

Concept Paper

Science-Driven Societal Transformation, Part I: Worldview

John C. Boik ^{1,†} 

¹ Courtesy Faculty, Environmental Sciences Graduate Program, Oregon State University; boikj@oregonstate.edu

† Principled Societies Project; <https://principledsocietiesproject.org/>

Version August 4, 2020 submitted to Sustainability

Abstract: Humanity faces serious social and environmental problems, including climate change and biodiversity loss. Risks are increasing and conditions deteriorating. Increasingly, scientists, global policy experts, and the general public conclude that incremental approaches are insufficient and transformative change is needed across all sectors of society. However, the meaning of transformation is still unsettled in the literature, as is the proper role of science in fostering it. This paper is the first in a three-part series that adds to the discussion by proposing a novel science-driven research-and-development program aimed at societal transformation. More than a proposal, it offers a perspective and conceptual framework from which societal transformation might be approached and understood. While acknowledging the necessity of reform to existing societal systems (e.g., governance, economic, and financial systems), the focus of the series is on transformation understood as systems change or systems migration—the de novo development of and migration to new societal systems. The series provides definitions, aims, reasoning, worldview, and a theory of change, and discusses fitness metrics and design principles for new systems. This first paper proposes a worldview built using ideas from evolutionary biology, complex systems science, cognitive sciences, and information theory that is intended to serve as the foundation for the R&D program.

Keywords: societal transformation; systems change; sustainability; complex systems; societal cognition; climate change; biodiversity loss; active inference; free energy principle; self-organized criticality; SAILS)

JEL Classification: B50, I3, O1, O30, P20, P41, P50, Q01, Q54, Q57

1. Introduction

In 1992, over 1,500 scientists, including 99 Nobel laureates, signed a document titled “World Scientists’ Warning to Humanity” (full text in supplement of [1]). The first sentence reads “Human beings and the natural world are on a collision course.” Unfortunately, nearly 30 years later, that statement is still true. Only one of the issues raised, ozone depletion, has been successfully addressed. Other issues, including climate change, biodiversity loss, aquifer depletion, and soil degradation, have grown more severe. A second urgent warning was published in 2017, currently signed by more than 21,000 scientists [1]. A third urgent warning, specific to climate change, was published in 2019 and signed by more than 13,000 scientists [2]. Two more scientists’ warnings to humanity, one on insect extinctions and the other on the freshwater biodiversity crisis, were published in 2020 [3,4].

To give some idea of risk, one rough estimate is that climate change could cause between 300 million to 3 billion additional premature deaths over the next one to two centuries (1.5 million to 30 million deaths annually, on average), even if the rise in global temperature is limited to 2 °C [5]. The higher estimate amounts to roughly 30 percent of an otherwise expected future world population. In

context of current events, the annual impacts of climate change could far exceed that of the Covid-19 pandemic, with associated social, economic, and environmental strains lasting for centuries.

Due to the lack of progress and limited time remaining to prevent severe damage and death, a growing portion of the global policymaking and scientific worlds, and general public, are calling for rapid transformational change [6–17]. That dramatic and rapid change may now be necessary implies a failure by societies to protect human health and the environment. Reasonable measures to mitigate risks have long been available, which suggests that the failure is not technical. Nor does it appear to be due to fatal (or near-fatal) flaws in the human psyche. Humans can and do make poor decisions, obviously, both as individuals and as groups. But as will be discussed, humans are also highly social animals that tend to value fairness and seek cooperation. Moreover, as individuals and as groups, humans can be and often are capable problem solvers.

Instead, failure to protect human health and the environment can reasonably be attributed to dysfunctional or maladaptive societal systems (e.g., economic, financial, legal, education, and governance systems), potentially involving policies, rules, laws, institutions, norms, priorities, power distributions, motivations, and core designs. These are the systems through which societies address collective problems, and they are implicated if the problem-solving process is thwarted or the results are flawed.

To the degree that societal systems are dysfunctional or maladaptive, transformation presents an attractive path forward. The meaning of societal transformation and the role of science in fostering it have, however, only recently become topics of discussion in the scientific literature. This paper is the first in a series of three that as a group adds to existing work by proposing a practical, affordable research-and-development program aimed at transformation [18,19]. More than just a proposal, it offers a perspective and conceptual framework from which societal transformation might be approached and understood. While acknowledging that reform of existing systems is important and necessary, especially in the short term (over the next few decades), the focus of the series is instead on transformation realized as systems change or systems migration, important in the long term (over the next several decades and beyond). That is, the focus is on the *de novo* development of and migration to new societal systems. The series provides definitions, aims, reasoning, worldview, and a theory of change. As well, it discusses organizing principles, fitness metrics, and design principles for new societal systems.

Societal systems refers here to nearly all aspects of societal organization (beyond the family level) that involve collective learning, decision making, and adaptation. Thus, the definition covers nearly every system of societal organization, including the six overarching ones: economic-financial-monetary; governance-political-electoral; legal-justice-criminal; public health-medical; census-analytical-forecasting; and education-research-science. For convenience, I refer to these as economic, governance, legal, health, analytical, and education systems. *Societal systems* also includes their associated institutions, rules, and policies. And it includes the coevolving social-cultural components (beliefs, values, norms, and worldviews) that shape and are shaped by systems [20]. Existing systems (those in use by societies today) are called *native systems*.

Society refers here to any sizable scale of social organization, from local community up through city, region, nation, and world. *Local community*, or *community* for short, which is a focus of the R&D program, refers to a fraction of a local population, held together by intent, social bonds, common interests, and shared concerns. A local community can be but is not necessarily (or usually) a formal political body, such as a city.

The definition of societal systems in terms of learning, decision making, and adaptation reflects three underlying propositions. First, a society of any scale can be viewed as a superorganism, composed of individuals (and groups, institutions, etc.) who coordinate behavior to some degree. Second, a society's complete set of systems can be viewed as the cognitive architecture by which it senses its world, processes information, directs attention, remembers, communicates, explains, learns, anticipates, decides, and orchestrates action. Third, a society's efforts to learn, decide, and adapt can be viewed as

being driven by an intrinsic purpose, which is to achieve and sustainably maintain vitality. The claim is that these propositions can be useful as guides to research, in generating testable hypotheses, and ultimately, in defining and assessing the fitness of alternative societal system designs.

Learning, decision making, and adaptation are, essentially, informational processes. Viewing them as such, and viewing societies as cognitive superorganisms, opens the door to modern perspectives and the use of analytical tools from fields such as information theory, computer science, cognitive science, complex systems science, and evolutionary biology. Advances in these and other fields allow, perhaps for the first time, the design of societal systems to fall under the purview of science.

The vitality of a society is broadly defined here. As will be discussed, a society is deeply integrated with its physical environment, and is connected with other societies and their environments through multiple, overlapping layers of organization and fields of influence, all the way up to the biosphere. A society's intrinsic purpose, as viewed here, is to achieve and maintain vitality within itself and in light of these interrelationships. Its purpose is to serve the common good, if you will.

The R&D program seeks answers to two key questions. First, out of all conceivable designs for societal systems, which ones might be among the best at demonstrably serving the common good? Second, by which viable strategies might new designs be tested, implemented, and monitored so as to be capable of positively impacting societies and environments on a near-global scale within a reasonable period of time?

The hypotheses underlying these questions are that new societal systems can be designed in a science-driven process to be fit for purpose; that defensible definitions of fitness can be constructed; that fitness can be theoretically and empirically assessed and compared across designs; that new system designs can be practically and viably implemented; and that new systems designed to be highly fit will be more effective than native ones at serving the common good.

This first paper in the series describes a worldview intended to serve as the foundation for the R&D program. A first-principles approach is taken, signifying the importance of worldview. Subsequent papers in the series use the worldview in discussing fitness metrics, system design principles, and a strategy for change.

To make initial progress, it is not necessary that a large percentage of any population engage with or support the R&D program. Support from just a tiny fraction of a local and the global population is adequate. Nor is it desirable that all who choose to engage—whether scientists, other professionals, or laypersons—subscribe to the worldview exactly as presented here. Rather, the worldview is offered as a start. The science behind it is quickly evolving, the topics are deep, and maintaining a variety of perspectives is helpful. As the program progresses, the goal is to cultivate enough common ground among those involved to keep the program focused, effective, and on scientific footing.

For convenience, I call the proposed worldview the *purpose-fitness* worldview, reflecting its intended use. We cannot hope to usefully assess the fitness of a system design if we do not know a system's purpose. The purpose of a societal system derives from the understood purpose of a society, and the understood purpose of a society is a product of worldview. The functional ordering considered here is *worldview* → *purpose* → *fitness metrics* → *system design*. Fitness metrics are positioned as a precursor to system design, to suggest a kind of reverse engineering process. We might wish for certain target levels of social and environmental wellbeing, for example. What kinds of designs would help us reach those targets within a reasonable length of time?

2. Worldview

2.1. A New Era of Science and Technology

The proposed R&D program spans a period approximately 50 years into the future (with early adopter communities seeing relatively rapid benefits). To stay relevant and maintain progress over this extensive horizon, the program must anticipate the directions of scientific advancement and operate

on that forward edge. Given that science and technology have been expanding exponentially fast, this is no small challenge.

Taking a bird's-eye view of recent history, science and technology began a transition to a new era starting roughly in the 1950s. That transition is not yet complete. The new era of science is qualitatively and quantitatively different from what came before, not just due to the emergence of phenomena like big data, artificial intelligence, robotics, the Internet, nanoparticles, and quantum computing, the latter of which is just around the corner. Differences go deeper. Relative to where science has been, and especially to where much of the general public still is, the new scientific era represents not just new technologies and methods, but a new worldview. This goes beyond what might be called an ecocentric (environmentalism) worldview and its extensions [21,22]. The new worldview emphasizes the cognitive and information processing characteristics of nested human societies and entwined ecologies, as well as the role of flexible self-identity in human cognition.

The power of worldviews should not be underestimated. They shape our cultural norms, behaviors, policies, institutions, and systems [23–28]. Worldviews help us make sense of the world and ourselves, including personal and shared identity and purpose. They color how we assign value and cause, who we view as enemies, and how and with whom we cooperate. The emerging scientific worldview will change us, fundamentally. It will alter our conception of who, what, and why we are, what we are doing, and where we are headed.

Consider recent speculations by Frank Wilczek, theoretical physicist and Nobel laureate, on where he thinks physics might be in 100 years [29]. Many of his musings are specific to quantum mechanics, but others touch on broader issues. A few highlights are given below.

1. The laws of physics will be reinterpreted as statements about information and its transformation.
2. Biological memory, cognitive processing, motivation, and emotion will be understood at the molecular level.
3. Calculation will increasingly replace experimentation in the design of products and services, leading to new opportunities for creativity.
4. Fault-tolerant, self-repairing computers will be developed. So will self-assembling, self-reproducing, and autonomously creative machines. Bootstrap engineering will become common, where machines build other sophisticated machines.
5. A substantial fraction of the sun's energy that the Earth receives will be captured for human use.
6. Quantum computers will become real and useful.
7. Empowered by these developments, the web of intelligence will expand faster than the web of human settlements. As it does, we will know much more about, and have vastly greater power over, the physical world. Individuals will expand their sense of identity and become more sensitive to gaps in knowledge and alternative interpretations. Collective ambition will expand. Along the way, a certain kind of humility will emerge that reflects not so much modesty, as largeness of vision.

Information, referenced in item #1, refers to a formal concept described in information theory and its extensions. It is deeply related to thermodynamic entropy, complexity, computation, and probability. Wilczek mentions that items #1 and #2 together close a circle of ideas regarding information and cognition. He speculates that mind will be understood as more matter-like and matter will be understood as more mind-like. A common language to link mind and matter, and many other seemingly disparate topics, is afforded by information theory [30–33]. Already, item #1 is well on its way to becoming true.

Reading between the lines, and taking some poetic license, the themes underlying Wilczek's comments are not just about computation, but about inference and cognition, not just in humans, but in intelligent systems. The implication is not just that humans will build advanced machines, but that we will understand advanced machines (and organisms, societies, and ecologies) as complex adaptive systems. The themes are not just about intelligence, but about distributed intelligence. Not just about

expanded human knowledge, but about expanded self-identify, and ultimately, greater wisdom. In short, Wilczek appears to be describing a turn toward holism and extended awareness and intelligence.

What Wilczek does not mention, perhaps because the paper's topic is physics, not sociology, is the possibility that societies will use their expanding knowledge, capabilities, and power to reorganize—that is, to transform. Why would societies not do so, given that: a) they can; b) collective wellbeing could dramatically improve; and c) failure to transform could lead to catastrophe or limit our ability to recover if catastrophe occurs. Finally, the impetus to explore, learn, and transform is, as we shall see, nearly irresistible.

