Introduction

Author for correspondence:
Adam Safron
e-mail: asafron@gmail.com

THE ROYAL SOCIETY
PUBLISHING

Making and breaking symmetries in mind and life

Adam Safron[1,2,3,†], Dalton A. R. Sakthivadivel[4,5,†], Zahra Sheikhbahaee[6], Magnus Bein[7], Adeel Razi[8,9,10,11] and Michael Levin[12,13]

[1]Center for Psychedelic and Consciousness Research, Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine, Baltimore, MD 21224, USA
[2]Institute for Advanced Consciousness Studies, Santa Monica, CA 90403, USA
[3]Cognitive Science Program, Indiana University, Bloomington, IN 47405, USA
[4]VERSES Research Lab, Los Angeles, CA 90016, USA
[5]Departments of Mathematics, Physics and Astronomy, and Biomedical Engineering, Stony Brook University, Stony Brook, NY 11794-3651, USA
[6]David R. Cheriton School of Computer Science, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1
[7]Department of Biology and Department of Psychiatry, McGill University, Montreal, Quebec, Canada H3A 0G4
[8]Turner Institute for Brain and Mental Health, Monash University, Clayton, VIC 3800, Australia
[9]Monash Biomedical Imaging, Monash University, Clayton, VIC 3168, Australia
[10]Wellcome Centre for Human Neuroimaging, University College London, London WC1N 3AR, UK
[11]CIFAR Azrieli Global Scholars Program, CIFAR, Toronto, Ontario, Canada M5G 1M1
[12]Tufts Center for Regenerative and Developmental Biology, and [13]Allen Discovery Center, Tufts University, Medford, MA 02155-5801, USA

Symmetry is a motif featuring in almost all areas of science. Symmetries appear throughout the natural world, making them particularly important in our quest to understand the structure of the world around us. Symmetries and invariances are often first principles pointing to some lawful description of an observation, with explanations being understood as both 'satisfying' and potentially useful in their regularity. The sense of aesthetic beauty accompanying such explanations is reminiscent of our understanding of intelligence in terms of the ability to efficiently predict (or compress) data; indeed, identifying and building on symmetry can offer a particularly elegant description of a physical situation. The study of symmetries is so fundamental to mathematics and physics that one might ask where else it proves useful. This theme issue poses the question: what does the study of symmetry, and symmetry breaking, have to offer for the study of life and the mind?

1. Introduction

> When old age shall this generation waste,
>
> Thou shalt remain, in midst of other woe
>
> Than ours, a friend to man, to whom thou say'st,
>
> 'Beauty is truth, truth beauty,—that is all
>
> Ye know on earth, and all ye need to know'.
>
> —from Ode on a Grecian Urn, by John Keats

Symmetry is a motif featuring in almost all areas of science. Symmetries appear throughout the natural world, making them particularly important in our quest to understand the structure of the world around us [1,2]. Symmetries and invariances are often first principles pointing to some lawful description of an observation, with explanations being understood as both 'satisfying' and potentially useful in their regularity [3–5]. The sense of aesthetic beauty accompanying such explanations is reminiscent of our understanding of intelligence in terms of the ability to efficiently predict (or compress) data [6,7]; indeed, identifying and building on symmetry can offer a particularly elegant

description of a physical situation. The study of symmetries is so fundamental to mathematics and physics that one might ask where else it proves useful. This theme issue poses the question: what does the study of symmetry, and symmetry breaking, have to offer for the study of life and the mind?

In this diverse collection of articles, we explored the roles of symmetry in complex adaptive systems across a multitude of scales—from the emergent dynamics of biophysical systems and their underlying mechanisms, to the shaping of adaptations through ontogenic and phylogenetic processes. Across these theoretical and empirical explorations, we hoped to demonstrate that the study of symmetries may illuminate fundamental properties of living and intelligent systems. Towards this end, we considered a broad range of perspectives on (a)symmetries, exploring the extent to which intersections may be found between seemingly disparate phenomena.

One of the most powerful applications of symmetry-related concepts is found in *gauge theory* [8]. Whenever a physical theory has a redundant quantity, meaning a quantity that leaves a system's dynamics invariant with respect to local changes in the value of that quantity (a 'frame of reference' or 'gauge'), we can understand that quantity as a kind of abstract symmetry recorded in what is called a gauge field [9–11]. Deformations of gauge fields are understood as 'fictitious' forces; these forces restore the local symmetry of quantities that are dynamically invariant, by recording the system's interactions with the field of possible gauges for that quantity. Gauge theories provide a general way of modelling physical systems. Notable use-cases include general relativity's handling of gravity as the curvature of space–time and models of the attractive and repulsive forces in electromagnetic fields. These theories are so far-reaching that the word 'fictitious' may potentially be left out of descriptions of these emergent forces, as it may be the case that there are no other kinds [12].

Gauge-theoretic perspectives on biophysics have been suggested in the past, especially in the context of brains as goal-seeking systems, guided by hierarchical information processing and prediction-error minimization. This is one aspect of the view presented by the free energy principle and active inference (FEP-AI) [9,13–15]: the attracting states of nervous systems are understood as entailing predictions, where the consistent realization of these predictions can be viewed as the preservation of goal states, contingent on particular symmetries being enforced by a gauge field governing those dynamics. The importance of symmetry reaches deep into the functional aspects of the brain: mental causation may be understood as a kind of 'fictitious' force over neural dynamics (especially with respect to perception and action) [16], and a 'symmetry theory of valence' has even been proposed, in which pleasure and pain may best be understood as the degree to which mental systems transition to more or less symmetric states in some appropriate sense [17–19]. In this issue, questions such as 'along which dimensions are symmetries most important for nervous system functioning (e.g. connectomic resting-state networks as reflecting harmonic functions) [20]?' and 'could such organizational principles be evidenced by responses to different forms of music [21], or fractal-structured visual stimuli [22], or potentially the phenomenology associated with psychedelic states [23,24]?' were asked and potentially answered.

Intriguingly, symmetries may play yet another distinct role in mental causation with respect to the phenomenon of symmetry breaking, where irreversible processes and arrows of time may be required for establishing the conditions for realizing cognitive work cycles [25,26] involving varying degrees of energy expenditure and various forms of either mental 'effort' or 'flow' [27–29]. These particular temporal (a)symmetries play out in the context of the principle of detailed balance [27,30] and the breaking of detailed balance. Given that many interesting biological systems do not satisfy the detailed balance property, and that breaking detailed balance (in particular, the presence of non-zero circulation) can be described using gauge theory [31], this is an interesting direction to consider in our search for symmetries [32].

