the internal, implicit memories encoded by the internal interaction H_C. As these memories are external, interactions with them are via $H_{C,E(C)}$; hence we can view them, for the purposes of modeling $C - E(C)$ information exchange, as sectors on the cellular boundary $\mathscr{B}_C$. All of these comments generalize trivially to the case of organism-scale reproduction and organism-scale lineages.

In the microbial world, lineages can interact via lateral transfer of genetic information [109–111]. At every scale, lineages can interact via niche construction, predation, parasitism, symbiosis, including endosymbiosis, and various forms of inter-individual signaling, including language. Hence information, and therefore memory, can be transferred between lineages at every scale in the temporal evo/devo/eco hierarchy, including the inter-individual "ego" scale (R. Watson, personal communication), and spread essentially arbitrarily across spatial scales. The lineage of information is, therefore, not tree-structured, but rather a backward- and forward-branching structure in which the past of any item of information may be as ambiguous as its future. This is particularly true of the overall internal interaction H_{Life} at any given time, the components of which have "flowed through" multiple cell-division and lateral-transfer pathways.

We are left, therefore, with an overall picture of "what evolves/develops" as a multiscale network of constraints and enabling conditions that trades off between variation and error correction along both temporal and spatial scales. We can think of temporal slices – snapshots – of this network as collections of individual cells or organisms that are distributed in space, but this time-slice view does not reveal the informational structure of the network. When viewed as a multiscale process extended through time, Life is an active inference agent that consumes both thermodynamic energy and abiotic degrees of freedom in order to increase its ability to predict the behavior of and adaptively influence its environment. As it grows, it engages in more and more elaborate niche construction, progressively remodeling abiotic degrees of freedom into a built environment with stable, predictable properties. Life's ability to colonize and remodel its environment depends critically on lateral information flows between its component lineages, from gene flow between otherwise unrelated individuals to the multimodal communication seen in microbial mats or, in even more elaborated form, in obligate multicellular holobionts or in plant, animal, or human communities.

4. Self models

The idea that what we humans experience as "our self" is a model constructed in real time and as needed by components of the default-mode network has become mainstream in cognitive neuroscience [112–116]. That many activities are performed "non self-consciously" – and in particular, that activities performed with expertise are, and can only be, performed non self-consciously – indicates that the self model is deployed in circumstances of uncertainly in which explicit metacognitive monitoring is needed [117–119]. Turning the experience of the self model off is a goal in some meditation practices, and neurocognitive studies of such practices confirm that default-mode network deactivation is involved in this experience [120–122].

From an active-inference perspective, the human experienced-self model is a component of the overall generative model employed to predict the consequences of actions, particularly actions in a social context. Dysfunctions in the model, particularly dysfunctional assignments of precision or relative importance to particular environmental inputs or prior beliefs, result in contextually-dysfunctional actions in predictable ways [123–125]. This suggests generalizing the notion of a "self-model" to any metacognitive component of an overall generative model that modulates precision assignments and hence effectively allocates attention and other cognitive resources, including memory, to some aspects of an action context at the expense of others [126]. Such metacognitive components can be expected in any system that experiences selective pressures with different characteristic timescales [81]. Properties of a context that a system assigns high precision to can be considered properties that the system cares about [127,128].[1] Positioning these items

---

[1] Indeed, one could argue that the competence of large language models based upon transformer architectures relies exactly upon this mechanism; namely, attention heads that afford a high precision to observations in the past, when determining predictions of the future.
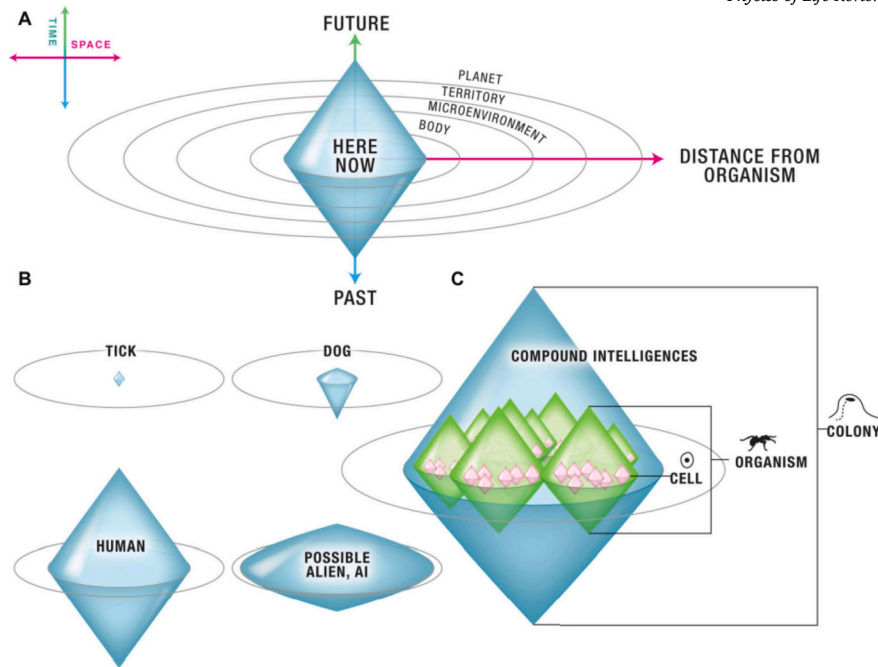
Fig. 5. The cognitive light cone (CLC) generalizes the idea of an experienced self-model to the collection of all "items of concern" that can influence a system's allocation of attention and other cognitive resources. The base of the cone represents what the agent can care about – the spatial extent of the agent's goals or concerns at any given time. The height of the cone represents the temporal reach of the agent's goals into the future, while its depth represents the temporal depth of the memories that contribute to the agent's pursuit of current goals. Systems at different scales, or at different developmental timepoints or lifecycle stages, can have CLCs of greatly varying spatial and temporal extents. Figure by Jeremy Guay of Peregrine Creative; from [127], CC-BY license.

of concern, which may include the system's own past and future selves, in space and time relative to the system's space-time location generates the system's "cognitive light cone" (CLC), a representation of the maximal extent in space and time of its concerns [127]; crucially, the CLC is not depicting the maximal reach of sensors or effectors, but rather the maximum size of states that the agent is capable of working towards as goals. As what is in the CLC is what can influence the system's metacognitive control of attention, we can view the CLC as a generalization of the experienced self-model that can be applied to any system, and in particular, to any organism.