I frame the emerging scientific worldview in a way not inconsistent with Wilczek's speculations. Some aspects of this framing are based on recent science, taking shape in just the last decade, even the last few years. Thus, I offer a snapshot of a work in progress. For any topic touched upon, multiple variations on a theme exist in the literature, with differences among them ranging from subtle to patent. Yet the framing offered provides a starting point, a basis for the R&D program to move forward.

Before summarizing assertions of the purpose-fitness worldview, it will help to clarify that *information* in this series of papers, unless mentioned with respect to information theory, means quality information. Information quality is a function of volume, completeness, timeliness, accuracy, reliability, relevance, and other factors [34]. Thus, *information* means signals, data, or communications that are useful in understanding some aspect of reality, or in predicting or inferring a past, current, or future state of the world. Rovelli coins the term *meaningful information* [35], based on the work of Kolchinsky and Wolpert [36].

Not surprisingly, addressing information quality means addressing foundational aspects of societal system design. In particular, it means addressing the motivations that a societal system engenders in individuals and groups to distort, corrupt, or otherwise degrade, obscure, or ignore information. More generally, it means addressing the motivations that a societal system engenders for prosocial behaviors (e.g., cooperation, honesty, and healthy communication) versus antisocial ones (e.g., selfishness, dishonesty, cheating, and dysfunctional communication). *Cheating*, sometimes called *free-riding* in the literature, means that an individual does not cooperate but benefits from the cooperation of others. Thus, the assessed fitness of new societal system designs will depend in part on design-engendered motivations for pro-social behavior.

That said, it would be unwise to encourage or expect unnatural behaviors (for example, excessive altruism). A design goal of the R&D program is to create societal cognitive architectures that facilitate health and wellbeing, taking into account the complicated nature of humans, societies, and environments. By analogy, a healthy nervous system serves as the cognitive architecture for a healthy person. It allows a person to learn, decide, and act as appropriate to internal and external conditions. As part of this, it facilitates health and repair in the body, including communication and cooperation between cells and tissues (see for example, [37]).

Short of catastrophe, human computational capacity is likely to expand over coming decades [38]. The advent of quantum computing is just one aspect. As computational power increases and technology advances, so too will our capacity to design, test, and implement new societal system designs, including development of new information processing tools for data collection, modeling, forecasting, and communications. However, existing technologies are sufficient for the R&D program to begin and move forward.

2.2. Assertions of the Purpose-Fitness Worldview

This section describes a set of assertions of the purpose-fitness worldview, discussed in non-mathematical terms and without reference to specific computational approaches. These form the background for a slightly more technical treatment in subsequent sections.

2.2.1. Evolution of complex systems

The biotic and abiotic world we see today can be understood as an immense collection of interacting, partially overlapping and hierarchical, partially self-similar (fractal) dynamic systems that evolved and continue to evolve in a way that is consistent with known physical laws. An ancient ancestor of all is the hot, dense, energetic plasma that existed moments after the Big Bang. While the exact evolution of systems from that early beginning to today is poorly understood, certain statistical regularities, organizing principles, and patterns of development are discernible by examining physical systems across domains, scales, and individual instances. Understanding these statistical regularities, organizing principles, and patterns of development and dynamics is key to developing high-quality societal systems. It allows us to see, for example, what a forest, amoeba, and financial system have in common, and to design and act accordingly.

Most (if not all) important natural systems can be categorized as *complex*, which implies that they contain a large number of interacting parts (many degrees of freedom) and that the parts display a substantial degree of correlated behavior. A large subset of complex systems self-organize, and may merge and reorganize, due to local interactions among parts. This can result in new, emergent forms, dynamics, characteristics, and/or capabilities, which are more than the aggregation of the parts, or which cannot be decomposed into the parts. Emergence brings forth something novel and nontrivial. The primary driver of self-organization can be described using the second law of thermodynamics (i.e., as being related to entropy) or, more abstractly and perhaps more usefully, as being related to information (as in information theory), and hence also to inference and computation. Information theory offers an expanded, probabilistic definition of entropy as uncertainty. See [39–41] for reviews on complex systems from a thermodynamic and information theory perspective.

2.2.2. Living organisms

A human society can be viewed as a superorganism [42]. As such, it shares important commonalities with all other living organisms—plant and animal, unicellular and multicellular. All are instances of *complex adaptive systems*, which are self-organizing systems.¹ Hereafter, reference to organisms implies also reference to superorganisms, including human societies, large and small.

Every organism exists under thermodynamic conditions far from equilibrium; if it is at thermodynamic equilibrium with its environment, it is dead. An organism must actively maintain a disequilibrium across its (permeable) boundary and maintain homeostasis within that boundary. This effort requires energy, which the organism obtains from the environment in the form of sunlight or chemicals (e.g., food). An organism is delicate relative to the highly uncertain, complex, and dynamic world it inhabits, and so is subject to numerous kinds of insults and damage, some of which can be fatal. It must act to avoid damage, and repair itself when damage occurs. Further, it must make decisions and take actions under uncertainty, constrained by energy, time, and ability [43,44].

An organism survives by being recursively self-maintaining, which means that it not only performs self-maintenance, it has some agency in how it does so [45,46]. It does not just reflexively react to environmental stimuli, like a simple machine. It is not just homeostatic, but allostatic, which means that it achieves homeostasis using agency and anticipation [47]. Every organism is cognitive and adaptive [46,48–51], even bacteria and plants [52–55]. *Cognition*, as used in this series, does not necessarily imply a central nervous system. Every organism is cognitive in the sense that it displays capacities typically associated with human cognition, such as sensing, learning, problem solving, memory storage and recall, anticipation, and attention.

Every organism has the capacity to evaluate its states as they pertain to its continued vitality (thrivability, viability), and to select actions based on those graded evaluations. Relatively simple

¹ Some systems, such as viruses, are edge cases of life.

organisms have relatively simple, perhaps even minimal cognitive capacity. More complex organisms have greater capacity. Importantly, the states that an organism evaluates include its own autonomy and cognition. Autonomy, aided by cognition, empowers an organism to use its own capacities to manage itself (via allostasis) and its interactions with the world (via behavior) in order to remain vital [46].

The products of an organism include itself (autopoiesis [56–58]), and typically, wastes, other chemicals, and heat. An organism also produces effects in the world, such as through action, including communication. Typically, an organism can produce or pass on *persistent information*, which is defined here as genetic, epigenetic, cultural, and/or symbolic information that impacts future generations.² Persistent information can include changes to the environment that are informational (a simple example is a footpath worn into a trail that informs future travelers where to walk). This series considers actions and events on four temporal scales, all of which can impact persistent information: mechanistic (real-time), ontogenic (developmental, having to do with individuals), phylogenic (inter-generational, having to do with groups), and evolutionary (having to do with species, for example via genetic evolution, speciation, or ecosystem transitions) [59].

Cognition in humans, and in many if not most other organisms, can be described by the 4E's: embodied (occurring in the body, not limited to a brain and central nervous system), enacted (occurring through actions on the environment), embedded (occurring in relation to an external environment), and extended (into the environment) [60]. Extended cognition occurs in humans, for example, through use of computers, through environmental markers, and by associating memories with physical objects (which improves later recall).

Emotions play critical roles in human cognition. They appear to be cognitive states themselves, resulting from the gathering and processing of information. That is, they are not reactions from primal, innately programmed neural systems, but part of the normal cognitive process [61,62], much like sensory cognition. Emotions may mediate the exercise of constitutive and behavioral autonomy [46], and play a role in evaluating options during decision making [63]. Neither emotion nor consciousness is uniquely human. Emotions occur in a wide range of species [64], and all organisms may be considered conscious to a degree [65]. The latter follows from the notion of graded consciousness, as described in Integrated Information Theory [66].

In conclusion, organisms possess natural agency; an intrinsic teleology, or purpose; sense-making capability; and causal powers [45,46,48]. The intrinsic purpose of an organism is to achieve and maintain vitality—sustainable flourishing of self—which occurs by sensing and evaluating states and choosing and implementing appropriate actions based on anticipation [67]. Anticipation necessarily implies some kind of internal modeling of the world.

We can call *wise* those choices and actions that adequately support an organism's continued vitality. Thus, the intrinsic purpose of an organism can be restated as the directive to express wisdom. The more complex the organism, the more complex the situational elements and the wider the spatial and temporal scales that it can and must consider.

2.2.3. Cooperation

That the intrinsic purpose of an organism is to maintain vitality does not mean that the individuals of a population always or even usually act antisocially, or selfishly. In all populations there is an ongoing tension between prosocial (e.g., cooperative) and antisocial (e.g., cheating) behaviors. This tension can occur over multiple levels of organization and temporal scales.

That nature tends to favor cooperation over cheating is evident from the very existence of multicellular animals, within which cooperation between cells is foundational [68,69]. Furthermore,

² Cultural information can include tools, artifacts, art, norms, and governance systems. Symbolic information can include language and mathematics.

individuals in most, if not all, species display some degree of intra- and/or inter-species cooperation [69–71], and cooperation between species helps to stabilize diverse ecosystems [72]. Trees, for example, depend on cooperative, symbiotic relationships with soil fungi (e.g., mycorrhizal associations) [73]. Human societies depend on a wide assortment of internal and external cooperative relationships. Advantages of cooperation include acquisition of otherwise inaccessible resources, more efficient acquisition of resources, and reduced uncertainty [71].

Humans are highly social animals that are inclined to cooperate when that is a viable option [74,75]. Cooperation serves the individual's intrinsic purpose. It's a path of least energy expenditure to improve security, increase opportunities, and solve shared problems. It stands to reason that the degree to which humans actually cooperate, and opportunities for cooperation, are strongly influenced by the design and function of societal systems. Studies in game theory and other fields identify key factors—rules, information processes, organizational structures, etc.—that promote or inhibit cooperation [69,75–79].

For example, a basic requirement for cooperation is communication. Generally speaking, limited communication supports limited cooperation, whereas rich communication makes rich cooperation possible. Cells in the human body, for example, richly communicate with one another via hormones, small molecules, and other types of signals. Insects in a colony also richly communicate [80].³ Other factors that can promote human cooperation include transparency, the capacity of a group to penalize cheaters, inclusiveness of the group, and fairness. Knowledge of these factors can be used to evaluate system designs as to their propensity to promote cooperation.

Quite likely, new system designs would only slightly resemble native ones. Native systems were not designed through a scientific process to promote cooperation or, more generally, to function as a cognitive architecture in achieving and maintaining societal vitality. That is, they were not designed to be fit for the intrinsic purpose described here. One can argue that modern capitalism, for example, generates a shallow form of cooperation relative to what humans are capable of. At the investor and C-suite level, the primary motivation is typically (short-term) profit, which is often at odds with societal and environmental wellbeing [81]. Thus, history is replete with examples of corporate deception and wrongdoing, spanning nearly all industries [82–88]. As one example, Exxon reportedly knew about climate change more than 40 years ago, yet funded disinformation campaigns to discourage public action [89,90]. Such deceptions and lack of cooperation (with the public) are more usefully viewed as symptoms of faulty system designs, rather than as isolated incidents that stem from the poor choices of a few individuals (i.e., the bad-apple theory).

2.2.4. Societal cognition

The very definition of a society as a group of individuals that coordinate behavior implies that a society is cognitive. For example, coordination implies that a society must anticipate the future, given a set of potential actions. It must make and evaluate plans. Cooperation itself is a cognitive act that requires communication and anticipation [91]. Anticipation requires models or explanations about how the world works. Thus, a society is defined by its cognitive ability in coordinating action. Cooperation and cognition may have coevolved together in humans [92].

Obviously, individual humans are cognitive. But the description here of a cognitive society goes beyond the sum of the parts. The claim is that societal cognition is an emergent characteristic, not reducible to the cognition of individuals. Examples can be seen in insect colonies, non-human primate societies, and human societies [80,93–97].

Societal cognition also goes beyond what has been called “wisdom of the crowd.” Early studies on crowd wisdom typically focused on the accuracy of averaged predictions to numerical questions

³ This is not to imply that hyper-connected systems, where every component communicates strongly with every other, is beneficial. Few systems like this exist in nature.