The role of symmetries as perceptual invariants and inductive biases has also been identified in machine learning [33,34]. Could these physics-inspired algorithms shed light on the computational principles underlying the remarkable intelligence of biological systems?

Could the functions of symmetries in human perception and learning help inspire advances in artificial intelligence? To what extent can the study of conservation laws of nature provide more powerful and interpretable approaches to machine learning [35,36]? Indeed, when we talk about the preconditions for 'System 2' cognition with the power of conceptual understanding through abstraction [37,38], or even our ability to generate stable percepts, are such phenomena best understood as kinds of informational symmetries? In a different direction, we may ask: to what extent does the preference for symmetry mentioned earlier come 'for free' via predictive coding, where symmetric structures may be easier to predict or compress, and where efficient prediction-error minimization or compression constitutes a foundation for valence for living organisms [39–41]?

We believe these same principles apply to morphogenesis and self-organization, based on their relations to predictive coding [42], potentially offering a way to understand the kinds of unusual causation observed in living organisms as complex adaptive systems. Perhaps we may even think of pre-theoretic intuitions relating to the nature of living phenomena, where the notion of 'elan vital' and 'life force' may receive some (limited) support from abstract formalisms [43,44]. Notably, the reliable creation (and regeneration) of particular forms over the course of development has been described in terms of 'morphogenetic fields' [45]. Some models suggest the ability of biophysical systems to construct and preserve their phenotypes can be understood as being governed by a kind of force field over an information geometry, generating particular phenotypic forms as attracting states [13]. The idea that gauge-theoretic forces could be understood as governing not just ontogeny, but also phylogeny as a free energy minimizing process, is an attractive one—especially where development is itself understood as a peculiar kind of evolution [46–50].

This collection was also inspired by the various roles symmetry can play in determining the functional properties of biological systems. Motivating questions in this direction are plentiful. How is it that symmetry breaking occurs with respect to laterality in biological systems [51–54]? How is molecular chirality, or asymmetry with respect to *direction* (present in almost all cells) amplified into body-wide asymmetry with respect to organ *position relative to the midline* in metazoan development? What is the functional significance of asymmetries in the organization of nervous systems [55–65], and do they sometimes reflect a *lack* of organismal fitness [66,67]?

The goal in asking such a broad range of questions was to look for invariances across the different ways in which the concept of symmetry can be used to characterize the natural world. While only some of these questions were directly addressed by contributors to this theme issue, the contributions we received set the stage to think much more deeply about the origins of life, consciousness and even the nature(s) of time. We are honoured to have received these contributions. Below we will attempt to provide some high-level summaries of the articles in this collection, largely drawing upon the authors' descriptions of their own work, with some of our own speculations and interpretation interspersed. Considering the profound nature(s) of these issues, there is no way we can do these authors justice with our editorial, and so we refer interested readers to the original papers.

2. Summary of contributions

With 'Symmetry-simplicity, broken symmetry-complexity', David Krakauer [68] provides a beautiful introduction to many of the core themes of this collection. He describes how complex phenomena are made possible when physical symmetries are broken and selected ground states perform mechanical work and store adaptive information. Meditating on the 50th anniversary of the groundbreaking article by Philip Anderson, 'More is different' [69], Krakauer describes how emergence, frustrated random functions, autonomy, and generalized rigidity characterize the nature(s) of complexity.

With these foundations in mind, 'A third transition in science?' by Stuart Kauffman and Andrea Roli [70] argues against the standard 'Newtonian paradigm' in which relevant variables and governing laws (or master equations) of systems may be clearly identified, with boundary conditions defining phase spaces over which all possible values (and their actions) are determined and fixed *a priori*, outside of time. These authors argue that such an approach is inadequate for the inherently time-dependent evolution of organisms in complex environments like our biosphere, wherein living cells exhibit ever-new adaptations, achieve constraint closure and construct themselves via evolutionary selection—leading to genuinely new possibilities that emerge from the edge of the adjacent possible. Their ultimate conclusion is that a 'true' phase space is necessarily undefinable using any mathematical or formal analytic tools, dashing hopes for theories of everything and the 'Pythagorean dream' which would attempt to capture the essence of all phenomena in terms of quantitative and symbolic terms. Based on these considerations, they suggest that this new major transition in the evolution of science may allow us to understand the nature(s) of emergence and the creativity of an evolving biosphere. In the context of this collection, by identifying natural systems as 'Kantian wholes', we may think of the evolution of *parts* to support the functions of *wholes* as self-organization in the sense described above: that of gauge forces and morphogenic fields.

In contrast with this anti-reductionist manifesto, with 'Bayesian mechanics: a physics of and by beliefs', Maxwell Ramstead *et al.* [71] introduce a program attempting to do what many would consider be impossible: creating a general systems theory capable of describing all 'things' (in the sense of objects individuated from an environment) within a single formal modelling framework. This landmark paper ushers in

a recasting of the FEP in terms of gauge theory, providing bridges between fundamental physics and dynamical systems perspectives on the FEP. Beginning from a model of a system in terms of a particular partition, wherein trajectories of the internal states of a system encode the parameters of beliefs about external states/processes, the core tenets of the FEP are rederived from the principle of maximum entropy, revealing a duality between the two. The authors discuss how, in this framework, mechanical theories can be specified for systems that 'look as if' they are estimating posterior probability distributions over the causes of their sensory states. The 'inferential dynamics' [72,73] of interacting systems are set up in different classes of model described as path-tracking, mode-tracking and mode-matching. The authors then formulate a gauge-theoretic description of those inferential dynamics by using the description from maximum entropy. While this new viewpoint on complex and interacting systems will likely inspire debate, the creation of such a formal modelling framework has implications which are difficult to overstate. Most relevant to this issue is its interdisciplinarity: offering a unifying theme to be found throughout these heterogeneous domains of knowledge presents the opportunity to generate new insights, and the means by which they may be realized [74,75].