Reconceptualizing the self-model as the CLC emphasizes its functional role as a resource-allocation tool. The CLC serves, in particular, as a context for event memories and a focus for planning. Both the spatial and temporal extents of the CLC are heavily dependent on both memory and computational processing resources, i.e. on the degrees of freedom available to implement the system's internal interaction and the size of its boundary. The CLCs of unicells and small multicellulars, e.g. the tick in Fig. 5, may be limited to the immediate vicinity of its physiological boundary and one or a few of its internally-counted time steps. The CLCs of humans, in contrast, can employ imagination as a resource to extend their active goals into interstellar space, and both backwards and forwards in time far beyond an individual's lifespan. Human CLCs, at least, can expand or contract depending on an individual's current state, resources, and ambitions; e.g. an infant might have a narrow cone focused on immediate survival, a child may care about friends or look forward to "what they want to be when they grow up", and a mature adult may have a broader cone encompassing concerns about the state of the world and the well-being of others even far into the future. Collectives of cells – e.g. of human cells – can clearly have CLCs much larger than those of individual cells, which represent anatomical states as target morphologies toward which they actively work in regulative development and regeneration [78,129]. Collectives of some social organisms – e.g. ants – appear to have CLCs larger than those of their individual members [130]. Whether collectives of humans have CLCs larger than those of individual humans remains an open question.

As with the experienced self-model as implemented by the default-mode network, the operational size of the CLC can be expanded or contracted in a task- or context-dependent way. Activities that require unwavering attentional focus on the immediate spatiotemporal environment – e.g. rock climbing or competitive driving – are incompatible with the re-examination of history or long-term planning. The farther spatial and temporal extents of the CLC can, moreover, be expected to be coarse-grained relative to the more proximal components. This coarse-graining reflects the fact that the CLC is, like the human experienced self-model, a construct that must be computed, in real time, from information available at some particular time and in some particular context. Hence the temporal depth of the CLC in both directions is critically dependent on the system's ability to track its own identity through imagined, or projected time, i.e. on its "mental time travel" ability [131]. It is also dependent, in any realistic, resource-limited setting, on the system's ability to forget unnecessary details of its past and make explicit only those elements of a plan that are assigned the greatest relevance. It is dependent, therefore, on the system implementing a heuristic solution to the Frame Problem [132], the problem of deciding what remains unchanged following an action [133].

A critical feature of the experienced self-model in humans is the felt sense of ownership of the self and its thoughts, actions, body, etc [134]. Failures of the sense of ownership manifest as various types of dissociation, which can be either pathological or therapeutic, depending on the mechanisms involved [135]. What is experienced as "owned" constitutes the experienced self, hence the distress caused by phenomena such as "thought insertion" in schizophrenia, in which some experiences are classified as "not mine" [136]. Humans are, however, also broadly susceptible to "false memories" that are considered "owned" [137]. Generation of the sense of ownership of experiences appears to be a component of the broader human "source" or "reality" monitoring function, which critically involves evolutionary-recent anterior-prefrontal cortical regions [138] and hence may not be widely shared even among primates. One can speculate that to the extent that memories are experienced by systems that lack source monitoring capabilities, they are experienced as anonymous. In this they would be on a par with non-experienced memories, such as those encoded in the genome, which in an important sense are anonymous.

These considerations raise, clearly, the question of whether Life as a whole or even substantial, multi-individual lineages have self models, CLCs, or senses of ownership. If a social group can have a CLC, a lineage of the same kind of organisms can presumably have a CLC. Examples of altruism or collective problem solving, even at the cellular level [139] suggest that collective CLCs may exist in diverse organisms and at multiple scales.

5. Procedural versus declarative

Drew McDermott [140] famously urged artificial intelligence researchers to avoid "any indulgence in the 'procedural-declarative' controversy," a controversy about data structures for memory that was, in the context of the standard von Neumann computer architecture, only about which virtual machine description [98,99] one prefers. The procedural-declarative dichotomy is just the process-object dichotomy for virtual implementations of memory. Memories encoded by updated model parameters, as in an artificial neural network, are procedural; memories encoded as discrete if-then rules, as in an expert system, are declarative. If S is a quantum system, memories encoded by H_S on the computational "bulk" of S are procedural, while memories encoded by H_{SE} on S's boundary are declarative. We know from Fig. 1, Eq. (1), and Fig. 4, however, that this is just a choice of interpretation. When we zoom out to the joint-system scale, the distinction between bulk and boundary disappears. When we zoom into the scale of components of S, the distinction does not disappear, but becomes fragmented. All of these descriptions are just alternative ways of representing the same dynamics. They enable different understandings, but do not describe different phenomena. They are, at bottom, just representations of a single phenomenon at different scales. They merely specify different virtual machines, as McDermott pointed out.

As we have emphasized throughout, physical interaction is *inter*action – there is no such thing, physically, as passive observation. It is, moreover, opportunities for action that make some observed difference "make a difference" [141] and hence give it meaning [142]. Any action on the environment potentially creates a stigmergic memory; whether the result of the action serves as a memory depends on the system's ability to re-identify it as such in a subsequent observation. That re-identification ability depends on the system's QRFs and, in particular, its ability to update its QRFs – via model refinement, i.e. learning – to correctly predict the consequences of its actions that will be subsequently observable. We humans tend to take this coupling between writing stigmergic memories and refining our models for granted, until the coupling fails and we forget where we have put our car keys or address book. Such failures reveal that procedural and declarative memories are mutually dependent: one without the other is useless. We can observe this at any scale; niche construction, for example, is useless if it cannot be remembered and the relevant environmental modifications identified later.

The apparent "hollowness" of language use by statistical models such as ChatGPT [143] provides an illustration of this point. Large Language Models (LLMs) can say what words mean, and if coupled with image recognizers and generators, can provide visual illustrations. However, their lack of manipulative abilities disconnect their knowledge of semantic relations from the world as we perceive it and describe it in language. At least at present, LLMs such as ChatGPT act only on their users, and aside from the interaction record in the context window, they do not update models of individual users to reflect what has been said to them in the past and predict how they will respond in future interactions. These systems have, in other words, very poor ability to re-identify the stigmergic memories they themselves have written on their world. They appear, therefore, to be very intelligent systems with severe anterograde amnesia; this renders their language use uncanny, and contributes to the sense that they use language without understanding it.