Core Need	Example Functionings
Affection	sharing love and friendship, having self-esteem
Creation	imagining, expressing art, building
Freedom	exercising agency, speaking freely, pursuing interests
Identity	self-assertion, self-distinction, self-understanding, belonging
Participation	cooperating, helping, interacting, engaging
Understanding	being curious, learning, communicating, exploring, asking
Leisure (Idleness)	sleeping, playing, exercising, resting, relaxing, rejuvenating
Protection	staying warm, dry, safe, and secure
Subsistence	eating food, drinking water, being healthy, receiving medical care

Table 1. Core human needs

[98]. More recent studies suggest that crowds can perform sophisticated data analysis, in particular, dimension reduction (and denoising) [99], and that a group can solve some challenging (sentence reconstruction) problems beyond the ability of any individual in the group [100]. But these are still relatively simple examples of group cognition. If we understand societal systems as constituting cognitive architectures, then we understand history as being replete with examples of societal cognition, in some cases functional, in others dysfunctional.

2.2.5. Core human needs

The long path of evolution over countless ancestral species has inserted core needs deep into human biology. All serve to promote continued vitality in one way or another. All promote prosocial behavior, at least under some circumstances. For example, different needs compel humans to give and receive love, care for children, engage with community, and cooperate with others in securing necessary resources. Emotions are involved. By and large, we feel good when helping others. We want to help others and be of use.

Investigators have proposed categories of core human needs. Max-Neef recognizes nine, listed in Table 1 [101]. Abraham Maslow proposes a somewhat different categorization [102]. Closely related, Shalom Schwartz proposes a list of ten universal human values [103].

Core needs, almost by definition, demand our attention. Our vitality suffers to the degree that we ignore them. Thus, *problems that matter* are the ones that impact core needs.⁴ These are the problems worthy of our attention and energy, and that hold meaning to us as societies. The distinction is important because not all problems that a society addresses are problems that matter. For example, excessive consumerism implies making and selling products that people do not need [105]. It's driven by advertising and related efforts to convince potential customers that they do need the products. This suggests that millions of employees worldwide are spending their precious days attending to problems that matter little to a society's vitality. In fact, solving this set of problems—convincing consumers to buy more, and supplying them with products they don't need—is destructive to a society and the environment.

Later we will consider human needs as pointing to *essential variables* that require attention in societal cognition.

2.2.6. Species–niche co-evolution and self-identity

US culture celebrates rugged individualism [106]. Capitalism too engenders a norm of individualism [107]. But individualism is largely incongruent with reality. Organisms, especially social organisms like humans, are independent entities in only a limited sense. A human body, for

⁴ The solutions to some problems are a process, not a specific set of actions. And some problems are not fully solvable, even if impacts can be mitigated. Moreover, some problems are more properly called challenges or opportunities. The term *solving problems that matter* [104] implies this larger view.

example, consists of an ecosystem of bacteria, viruses, and human cells, all of which interact and are necessary for health and evolution [108,109]. What we think of as a human is not entirely or even mostly human—a body contains roughly the same number of bacterial and human cells [110]. Among other things, their interplay affects our ability to learn and think, as well as our moods and social interactions [111,112].

Further, we and other organisms are not separate from our physical environments. Bodies are materially and energetically dependent on the viability of larger systems. These include ecosystems, climate systems, hydrologic systems, and nitrogen and carbon cycles. Further, bodies are permeable. Nitrogen, hydrologic, and other cycles flow through us.

The separation between self and world also blurs across time. In a process called *niche construction*, a species co-evolves with its niche—its surrounding abiotic and biotic environment, including sociocultural aspects [113–117]. Species and niche are causal factors in the evolution of each other (circular causation). As an example, environmental conditions can affect rates of genetic evolution [118,119].

Organisms become a model of niche over time and niche reflects organisms. Skeletons are a model of gravity, for example. Lungs are a model of oxygen dynamics and rate-driven diffusion processes. Further, the actions of one species can create niche opportunities for others. Phytoplankton in the ocean create oxygen that animals on land can breathe. Dam construction by beavers creates local niches in which other animals and plants can thrive.

The separation between self and other also blurs across social organization. Human culture and language are collective cognitive processes and artifacts of niche construction [120–123]. Personhood, or self-identity, spanning such characteristics as personality, preferences, identity, and worldview, is a social phenomenon molded by culture [124]. As Bentley et al. describe it, “the very substance of selfhood is shaped by the groups to which people feel they belong” [125]. People form preferences based in part on expectations about what others think and about what others might view as acceptable and appropriate [126]. What others view as acceptable changes over time as culture changes.

The concept of niche construction is one element of a larger turn toward an extended evolutionary synthesis that overcomes limitations in the gene-centric modern synthesis of evolutionary biology, established in the 1940s [127]. As noted, physical and cultural environments act on individuals, and, reciprocally, individuals act on physical and cultural environments. This plays out over mechanistic, ontogenic, phylogenic, and evolutionary time. But beyond this, the extended synthesis casts agency as an active force in evolution. Individuals, groups, and societies choose behaviors, based on predilections and affordances offered by niche. In this sense, adaptation and evolution are creative and opportunistic, not just random or reactive. The goals, intentions, and desires of individuals in a population impact the evolution of niche, which in turn impacts the evolution of the population.

All this raises the question, what is an individual? It turns out that the question is not easily answered. Any proposed individual exists within, is affected by, and affects its environment. In looking at the world, how does one decide that a dynamic pattern is somehow a distinct unit, something separate from its background? Krakauer et al. suggest that information theory can help [128]. They define an individual as a process that propagates information from the past into the future, which has the effect of reducing uncertainty. Their approach, Information Theory of Individuality (ITI), allows for a continuum of individuality, not just a binary categorization, and allows for nested individuality. Individuals can be composed of aggregate parts that are themselves individuals. ITI allows for individuality to arise at any level of organization.

A community can be understood as an individual that is one part of larger nation, also an individual. Nations can be understood as parts of a larger civilization, also an individual. Thus, the human niche spans local to global. We communicate with and care about others in our neighborhood and around the world, and we understand that what happens in far-away areas can affect local conditions, such as through trade and the spread of ideas, disease, and pollution. Further, we compute

and share information globally, over the Internet for example, and based on information transferred via satellites.

In all these ways, we are intimate with our greater world. We are individuals, but of the nested, overlapping kind—individual cells, bodies, groups, communities, ecologies, nations, and civilization. We are not separate from our world in any absolute sense, and there is no privileged level or scale at which individuality is viewed. Nor are we passive bystanders in the unfolding story of evolution. The desires and actions of individuals, be they societies or otherwise, generate waves of effects that radiate outward through the biosphere and reflect backward to impact the source. Indeed, the reach of a society now extends beyond Earth, and in time our actions in space will likely reflect backward from there too.

2.2.7. Societal intrinsic purpose

In summary, the intrinsic purpose of a society is to achieve and maintain vitality, which is accomplished through cognition and cooperation. Cooperation is akin to self-organization (into larger or wider scales of influence). Vitality applies to a society's members, its local environment, and the larger overlapping spheres of societies and environments to which it belongs and from which its actions reflect backward. The self that must be kept vital is the extended self.

It follows that the intrinsic purpose of a societal system (e.g., a financial system) is to serve the intrinsic purpose of a society. The purpose of each system and the set of all societal systems that a society employs—its cognitive architecture—is to achieve and maintain vitality of the extended self—the extended superorganism.

By adopting the notion of extended and flexible self-identity we become a more complete and accurate reflection of our immense world. We become a better model of our world's regularities, affordances, and dangers. At the same time, in the face of that immensity, we become more aware of our own ignorance and limitations. Perhaps this is what Wilczek meant by his speculation of a future emergent humility and largeness of vision.

3. Organizing Principles

Cognition, societal or otherwise, implies sensing, learning, problem solving, memory storage, memory recall, anticipation, and attention. Moreover it implies organization, and organization implies cooperation or coordination of aggregate parts. Brain tissue, for example, is a highly organized aggregate of cooperating cells that displays a dynamic structure. This section examines how complexity relates to structure and cognition, and identifies some organizing principles that tie the characteristics of cognition together. These principles can serve as guides—technical and/or philosophical, depending on circumstances—when designing new societal systems.

Intuitively, simple organisms (e.g., bacteria) conduct simple forms cognition. Organisms that are more complex (e.g., a human or human society) can conduct more sophisticated forms and so can address a larger range of problems over wider timescales. Thus, the complexity of an adaptive system must in some way relate to the difficulty class of problems that it can solve. And it does, at least by some definitions of complexity. There is, however, no single, accepted definition. Many have been proposed, and each may be more or less appropriate, depending on the research question and setting [129–135].

For our purposes, we can consider two meanings of complexity: as a measure of flexibility and as a measure of problem-solving capacity. The former stems from probability, statistics, and information theory, where complexity can refer to the flexibility of a model, or similarly, to the amount of extra information in one random variable relative to another. A model that is too complex (too flexible, too many parameters) can simply memorize training data and thus be incapable of generalizing what it learns to new settings (i.e., it overfits the data). The second meaning of complexity, as a measure of problem solving capacity, stems from the notion of a complex adaptive system as being both complex

(flexible, with many parts) and adaptive (capable of solving problems). Hereafter, context should make clear which meaning is implied.

Much has been written about the potential for complex adaptive systems to catastrophically fail. It is not necessarily true, however, that simpler systems are less prone to failure. Nor is it necessarily true that robustness and resilience decline as complexity increases. If complexity is a measure of problem-solving capacity, then to some degree the opposites are true. An organism or system “complexifies” over time when it faces challenges, opportunities, or environments that it’s ill equipped to handle [136–138]. With greater complexity come new options, additional flexibility, and expanded information processing. Four concepts help formalize these ideas: good regulators, requisite variety, self-organized criticality, and the free energy principle.

3.1. Good Regulators and Requisite Variety

A well-adapted organism is equipped to handle those aspects of its environment that it must handle. For example, a rabbit is equipped to run from predators. Its legs (and general morphology) allow for running, and as well its cognition allows for running at the appropriate moments. A rabbit evades its predator in part by anticipating how its predator will behave. This simple example demonstrates two key ideas. First, an organism has an internal model of its world that allows for making predictions. More accurately, over mechanistic, ontogenic, phylogenetic, and evolutionary time, the organism becomes a good model of those aspects of its environment that it must control. An organism’s variable world eventually presents it with a wide range of challenges. Over time, it comes to reflect a wider swath, or spectrum, of the regularities present in its world. In information theory terms, it comes to share more mutual information with its world. This notion is expressed in the good regulator theorem from cybernetics, which holds that every good regulator of a system must *be* (as opposed to *have*) a model of that system [139].

Second, to be a good regulator means that an organism must have available enough options, or flexibility, to control its environment. A rabbit has legs, muscles, and nerves that allow it to move quickly and with dexterity. Another way to say this is that an organism must be complex enough to control its environment. Yet at the same time, the organism should not be overly complex. Complexity is costly, and can interfere with an organism’s ability to generalize previously learned information to new settings. Thus, being a good regulator also means matching the complexity of that which must be controlled. This idea is captured in Ashby’s Law of requisite variety [140]. Here, variety refers to the number of tunable states in the system that is being controlled. The controller must have similar variety to that which it controls.

Organisms complexify in the face of challenges, but considerable decision-making power is typically retained by its subparts, rather than shifting to higher levels of organization. Local processing is faster and more efficient. For example, human immune cells act largely (but not completely) independent of the brain. In general, biological systems self-organize in a hierarchical fashion, such that at each level, processing occurs for information at that level and from levels just above and below. In this way, local levels are empowered (they retain flexibility and agency), but obtain some feedback from levels higher up and below.

3.2. Self-Organized Criticality

A core purpose of cognition is to anticipate and solve problems. But how does an organism solve problems, especially difficult ones?

In general, an organism solves problems by using a combination of exploitation and exploration [141]. Exploitation here refers to using available knowledge about a problem or situation to improve a solution. Exploration, in contrast, refers to discovering new knowledge. For example, a flock of birds might explore many neighborhoods in search of food. When it finds an area that shows potential, it might focus its attention in that location, exploiting local features to hone in on the best spot. Of course,

even better feeding areas might exist beyond the neighborhoods that were explored. To find them, the flock would have to conduct more exploration.

For many organisms and complex systems (and computational models in machine learning), the best strategy for solving difficult problems is to balance exploration with exploitation. In well-tuned systems, the optimal dynamic balance tends to have a particular characteristic. It places the system near critical thresholds. Coined by Bak et al. in 1987 [142], *self-organized criticality* describes the phenomenon of a system self-organizing such that it tends to maintain operation near a phase change threshold (a critical state, in physics). A system exhibiting criticality is on the verge of transition from one state to the next. Thus, its future trajectory is particularly sensitive to small perturbations or other inputs. Sometimes, just a small bump can set off a wave of events that can cascade through a system, leading to massive change. Typically in biological systems, such a cascade does not result in death, but rather in better solutions to difficult problems, or perhaps to changes that increase problem-solving capacity.