The potential scope (and impact) of such all-encompassing theories is directly supported by 'Free energy and inference in living systems' by Chang Sub Kim [76], which describes organisms as non-equilibrium steady state systems, self-organizing via spontaneous symmetry breaking and undergoing metabolic cycles with broken detailed balance in the environment. He goes on to suggest bridges between the principle of homeostasis as the regulation of biochemical work constrained by physical free energy costs—cf. flux balance analysis [77]—and allostasis as Bayesian inference facilitated by informational free energy, with perception and action understood in terms of the FEP. Here, brains act as a 'Schrödinger's machine' that minimizes sensory uncertainty. Further, in line with the aforementioned 'Bayesian mechanics', the author describes how optimal trajectories in neural manifolds may induce dynamic bifurcations between neural attractors in the process of active inference.

The potential explanatory (and clinical) utility of dynamical perspectives on minds in terms of bifurcating attractors is clearly expressed in 'The lack of temporal brain dynamics asymmetry as a signature of impaired consciousness states' by Elvira García Guzmán *et al.* [78] Beginning with an intriguing discussion, the ways in which complex adaptive systems must find ways to maintain themselves far from thermodynamic equilibrium, the authors go on to present a framework based on temporal asymmetry as a measure of non-equilibrium dynamics. The authors further detail how machine learning techniques can be used to establish an 'arrow of time' based on the reversibility of empirically measured time-series. Fascinatingly, decreases in asymmetry and non-stationarity of brain signals were found to be characteristics of impaired consciousness states, demonstrating how highly abstract conceptual frameworks can also end up being highly practical. In this case, a deep understanding of the nature(s) of (a)symmetries in mind may yield powerful tools for studying consciousness in both fundamental and translational research.

A further beautiful example of the power of dynamical systems perspectives on brain and mind can be found in

'Neuromodulatory control of complex adaptive dynamics in the brain' by James Mac Shine [79], who asks a key question: how can the massive complexity of nervous systems be brought under sufficiently tight control to coordinate adaptive behaviour? Shine suggests neurons are balanced close to critical point phase transitions, where small perturbations to neuronal excitability lead to nonlinear changes in overall neural activity, in order to realize this capacity to be coordinated. He describes how various portions of the brain's ascending arousal system provide a diverse set of heterogeneous control parameters that can be used to modulate the excitability and receptivity of target neurons. These mechanisms provide control variables and critical order parameters with respect to the topological complexity of neural networks and their dynamics, which govern complex adaptive behaviours. Once again, in addition to its theoretical import, this work likely has practical consequences—describing ways in which nervous systems may (or may not) be able to operate cohesively, as integrated systems capable of adaptatively responding to a wide range of events. Criticality, and self-organization to near-critical regimes, is important for more than just flexible brain functioning [80], but potentially represents a 'universality class' in that it may be a hallmark of all complex adaptive systems capable of persisting in a complex and changing world. Thus, such principles of (near-)critical control may apply to more than just brains, but to all cybernetic systems, ranging from the intelligent functioning of multicellular organisms as wholes to the functioning of individual cells.

Along these lines, 'The scaling of goals via homeostasis: an evolutionary simulation, experiment and analysis' by Léo Pio-Lopez *et al.* [81] demonstrates how fluctuating stress levels—cf. stochastic resonance and low rattling [82,83]—may allow for surprising levels of intelligence from coordinating sub-agents (here, in the context of a morphogenic process). The authors ask the question: what evolutionary dynamics enable individual cells to integrate their activities, resulting in the emergence of a novel, higher level intelligence with goals and competencies that belong to *it* and not to its parts? They describe a system consistent with the 'TAME' framework, in which evolution harnesses the collective intelligence of cells during morphogenesis of the body to develop traditional behavioural intelligence, scaling up goal states at the centre of homeostatic processes [84]. Using the classic French flag problem as a model of the establishment and maintenance of a particular symmetry (a body-wide positional axis), the authors confirmed predictions that emergent morphogenetic agents could use a combination of local and global signalling to achieve target morphologies, recover from perturbation, achieve long-term stability and even suddenly remodel long after the system stabilizes (a hallmark of near-critical systems). They further tested these predictions on a genuine biological system, observing similar phenomena in regenerating *Planaria* (flatworms), suggesting that these principles of intelligence arising from multi-scale competent sub-agents may provide a powerful explanatory framework for understanding the principles of intelligence underlying all living systems, and potentially our attempts to engineer systems that operate according to similar principles.

In 'Embodied cognitive morphogenesis as a route to intelligent systems', Bradly Alicea *et al.* [85] provide a powerfully compelling account of the ways in which morphogenetic symmetry breaking produces specialized organismal subsystems, and how this serves as a substrate for the emergence of autonomous behaviours with properties related to acquisition, generativity and transformation. They describe this embodied cognitive morphogenesis as providing a means of bridging an 'embryological view' (emphasizing coordinated gene expression, cellular physics, and migration as the basis for phenotypic complexity) with a more enactivist perspective (centring on informational feedback between organisms and their environment is key to the emergence of intelligent behaviours). They further outline their work with generic organismal agent modelling—involving tensegrity networks, differentiation trees and embodied hypernetworks—providing a means to identify the context for various symmetry-breaking events in developmental time. Theirs is a richly detailed framework that integrates a diversity of concepts, including modularity, homeostasis, 4E (embodied, enactive, embedded and extended) cognition and more.

Contemplating the asymmetric production of forms in nature, with 'Chiral conformity emerges from the least-time free energy consumption', Arto Annila [86] describes chiral symmetry breaking ('handedness', or geometric rotational asymmetry) across multiple systems, ranging from biophysical processes to the disproportional generation of matter and antimatter in the creation of the universe. The author argues that such symmetry breaking may not necessarily involve initial biases with respect to some underlying generative process. Rather, an analogy may be drawn with some aspects of handedness standards in societies, which have 'evolved over time to make things work'. The author goes on to draw connections between work as the universal measure of transferred energy, with universal principles of free energy minimization, and its relationships with the principle of maximum entropy production (as described earlier in this collection). The author further argues for an ontology that describes everything in terms of 'quanta of action', providing further bridges to the FEP (as a framework of systems actively inferring themselves into being), and for a universal law in which flows of energy naturally select some structures over others, based on their capacity to consume free energy in the least time [87]. This may be thought of as another convergent line of support for the approach of Bayesian mechanics, based on its connections to the principle of maximum calibre [88]. Ultimately the author concludes that it is 'meaningless' to speculate about life's origins, since 'thermodynamics makes no distinction between animate and inanimate'.