A second, more subtle illustration is provided by any agent's relationship to its memories, whether these are encoded on its environment or by its own body. Human memories, particularly episodic memories, are now widely recognized as reconstructed each time they are accessed [144,145]; see [146] for a formal model of this process. If we view the encoded engram as the system S of interest, its state is being modified whenever it interacts with its environment E, the (rest of the) agent that experiences the memory. Reading and rewriting a memory is, from this perspective, an interaction and hence an exchange of information – a conversation between an engram and an agent that employs it as a representation of some fact or event from the past.

As with the idea of a self-model or a CLC, we can examine this co-dependence between stigmergic memory and model refinement – and hence between declarative and procedural memories – at scales other than that of the individual organism. Federated learning systems with cognitive division of labor, from microbial mats to universities, are networks of mutually-dependent cognitive niches in which all participants perform both stigmergic memory recording – local niche construction – and model refinement.

> People don't possess ideas, ideas possess people.
>
> [Carl Jung]

6. Conclusion

We have discussed here how both process philosophy and traditional ontology objectify phenomena that can only be defined relative to a stipulated boundary that must be assumed, *a priori*, to remain stable (or slowly varying) over some time period of interest. Basic physics and the FEP provide relatively simple, well-developed formal constructs that allow us to view objects (from a given observer's perspective, encodings by its environment on its boundary) and processes (information flows within or between observable objects) as complementary while making only minimal, generic assumptions. This approach rejects the traditional object-process dichotomy as unhelpful and indeed ill-construed. By recasting all objects as processes – and vice-versa – it allows us to describe evo/devo/eco processes in terms of multi-scale constraint networks, redundancy generation, and error detection and correction. They thus give rise to a formally well-specified, scale-free biology.

These ideas have many implications beyond theory, specifically in areas which can benefit from an improved understanding of how to engineer with agential materials [147]: regenerative medicine and synthetic morphology [148–151]. While molecular biology has made enormous strides, we believe that the micromanagement approach to forcing molecular states will reach natural limitations beyond monogenic diseases. Inducing the complex morphological changes needed for repair of large-scale traumatic injuries and birth defects, reprogramming cancer, and creating synthetic living machines of desired form and function cannot be done exclusively bottom-up. Determining which genetic information needs to be changed (and how) to drive desired organ-level outcomes faces intractable inverse problems [78,152]. The path forward must include top-down approaches which exploit the competency of cells, tissues, and molecular networks to solve problems: pushing the complexity onto the system itself by re-writing its homeodynamic goal states, modifying its priors, and recognizing the valence of stimuli that control how it navigates its various problem spaces [129].

We have recently argued that laboratory and conceptual tools from the neuro- and cognitive sciences are useful far beyond the neural control of motion in conventional "behavior". Training and communication – ways to alter memories and thus behavior – are being used to control outcomes in physiological, transcriptional, and anatomical spaces [153,154]. Indeed, frameworks that seek to affect morphogenesis by modeling it as a collective intelligence of cells have led to numerous new discoveries (reviewed in [26,150]). For example, promising cancer therapeutic methods have been discovered by considering oncogenic phenotypes as resulting from a shrinkage of the CLC (a "software defect" not necessarily requiring any genetic lesion as a cause) [155]. This pathogenic reduction of the scale of the border between Self and World down-shifts the grandiose organ-maintenance goal memories of large cell groups into the ancient, tiny physiological and metabolic goals of unicellular organisms. They then view the rest of the body as just external environment [127,128]. From a biomedical perspective, it is interesting that bioelectrically reconnecting – and hence electrically realigning – cells via gap junctions, despite their very strong expression of powerful human oncogenes, abolishes the cancer phenotype in vivo [156–159]. Regulating gap junction connectivity regulates the extent to which cells act as biochemically-distinct units, and hence regulates the extent to which they have distinct biochemically-implemented memories and goals. Bioelectric networks coupled by such highly regulatable electrical synapses encompass most cell types, and can store larger and more complex states as setpoints for goal-directed processes that navigate diverse problem spaces, e.g. the anatomical space of tissues, at larger scales than the physiological and metabolic goal states of single cells [153].

From a more fundamental (evolutionary) perspective, one can view direct physiological coupling (via gap junctions) as a way to use anonymization of memories, and the fluidity of the size of the "owner" of those memories, to scale up a Self from that of single cells to that of a whole organism. For example, the molecular consequences (i.e., memory engram) of an event in cell A could propagate directly into cell B via the gap junctional connections. Cell B cannot distinguish that this memory trace is, in effect, a false memory for it. And while it is a false memory for cell B, it is a veridical memory for the newly-formed being AB, because the cognitive hacking induced by this memory sharing process produces a larger system with shared memories – a mind-meld in which it is difficult to maintain unique identities for the parts or to cheat or defect (in the game theory sense). The ability of multiple cells to share the same memories (i.e., world-model and priors) is what enables the collective to work together on enormous goals such as the reconstruction of a salamander limb after amputation. The sharing of anonymous memories is one of the key mechanisms for the scaling up of minds, and for combining smaller Selves into larger networks with greater intelligence in new problem spaces.

The fact that memories do not come with metadata indicating whose memories they are is critical – this is why cell B cannot maintain its own identity once it is directly coupled to cell A – as it prevents recipients of memories from rejecting the new memories as foreign. Thus, the nature of memories as a sort of autonomous, promiscuous cognitive Ronin is essential to their function as part of the glue that partially resolves the combination problem in philosophy of mind and the evolutionary origin of anatomical plasticity and robustness from simple metabolic homeostatic loops [27,160]. This gives rise to a pleasing, if unconventional, "strange loop" [161]: memories are patterns interpreted by agents (and have their reality within the cognitive medium of an existing agent), but they are also agents in their own right which contribute to the autopoiesis and robustness of the greater cognitive system around them, as a kind of niche construction. The causal influence seems to dynamically flow in both directions across the hierarchy of scale and solidity: downward and upward.

Taking seriously the potential uses of deep concepts in cognitive science to implement transformative biomedicine and bioengineering requires us to develop new frameworks for better understanding Selves, boundaries, memories, and their dynamic relationships across scales and material implementations. We must move beyond the constraining dichotomy of "real objects" (e.g., cells, tissues) and "patterns" (e.g., memories and signals within them) and investigate the symmetries and invariants that allow each to be both kinds of thing, depending on perspective, and amenable to the same unifying framework. These will be essential for developing practical tools to guide the up- and down-scaling of Selves and the movement and transformation of functional memories across tissues in therapeutic contexts [162,163]. It is also likely that a better understanding of the ability of memories to reify novel Selves, and

modify existing ones, will be of significant importance in (trans)personal psychology and efforts to understand social systems and other ways in which we engineer our reality.