The existence of critical states in physical systems is not controversial. The idea that natural systems can self-regulate their structures and/or dynamics to maintain operations near critical states is more so. Many questions remain, but in the roughly 30 years since Bak's proposal, a large volume of work suggests that many biological systems, including the human brain, display self-organized criticality or something close to it [143–147].

Somewhat poetically, systems at criticality are said to operate “on the edge of chaos” where a balance is achieved between system stability and agility. Stability stems from the continuance or improvement of existing patterns, while agility stems from exploration and adoption of new patterns [148,149]. Thus, we can also refer to the exploration–exploitation tradeoff as the stability–agility tradeoff. Viewed another way, criticality occurs when the use of stored information is balanced with the use of new information—the old–new tradeoff. At one extreme, new information is blocked and the system is rigid and unable to adapt to changing conditions. At the other extreme, the flow of new information is an incoherent flood that swamps the dynamics of the system. Yet another way to think of the stability–agility tradeoff is as the constraint–complexity tradeoff, where constraint here reflects a lack of flexibility. Similarly, one can speak of the compressibility–expressivity tradeoff, where compressibility refers to a system that is optimally simple or compact, and expressivity refers to a system that is complex enough to adequately reflect a situation or process [150].

However we phrase it, the idea is that information processing, responsiveness, and problem-solving capacity are optimal when a system is on the edge of instability, near a critical state [151–153]. Moreover, long-range correlations (e.g., cooperation between distant, non-neighboring parts) increase near criticality [147]. At criticality, new information can enter and alter a system, sometimes producing dramatic changes, and yet the system is not so sensitive or receptive that it is overwhelmed to the point of collapse. The system remains optimally robust and resilient. As such, the benefits of criticality are not unlike the hoped-for benefits of democracy—enhanced cooperation and maximal learning and problem-solving capacity, where each voice can make a difference, sometimes even a big difference. Self-organized criticality can serve as a biological inspiration for new forms of deep democracy.

Criticality is a dynamic process in which structures reorganize as is appropriate to context and need. The human brain, for instance, is highly neuroplastic. It dynamically rearranges its anatomical connectivity and alters the strength of (synaptic) interactions between neurons [154]. As is common throughout nature, the brain is arranged in hierarchical layers, is modular, and demonstrates self-similar arrangements over spatial scales. These properties allow a brain, or any system, to achieve different degrees of stability-agility when solving different kinds of problems. Modularity aids computation [155]. Indeed, even different phases of a single problem-solving process might benefit from different associative and amplification dynamics [96]. As we will see, societal cognition mimics at a large scale what occurs in the brain at a small scale (see also [19]).

3.3. Free Energy Principle

How does an organism balance exploitation with exploration in problem solving? For that matter, how does it know which problems to solve, or what information to pay attention to? And how does it plan? These are not trivial questions, and also arise outside of biology. For example, much of modern artificial intelligence relies on reinforcement learning, an approach in which an intelligent agent searches for policies—usually understood as a dictionary of context-dependent actions—by which it can best achieve goals. Typically, the analyst (a force outside the agent) must program the agent both with goals and a fitness function, the latter of which is used to evaluate trial policies. Moreover, the analyst must program the agent to balance exploitation with exploration, typically using a heuristic method [156]. But a living organism must, to some degree, set goals, evaluate policies, and solve problems without being externally programmed to do so. For example, an infant must act and learn. How does an organism accomplish these tasks? In looking for answers, a good place to start is by examining cognition in human adults.

The perceptual, cognitive, and executive processes by which learning and adaptation occur in individuals has received substantial research attention in recent years, in part because of potential applications to medicine and artificial intelligence. A promising framework is the free energy principle, pioneered by the neuroscientist Karl Friston [157]. It potentially explains how a system sets goals, evaluates actions and policies, balances exploitation with exploration, focuses attention, and more generally, how it self-organizes [158]. As applied to enactive cognition in biological systems, its corollary is called *active inference*. The technical details are substantial and I give only a superficial overview, focusing on some main points pertinent to a discussion on societal systems.

Our starting point is a generic organism or superorganism, which we know is allostatic, and which we will call an agent. The agent can exist only within a certain range of internal (e.g., physiologic) and external states, such that a healthy agent will revisit a certain neighborhood of (characteristic) states more often than others; mathematically, its dynamics are ergodic. The agent's intrinsic purpose, or goal, is to achieve and maintain vitality, which means it must achieve and maintain those characteristic states that confer it with vitality. To fulfill its goal, it must ingest energy, export wastes, avoid danger, make repairs, and in general act against the forces of entropy; if it does nothing it will reach equilibrium with its environment, which is death.

We cannot speak of the agent as being different from its environment unless it is independent from its environment in some sense. The sense we focus on is conditional independence due to what is called a *Markov blanket*, a statistical boundary that partitions states into internal and external. The blanket is a set of states that separates—or better yet, mediates the connection between—internal and external. The blanket consists of active and sensory states that arise from the agent's actuators/ effectors and sensory organs, respectively [157].

While the Markov blanket distinguishes an agent from its surroundings (i.e., it provides an identity), it also causes an immediate problem. The agent cannot know its world in an absolute sense. It can only poke at its world through action, and (imperfectly) sense the world's response. Internal and external influence each other through the blanket's active and sensory states. For example, a person cannot know the mind of another person, but can take actions (e.g., asking questions, watching behaviors) that lead to information on which a better model, or understanding, of the other can be constructed. That an agent is distinguished by a Markov blanket does not mean that it is disengaged from its world. On the contrary, it actively engages with its world via action.

Consistent with the good regulator theorem, the agent becomes a model of its world over time. For convenience, we will say that the agent has a model, even though the agent as a whole is a model. In actuality, the agent represents numerous models that can potentially overlap and/or compete, and that can span a range of temporal and spatial scales. The description here is primarily normative. An agent acts as if it has mathematical models and as if it is optimizing them, and these cause the agent to act as if it is purposeful. Nevertheless, the free energy principle is not inconsistent with what is known

about the physiology, function, and behavior of organisms [159–161]. Further, a more complete view casts active inference as a process theory (of development and change) [159,162].

The agent chooses actions and optimizes its internal model by minimizing surprise. Surprise occurs due to unexpected events. It is surprising when observations differ from beliefs (expectations). Because an agent exists in time, it more correctly minimizes expected (estimated long-term average) surprise, which takes into account the potential results of policies. Here, a policy is viewed as an explicit series of actions to be followed in the future. Not all surprises are equally important, however, and an agent must pay more attention to some and less to others.

Surprises that involve *essential variables* are, objectively, the most important. Essential variables are a minimal set of variables that determine a system's state (of vitality) and its development and trajectory. For a human, some essential variables reflect core physiological and psychological needs. Viewed at a high level, these constitute core human needs, discussed previously. Other essential variables reflect environmental and ecological needs, as per extended self-identity. See Lehmann et al. for a discussion of essential variables that relate to sustainability [163].

Making predictions and minimizing surprise (related to essential variables) are the agent's central activities. The human brain, for example, is massively predictive in nature [164]. Yet, a conscious agent, such as a human, might be largely (or completely) unaware that it is modeling its world or that it is minimizing surprise. Moreover, any specific agent might be more or less functional, even dysfunctional, in doing either.

3.3.1. Surprise

The term *surprise* has a formal, Bayesian definition. Without going far into mathematical details, the probability of an event in Bayesian statistics is given by Bayes's theorem:

$$\begin{aligned} \text{posterior} &= \frac{\text{prior} \times \text{likelihood}}{\text{evidence}} \\ p(\theta | D) &= \frac{p(\theta) p(D | \theta)}{p(D)}. \end{aligned} \quad (1)$$

This formula (informal version on top, symbolic below) describes how an agent updates its beliefs after seeing some data. In the data science setting, an analyst is typically interested in learning the distribution of an unknown parameter, call it θ (theta), after seeing some data, D .⁵ As an example, θ might be the mean body temperature of healthy humans (which should be about 97.8 °F). The data are typically samples from a population. The posterior distribution, what the analyst seeks, is the conditional distribution of θ , given the data. The *prior* is the distribution of θ before seeing the data, which we can also call a *belief*. *Likelihood* is the conditional distribution of the data given θ . Finally, *evidence* is the probability of the data, which is obtained by integrating prior \times likelihood over the full distribution of θ (i.e., integrating out θ , leaving just the probability of the data). Thus, evidence is sometimes called the marginal likelihood, as θ is marginalized out.

Intuitively, if we hold a belief (a prior) and then see some data, and the data reasonably match our belief, then our prior was of good quality and any update to it will be small; the posterior will be similar to the prior. If, on the other hand, our belief was of low quality, the update will be larger; the posterior will be substantially different from the prior.

Given a prior belief about θ , some data, and a likelihood function, and setting aside any computational challenges, one can plug values into Equation 1 to obtain a posterior distribution. That posterior will represent a kind of compromise between likelihood and prior. The denominator

⁵ The symbol $|$ in the bottom formula is read as "given," and $p()$ is read as "probability of," so for example, $p(D | \theta)$ is read as "probability of D given theta."

acts as a normalization factor. Another approach is to seek the prior that makes the data most probable. This is done by maximizing the evidence (the denominator in 1), which is an inference procedure called “maximum likelihood II” in the statistics literature.

Our agent does a form of maximum likelihood II. Given observations, it computes a belief about θ that best explains its world. It chooses a θ that (approximately) maximizes the evidence. The negative log evidence is called *surprise* in Bayesian statistics. Thus, maximizing the evidence means minimizing (Bayesian) surprise.

Because the agent seeks to minimize surprise, or maximize evidence, we can say that the agent is self-evidencing. Its original, fundamental belief is that it exists, which means it prefers to, or believes that it will, occupy a certain range of states that are conducive to its vitality. States of sickness or decay are very surprising by this definition, and an agent will act to avoid them.

3.3.2. Variational inference

The agent can’t minimize surprise (maximize evidence) directly because it can’t know the world directly. It knows only its own beliefs and the information conveyed by the Markov blanket. So instead, the agent approximately minimizes surprise by minimizing an upper bound on it. That bound is called *free energy*. The free energy principle is, in short, that systems will minimize free energy, a function that is a bound on surprise:

$$\begin{aligned} \text{free energy} &= \text{surprise} + \text{divergence} \\ &= \text{complexity} - \text{accuracy} . \end{aligned} \quad (2)$$

More accurately, free energy is a functional. The right hand side (surprise, divergence, complexity, accuracy) are functions that take functions—parameterized distributions, including approximate posterior beliefs—as arguments [165]. Divergence is the distance between the actual posterior distribution and a simpler approximation to it (called a variational approximation). Because divergence is a distance measure, it’s always equal to or greater than zero. That’s why free energy is an upper bound on surprise; surprise cannot be larger than free energy. By minimizing free energy, the agent minimizes surprise.

The second line in the equation shows that the agent can minimize the free energy functional by minimizing the complexity (flexibility) of its model and by maximizing accuracy. This is an instance of Occam’s razor. The simplest model that predicts observations with sufficient accuracy is usually the best one. Recall that an agent minimizes (future) expected surprise, and so also expected free energy. The latter can be divided into terms for epistemic value (learning) and extrinsic value (goal attainment) [166]. As such, free energy minimization naturally balances exploitation with exploration, without need for heuristics. Further, curiosity arises as part of the search for epistemic value.

Free energy minimization is consistent with Jayne’s maximum entropy principle, which is, loosely speaking, an information theory version of Occam’s razor [167,168]. It says that when predicting a probability distribution, the best choice is the one that is constrained no more than necessary to be consistent with known information or factors; the best choice is the one that keeps options open, the one that’s least committal. An agent that is least committal when preferences are equal is maximally flexible and has low barriers to change. Just a bump can, at times, cascade through its parts to produce new configurations. This is, of course, a description of criticality; free energy minimization is a path to self-organized criticality [169].

As noted, an agent acts to minimize surprise. Viewed differently, an agent acts to minimize uncertainty. Expected free energy is equal to expected cost (the divergence between predicted and preferred outcomes) plus expected entropy (ambiguity, or uncertainty) under predicted states. [162]. Minimization of free energy ensures that entropy (uncertainty) over sensations is bounded [157]. The imperative to reduce uncertainty is evident in the human experience. Uncertainty usually makes us uncomfortable. It causes stress. If uncertainty grows large enough in individuals or societies, they

reach a tipping point that can precede major changes. Arguably, this may be the situation today with respect to climate change and other pressing problems.