However, if approached with care, perhaps such questions may not be as meaningless as they might seem. In 'Mixed anhydrides at the intersection between peptide and RNA autocatalytic sets: evolution of biological coding', Stuart Kauffman and Niles Lehman [89] offer a proposal for the origins of biological coding as a semiotic relationship between chemical information stored in one location that links to chemical information stored in a separate location. In their proposal, coding originates from cooperation between two, originally separate, collectively autocatalytic sets, which (via pressures to eliminate energetic waste) eventually results in a 1 : 1 relationship between single amino acids and short RNA pieces—establishing the 'genetic code'. The idea of Kantian wholes is introduced again, where every stage in the evolution of coding is driven by the downward selection on the components of a system,

according to a kind of criterial causation [90]. On this view, the separation between *code* and *coded* was a prerequisite for robust cumulative evolution [91–93] and thus may be 'synonymous with life as we know it'.

In 'As without, so within: how the brain's temporo-spatial alignment to the environment shapes consciousness', Georg Northoff *et al.* [94] consider symmetries with respect to modes of synchronization between brain and environment. With the 'temporo-spatial theory of consciousness', temporo-spatial alignment provides a mechanism by which the brain adapts to—and coordinates its neuronal activity with—various intero-ceptive (bodily) and exteroceptive (environmental) stimuli. A three-layer, conceptual, neuro-phenomenal model for con-sciousness is proposed; this is constituted by various lengths of the brain's intrinsic neuronal timescales, each corresponding to phenomenal layers of consciousness, such as the environ-mental background and specific contents in the foreground of awareness. These layers are shared between the brain's neur-onal activity and consciousness, providing their 'common currency' on dynamical grounds coupled to the environment. Through temporo-spatial alignment this suggests ways in which various forms of mental representation may emerge as a kind of resonant matching (or entrainment), which offers ana-logies to both thermodynamic and informational free energy gradients. These ideas are also considered in terms of the closely related integrated world modelling theory of consciousness, which attempts to bring together multiple models within the overarching framework of the FEP [95,96].

In 'Symmetry and complexity in object-centric deep active inference models', Stefano Ferraro *et al.* [97] describe auton-omous robotic systems designed according to the principles of the FEP, wherein agents achieve coherent perception—and take intelligent actions—by employing mental models of objects that exploit symmetries in shape and appearance. Such 'object-ness' is often described as part of inborn 'core knowledge', by which biological learners are capable of effi-ciently bootstrapping sophisticated mental functions with so little training data [98,99]. This work, however, demonstrates how such capacities may be learned without clear innate induc-tive biases directly related to learning object models. Rather, these agents learn and act by minimizing an upper bound on their surprisal (i.e. their free energy) with respect to a genera-tive model describing (and governing) interactions with their environment. This free energy objective functional decomposes into accuracy and complexity terms, inducing pressure on agents to favour simpler models that can accurately explain sensory observations. As a result of this set-up, inherent symmetries of particular objects also emerge as (complexity-minimizing) symmetries in latent state spaces of generative models that minimize free energy. These object-centric representations further allow for novel object views to be pre-dicted as the agent moves and changes its perspective on the world. The principal axes of such symmetries were observed as principal components explaining significant amounts of var-iance for objects within latent space, which were not just elegantly explanatory, but also functional, in that exploiting these symmetries allowed for better generalization in the con-text of robotic manipulation. Intriguingly, shared latent space representations have also been associated with consciousness theories based on world modelling [5,95,96]. There, per-ceptions are generated as the iterative prediction of likely system-world configurations, conditioned on causal world models whereby embodied agents select actions predicted to

realize value (or minimize expected free energy). Perhaps more relevant to this theme issue, if the identification of sym-metric forms is an essential part of autonomous functioning (and a means of reducing the complexity of internal models), then it may be no surprise that we often search for (and construct) such forms in the world.

The power of identifying symmetries for adaptive functioning is compellingly demonstrated with 'Emergence of common concepts, symmetries, and conformity in agent groups—an information-theoretic model', in which Marco Möller and Daniel Polani [100] argue that leveraging symmetry is the principle underlying the efficient and accurate formation of shared representations of the world. The authors demon-strate, in a simulation of agents capable of perception–action loops, that when some common structure is appreciated by all agents at once, it is usually one rooted in some symmetry of the environment. In these models of a simple environment, individuals extract representations through an information maximization principle, which differ across agents to varying degrees. Drawing upon a variant of the information bottleneck principle [101–103], they extract a 'common concep-tualization' of the world for this group of agents, which is shown to emphasize higher regularities (or symmetries) of the environment, as compared to the individual represen-tations. Möller and Polani further examine the identification of symmetries in the environment both with respect to 'extrin-sic' (allocentric) operations on the environment as well as with respect to 'intrinsic' operations corresponding to reconfigura-tions from the point-of-view of the agent's particular embodiment-embedding. The authors note how the intrinsic perspective supports more efficient coordination of concepts, which further support generalization and transfer across learn-ing environments. In the light of the effectiveness of group equivariant neural networks, which are known to exploit rep-resentational symmetries in data [104,105], this is a particularly relevant contribution to our issue. It is tempting to think of its potential extensions to modelling groups as collective minds [106,107], as well as minds as collections of sub-agents [108–110]. One may even wonder whether identifying sym-metries, as points of coordination among competing and cooperating processes, may provide one of the major functional or adaptive roles of consciousness—especially in the context of the information bottleneck or cybernetic control [111–113].

Finally, we come full circle, examining the question of how there may be space for our meanings in the physical world [114]. In 'Reflections on the asymmetry of causation', Jenann Ismael [115] helps to guide us through a vision in which scientific understanding and the 'manifest image' may be not just reconcilable [16,116], but mutually enriching. She begins by reminding us how causation is a paradigmatic example of an objective worldly relation, even *the* fundamen-tal ordering relation of the world. She then considers Bertrand Russell's suggestions that the concept of causation ought to be replaced with time-symmetric laws of temporal evolution and goes on to explore the question of how time asymmetries emerge from otherwise symmetric underlying dynamics. While there may be no place for notions of time or cause in fundamental physics, the asymmetry of causation is arguably the most important kind of symmetry breaking for our experience of the natural world. Ismael asks the question '[W]hat precisely, is the status of the causal arrow, assuming a thermodynamic gradient and the interventionist account of causation?' She then provides a compelling

answer to this question, by showing how an objective asymmetry can be rooted in interventionist causal pathways that propagate influence into the future, but not into the past. She supports this argument by suggesting that the present macrostate of the world screens off probabilistic correlations to the past, in the presence of a low entropy boundary condition and under a particular agent-centred coarse-graining of the world. However, this agent-dependent carving up of the world does not leave us with arbitrary perspectivism, but instead demands a certain shared sensemaking. It is neither the singular vision of the Newtonian paradigm, nor the view-from-nowhere of Laplace's demon, nor a capricious relativism. Rather, it is an ecumenical, human place, where we may find space for things such as beauty and meanings, and—perhaps—even truth.