Declaration of competing interest

Acknowledgements

References

[1] James W. The principles of psychology, vol. I. New York: Henry Holt; 1890.

[2] Whitehead AN. Process and reality. New York: Free Press; 1929.

[3] Adámek J, Herrlich H, Strecker GE. Abstract and concrete categories – the joy of cats. New York: Wiley; 2004. Current edition available from http://katmat.math.uni-bremen.de/acc.

[4] Feynman RP, Leighton RB, Sands M. The Feynman lectures on physics, vol. III. Boston: Addison-Wesley; 1963. Available from https://www.feynmanlectures.caltech.edu/.

[5] Fields C, Glazebrook JF, Marcianò A. Sequential measurements, topological quantum field theories, and topological quantum neural networks. Fortschr Phys 2022;70:202200104. https://doi.org/10.1002/prop.202200104.

[6] von Foerster H. Objects: tokens for (eigen-)behaviors. ASC Cybernet Forum 1976;8(3–4):91–6.

[7] Kauffman L. Eigenforms – objects as tokens for eigenbehaviors. Cybern Hum Knowing 2003;10(3–4):73–90.

[8] Maturana HR, Varela FJ. Autopoiesis and cognition: the realization of the living. Boston: Reidel; 1980.

[9] Gibson JJ. The ecological approach to visual perception. Boston: Houghton-Mifflin; 1979.

[10] Bunge M. The Furniture of the world. Dordrecht: Reidel; 1977.

[11] Peres A. Unperformed experiments have no results. Am J Phys 1978;46(7):745–7.

[12] Friston KJ. The free-energy principle: a unified brain theory? Nat Rev Neurosci 2010;11:127–38. https://doi.org/10.1038/nrn2787.

[13] Friston KJ. Life as we know it. J R Soc Interface 2013;10:20130475. https://doi.org/10.1098/rsif.2013.0475.

[14] Friston K. A free energy principle for a particular physics. Preprint. arXiv:1906.10184, 2019. https://doi.org/10.48550/arXiv.1906.10184.

[15] Fields C, Friston KJ, Glazebrook JF, Levin M. A free energy principle for generic quantum systems. Prog Biophys Mol Biol 2022;173:36–59. https://doi.org/10.1016/j.pbiomolbio.2022.05.006.

[16] Fields C, Fabrocini F, Friston KJ, Glazebrook JF, Hazan H, Levin M, et al. Control flow in active inference systems, Part I: Classical and quantum formulations of active inference. IEEE Trans Mol Biol Multi-Scale Commun 2023;9:235–45. https://doi.org/10.1109/TMBMC.2023.3272150.

[17] Ramstead MJ, Sakthivadivel DAR, Heins C, Koudahl M, Millidge B, Da Costa L, et al. On Bayesian mechanics: a physics of and by beliefs. Interface Focus 2022;13:2022.0029. https://doi.org/10.1098/rsfs.2022.0029.

[18] Friston K, Da Costa L, Sakthivadivel DAR, Heins C, Pavliotis GA, Ramstead MJ, et al. Path integrals, particular kinds, and strange things. Phys Life Rev 2023;47:35–62. https://doi.org/10.1016/j.plrev.2023.08.016.

[19] 't Hooft G. Dimensional reduction in quantum gravity. In: Ali A, Ellis J, Randjbar-Daemi S, editors. Salamfestschrift. Singapore: World Scientific; 1993. p. 284–96.

[20] Susskind L. The world as a hologram. J Math Phys 1995;36:6377–96. https://doi.org/10.1063/1.531249.

[21] Bousso R. The holographic principle. Rev Mod Phys 2002;74:825–74. https://doi.org/10.1103/RevModPhys.74.825.

[22] Fields C, Glazebrook JF, Marcianò A. The physical meaning of the holographic principle. Quanta 2022;11:72–96. https://doi.org/10.12743/quanta.v11i1.206.

[23] Stolze CH. A history of the divergence theorem. Hist Math 1978;5(4):437–42. https://doi.org/10.1016/0315-0860(78)90212-4.

[24] Friston KJ, FitzGerald T, Rigoli F, Schwartenbeck P, Pezzulo G. Active inference: a process theory. Neural Comput 2017;29:1–49. https://doi.org/10.1162/NECO_a_00912.

[25] Ramstead MJD, Constant A, Badcock PB, Friston KJ. Variational ecology and the physics of sentient systems. Phys Life Rev 2019;31:188–205. https://doi.org/10.1016/j.plrev.2018.12.002.

[26] Levin M. Bioelectric signaling: reprogrammable circuits underlying embryogenesis, regeneration, and cancer. Cell 2021;184:1971–89. https://doi.org/10.1016/j.cell.2021.02.034.

[27] Levin M. Technological approach to mind everywhere: an experimentally-grounded framework for understanding diverse bodies and minds. Front Syst Neurosci 2023;16:768201. https://doi.org/10.3389/fnsys.2022.768201.

[28] Biswas S, Manika S, Hoel E, Levin M. Gene regulatory networks exhibit several kinds of memory: quantification of memory in biological and random transcriptional networks. iScience 2021;24:102131. https://doi.org/10.1016/j.isci.2021.102131.

[29] Biswas S, Clawson W, Levin M. Learning in transcriptional network models: computational discovery of pathway-level memory and effective interventions. Int J Mol Sci 2023;4:285. https://doi.org/10.3390/ijms24010285.

[30] Levin M. Self-improvising memory: a perspective on memories as agential, dynamically reinterpreting cognitive glue. Entropy 2024;26:481. https://doi.org/10.3390/e26060481.

[31] Landauer R. Irreversibility and heat generation in the computing process. IBM J Res Dev 1961;5:183–95. https://doi.org/10.1147/rd.53.0183.

[32] Landauer R. Information is a physical entity. Physica A 1999;263:63–7. https://doi.org/10.1016/S0378-4371(98)00513-5.

[33] Bennett CH. The thermodynamics of computation. Int J Theor Phys 1982;21:905–40. https://doi.org/10.1007/BF02084158.