Active inference (free energy minimization in biological systems) might more aptly be called *enactive inference*, emphasizing that cognition occurs partly through action [170]. An agent minimizes free energy (and reduces uncertainty) in two ways. The first is perception—sensing the outcomes of actions and updating beliefs accordingly. Perception reduces divergence (Equation 2) and is akin to approximate Bayesian inference and the Bayesian brain hypothesis [171]. The second is action, which is a selective sampling of the world [169,172]. As mentioned, action can aim at goal attainment or epistemic gain.

3.3.3. Attention

An agent's intrinsic purpose is to achieve and maintain those characteristic states that convey vitality (i.e., to choose actions that produce low surprise, given evolutionary, cultural, developmental, and other kinds of expectations). To fulfill its purpose, the agent must decide and act, but is constrained in doing so. It is uncertain about its world and has limited time, energy, and ability to resolve that uncertainty. Moreover, much of the information available from the world does not directly pertain to the agent's purpose. Given these constraints, the agent must focus in on what is most important and ignore the rest. That focus is called attention.

In active inference, attention is the process by which prediction errors are weighted by their precision (inverse variance). Precision of incoming sensory data also plays a role. The agent learns the precision of predictions along with other model parameters. Thus, for example, if an agent expects to see a dog in the yard and instead sees a horse, but is uncertain about the dog prediction, then the precision-weighted surprise of seeing a horse will be small. If instead the agent is sure about the dog prediction, the precision-weighted surprise of seeing a horse will be large. If it's dark, and the precision of incoming data is low (i.e., the agent is unsure what it is looking at), precision-weighted surprise due to the presence of a horse or dog will be diminished, all else being equal.

Recall that humans are highly social and cognition is a social affair. Because attention (moderated by precision) is learned, it is learned in part, perhaps even in large part, through societal cognition. The norms, rules, and organization of a society, including its societal systems, greatly influence the precision placed on predictions, and thus also the importance and attention that individuals attribute to information. On the one hand, societal cognition and cooperation rest upon the capacity to develop shared patterns of attention. During human development, for example, we learn from our parents and culture about what is important, what is dangerous, and what can be safely ignored. Indeed a primary role of teachers, and more broadly, of culture, might be to direct attention to that which is most important [126].

On the other hand, shared attention can go awry, leading to maladaptive societal cognition. The attention of a society can be manipulated in a variety of ways, and those who wield financial, political, and other forms of power have more opportunity to manipulate societal attention than do others. As explored in the third paper in this series, power and societal attention are closely related. An example of societal manipulation is the Exxon case of apparent climate change deception, already mentioned. Exxon is not unique, of course. Many large corporations have motivation to manipulate public attention at the public's expense, and do so [173].

3.3.4. Self-similarity, cooperation, cognition, and communication

We have considered both organisms and superorganisms as agents. Obviously, these systems exist at two different scales. But they are not the only scales to consider, nor are they privileged ones. Agents occur at every spatial scale, if we adopt a more broad definition of an agent as an enactive adaptive system far from equilibrium, distinguished from its environment by a Markov blanket (or likewise, if we adopt the definition of an individual from Information Theory of Individuality [128]). Then all of nature can be seen as blankets within blankets (or individuals within individuals), nested

and potentially overlapping, spanning a range of temporal and spatial scales and complexities. A society is an independent agent at the kilometer scale. A person is an independent agent at the meter scale. A cell is an independent agent at the micrometer scale. And so on down to the molecular and atomic levels [59]. Molecules vibrate, for example, and switch conformations (shapes) in response to environmental conditions. They act (reflexively) to reduce stress, and have memory. Some molecular species can even be said to self-replicate and evolve over time, and to engage in metabolism and cooperation [174–178]. Moving in the opposite direction, the random and intrinsic fluctuations at scales below a given Markov blanket are progressively suppressed [33]. To understand a society, for example, we need not focus on the behavior of a particular cell in a particular person.

Blankets within blankets, driven by the minimization of free energy, or similarly, the minimization of uncertainty, results in fractal, or self-similar patterns. Fractal, hierarchical patterns are seen throughout nature, including within the human brain. These patterns are dynamic. New ones form while old ones dissipate, change, or collapse. Restated, self-similar patterns arise because the same organizing principles are active at each scale.

At a given scale (e.g., a human body), individuals in the micro level below (e.g., cells) interact locally, largely through cooperative relationships. Cooperation (or coordination) occurs at the micro level because it reduces free energy there. The dynamics of local interactions at the micro level lead to long-distance correlations, which alter or form the macro level.

A good artificial example is robotic “swimmers” in a pool of water [179]. If the density of swimmers is not too high, each will swim in random directions. If the density of swimmers is increased, local interactions between them will cause some to begin swimming in unison, either clockwise or counterclockwise in a circular pattern around the pool. If the density is increased still further, more will join the pattern until all are swimming circularly. In this example, the number of choices dropped from many (any direction of swimming) to two, swimming either clockwise or counterclockwise. The circular macro pattern arises from a reduction in stress and uncertainty at the micro level. Once the macro pattern is established, providing fewer degrees of freedom, it is said to “enslave” agents in the level below (here, the swimmers). But causal influence goes both ways, and enslavement is just another pattern that can dissipate, change, or collapse.

In biological systems, the cooperation between agents at a given level exists as a dynamic if not tenuous balance between the preservation of individual autonomy and the benefits that coordination brings. Thus, for example, each cell within a human is capable of cooperating with neighboring cells, such that all cells, and the body, benefit. At the same time, each cell or subgroup of cells is capable of failing to cooperate with others and of selfishly promoting its own existence and prodigy. When this occurs, the resulting disease is called cancer.

The tension between independence and coordination is ever present. A meta pattern remains stable only to the degree that it benefits the micro level and overcomes the tendency of agents in the micro level to act independently. If new challenges render the benefits of coordination insufficient, or new patterns of behavior arise that better reduce uncertainty, then agitation in the micro level will eventually, sometimes quickly and dramatically, result in new macro patterns.

It is worth emphasizing that cooperation, communication, societal cognition, culture, and cognition at the individual level are all deeply entwined. Further, their dynamics can be cast in terms of active inference and the self-organization that it drives. Shared expectations, selective patterning of attention and behavior, cultural evolution, the formation of joint goals, implicit learning, and cultural inheritance underpin an active inference account of cognition [126]. Humans appear predisposed to align their mental states with others of the same kind, and human communication can largely be cast as a cooperative activity that serves to reduce uncertainty [180]. In the process, communication and the culture that it spawns offloads some of the computational demand that an individual would otherwise face onto the larger cultural (and physical) niche. Moreover, shared understandings, once established, reduce the effort and time necessary to sufficiently convey an idea or need. As such, communication, cooperation, and alignment through enculturation minimize the amount of energy necessary for a

human to survive, minimize experienced surprise, increase problem-solving capacity, and increase the set of affordances offered by niche. An individual cannot easily travel long distances (e.g., from New York to London) without help from past and current generations, for example. One affordance of the human niche is extensive travel.

4. SAILS as an Umbrella Term

The complete set of societal systems—economic, governance, legal, health, analytical, and education systems—by which a society senses, learns, makes decisions, and adapts is a new concept and has no standard name. A standard name would facilitate categorization and retrieval of pertinent literature. In previous documents, I refer to such systems alternatively as *social choice systems*, *decision-making systems*, and *learning systems* [104,181,182]. Emphasizing society as a superorganism, one could call the set a *societal nervous system*. In this series, I call it a *societal cognitive architecture*.

I propose that the term *societal cognitive architecture* be used to describe a generic set of societal systems, and that *societal active inference and learning system* (SAILS) be used to describe new sets of systems designed and developed in a science-driven process.⁶ SAILS is a nod to the work of Karl Friston on active inference. SAILS would apply to scientific research and other works (e.g., literature, art, music, education) that build on the following positions:

1. The design, testing, and implementation of new, integrated societal systems falls (non exclusively) under the purview of science.
2. Societal systems can be usefully viewed as learning, problem-solving, and adaptation systems (i.e, cognitive systems). They can perform well or poorly in this capacity. Their quality is a function of design, and design and expected and actual operation can be theoretically and empirically assessed. New systems can be designed to be fit for purpose, where that purpose is to serve and improve the common good.
3. Societal cognition occurs through an enactive process, and societies are self-evidencing in the sense that they naturally seek (appropriately simple) models to explain their lived, uncertain worlds. Quality societal cognition, involving rich communication, wise action, and sufficiently accurate models and explanations of the world, formal and casual, reduce experienced uncertainty. This, in turn, makes thriving and survival more likely.

Even though the phrase *active inference* appears in SAILS, I do not intend that SAILS applies only to works involving active inference proper. Active inference is but one promising approach to understanding cognition, and like others can be helpful as a philosophical guide and/or as concrete implementations (e.g., as predictive computational models). If societies are viewed as cognitive organisms, then theories of cognition are necessary to make progress. As of yet, active inference has barely been applied at the group level. Much work remains to demonstrate its utility and ability to predict social outcomes and behaviors in the real world. Many opportunities for testing exist, at different scales of sophistication, realism, and group size. The R&D program can help move this effort forward. In short, active inference and related theories and principles of cognition can serve as guides to discovery.

The proposed R&D program is but one approach to developing SAILS. Other programs are conceivable and could be conducted concurrently. If pursued, they might develop and implement new kinds of SAILS designs, or designs unrelated to the SAILS positions. All such programs could be viewed as components within a larger project, which is engaging the science community and public in the study and application of intentional societal self-organization.

⁶ SAILS is pronounced *sails*, as in a ship's sails, and is either singular or plural according to context.

5. Conclusions

Given the R&D program's aim of developing and testing new societal systems, a first logical step is to define a worldview that new systems can embody. Worldview begets purpose, and purpose suggests fitness metrics and system design. The general idea is that new systems can be designed to be fit for purpose and to reflect value and meaning as per worldview. This means, in part, being fit to solve the difficult problems that modern societies must solve. Fortunately, nature provides an abundance of examples of systems that excel at solving difficult problems. Some occur within the human body.

While certain ideas discussed in this paper might seem obvious to some readers—for example, that our wellbeing depends on the health of the environment; that self-identification can extend out to the social and environmental setting; or that societal systems function as learning, decision-making, and adaptation systems—it is not at all obvious, or at least not widely accepted, that new societal systems can or should be designed from the ground up to reflect these ideas. Science has only recently matured to the point where societal self-organization can fall under its purview. As yet, there are no societal systems that are designed in a modern, science-driven process to be fit for purpose. Humanity has a window of opportunity now to conduct this work. Fortunately, the prospect is attractive. Humans are naturally curious and compelled to learn and explore. In this case, the benefits of doing so are potentially massive, giving ample reason to act. Moreover, humans are compelled to reduce uncertainty. Given that uncertainty and associated anxiety are on track to greatly increase over coming decades, due to climate change, biodiversity loss, and other unsolved problems, humans may feel driven to explore new ways of organizing societies that ease uncertainty, fear, and strain.

Funding: This research received no external funding.

Acknowledgments: I would like to thank Ning Yan for reading the manuscript and offering helpful comments and suggestions.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Ripple, W.J.; Wolf, C.; Newsome, T.M.; Galetti, M.; Alamgir, M.; Crist, E.; Mahmoud, M.I.; Laurance, W.F. World Scientists' Warning to Humanity: A Second Notice. *BioScience* **2017**, *67*, 1026–1028. doi:10.1093/biosci/bix125.
2. Ripple, W.J.; Wolf, C.; Newsome, T.M.; Barnard, P.; Moomaw, W.R. World Scientists' Warning of a Climate Emergency. *BioScience* **2019**. doi:10.1093/biosci/biz088.
3. Albert, J.S.; Destouni, G.; Duke-Sylvester, S.M.; Magurran, A.E.; Oberdorff, T.; Reis, R.E.; Winemiller, K.O.; Ripple, W.J. Scientists' Warning to Humanity on the Freshwater Biodiversity Crisis. *Ambio* **2020**. doi:10.1007/s13280-020-01318-8.
4. Cardoso, P.; Barton, P.S.; Birkhofer, K.; Chichorro, F.; Deacon, C.; Fartmann, T.; Fukushima, C.S.; Gaigher, R.; Habel, J.C.; Hallmann, C.A.; Hill, M.J.; Hochkirch, A.; Kwak, M.L.; Mammola, S.; Ari Noriega, J.; Orfinger, A.B.; Pedraza, F.; Pryke, J.S.; Roque, F.O.; Settele, J.; Simaika, J.P.; Stork, N.E.; Suhling, F.; Vorster, C.; Samways, M.J. Scientists' Warning to Humanity on Insect Extinctions. *Biological Conservation* **2020**, *242*, 108426. doi:10.1016/j.biocon.2020.108426.
5. Parncutt, R. The Human Cost of Anthropogenic Global Warming: Semi-Quantitative Prediction and the 1,000-Tonne Rule. *Frontiers in Psychology* **2019**, *10*. doi:10.3389/fpsyg.2019.02323.
6. Gillard, R.; Gouldson, A.; Paavola, J.; Van Alstine, J. Transformational Responses to Climate Change: Beyond a Systems Perspective of Social Change in Mitigation and Adaptation. *Wiley Interdisciplinary Reviews: Climate Change* **2016**, *7*, 251–265, [<https://onlinelibrary.wiley.com/doi/pdf/10.1002/wcc.384>]. doi:10.1002/wcc.384.
7. IPCC. Summary for Policymakers. In *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation*; Cambridge University Press: Cambridge, UK, 2012.