3. Conclusion

Each of these articles presents an interesting take on a symmetry or asymmetry relevant to complex adaptive systems. We believe that the study of such systems, which includes many of the most interesting questions in twenty-first-century physics—ranging from soft matter and active matter, to living and intelligent systems, to neural dynamics and consciousness—is unexpectedly rich in symmetries. While not all these articles are conceptualized in terms of the FEP, we nonetheless believe that it provides a golden thread stretching through this collection, complementary and deeply entwined with the application of principles of symmetry to the world of complex systems. However, no matter the particular formalism within which we ask and answer these questions, one thing is clear: the search for symmetry is a productive one that will likely prove a guiding principle in these areas—just as it has been throughout history in other areas of mathematics and physics.

References

1. Gross DJ. 1996 The role of symmetry in fundamental physics. *Proc. Natl Acad. Sci. USA* **93**, 14 256–14 259. (doi:10.1073/pnas.93.25.14256)

2. Atiyah M, Dijkgraaf R, Hitchin N. 2010 Geometry and physics. *Phil. Trans. R. Soc. A* **368**, 913–926. (doi:10.1098/rsta.2009.0227)

3. Gopnik A. 1998 Explanation as orgasm. *Minds Mach.* **8**, 101–118. (doi:10.1023/A:1008290415597)

4. Oh Y, Chesebrough C, Erickson B, Zhang F, Kounios J. 2020 An insight-related neural reward signal. *Neuroimage* **214**, 116757. (doi:10.1016/j.neuroimage.2020.116757)

5. Safron A, Çatal O, Verbelen T. 2022 Generalized simultaneous localization and mapping (G-SLAM) as unification framework for natural and artificial intelligences: towards reverse engineering the hippocampal/entorhinal system and principles of high-level cognition. *Front. Syst. Neurosci.* **16**, 787659. (doi:10.3389/fnsys.2022.787659)

6. Safron A. 2023 AIXI, FEP-AI, and integrated world models: towards a unified understanding of intelligence and consciousness. In *Active inference* (ed. CL Buckley *et al.*), pp. 251–273. Cham, Switzerland: Springer. (doi:10.1007/978-3-031-28719-0_18)

7. Schmidhuber J. 2010 Formal theory of creativity, fun, and intrinsic motivation (1990–2010). *IEEE Trans. Auton. Ment. Dev.* **2**, 230–247. (doi:10.1109/TAMD.2010.2056368)

8. Bomark NE. 2020 Teaching gauge theory to first year students. (https://arxiv.org/abs/2009.02162)

9. Fagerholm ED, Foulkes WM, Gallero-Salas Y, Helmchen F, Friston KJ, Moran RJ, Leech R. 2020 Conservation laws by virtue of scale symmetries in neural systems. *PLoS Comput. Biol.* **16**, e1007865. (doi:10.1371/journal.pcbi.1007865)

10. Maldacena J. 2015 The symmetry and simplicity of the laws of physics and the Higgs boson. *Eur. J. Phys.* **37**, 015802. (doi:10.1088/0143-0807/37/1/015802)

11. Rovelli C. 2020 Gauge is more than mathematical redundancy. In *One hundred years of gauge theory: past, present and future perspectives* (eds S De Bianchi, C Kiefer), pp. 107–110. Cham, Switzerland: Springer International Publishing.

12. Carroll S. 2016 *The big picture: on the origins of life, meaning, and the universe itself*. New York, NY: Penguin.

13. Sengupta B, Tozzi A, Cooray GK, Douglas PK, Friston KJ. 2016 Towards a neuronal gauge theory. *PLoS Biol.* **14**, e1002400. (doi:10.1371/journal.pbio.1002400)

14. Friston KJ. 2010 The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* **11**, 127–138. (doi:10.1038/nrn2787)

15. Tozzi A, Sengupta B, Peters JF, Friston KJ. 2017 Gauge fields in the central nervous system. In *The physics of the mind and brain disorders: integrated neural circuits supporting the emergence of mind* (eds I Opris, MF Casanova), pp. 193–212. Cham, Switzerland: Springer International Publishing.

16. Safron A. 2021 The radically embodied conscious cybernetic Bayesian brain: from free energy to free will and back again. *Entropy* **23**, 6. (doi:10.3390/e23060783)

17. Johnson ME. 2017 *Principia qualia: volume I: framework & valence*, 1st edn. Scotts Valley, CA: CreateSpace Independent Publishing Platform.

18. Joffily M, Coricelli G. 2013 Emotional valence and the free-energy principle. *PLoS Comput. Biol.* **9**, e1003094. (doi:10.1371/journal.pcbi.1003094)

19. Hesp C, Smith R, Allen M, Friston K, Ramstead M. 2021 Deeply felt affect: the emergence of valence in deep active inference. *Neural Comput.* **33**, 398–446. (doi:10.1162/neco_a_01341)

20. Atasoy S, Donnelly I, Pearson J. 2016 Human brain networks function in connectome-specific harmonic waves. *Nat. Commun.* **7**, 10340. (doi:10.1038/ncomms10340)

21. Atasoy S, Roseman L, Kaelen M, Kringelbach ML, Deco G, Carhart-Harris RL. 2017 Connectome-harmonic decomposition of human brain activity reveals dynamical repertoire re-organization under LSD. *Sci. Rep.* **7**, 17661. (doi:10.1038/s41598-017-17546-0)

22. Robles KE, Liaw NA, Taylor RP, Baldwin DA, Sereno ME. 2020 A shared fractal aesthetic across development. *Human. Soc. Sci. Commun.* **7**, 158. (doi:10.1057/s41599-020-00648-y)