[34] Parrondo JMR, Horowitz JM, Sagawa T. Thermodynamics of information. Nat Phys 2015;11:131–9. https://doi.org/10.1038/nphys3230.

[35] Zurek WH. Decoherence, einselection, and the quantum origins of the classical. Rev Mod Phys 2003;75:715–75. https://doi.org/10.1103/RevModPhys.75.715.

[36] Bassi A, Grossart A, Ulbricht H. Gravitational decoherence. Class Quantum Gravity 2017;34:193002. https://doi.org/10.1088/1361-6382/aa864f.

[37] Melnychuk O, Giaccone B, Bornman N, Cervantes R, Grassellino A, Harnik R, et al. An improved bound on nonlinear quantum mechanics using a cryogenic radio frequency experiment. arXiv:2411.09611, 2024. https://doi.org/10.48550/arXiv.2411.09611.

[38] Smolin L. The case for background independence. In: Rickles D, French S, Saatsi J, editors. The Structural Foundations of Quantum Gravity. Oxford, UK: Oxford University Press; 2006. p. 196–239.

[39] Horsman C, Stepney S, Wagner RC, Kendon V. When does a physical system compute? Proc R Soc A 2014;470:20140182. https://doi.org/10.1098/rspa.2014.0182.

[40] Bongard J, Levin M. There's plenty of room right here: biological systems as evolved, overloaded, multi-scale machines. Biomimetics 2023;8(1):110. https://doi.org/10.3390/biomimetics8010110.

[41] Fields C. The free energy principle induces intracellular compartmentalization. Biochem Biophys Res Commun 2024;723:150070. https://doi.org/10.1016/j.bbrc.2024.150070.

[42] Di Biagio A, Donà P, Rovelli C. The arrow of time in operational formulations of quantum theory. Quantum 2021;5:520. https://doi.org/10.22331/q-2021-08-09-520.

[43] Wheeler JA. Information, physics, quantum: the search for links. In: Zurek W, editor. Complexity, entropy, and the physics of information. Boca Raton, FL: CRC Press; 1989. p. 3–28.

[44] Simons DJ, Ambinder MS. Change blindness: theory and consequences. Curr Dir Psychol Sci 2005;14:44–8. https://doi.org/10.1111/j.0963-7214.2005.00332.x.

[45] Gibson JJ. The senses considered as perceptual systems. Boston: Houghton-Miffin; 1966.

[46] Michaels C, Carello C. Direct perception. Englewood Cliffs, NJ: Prentice-Hall; 1981.

[47] Chemero A. Radical embodied cognitive science. Rev Gen Psychol 2013;17:145–50. https://doi.org/10.1037/a0032923.

[48] Fields C, Glazebrook JF. Representing measurement as a thermodynamic symmetry breaking. Symmetry 2020;12:810. https://doi.org/10.3390/sym12050810.

[49] Aharonov Y, Kaufherr T. Quantum frames of reference. Phys Rev D 1984;30:368–85. https://doi.org/10.1103/PhysRevD.30.368.

[50] Bartlett SD, Rudolph T, Spekkens RW. Reference frames, superselection rules, and quantum information. Rev Mod Phys 2007;79:555–609. https://doi.org/10.1103/RevModPhys.79.555.

[51] Moore EF. Gedanken-experiments on sequential machines. In: Shannon CW, McCarthy J, editors. Automata studies. Princeton, NJ: Princeton University Press; 1956. p. 129–55.

[52] Fields C, Glazebrook JF, Levin M. Principled limitations on self-representation for generic physical systems. Entropy 2024;26:194. https://doi.org/10.3390/e26030194.

[53] Flombaum JI, Scholl BJ, Santos LR. Spatiotemporal priority as a fundamental principle of object persistence. In: Hood B, Santos L, editors. The origins of object knowledge. New York: Oxford University Press; 2008. p. 135–64.

[54] Fields C. Trajectory recognition as the basis for object individuation: a functional model of object file instantiation and object-token encoding. Front Psychol 2011;2:49. https://doi.org/10.3389/fpsyg.2011.00049.

[55] Fields C. The very same thing: extending the object token concept to incorporate causal constraints on individual identity. Adv Cogn Psychol 2012;8:234–47. https://doi.org/10.2478/v10053-008-0119-8.

[56] Thompson E, Stapleton M. Making sense of sense-making: reflections on enactive and extended mind theories. Topoi 2009;28:23–30. https://doi.org/10.1007/s11245-008-9043-2.

[57] Pais A. Einstein and the quantum theory. Rev Mod Phys 1979;51:863–914. https://doi.org/10.1103/RevModPhys.51.863.

[58] Clark A. How to knit your own Markov blanket: resisting the second law with metamorphic minds. In: Metzinger T, Wiese W, editors. Philosophy and predictive processing: vol 3. Frankfurt: MIND Group; 2017 (unpaginated).

[59] Clark A, Chalmers D. The extended mind. Analysis 1998;58:7–19.

[60] Maravita A, Iriki A. Tools for the body (schema). Trends Cogn Sci 2004;8:79–86. https://doi.org/10.1016/j.tics.2003.12.008.

[61] Clawson W, Levin M. Endless forms most beautiful 2.0: teleonomy and the bioengineering of chimaeric and synthetic organisms. Biol J Linn Soc 2023;139:457–86. https://doi.org/10.1093/biolinnean/blac073.

[62] Zanardi P. Virtual quantum subsystems. Phys Rev Lett 2001;87:077901. https://doi.org/10.1103/PhysRevLett.87.077901.

[63] Zanardi P, Lidar DA, Lloyd S. Quantum tensor product structures are observable-induced. Phys Rev Lett 2004;92:060402. https://doi.org/10.1103/PhysRevLett.92.060402.

[64] Friston KJ, Thornton C, Clark A. Free-energy minimization and the dark-room problem. Front Psychol 2012;3:130. https://doi.org/10.3389/fpsyg.2012.00130.

[65] Gopnik A. Explanation as orgasm and the drive for causal understanding: the evolution, function and phenomenology of the theory-formation system. In: Keil F, Wilson R, editors. Cognition and explanation. Cambridge, MA: MIT Press; 2000. p. 299–323.

[66] Ororbia A, Friston KJ. Mortal computation: a foundation for biomimetic intelligence. Preprint. arXiv:2311.09589v2. https://doi.org/10.48550/arXiv.2311.0958.