8. Termeer, C.J.A.M.; Dewulf, A.; Biesbroek, G.R. Transformational Change: Governance Interventions for Climate Change Adaptation from a Continuous Change Perspective. *Journal of Environmental Planning and Management* **2017**, *60*, 558–576. doi:10.1080/09640568.2016.1168288.
9. McAlpine, C.; Seabrook, L.; Ryan, J.; Feeney, B.; Ripple, W.; Ehrlich, A.; Ehrlich, P. Transformational Change: Creating a Safe Operating Space for Humanity. *Ecology and Society* **2015**, *20*. doi:10.5751/ES-07181-200156.
10. Tàbara, D.J.; Jäger, J.; Mangalagiu, D.; Grasso, M. Defining Transformative Climate Science to Address High-End Climate Change. *Regional Environmental Change* **2019**, *19*, 807–818. doi:10.1007/s10113-018-1288-8.
11. Bennett, N.J.; Blythe, J.; Cisneros-Montemayor, A.M.; Singh, G.G.; Sumaila, U.R. Just Transformations to Sustainability. *Sustainability* **2019**, *11*, 3881. doi:10.3390/su11143881.
12. Monbiot, G. Dare to Declare Capitalism Dead – before It Takes Us All down with It | George Monbiot. *The Guardian* **2019**.
13. Savaget, P.; Geissdoerfer, M.; Kharrazi, A.; Evans, S. The Theoretical Foundations of Sociotechnical Systems Change for Sustainability: A Systematic Literature Review. *Journal of Cleaner Production* **2019**, *206*, 878–892. doi:10.1016/j.jclepro.2018.09.208.
14. Few, R.; Morchain, D.; Spear, D.; Mensah, A.; Bendapudi, R. Transformation, Adaptation and Development: Relating Concepts to Practice. *Palgrave Communications* **2017**, *3*, 1–9. doi:10.1057/palcomms.2017.92.
15. Kates, R.W.; Travis, W.R.; Wilbanks, T.J. Transformational Adaptation When Incremental Adaptations to Climate Change Are Insufficient. *Proceedings of the National Academy of Sciences of the United States of America* **2012**, *109*, 7156–7161.
16. O'Brien, K. Climate Change Adaptation and Social Transformation. In *International Encyclopedia of Geography*; American Cancer Society, 2017; pp. 1–8. doi:10.1002/9781118786352.wbieg0987.
17. Brand, U.; Görg, C.; Wissen, M. Overcoming Neoliberal Globalization: Social-Ecological Transformation from a Polanyian Perspective and Beyond. *Globalizations* **2020**, *17*, 161–176. doi:10.1080/14747731.2019.1644708.
18. Boik, J. Science-Driven Societal Transformation, Part II: Motivation and Strategy. *IDEAS [preprint]* **2020**.
19. Boik, J. Science-Driven Societal Transformation, Part III: Design. *IDEAS [preprint]* **2020**.
20. Muhar, A.; Raymond, C.M.; van den Born, R.J.G.; Bauer, N.; Böck, K.; Braitto, M.; Buijs, A.; Flint, C.; de Groot, W.T.; Ives, C.D.; Mitrofanenko, T.; Plieninger, T.; Tucker, C.; van Riper, C.J. A Model Integrating Social-Cultural Concepts of Nature into Frameworks of Interaction between Social and Natural Systems. *Journal of Environmental Planning and Management* **2018**, *61*, 756–777. doi:10.1080/09640568.2017.1327424.
21. Dunlap, R.E.; Liere, K.D.V. The “New Environmental Paradigm”. *The Journal of Environmental Education* **1978**, *9*, 10–19. doi:10.1080/00958964.1978.10801875.
22. Corral-Verdugo, V.; Carrus, G.; Bonnes, M.; Moser, G.; Sinha, J.B.P. Environmental Beliefs and Endorsement of Sustainable Development Principles in Water Conservation: Toward a New Human Interdependence Paradigm Scale. *Environment and Behavior* **2008**. doi:10.1177/0013916507308786.
23. Shi, J.; Visschers, V.H.M.; Siegrist, M. Public Perception of Climate Change: The Importance of Knowledge and Cultural Worldviews. *Risk Analysis* **2015**, *35*, 2183–2201. doi:10.1111/risa.12406.
24. Spangenberg, J.H. The World We See Shapes the World We Create: How the Underlying Worldviews Lead to Different Recommendations from Environmental and Ecological Economics - the Green Economy Example. *International Journal of Sustainable Development* **2016**, *19*, 127–146. doi:10.1504/IJSD.2016.077208.
25. Kohler, F.; Holland, T.G.; Kotiaho, J.S.; Desrousseaux, M.; Potts, M.D. Embracing Diverse Worldviews to Share Planet Earth. *Conservation Biology* **2019**, *0*. doi:10.1111/cobi.13304.
26. Cherry, T.L.; McEvoy, D.M.; Sælen, H. Conditional Cooperation and Cultural Worldviews. *Economics Letters* **2017**, *158*, 51–53. doi:10.1016/j.econlet.2017.06.034.
27. Rogoff, B. Culture and Participation: A Paradigm Shift. *Current Opinion in Psychology* **2016**, *8*, 182–189. doi:10.1016/j.copsyc.2015.12.002.
28. Kuhn, T.S. *The Structure of Scientific Revolutions*, 3rd edition ed.; University of Chicago Press: Chicago, IL, 1996.
29. Wilczek, F. Physics in 100 Years. *arXiv* **2015**, [1503.07735].
30. Linson, A.; Clark, A.; Ramamoorthy, S.; Friston, K. The Active Inference Approach to Ecological Perception: General Information Dynamics for Natural and Artificial Embodied Cognition. *Frontiers in Robotics and AI* **2018**, *5*. doi:10.3389/frobt.2018.00021.

31. Caticha, A. Entropic Inference and the Foundations of Physics. Technical report, Brazilian Chapter of the International Society for Bayesian Analysis, 2012.
32. Jaynes, E.T. The Evolution of Carnot's Principle. In *Maximum-Entropy and Bayesian Methods in Science and Engineering: Foundations*; Erickson, G.J.; Smith, C.R., Eds.; Fundamental Theories of Physics, Springer Netherlands: Dordrecht, 1988; pp. 267–281. doi:10.1007/978-94-009-3049-0_15.
33. Friston, K. A Free Energy Principle for a Particular Physics. *arXiv:1906.10184 [q-bio]* **2019**, [arXiv:q-bio/1906.10184].
34. Firmani, D.; Mecella, M.; Scannapieco, M.; Batini, C. On the Meaningfulness of “Big Data Quality” (Invited Paper). *Data Science and Engineering* **2016**, *1*, 6–20. doi:10.1007/s41019-015-0004-7.
35. Rovelli, C. Meaning = Information + Evolution. *arXiv:1611.02420 [physics]* **2016**, [arXiv:physics/1611.02420].
36. Kolchinsky, A.; Wolpert, D.H. Semantic Information, Autonomous Agency, and Nonequilibrium Statistical Physics. *Interface Focus* **2018**, *8*, 20180041, [1806.08053]. doi:10.1098/rsfs.2018.0041.
37. Dantzer, R. Neuroimmune Interactions: From the Brain to the Immune System and Vice Versa. *Physiological Reviews* **2017**, *98*, 477–504. doi:10.1152/physrev.00039.2016.
38. Herrmann-Pillath, C. Energy, Growth, and Evolution: Towards a Naturalistic Ontology of Economics. *Ecological Economics* **2015**, *119*, 432–442. doi:10.1016/j.ecolecon.2014.11.014.
39. Chapman, E.J.; Childers, D.L.; Vallino, J.J. How the Second Law of Thermodynamics Has Informed Ecosystem Ecology through Its History. *BioScience* **2016**, *66*, 27–39. doi:10.1093/biosci/biv166.
40. Skene, K.R. Life's a Gas: A Thermodynamic Theory of Biological Evolution. *Entropy* **2015**, *17*, 5522–5548. doi:10.3390/e17085522.
41. Siegenfeld, A.F.; Bar-Yam, Y. An Introduction to Complex Systems Science and Its Applications. *arXiv:1912.05088 [physics]* **2019**, [arXiv:physics/1912.05088].
42. Kesebir, S. The Superorganism Account of Human Sociality: How and When Human Groups Are Like Beehives. *Personality and Social Psychology Review* **2012**, *16*, 233–261. doi:10.1177/1088868311430834.
43. Lebiere, C.; Anderson, J.R. Cognitive Constraints on Decision Making under Uncertainty. *Frontiers in Psychology* **2011**, *2*. doi:10.3389/fpsyg.2011.00305.
44. Gershman, S.J.; Horvitz, E.J.; Tenenbaum, J.B. Computational Rationality: A Converging Paradigm for Intelligence in Brains, Minds, and Machines. *Science* **2015**, *349*, 273–278. doi:10.1126/science.aac6076.
45. Campbell, R. *The Metaphysics of Emergence*; Palgrave Macmillan, 2015.
46. Vernon, D.; Lowe, R.; Thill, S.; Ziemke, T. Embodied Cognition and Circular Causality: On the Role of Constitutive Autonomy in the Reciprocal Coupling of Perception and Action. *Frontiers in Psychology* **2015**, *6*. doi:10.3389/fpsyg.2015.01660.
47. Ramsay, D.S.; Woods, S.C. Clarifying the Roles of Homeostasis and Allostasis in Physiological Regulation. *Psychological review* **2014**, *121*, 225–247. doi:10.1037/a0035942.
48. Di Paolo, E.A. Autopoiesis, Adaptivity, Teleology, Agency. *Phenomenology and the Cognitive Sciences* **2005**, *4*, 429–452. doi:10.1007/s11097-005-9002-y.
49. Miłkowski, M.; Clowes, R.; Rucińska, Z.; Przegalińska, A.; Zawidzki, T.; Krueger, J.; Gies, A.; McGann, M.; Afeltowicz, Ł.; Wachowski, W.; Stjernberg, F.; Loughlin, V.; Hohol, M. From Wide Cognition to Mechanisms: A Silent Revolution. *Frontiers in Psychology* **2018**, *9*. doi:10.3389/fpsyg.2018.02393.
50. Bettinger, J.S.; Eastman, T.E. Foundations of Anticipatory Logic in Biology and Physics. *Progress in Biophysics and Molecular Biology* **2017**, *131*, 108–120. doi:10.1016/j.pbiomolbio.2017.09.009.
51. Miller, W.B. Biological Information Systems: Evolution as Cognition-Based Information Management. *Progress in Biophysics and Molecular Biology* **2018**, *134*, 1–26. doi:10.1016/j.pbiomolbio.2017.11.005.
52. Lyon, P. The Cognitive Cell: Bacterial Behavior Reconsidered. *Frontiers in Microbiology* **2015**, *6*. doi:10.3389/fmicb.2015.00264.
53. Segundo-Ortin, M.; Calvo, P. Are Plants Cognitive? A Reply to Adams. *Studies in History and Philosophy of Science Part A* **2019**, *73*, 64–71. doi:10.1016/j.shpsa.2018.12.001.
54. Keijzer, F.A. Evolutionary Convergence and Biologically Embodied Cognition. *Interface Focus* **2017**, *7*. doi:10.1098/rsfs.2016.0123.
55. Ross-Gillespie, A.; Kümmerli, R. Collective Decision-Making in Microbes. *Frontiers in Microbiology* **2014**, *5*. doi:10.3389/fmicb.2014.00054.
56. Varela, F.G.; Maturana, H.R.; Uribe, R. Autopoiesis: The Organization of Living Systems, Its Characterization and a Model. *Biosystems* **1974**, *5*, 187–196. doi:10.1016/0303-2647(74)90031-8.