23. Bressloff PC, Cowan JD, Golubitsky M, Thomas PJ, Wiener MC. 2002 What geometric visual hallucinations tell us about the visual cortex. *Neural Comput.* **14**, 473–491. (doi:10.1162/089976602317250861)

24. Varley TF, Carhart-Harris R, Roseman L, Menon DK, Stamatakis EA. 2020 Serotonergic psychedelics LSD & psilocybin increase the fractal dimension of cortical brain activity in spatial and temporal domains. *Neuroimage* **220**, 117049. (doi:10.1016/j.neuroimage.2020.117049)

25. Smith R. 2014 Do brains have an arrow of time? *Phil. Sci.* **81**, 265–275. (doi:10.1086/675644)

26. Fry RL. 2017 Physical intelligence and thermodynamic computing. *Entropy* **19**, 3. (doi:10.3390/e19030107)

27. Lynn CW, Cornblath EJ, Papadopoulos L, Bertolero MA, Bassett DS. 2021 Broken detailed balance and entropy production in the human brain. *Proc. Natl Acad. Sci. USA* **118**, e2109889118. (doi:10.1073/pnas.2109889118)

28. Tardiff N, Medaglia JD, Bassett DS, Thompson-Schill SL. 2021 The modulation of brain network integration and arousal during exploration. *Neuroimage* **240**, 118369. (doi:10.1016/j.neuroimage.2021.118369)

29. Csíkszentmihályi M. 1991 *Flow*. New York, NY: Harper Collins.

30. Wang J, Xu L, Wang E. 2008 Potential landscape and flux framework of nonequilibrium networks: robustness, dissipation, and coherence of biochemical oscillations. *Proc. Natl Acad. Sci. USA* **105**, 12 271–12 276. (doi:10.1073/pnas.0800579105)

31. Polettini M. 2012 Nonequilibrium thermodynamics as a gauge theory. *EPL* **97**, 30003. (doi:10.1209/0295-5075/97/30003)

32. Friston KJ, Fagerholm ED, Zarghami TS, Parr T, Hipólito I, Magrou L, Razi A. 2021 Parcels and particles: Markov blankets in the brain. *Network Neurosci.* **5**, 211–251. (doi:10.1162/netn_a_00175)

33. Cohen TS, Weiler M, Kicanaoglu B, Welling M. 2019 Gauge equivariant convolutional networks and the icosahedral CNN. (http://arxiv.org/abs/1902.04615)

34. Liu Z, Chen Y, Du Y, Tegmark M. 2021 Physics-augmented learning: a new paradigm beyond physics-informed learning. (https://arxiv.org/abs/2109.13901)

35. Liu Z, Tegmark M. 2021 Machine learning conservation laws from trajectories. *Phys. Rev. Lett.* **126**, 180604. (doi:10.1103/PhysRevLett.126.180604)

36. Cranmer M, Sanchez-Gonzalez A, Battaglia P, Xu R, Cranmer K, Spergel D, Ho S. 2020 Discovering symbolic models from deep learning with inductive biases. (http://arxiv.org/abs/2006.11287)

37. Kahneman D. 2011 *Thinking, fast and slow*, 1st edn. New York, NY: Farrar, Straus and Giroux.

38. Bengio Y. 2017 The consciousness prior. (http://arxiv.org/abs/1709.08568)

39. Dresp-Langley B, Wandeto JM. 2021 Human symmetry uncertainty detected by a self-organizing neural network map. *Symmetry* **13**, 299. (doi:10.3390/sym13020299)

40. Renoult JP, Bovet J, Raymond M. 2016 Beauty is in the efficient coding of the beholder. *R. Soc. Open Sci.* **3**, 160027. (doi:10.1098/rsos.160027)

41. Schmidhuber J. 2007 Simple algorithmic principles of discovery, subjective beauty, selective attention, curiosity & creativity. (http://arxiv.org/abs/0709.0674)

42. Fields C, Levin M. 2020 Scale-free biology: integrating evolutionary and developmental thinking. *Bioessays* **42**, 1900228. (doi:10.1002/bies.201900228)

43. Friston KJ, Wiese W, Hobson JA. 2020 Sentience and the origins of consciousness: from Cartesian duality to Markovian monism. *Entropy* **22**, 5. (doi:10.3390/e22050516)

44. Koch C. 2019 *The feeling of life itself: why consciousness is widespread but can't be computed*. Cambridge, MA: MIT Press.

45. Jaeger J, Monk N. 2021 Dynamical modules in metabolism, cell and developmental biology. *Interface Focus* **11**, 20210011. (doi:10.1098/rsfs.2021.0011)

46. Campbell JO. 2016 Universal Darwinism as a process of Bayesian inference. *Front. Syst. Neurosci.* **10**, 49. (doi:10.3389/fnsys.2016.00049)

47. Kaila V, Annila A. 2008 Natural selection for least action. *Proc. R. Soc. A* **464**, 3055–3070. (doi:10.1098/rspa.2008.0178)

48. Safron A. 2019 Multilevel evolutionary developmental optimization (MEDO): a theoretical framework for understanding preferences and selection dynamics. (http://arxiv.org/abs/1910.13443)

49. Stewart I, Elmhirst T, Cohen J. 2003 Symmetry-breaking as an origin of species. In *Bifurcation, symmetry and patterns* (eds J Buescu, SBSD Castro, AP da Silva Dias, IS Labouriau), pp. 3–54. Basel, Switzerland: Birkhäuser.