[67] Tishby N, Pereira FC, Bialek W. The information bottleneck method. In: Proc. 37th annual allerton conf communication, control, and computing; 1999. p. 368–77.

[68] Still S, Sivak DA, Bell AJ, Crooks GE. Thermodynamics of prediction. Phys Rev Lett 2012;109:120604. https://doi.org/10.1103/PhysRevLett.109.120604.

[69] Monod J. Chance and necessity. New York: Random House; 1972.

[70] Dawkins R. Replicators and vehicles. In: Brandon RN, Burian RM, editors. Genes, organisms, populations: controversies over the units of selection. Cambridge, MA: MIT Press; 1984. p. 161–80.

[71] Szathmáry E, Maynard Smith J. The major evolutionary transitions. Nature 1995;374:227–32. https://doi.org/10.1038/374227a0.

[72] Michod RE. Darwinian dynamics. Princeton, NJ: Princeton University Press; 1999.

[73] Pigliucci M, Müller CB. Elements of an extended evolutionary synthesis. In: Pigliucci M, Müller CB, editors. Evolution - the extended synthesis. Cambridge, MA: MIT Press; 2010. p. 3–17.

[74] Danchin E, Charmantier A, Champagne FA, Mesoudi A, Pujol B, Blanchet S. Beyond DNA: integrating inclusive inheritance into an extended theory of evolution. Nat Rev Genet 2011;12:475–86. https://doi.org/10.1038/nrg3028.

[75] Laland KN, Uller T, Feldman MW, Sterelny K, Müller GB, Moczek A, et al. The extended evolutionary synthesis: its structure, assumptions and predictions. Proc R Soc B 2015;282:20151019. https://doi.org/10.1098/rspb.2015.1019.

[76] Noble D. Evolution beyond neo-Darwinism: a new conceptual framework. J Exp Biol 2015;218:7–13. https://doi.org/10.1242/jeb.106310.

[77] Fields C, Levin M. Multiscale memory and bioelectric error correction in the cytoplasm-cytoskeleton-membrane system. WIRES Syst Biol Med 2018;10:e1410. https://doi.org/10.1002/wsbm.1410.

[78] Pezzulo G, Levin M. Top-down models in biology: explanation and control of complex living systems above the molecular level. J R Soc Interface 2016;13:20160555. https://doi.org/10.1098/rsif.2016.0555.

[79] Kirchhoff M, Parr T, Palacios E, Friston K, Kiverstein J. The Markov blankets of life: autonomy, active inference and the free energy principle. J R Soc Interface 2018;15:0792. https://doi.org/10.1098/rsif.2017.0792.

[80] Kuchling F, Friston K, Georgiev G, Levin M. Morphogenesis as Bayesian inference: a variational approach to pattern formation and control in complex biological systems. Phys Life Rev 2020;33:88–108. https://doi.org/10.1016/j.plrev.2019.06.001.

[81] Kuchling F, Fields C, Levin M. Metacognition as a consequence of competing evolutionary time scales. Entropy 2022;24:601. https://doi.org/10.3390/e24050601.

[82] Friston K, Friedman DA, Constant A, Knight VB, Fields C, Parr T, et al. A variational synthesis of evolutionary and developmental dynamics. Entropy 2023;25:964. https://doi.org/10.3390/e25070964.

[83] Noble D. The music of life – biology beyond the genome. Oxford, UK: Oxford University Press; 2006.

[84] Sultan SE. Developmental plasticity: re-conceiving the genotype. Interface Focus 2017:7:2017.0009. https://doi.org/10.1098/rsfs.2017.0009.

[85] Walsh DM, Rupik G. The agential perspective: countermapping the modern synthesis. Evol Develop 2023;25:335–52. https://doi.org/10.1111/ede.12448.

[86] Fields C, Levin M. Integrating evolutionary and developmental thinking into a scale-free biology. BioEssays 2020;42:1900228. https://doi.org/10.1002/bies.201900228.

[87] Fields C, Levin M. Does evolution have a target morphology? Organisms 2020;4:57–76. https://doi.org/10.13133/2532-5876/16961.

[88] Guerrero R, Margulis L, Berlanga M. Symbiogenesis: the holobiont as a unit of evolution. Int Microbiol 2013;16:133–43. https://doi.org/10.2436/20.1501.01.188.

[89] Gilbert SF. A holobiont birth narrative: the epigenetic transmission of the human microbiome. Front Genet 2014;5:282. https://doi.org/10.3389/fgene.2014.00282.

[90] Bordenstein SR, Theis KR. Host biology in light of the microbiome: ten principles of holobionts and hologenomes. PLoS Biol 2015;13(8):e1002226. https://doi.org/10.1371/journal.pbio.1002226.

[91] Fields C, Friston K, Glazebrook JF, Levin M, Marcianò A. The free energy principle induces neuromorphic development. Neuromorph Comput Eng 2022;2:042002. https://doi.org/10.1088/2634-4386/aca7de.

[92] Polanyi M. Life's irreducible structure. Live mechanisms and information in DNA are boundary conditions with a sequence of boundaries above them. Science 1968;160:1308–12. https://doi.org/10.1126/science.160.3834.1308.

[93] Rosen R. On information and complexity. In: Casti JL, Karlqvist A, editors. Complexity, language, and life: mathematical approaches. Berlin: Springer; 1986. p. 174–96.

[94] Noble D. Genes and causation. Philos Trans R Soc A 2008;366:3001–15. https://doi.org/10.1098/rsta.2008.0086.

[95] Fields C, Levin M. Regulative development as a model for origin of life and artificial life studies. BioSystems 2023;229:104927. https://doi.org/10.1016/j.biosystems.2023.104927.

[96] Conway JH, Kochen S. The strong free will theorem. Not Am Math Soc 2009;56:226–32.

[97] Hoel EP. When the map is better than the territory. Entropy 2017;19:188. https://doi.org/10.3390/e19050188.

[98] Tanenbaum AS. Structured computer organization. Upper Saddle River, NJ: Prentice Hall; 1976.

[99] Smith JE, Nair R. The architecture of virtual machines. IEEE Comput 2005;38(5):32–8. https://doi.org/10.1109/MC.2005.173.

[100] Bedau MA. Weak emergence. In: Tomberlin J, editor. Philosophical perspectives: mind, causation, and world, vol. 11. Malden, MA: Blackwell; 1997. p. 375–99.