57. Mingers, J. An Introduction to Autopoiesis—Implications and Applications. *Systems practice* **1989**, *2*, 159–180. doi:10.1007/BF01059497.
58. Maturana, H.R. The Organization of the Living: A Theory of the Living Organization. *International Journal of Human-Computer Studies* **1999**, *51*, 149–168. doi:10.1006/ijhc.1974.0304.
59. Ramstead, M.J.D.; Badcock, P.B.; Friston, K.J. Answering Schrödinger’s Question: A Free-Energy Formulation. *Physics of Life Reviews* **2018**, *24*, 1–16. doi:10.1016/j.plrev.2017.09.001.
60. Newen, A.; Gallagher, S.; Bruin, L.D. 4E Cognition: Historical Roots, Key Concepts, and Central Issues. In *The Oxford Handbook of 4E Cognition*; 2018.
61. LeDoux, J.E.; Brown, R. A Higher-Order Theory of Emotional Consciousness. *Proceedings of the National Academy of Sciences* **2017**, *114*, E2016–E2025. doi:10.1073/pnas.1619316114.
62. Barrett, L.F. The Theory of Constructed Emotion: An Active Inference Account of Interoception and Categorization. *Social Cognitive and Affective Neuroscience* **2017**, *12*, 1–23. doi:10.1093/scan/nsw154.
63. Garrigan, B.; Adlam, A.L.R.; Langdon, P.E. Moral Decision-Making and Moral Development: Toward an Integrative Framework. *Developmental Review* **2018**, *49*, 80–100. doi:10.1016/j.dr.2018.06.001.
64. Burghardt, G.M. A Place for Emotions in Behavior Systems Research. *Behavioural Processes* **2019**, *166*, 103881. doi:10.1016/j.beproc.2019.06.004.
65. Lamme, V.A.F. Challenges for Theories of Consciousness: Seeing or Knowing, the Missing Ingredient and How to Deal with Panpsychism. *Philosophical Transactions of the Royal Society B: Biological Sciences* **2018**, *373*, 20170344. doi:10.1098/rstb.2017.0344.
66. Tononi, G.; Boly, M.; Massimini, M.; Koch, C. Integrated Information Theory: From Consciousness to Its Physical Substrate. *Nature Reviews. Neuroscience* **2016**, *17*, 450–461. doi:10.1038/nrn.2016.44.
67. Miller, W.B.; Torday, J.S.; Baluška, F. Biological Evolution as Defense of ‘Self’. *Progress in Biophysics and Molecular Biology* **2019**, *142*, 54–74. doi:10.1016/j.pbiomolbio.2018.10.002.
68. Aktipis, A.; Maley, C.C. Cooperation and Cheating as Innovation: Insights from Cellular Societies. *Philosophical Transactions of the Royal Society B: Biological Sciences* **2017**, *372*. doi:10.1098/rstb.2016.0421.
69. Aktipis, A. Principles of Cooperation across Systems: From Human Sharing to Multicellularity and Cancer. *Evolutionary Applications* **2016**, *9*, 17–36. doi:10.1111/eva.12303.
70. Bruger, E.; Waters, C. Sharing the Sandbox: Evolutionary Mechanisms That Maintain Bacterial Cooperation. *F1000Research* **2015**, *4*. doi:10.12688/f1000research.7363.1.
71. Barker, J.L.; Bronstein, J.L.; Friesen, M.L.; Jones, E.I.; Reeve, H.K.; Zink, A.G.; Frederickson, M.E. Synthesizing Perspectives on the Evolution of Cooperation within and between Species. *Evolution* **2017**, *71*, 814–825. doi:10.1111/evo.13174.
72. Tu, C.; Suweis, S.; Grilli, J.; Formentin, M.; Maritan, A. Reconciling Cooperation, Biodiversity and Stability in Complex Ecological Communities. *Scientific Reports* **2019**, *9*, 1–10. doi:10.1038/s41598-019-41614-2.
73. Ferlian, O.; Cesarz, S.; Craven, D.; Hines, J.; Barry, K.E.; Bruelheide, H.; Buscot, F.; Haider, S.; Heklau, H.; Herrmann, S.; Kühn, P.; Pruschitzki, U.; Schädler, M.; Wagg, C.; Weigelt, A.; Wubet, T.; Eisenhauer, N. Mycorrhiza in Tree Diversity–Ecosystem Function Relationships: Conceptual Framework and Experimental Implementation. *Ecosphere* **2018**, *9*, e02226. doi:10.1002/ecs2.2226.
74. Apicella, C.L.; Silk, J.B. The Evolution of Human Cooperation. *Current Biology* **2019**, *29*, R447–R450. doi:10.1016/j.cub.2019.03.036.
75. Gavrilets, S.; Richerson, P.J. Collective Action and the Evolution of Social Norm Internalization. *Proceedings of the National Academy of Sciences of the United States of America* **2017**, *114*, 6068–6073. doi:10.1073/pnas.1703857114.
76. Ågren, J.A.; Davies, N.G.; Foster, K.R. Enforcement Is Central to the Evolution of Cooperation. *Nature Ecology & Evolution* **2019**, *3*, 1018–1029. doi:10.1038/s41559-019-0907-1.
77. Liu, L.; Chen, X.; Szolnoki, A. Competitions between Prosocial Exclusions and Punishments in Finite Populations. *Scientific Reports* **2017**, *7*, 46634. doi:10.1038/srep46634.
78. Fiala, L.; Suetens, S. Transparency and Cooperation in Repeated Dilemma Games: A Meta Study. *Experimental Economics* **2017**, *20*, 755–771. doi:10.1007/s10683-017-9517-4.
79. Unakafov, A.M.; Schultze, T.; Kagan, I.; Möller, S.; Eule, S.; Wolf, F. Emergence and Suppression of Cooperation by Action Visibility in Transparent Games. *bioRxiv* **2018**, p. 314500. doi:10.1101/314500.
80. Feinerman, O.; Korman, A. Individual versus Collective Cognition in Social Insects. *Journal of Experimental Biology* **2017**, *220*, 73–82. doi:10.1242/jeb.143891.

81. Admati, A.R. A Skeptical View of Financialized Corporate Governance. *Journal of Economic Perspectives* **2017**, *31*, 131–150. doi:10.1257/jep.31.3.131.
82. den Nieuwenboer, N.A.; Kaptein, M. Spiraling down into Corruption: A Dynamic Analysis of the Social Identity Processes That Cause Corruption in Organizations to Grow. *Journal of Business Ethics* **2008**, *83*, 133–146.
83. Solas, J. Conscientious Objections to Corporate Wrongdoing. *Business and Society Review* **2019**, *124*, 43–62. doi:10.1111/basr.12162.
84. Wiegatz, J. 'They'Re All in It Together': The Social Production of Fraud in Capitalist Africa. *Review of African Political Economy* **2019**, *46*, 357–368. doi:10.1080/03056244.2019.1682297.
85. Giuliani, E.; Nieri, F. Fixing "Bad" Capitalism: Why CSR and pro-Good Initiatives May Not Be Enough. *critical perspectives on international business* **2019**, ahead-of-print. doi:10.1108/cpoib-07-2019-0055.
86. Leon, K.S.; Ken, I. Legitimized Fraud and the State-Corporate Criminology of Food – a Spectrum-Based Theory. *Crime, Law and Social Change* **2019**, *71*, 25–46. doi:10.1007/s10611-018-9787-6.
87. Chaikin, D. Money Laundering and Tax Evasion – The Assisting of the Banking Sector. In *The Handbook of Business and Corruption*; S. Aßländer, M.; Hudson, S., Eds.; Emerald Publishing Limited, 2017; pp. 237–254. doi:10.1108/978-1-78635-445-720161012.
88. Barkemeyer, R.; Faugère, C.; Gergaud, O.; Preuss, L. Media Attention to Large-Scale Corporate Scandals: Hype and Boredom in the Age of Social Media. *Journal of Business Research* **2020**, *109*, 385–398. doi:10.1016/j.jbusres.2019.12.011.
89. Hall, S. Exxon Knew about Climate Change Almost 40 Years Ago, 2015.
90. Bush, M.J. Denial and Deception. In *Climate Change and Renewable Energy: How to End the Climate Crisis*; Bush, M.J., Ed.; Springer International Publishing: Cham, 2020; pp. 373–420. doi:10.1007/978-3-030-15424-0_8.
91. Pezzulo, G.; Dindo, H. What Should I Do next? Using Shared Representations to Solve Interaction Problems. *Experimental Brain Research* **2011**, *211*, 613–630. doi:10.1007/s00221-011-2712-1.
92. dos Santos, M.; West, S.A. The Coevolution of Cooperation and Cognition in Humans. *Proceedings of the Royal Society B: Biological Sciences* **2018**, *285*, 20180723. doi:10.1098/rspb.2018.0723.
93. Flack, J.C.; Krakauer, D.C. Challenges for Complexity Measures: A Perspective from Social Dynamics and Collective Social Computation. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **2011**, *21*, 037108. doi:10.1063/1.3643063.
94. Flack, J.C. Coarse-Graining as a Downward Causation Mechanism. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **2017**, *375*, 20160338. doi:10.1098/rsta.2016.0338.
95. Theiner, G.; Allen, C.; Goldstone, R.L. Recognizing Group Cognition. *Cognitive Systems Research* **2010**, *11*, 378–395. doi:10.1016/j.cogsys.2010.07.002.
96. Smart, P.R. Mandevillian Intelligence. *Synthese* **2018**, *195*, 4169–4200. doi:10.1007/s11229-017-1414-z.
97. Kao, A.B.; Miller, N.; Torney, C.; Hartnett, A.; Couzin, I.D. Collective Learning and Optimal Consensus Decisions in Social Animal Groups. *PLoS Computational Biology* **2014**, *10*. doi:10.1371/journal.pcbi.1003762.
98. Sheshadri, A.; Lease, M. SQUARE: A Benchmark for Research on Computing Crowd Consensus. First AAAI Conference on Human Computation and Crowdsourcing, 2013.
99. Wang, L.; Michoel, T. Accurate Wisdom of the Crowd from Unsupervised Dimension Reduction. *Royal Society Open Science* **2019**, *6*, 181806. doi:10.1098/rsos.181806.
100. Clément, R.J.G.; Krause, S.; von Engelhardt, N.; Faria, J.J.; Krause, J.; Kurvers, R.H.J.M. Collective Cognition in Humans: Groups Outperform Their Best Members in a Sentence Reconstruction Task. *PLOS ONE* **2013**, *8*, e77943. doi:10.1371/journal.pone.0077943.
101. Max-Neef, M. The World on a Collision Course and the Need for a New Economy. *Ambio* **2010**, *39*, 200–210.
102. Maslow, A.H. A Theory of Human Motivation. *Psychological Review* **1943**, *50*, 370–396. doi:10.1037/h0054346.
103. Schwartz, S.H. Are There Universal Aspects in the Structure and Contents of Human Values? *Journal of Social Issues* **1994**, *50*, 19–45. doi:10.1111/j.1540-4560.1994.tb01196.x.
104. Boik, J.C. Solving Problems That Matter Could Be the Next Big Thing, 2017.
105. Brown, H.S.; Vergragt, P.J. From Consumerism to Wellbeing: Toward a Cultural Transition? *Journal of Cleaner Production* **2016**, *132*, 308–317. doi:10.1016/j.jclepro.2015.04.107.
106. Eppard, L.; Rank, M.R.; Bullock, H.E. *Rugged Individualism and the Misunderstanding of American Inequality*; University Press Copublishing Division / Lehigh University Press, 2020.