50. Friston K, Friedman DA, Constant A, Knight VB, Parr T, Campbell JO. 2023 A variational synthesis of evolutionary and developmental dynamics. (https://arxiv.org/abs/2303.04898)

51. Larsson M. 2017 Did heart asymmetry play a role in the evolution of human handedness? *J. Cult. Cogn. Sci.* **1**, 65–76. (doi:10.1007/s41809-017-0009-z)

52. Levin M. 2005 Left–right asymmetry in embryonic development: a comprehensive review. *Mech. Dev.* **122**, 3–25. (doi:10.1016/j.mod.2004.08.006)

53. Levin M, Johnson RL, Sterna CD, Kuehn M, Tabin C. 1995 A molecular pathway determining left-right asymmetry in chick embryogenesis. *Cell* **82**, 803–814. (doi:10.1016/0092-8674(95)90477-8)

54. Levin M, Thorlin T, Robinson KR, Nogi T, Mercola M. 2002 Asymmetries in H+/K+-ATPase and cell membrane potentials comprise a very early step in left-right patterning. *Cell* **111**, 77–89. (doi:10.1016/S0092-8674(02)00939-X)

55. Beeman MJ, Bowden EM. 2000 The right hemisphere maintains solution-related activation for yet-to-be-solved problems. *Mem. Cognit.* **28**, 1231–1241. (doi:10.3758/BF03211823)

56. Bowden EM, Jung-Beeman M. 2003 Aha! Insight experience correlates with solution activation in the right hemisphere. *Psychon. Bull. Rev.* **10**, 730–737. (doi:10.3758/BF03196539)

57. Brookshire G, Casasanto D. 2012 Motivation and motor control: hemispheric specialization for approach motivation reverses with handedness. *PLoS ONE* **7**, e36036. (doi:10.1371/journal.pone.0036036)

58. Craig AD. 2014 *How do you feel? An interoceptive moment with your neurobiological self*, illustrated edn. Princeton, NJ: Princeton University Press.

59. Davidson RJ, Ekman P, Saron CD, Senulis JA, Friesen WV. 1990 Approach-withdrawal and cerebral asymmetry: emotional expression and brain physiology. I. *J. Pers. Soc. Psychol.* **58**, 330–341. (doi:10.1037/0022-3514.58.2.330)

60. Gazzaniga MS. 2018 *The consciousness instinct: unraveling the mystery of how the brain makes the mind*. New York, NY: Farrar, Straus and Giroux.

61. Kaur A, Chinnadurai V, Chaujar R. 2020 Microstates-based resting frontal alpha asymmetry approach for understanding affect and approach/withdrawal behavior. *Sci. Rep.* **10**, 1. (doi:10.1038/s41598-020-61119-7)

62. McGilchrist I. 2019 *The master and his emissary: the divided brain and the making of the western world*. New Haven, CT: Yale University Press.

63. Pinto Y, Neville DA, Otten M, Corballis PM, Lamme VA, de Haan EH, Foschi N, Fabri M. 2017 Split brain: divided perception but undivided consciousness. *Brain* **140**, 1231–1237. (doi:10.1093/brain/awx235)

64. Sasai S, Boly M, Mensen A, Tononi G. 2016 Functional split brain in a driving/listening paradigm. *Proc. Natl Acad. Sci. USA* **113**, 14 444–14 449. (doi:10.1073/pnas.1613200113)

65. Tommasi L. 2009 Mechanisms and functions of brain and behavioural asymmetries. *Phil. Trans. R. Soc. B* **364**, 855–859. (doi:10.1098/rstb.2008.0293)

66. Lens L, Van Dongen S, Kark S, Matthysen E. 2002 Fluctuating asymmetry as an indicator of fitness: can we bridge the gap between studies? *Biol. Rev. Camb. Philos. Soc.* **77**, 27–38. (doi:10.1017/S1464793101005796)

67. Wiberg A et al. 2019 Handedness, language areas and neuropsychiatric diseases: insights from brain imaging and genetics. *Brain* **142**, 2938–2947. (doi:10.1093/brain/awz257)

68. Krakauer DC. 2023 Symmetry–simplicity, broken symmetry–complexity. *Interface Focus* **13**, 20220075. (doi:10.1098/rsfs.2022.0075)

69. Anderson PW. 1972 More is different. *Science* **177**, 393–396. (doi:10.1126/science.177.4047.393)

70. Kauffman SA, Roli A. 2023 A third transition in science? *Interface Focus* **13**, 20220063. (doi:10.1098/rsfs.2022.0063)

71. Ramstead MJD, Sakthivadivel DAR, Heins C, Koudahl M, Millidge B, Da Costa L, Klein B, Friston KJ. 2023 On Bayesian mechanics: a physics of and by beliefs. *Interface Focus* **13**, 20220029. (doi:10.1098/rsfs.2022.0029)

72. Parr T, Da Costa L, Friston K. 2020 Markov blankets, information geometry and stochastic thermodynamics. *Phil. Trans. R. Soc. A* **378**, 20190159. (doi:10.1098/rsta.2019.0159)

73. Parr T. 2022 Inferential dynamics: Comment on: how particular is the physics of the free energy principle? by Aguilera et al. *Phys. Life Rev.* **42**, 1–3. (doi:10.1016/j.plrev.2022.05.006)

74. Hofstadter D, Sander E. 2013 *Surfaces and essences: analogy as the fuel and fire of thinking*, 1st edn. New York, NY: Basic Books.

75. Safron A. 2019 Bayesian analogical cybernetics. (http://arxiv.org/abs/1911.02362)

76. Kim CS. 2023 Free energy and inference in living systems. *Interface Focus* **13**, 20220041. (doi:10.1098/rsfs.2022.0041)

77. Raman K, Chandra N. 2009 Flux balance analysis of biological systems: applications and challenges. *Brief. Bioinform.* **10**, 435–449. (doi:10.1093/bib/bbp011)

78. G-Guzmán E *et al.* 2023 The lack of temporal brain dynamics asymmetry as a signature of impaired consciousness states. *Interface Focus* **13**, 20220086. (doi:10.1098/rsfs.2022.0086)

79. Shine JM. 2023 Neuromodulatory control of complex adaptive dynamics in the brain. *Interface Focus* **13**, 20220079. (doi:10.1098/rsfs.2022.0079)

80. Safron A, Klimaj V, Hipólito I. 2022 On the importance of being flexible: dynamic brain networks and their potential functional significances. *Front. Syst. Neurosci.* **15**, 688424. (doi:10.3389/fnsys.2021.688424)

81. Pio-Lopez L, Bischof J, LaPalme JV, Levin M. 2023 The scaling of goals from cellular to anatomical homeostasis: an evolutionary simulation, experiment and analysis. *Interface Focus* **13**, 20220072. (doi:10.1098/rsfs.2022.0072)

82. Vázquez-Rodríguez B, Avena-Koenigsberger A, Sporns O, Griffa A, Hagmann P, Larralde H. 2017 Stochastic resonance at criticality in a network model of the human cortex. *Sci. Rep.* **7**, 13020. (doi:10.1038/s41598-017-13400-5)

83. Chvykov P, Berrueta TA, Vardhan A, Savoie W, Samland A, Murphey TD, Wiesenfeld K, Goldman DI, England JL. 2021 Low rattling: a predictive principle for self-organization in active collectives. *Science* **371**, 90–95. (doi:10.1126/science.abc6182)