[101] Clayton P. Conceptual foundations of emergence theory. In: Clayton P, Davies PCW, editors. The re-emergence of emergence. Oxford, UK: Oxford University Press; 2006. p. 1–31.

[102] Ellis GFR, Kopel J. The dynamical emergence of biology from physics: branching causation via biomolecules. Front Physiol 2019;9:1966. https://doi.org/10.3389/fphys.2018.01966.

[103] Friston K, Heins C, Verbelen T, Da Costa L, Salvatori T, Markovic D, et al. From pixels to planning: scale-free active inference. arXiv:2407.20292, 2024.

[104] Kukushkin NV, Carew TJ. Memory takes time. Neuron 2017;95:259–79. https://doi.org/10.1016/j.neuron.2017.05.029.

[105] Fields C, Levin M. Why isn't sex optional? Stem-cell competition, loss of regenerative capacity, and cancer in metazoan evolution. Commun Integr Biol 2020;13:170–83. https://doi.org/10.1080/19420889.2020.1838809.

[106] Wootters WK, Zurek WH. A single quantum cannot be cloned. Nature 1982;299:802–3. https://doi.org/10.1038/299802a0.

[107] Ashby WR. Introduction to cybernetics. London: Chapman and Hall; 1956.

[108] Fields C, Levin M. Somatic multicellularity as a satisficing solution to the prediction-error minimization problem. Commun Integr Biol 2019;12:119–32. https://doi.org/10.1080/19420889.2019.1643666.

[109] Paul JH. Microbial gene transfer: an ecological perspective. J Mol Microbiol Biotechnol 1999;1(1):45–50. https://europepmc.org/article/med/10941783.

[110] Anderson JO. Gene transfer and diversification of microbial eukaryotes. Annu Rev Microbiol 2009;63:177–93. https://doi.org/10.1146/annurev.micro.091208.073203.

[111] Brito IL. Examining horizontal gene transfer in microbial communities. Nat Rev Microbiol 2021;19:422–53. https://doi.org/10.1038/s41579-021-00534-7.

[112] Metzinger T. Being no one: the self-model theory of subjectivity. Cambridge, MA: MIT Press; 2003.

[113] Qin P, Northoff G. How is our self related to midline regions and the default-mode network? NeuroImage 2011;57:1221–33. https://doi.org/10.1016/j.neuroimage.2011.05.028.

[114] Seth AK. Interoceptive inference, emotion, and the embodied self. Trends Cogn Sci 2013;17:565–73. https://doi.org/10.1016/j.tics.2013.09.007.

[115] Andrews-Hanna JR, Smallwood J, Spreng RN. The default network and self-generated thought: component processes, dynamic control, and clinical relevance. Ann NY Acad Sci 2014;1316:29–52. https://doi.org/10.1111/nyas.1236.

[116] Chater N. The mind is flat. London: Allen Lane; 2018.

[117] Csikszentmihályi M. Flow: the psychology of optimal experience. New York: Harper and Row; 1990.

[118] Bargh JA, Ferguson MJ. Beyond behaviorism: on the automaticity of higher mental processes. Psychol Bull 2000;126:925–45. https://doi.org/10.1037/0033-2909.126.6.925.

[119] Bargh JA, Schwader KL, Hailey SE, Dyer RL, Boothby EJ. Automaticity in social-cognitive processes. Trends Cogn Sci 2012;16:593–605. https://doi.org/10.1016/j.tics.2012.10.002.

[120] Dahl CJ, Lutz A, Davidson RJ. Reconstructing and deconstructing the self: cognitive mechanisms in meditation practice. Trends Cogn Sci 2015;19:515–23. https://doi.org/10.1016/j.tics.2015.07.001.

[121] Letheby C, Gerrans P. Self unbound: ego dissolution in psychedelic experience. Neurosci Conscious 2017;2017:nix016. https://doi.org/10.1093/nc/nix016.

[122] Lindahl JR, Britten WB. 'I have this feeling of not really being here': Buddhist meditation and changes in sense of self. J Conscious Stud 2019;26:157–83.

[123] Adams RA, Stephan KE, Brown HR, Frith CD, Friston KJ. The computational anatomy of psychosis. Front Psychiatry 2013;4:47. https://doi.org/10.3389/fpsyt.2013.00047.

[124] Lawson RP, Rees G, Friston KJ. An aberrant precision account of autism. Front Human Neurosci 2014;8:302. https://doi.org/10.3389/fnhum.2014.00302.

[125] Smith R, Badcock P, Friston KJ. Recent advances in the application of predictive coding and active inference models within clinical neuroscience. Psychiatry Clin Neurosci 2021;75:3–13. https://doi.org/10.1111/pcn.13138.

[126] Fields C, Glazebrook JF, Levin M. Minimal physicalism as a scale-free substrate for cognition and consciousness. Neurosci Conscious 2021;2021:niab013. https://doi.org/10.1093/nc/niab013.

[127] Levin M. The computational boundary of a "self": developmental bioelectricity drives multicellularity and scale-free cognition. Front Psychol 2019;10:2688. https://doi.org/10.3389/fpsyg.2019.02688.

[128] Levin M. Life, death, and self: fundamental questions of primitive cognition viewed through the lens M body plasticity and synthetic organisms. Biochem Biophys Res Commun 2021;564:114–33. https://doi.org/10.1016/j.bbrc.2020.10.077.

[129] Pezzulo G, Levin M. Re-membering the body: applications of computational neuroscience to the top-down control of regeneration of limbs and other complex organs. Integr Biol 2015;7:1487–517. https://doi.org/10.1039/c5ib00221d.

[130] Friedman D, Tschantz A, Ramstead MJD, Friston KJ, Constant A. Active inferants: an active inference framework for ant colony behavior. Front Behav Neurosci 2021;15:647732. https://doi.org/10.3389/fnbeh.2021.647732.

[131] Boyer P. Evolutionary economics of mental time travel? Trends Cogn Sci 2008;12(6):219–24. https://doi.org/10.1016/j.tics.2008.03.003.

[132] Fields C. How humans solve the frame problem. J Exp Theor Artif Intell 2013;25:441–56. https://doi.org/10.1080/0952813X.2012.741624.

[133] McCarthy J, Hayes PJ. Some philosophical problems from the standpoint of artificial intelligence. In: Michie DD, Meltzer B, editors. Machine intelligence. Edinburgh, UK: Edinburgh University Press; 1969. p. 463–502.