84. Levin M. 2022 Technological approach to mind everywhere: an experimentally-grounded framework for understanding diverse bodies and minds. *Front. Syst. Neurosci.* **16**, 768201. (doi:10.3389/fnsys.2022.768201)

85. Alicea B, Gordon R, Parent J. 2023 Embodied cognitive morphogenesis as a route to intelligent systems. *Interface Focus* **13**, 20220067. (doi:10.1098/rsfs.2022.0067)

86. Annila A. 2023 Chiral conformity emerges from the least-time free energy consumption. *Interface Focus* **13**, 20220074. (doi:10.1098/rsfs.2022.0074)

87. Bejan A, Lorente S. 2010 The constructal law of design and evolution in nature. *Phil. Trans. R Soc. Lond. B* **365**, 1335–1347. (doi:10.1098/rstb.2009.0302)

88. Sakthivadivel DAR. 2022 Towards a geometry and analysis for Bayesian mechanics. (https://arxiv.org/abs/2204.11900)

89. Kauffman SA, Lehman N. 2023 Mixed anhydrides at the intersection between peptide and RNA autocatalytic sets: evolution of biological coding.

90. Tse P. 2013 *The neural basis of free will: criterial causation*. Cambridge, MA: MIT Press.

91. Rocha LM. 2000 Syntactic autonomy. Why there is no autonomy without symbols and how self-organizing systems might evolve them. *Ann. N. Y. Acad. Sci.* **901**, 207–223. (doi:10.1111/j.1749-6632.2000.tb06280.x)

92. Pattee HH. 2001 The physics of symbols: bridging the epistemic cut. *Biosystems* **60**, 5–21. (doi:10.1016/S0303-2647(01)00104-6)

93. Deacon TW. 2011 *Incomplete nature: how mind emerged from matter*, 1st edn. New York, NY: W. W. Norton & Company.

94. Northoff G, Klar P, Bein M, Safron A. 2023 As without, so within: how the brain's temporo-spatial alignment to the environment shapes consciousness. *Interface Focus* **13**, 20220076. (doi:10.1098/rsfs.2022.0076)

95. Safron A. 2020 An Integrated World Modeling Theory (IWMT) of consciousness: combining integrated information and global neuronal workspace theories with the free energy principle and active inference framework; toward solving the hard problem and characterizing agentic causation. *Front. Artif. Intell.* **3**, 30. (doi:10.3389/frai.2020.00030)

96. Safron A. 2022 Integrated world modeling theory expanded: implications for the future of consciousness. *Front. Comput. Neurosci.* **16**, 642397. (doi:10.3389/fncom.2022.642397)

97. Ferraro S, Van de Maele T, Verbelen T, Dhoedt B. 2023 Symmetry and complexity in object-centric deep active inference models. *Interface Focus* **13**, 20220077. (doi:10.1098/rsfs.2022.0077)

98. Spelke ES. 2000 Core knowledge. *Am. Psychol.* **55**, 1233–1243. (doi:10.1037/0003-066X.55.11.1233)

99. von Hofsten C, Feng Q, Spelke ES. 2000 Object representation and predictive action in infancy. *Dev. Sci.* **3**, 193–205. (doi:10.1111/1467-7687.00113)

100. Möller M, Polani D. 2023 Emergence of common concepts, symmetries and conformity in agent groups—an information-theoretic model. *Interface Focus* **13**, 20230006. (doi:10.1098/rsfs.2023.0006)

101. Tishby N, Pereira FC, Bialek W. 2000 The information bottleneck method. (https://arxiv.org/abs/physics/0004057)

102. Tishby N, Zaslavsky N. 2015 Deep learning and the information bottleneck principle. (http://arxiv.org/abs/1503.02406)

103. Murphy KA, Bassett DS. 2022 The distributed information bottleneck reveals the explanatory structure of complex systems. (https://arxiv.org/abs/2204.07576)

104. Cohen T, Welling M. 2016 Group equivariant convolutional networks. In *Proc. 33rd Int. Conf. on Machine Learning, New York, NY, USA, 19–24 June 2016*, pp. 2990–2999. See https://proceedings.mlr.press/v48/cohenc16.html.

105. Veličković P. 2023 Everything is connected: graph neural networks. *Curr. Opin. Struct. Biol.* **79**, 102538. (doi:10.1016/j.sbi.2023.102538)

106. Engel D, Malone TW. 2018 Integrated information as a metric for group interaction. *PLoS ONE* **13**, e0205335. (doi:10.1371/journal.pone.0205335)

107. Koriat A. 2012 When are two heads better than one and why? *Science* **336**, 360–362. (doi:10.1126/science.1216549)

108. Minsky M. 1988 *Society of mind*. New York, NY: Simon and Schuster.

109. Ainslie G. 2014 Selfish goals must compete for the common currency of reward. *Behav. Brain Sci.* **37**, 135–136. (doi:10.1017/S0140525X13001933)

110. Levin M. 2019 The computational boundary of a 'Self': developmental bioelectricity drives multicellularity and scale-free cognition. *Front. Psychol.* **10**, 2688. (doi:10.3389/fpsyg.2019.02688)

111. Graziano MSA. 2017 The attention schema theory: a foundation for engineering artificial consciousness. *Front. Robot. AI* **4**, 60. (doi:10.3389/frobt.2017.00060)

112. Thomas V, Pondard J, Bengio E, Sarfati M, Beaudoin P, Meurs MJ, Pineau J, Precup D, Bengio Y. 2017 Independently controllable factors. (https://arxiv.org/abs/1708.01289)

113. Thomas V, Bengio E, Fedus W, Pondard J, Beaudoin P, Larochelle H, Pineau J, Precup D, Bengio Y. 2018 Disentangling the independently controllable factors of variation by interacting with the world. (https://arxiv.org/abs/1802.09484)

114. Ismael J. 2016 *How physics makes us free*. Oxford, UK: Oxford University Press.

115. Ismael J. 2023 Reflections on the asymmetry of causation. *Interface Focus* **13**, 20220081. (doi:10.1098/rsfs.2022.0081)

116. Dennett D. 2017 *From bacteria to Bach and back: the evolution of minds*, 1st edn. New York, NY: W. W. Norton & Company.