[134] Braun N, Debener S, Spychala N, Bongartz E, Sörös P, Müller HHO, et al. The senses of agency and ownership: a review. Front Psychol 2018;9:535. https://doi.org/10.3389/fpsyg.2018.00535.

[135] Deane G, Miller M, Wilkinson S. Losing ourselves: active inference, depersonalization, and meditation. Front Psychol 2020;11:539726. https://doi.org/10.3389/fpsyg.2020.539726.

[136] Humpston C, Broome M. Delusional beliefs and thought insertion. In: López-Silva P, McClelland T, editors. Intruders in the mind: interdisciplinary perspectives on thought insertion. Oxford, UK: Oxford University Press; 2023. p. 7–26.

[137] Laney C, Loftus EF. False memory. In: Brown JM, Campbell EA, editors. The Cambridge handbook of forensic psychology. Cambridge, UK: Cambridge University Press; 2010. p. 187–94.

[138] Simons JS, Garrison JR, Johnson MK. Brain mechanisms of reality monitoring. Trends Cogn Sci 2017;21:462–73. https://doi.org/10.1016/j.tics.2017.03.012.

[139] Tung A, Sperry MM, Clawson W, Pavuluri A, Bulatao S, Yue M, et al. Embryos assist morphogenesis of others through calcium and ATP signaling mechanisms in collective teratogen resistance. Nat Commun 2024;15:535. https://doi.org/10.1038/s41467-023-44522-2.

[140] McDermott D. Artificial intelligence meets natural stupidity. ACM SIGART Newsl 1976:4–9.

[141] Bateson G. Steps to an ecology of mind: collected essays in anthropology, psychiatry, evolution, and epistemology. Northvale, NJ: Jason Aronson; 1972.

[142] Fields C, Levin M. How do living things create meaning? Philosophies 2020;5:36. https://doi.org/10.3390/philosophies5040036.

[143] Mitchell M, Krakauer DC. The debate over understanding in AI's large language models. Proc Natl Acad Sci USA 2023:120:e2215907120. https://doi.org/10.1073/pnas.2215907120.

[144] Nadel L, Hupbach A, Gomez R, Newman-Smith K. Memory formation, consolidation and transformation. Neurosci Biobehav Rev 2012;36(7):1640–5. https://doi.org/10.1016/j.neubiorev.2012.03.001.

[145] Schwabe L, Nader K, Pruessner JC. Reconsolidation of human memory: brain mechanisms and clinical relevance. Biol Psychiatry 2014;76(4):274–80. https://doi.org/10.1016/j.biopsych.2014.03.008.

[146] Dobson S, Fields C. Constructing condensed memories in functorial time. J Exp Theor Artif Intell 2023. https://doi.org/10.1080/0952813X.2023.2222374. In press.

[147] Davies J, Levin M. Synthetic morphology with agential materials. Nature Rev Bioeng 2023;1:46–59. https://doi.org/10.1038/s44222-022-00001-9.

[148] Ebrahimkhani MR, Levin M. Synthetic living machines: a new window on life. iScience 2021;24:102505. https://doi.org/10.1016/j.isci.2021.102505.

[149] Kamm RD, Bashir R, Arora N, Dar RD, Gillette MU, Griffith LG, et al. The promise of multi-cellular engineered living systems. Apl Bioeng 2018;2:040901. https://doi.org/10.1063/1.5038337.

[150] Lagasse E, Levin M. Future medicine: from molecular pathways to the collective intelligence of the body. Trends Mol Med 2023;29:687–710. https://doi.org/10.1016/j.molmed.2023.06.007.

[151] Teague BP, Guye P, Weiss R. Synthetic morphogenesis. Cold Spring Harbor Perspect Biol; 2016. p. 8:a023929.

[152] Lobo D, Solano M, Bubenik GA, Levin M. A linear-encoding model explains the variability of the target morphology in regeneration. J R Soc Interface 2014;11:20130918. https://doi.org/10.1098/rsif.2013.0918.

[153] Fields C, Levin M. Competency in navigating arbitrary spaces as an invariant for analyzing cognition in diverse embodiments. Entropy 2022;24:819. https://doi.org/10.3390/e24060819.

[154] Mathews J, Chang AJ, Devlin L, Levin M. Cellular signaling pathways as plastic, proto-cognitive systems: implications for biomedicine. Patterns (N Y) 2023;4:100737. https://doi.org/10.1016/j.patter.2023.100737.

[155] Levin M. Bioelectrical approaches to cancer as a problem of the scaling of the cellular self. Prog Biophys Mol Biol 2021;165:102–13. https://doi.org/10.1016/j.pbiomolbio.2021.04.007.

[156] Chernet BT, Adams DS, Lobikin M, Levin M. Use of genetically encoded, light-gated ion translocators to control tumorigenesis. Oncotarget 2016;7:19575–88. https://doi.org/10.18632/oncotarget.8036.

[157] Chernet BT, Levin M. Endogenous voltage potentials and the microenvironment: bioelectric signals that reveal, induce and normalize cancer. J Clin Exp Oncol 2013;(Suppl 1). S1-002. https://doi.org/10.4172/2324-9110.S1-002.

[158] Chernet BT, Levin M. Transmembrane voltage potential is an essential cellular parameter for the detection and control of tumor development in a Xenopus model. Dis Models Mech 2013;6:595–607. https://doi.org/10.1242/dmm.010835.

[159] Chernet BT, Levin M. Transmembrane voltage potential of somatic cells controls oncogene-mediated tumorigenesis at long-range. Oncotarget 2014;5:3287–306. https://doi.org/10.18632/oncotarget.1935.

[160] Levin M. Bioelectric networks: the cognitive glue enabling evolutionary scaling from physiology to mind. Anim Cogn 2023;26:1865–91. https://doi.org/10.1007/s10071-023-01780-3.

[161] Hofstadter DR. I am a strange loop. New York: Basic Books; 2007.

[162] Blackiston D, Shomrat T, Levin M. The stability of memories during brain remodeling: a perspective. Commun Integr Biol 2015;8:e1073424. https://doi.org/10.1080/19420889.2015.1073424.

[163] Levin M. The new wisdom of the body: collective intelligence as a tractable interface for next-generation biomedicine. BioEssays 2024. In press. https://osf.io/preprints/osf/yq9xe.