107. Feola, G. Capitalism in Sustainability Transitions Research: Time for a Critical Turn? *Environmental Innovation and Societal Transitions* **2019**. doi:10.1016/j.eist.2019.02.005.
108. Blum, H.E. The Human Microbiome. *Advances in Medical Sciences* **2017**, *62*, 414–420. doi:10.1016/j.advms.2017.04.005.
109. Rakoff-Nahoum, S.; Foster, K.R.; Comstock, L.E. The Evolution of Cooperation within the Gut Microbiota. *Nature* **2016**, *533*, 255–259. doi:10.1038/nature17626.
110. Sender, R.; Fuchs, S.; Milo, R. Revised Estimates for the Number of Human and Bacteria Cells in the Body. *PLOS Biology* **2016**, *14*, e1002533. doi:10.1371/journal.pbio.1002533.
111. Smith, L.K.; Wissel, E.F. Microbes and the Mind: How Bacteria Shape Affect, Neurological Processes, Cognition, Social Relationships, Development, and Pathology. *Perspectives on Psychological Science* **2019**, *14*, 397–418. doi:10.1177/1745691618809379.
112. Davidson, G.L.; Cooke, A.C.; Johnson, C.N.; Quinn, J.L. The Gut Microbiome as a Driver of Individual Variation in Cognition and Functional Behaviour. *Philosophical Transactions of the Royal Society B: Biological Sciences* **2018**, *373*. doi:10.1098/rstb.2017.0286.
113. Magliocca, N.R.; Ellis, E.C. Evolving Human Landscapes: A Virtual Laboratory Approach. *Journal of Land Use Science* **2016**, *11*, 642–671. doi:10.1080/1747423X.2016.1241314.
114. Heras-Escribano, M.; De Pinedo-García, M. Affordances and Landscapes: Overcoming the Nature–Culture Dichotomy through Niche Construction Theory. *Frontiers in Psychology* **2018**, *8*. doi:10.3389/fpsyg.2017.02294.
115. Constant, A.; Ramstead, M.J.D.; Veissière, S.P.L.; Campbell, J.O.; Friston, K.J. A Variational Approach to Niche Construction. *Journal of the Royal Society Interface* **2018**, *15*. doi:10.1098/rsif.2017.0685.
116. Laland, K.N.; O'Brien, M.J. Niche Construction: Implications for Human Sciences. In *Emerging Trends in the Social and Behavioral Sciences*; American Cancer Society, 2015; pp. 1–10. doi:10.1002/9781118900772.etrds0242.
117. Laland, K.N.; Odling-Smee, J.; Feldman, M.W. Niche Construction, Biological Evolution, and Cultural Change. *The Behavioral and Brain Sciences* **2000**, *23*, 131–146; discussion 146–175. doi:10.1017/s0140525x00002417.
118. Canino-Koning, R.; Wiser, M.J.; Ofria, C. Fluctuating Environments Select for Short-Term Phenotypic Variation Leading to Long-Term Exploration. *PLOS Computational Biology* **2019**, *15*, e1006445. doi:10.1371/journal.pcbi.1006445.
119. Xue, C.; Liu, Z.; Goldenfeld, N. Scale-Invariant Topology and Bursty Branching of Evolutionary Trees Emerge from Niche Construction. *Proceedings of the National Academy of Sciences* **2020**, *117*, 7879–7887. doi:10.1073/pnas.1915088117.
120. Whiten, A.; Ayala, F.J.; Feldman, M.W.; Laland, K.N. The Extension of Biology through Culture. *Proceedings of the National Academy of Sciences* **2017**, *114*, 7775–7781. doi:10.1073/pnas.1707630114.
121. Smith, C.; Gabora, L.; Gardner-O'Kearney, W. The Extended Evolutionary Synthesis Facilitates Evolutionary Models of Culture Change. *Cliodynamics: The Journal of Quantitative History and Cultural Evolution* **2018**, *9*, 84–107. doi:10.21237/C7clio9235187.
122. Sinha, C. Language and Other Artifacts: Socio-Cultural Dynamics of Niche Construction. *Frontiers in Psychology* **2015**, *6*. doi:10.3389/fpsyg.2015.01601.
123. Ramstead, M.J.D.; Constant, A.; Badcock, P.B.; Friston, K.J. Variational Ecology and the Physics of Sentient Systems. *Physics of Life Reviews* **2019**, *31*, 188–205. doi:10.1016/j.plrev.2018.12.002.
124. Bolis, D.; Schilbach, L. 'I Interact Therefore I Am': The Self as a Historical Product of Dialectical Attunement. *Topoi* **2018**. doi:10.1007/s11245-018-9574-0.
125. Bentley, S.V.; Greenaway, K.H.; Haslam, S.A. Cognition in Context: Social Inclusion Attenuates the Psychological Boundary between Self and Other. *Journal of Experimental Social Psychology* **2017**, *73*, 42–49. doi:10.1016/j.jesp.2017.06.008.
126. Veissière, S.P.L.; Constant, A.; Ramstead, M.J.D.; Friston, K.J.; Kirmayer, L.J. Thinking through Other Minds: A Variational Approach to Cognition and Culture. *Behavioral and Brain Sciences* **2020**, *ed*, 43. doi:10.1017/S0140525X19001213.
127. Müller, G.B. Why an Extended Evolutionary Synthesis Is Necessary. *Interface Focus* **2017**, *7*, 20170015. doi:10.1098/rsfs.2017.0015.

128. Krakauer, D.; Bertschinger, N.; Olbrich, E.; Flack, J.C.; Ay, N. The Information Theory of Individuality. *Theory in Biosciences* **2020**. doi:10.1007/s12064-020-00313-7.
129. Crutchfield, J.P. The Origins of Computational Mechanics: A Brief Intellectual History and Several Clarifications. *arXiv:1710.06832 [cond-mat, physics:nlin]* **2017**, [arXiv:cond-mat, physics:nlin/1710.06832].
130. Wolpert, D.H. Physical Limits of Inference. *Physica D: Nonlinear Phenomena* **2008**, *237*, 1257–1281. doi:10.1016/j.physd.2008.03.040.
131. Clark, J.B.; Jacques, D.R. Practical Measurement of Complexity in Dynamic Systems. *Procedia Computer Science* **2012**, *8*, 14–21. doi:10.1016/j.procs.2012.01.008.
132. Efatmaneshnik, M.; Ryan, M.J. A General Framework for Measuring System Complexity. *Complexity* **2016**, *21*, 533–546. doi:10.1002/cplx.21767.
133. Lui, L.T.; Terrazas, G.; Zenil, H.; Alexander, C.; Krasnogor, N. Complexity Measurement Based on Information Theory and Kolmogorov Complexity. *Artificial Life* **2015**, *21*, 205–224. doi:10.1162/ARTL_a_00157.
134. Kanwal, M.S.; Grochow, J.A.; Ay, N. Comparing Information-Theoretic Measures of Complexity in Boltzmann Machines. *Entropy* **2017**, *19*, 310. doi:10.3390/e19070310.
135. Riechers, P.M.; Crutchfield, J.P. Spectral Simplicity of Apparent Complexity, Part I: The Nondiagonalizable Metadynamics of Prediction. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **2018**, *28*, 033115, [1705.08042]. doi:10.1063/1.4985199.
136. Levin, M. The Computational Boundary of a “Self”: Developmental Bioelectricity Drives Multicellularity and Scale-Free Cognition. *Frontiers in Psychology* **2019**, *10*. doi:10.3389/fpsyg.2019.02688.
137. Auerbach, J.E.; Bongard, J.C. On the Relationship between Environmental and Morphological Complexity in Evolved Robots. Proceedings of the 14th Annual Conference on Genetic and Evolutionary Computation; Association for Computing Machinery: Philadelphia, Pennsylvania, USA, 2012; GECCO '12, pp. 521–528. doi:10.1145/2330163.2330238.
138. Auerbach, J.E.; Bongard, J.C. Environmental Influence on the Evolution of Morphological Complexity in Machines. *PLOS Computational Biology* **2014**, *10*, e1003399. doi:10.1371/journal.pcbi.1003399.
139. CONANT, R.C.; ASHBY, W.R. Every Good Regulator of a System Must Be a Model of That System. *International Journal of Systems Science* **1970**, *1*, 89–97. doi:10.1080/00207727008920220.
140. Ashby, W.R. *An Introduction to Cybernetics*; Chapman & Hall, London, 1956.
141. Hills, T.T.; Todd, P.M.; Lazer, D.; Redish, A.D.; Couzin, I.D. Exploration versus Exploitation in Space, Mind, and Society. *Trends in Cognitive Sciences* **2015**, *19*, 46–54. doi:10.1016/j.tics.2014.10.004.
142. Bak, P.; Tang, C.; Wiesenfeld, K. Self-Organized Criticality: An Explanation of the 1/f Noise. *Physical Review Letters* **1987**, *59*, 381–384. doi:10.1103/PhysRevLett.59.381.
143. Timme, N.M.; Marshall, N.J.; Bennett, N.; Ripp, M.; Lautzenhiser, E.; Beggs, J.M. Criticality Maximizes Complexity in Neural Tissue. *Frontiers in Physiology* **2016**, *7*. doi:10.3389/fphys.2016.00425.
144. Hesse, J.; Gross, T. Self-Organized Criticality as a Fundamental Property of Neural Systems. *Frontiers in Systems Neuroscience* **2014**, *8*. doi:10.3389/fnsys.2014.00166.
145. Valverde, S.; Ohse, S.; Turalska, M.; West, B.J.; Garcia-Ojalvo, J. Structural Determinants of Criticality in Biological Networks. *Frontiers in Physiology* **2015**, *6*. doi:10.3389/fphys.2015.00127.
146. Daniels, B.C.; Kim, H.; Moore, D.; Zhou, S.; Smith, H.B.; Karas, B.; Kauffman, S.A.; Walker, S.I. Criticality Distinguishes the Ensemble of Biological Regulatory Networks. *Physical Review Letters* **2018**, *121*, 138102. doi:10.1103/PhysRevLett.121.138102.
147. Bäuerle, T.; Löffler, R.C.; Bechinger, C. Formation of Stable and Responsive Collective States in Suspensions of Active Colloids. *Nature Communications* **2020**, *11*, 2547. doi:10.1038/s41467-020-16161-4.
148. Bernardes, A.T.; Zorzenon dos Santos, R.M. Immune Network at the Edge of Chaos. *Journal of Theoretical Biology* **1997**, *186*, 173–187. doi:10.1006/jtbi.1996.0316.
149. Kauffman, S. *The Origins of Order: Self-Organization and Selection in Evolution*; Oxford University Press: Oxford, New York, 1993.
150. Kirby, S.; Tamariz, M.; Cornish, H.; Smith, K. Compression and Communication in the Cultural Evolution of Linguistic Structure. *Cognition* **2015**, *141*, 87–102. doi:10.1016/j.cognition.2015.03.016.
151. Khajehabdollahi, S.; Abeyasinghe, P.M.; Owen, A.M.; Soddu, A. The Emergence of Integrated Information, Complexity, and Consciousness at Criticality. *bioRxiv* **2019**, p. 521567. doi:10.1101/521567.

152. Goudarzi, A.; Teuscher, C.; Gulbahce, N.; Rohlf, T. Emergent Criticality through Adaptive Information Processing in Boolean Networks. *Physical Review Letters* **2012**, *108*, 128702. doi:10.1103/PhysRevLett.108.128702.
153. Haruna, T.; Nakajima, K. Maximizing Local Information Transfer in Boolean Networks. *New Journal of Physics* **2018**, *20*, 083046. doi:10.1088/1367-2630/aadbc3.
154. Friston, K.J. Functional and Effective Connectivity: A Review. *Brain Connectivity* **2011**, *1*, 13–36. doi:10.1089/brain.2011.0008.
155. Boyd, A.B.; Mandal, D.; Crutchfield, J.P. Thermodynamics of Modularity: Structural Costs Beyond the Landauer Bound. *Physical Review X* **2018**, *8*, 031036. doi:10.1103/PhysRevX.8.031036.
156. Sajid, N.; Ball, P.J.; Friston, K.J. Active Inference: Demystified and Compared. *arXiv:1909.10863 [cs, q-bio]* **2020**, [arXiv:cs, q-bio/1909.10863].
157. Kirchhoff, M.; Parr, T.; Palacios, E.; Friston, K.; Kiverstein, J. The Markov Blankets of Life: Autonomy, Active Inference and the Free Energy Principle. *Journal of The Royal Society Interface* **2018**, *15*, 20170792. doi:10.1098/rsif.2017.0792.
158. Friston, K. The Free-Energy Principle: A Unified Brain Theory? *Nature Reviews Neuroscience* **2010**, *11*, 127–138. doi:10.1038/nrn2787.
159. Friston, K.J.; Parr, T.; de Vries, B. The Graphical Brain: Belief Propagation and Active Inference. *Network Neuroscience* **2017**, *1*, 381–414. doi:10.1162/NETN_a_00018.
160. Brown, H.; Adams, R.A.; Parees, I.; Edwards, M.; Friston, K. Active Inference, Sensory Attenuation and Illusions. *Cognitive Processing* **2013**, *14*, 411–427. doi:10.1007/s10339-013-0571-3.
161. Friston, K. Hierarchical Models in the Brain. *PLOS Computational Biology* **2008**, *4*, e1000211. doi:10.1371/journal.pcbi.1000211.
162. Friston, K.; FitzGerald, T.; Rigoli, F.; Schwartenbeck, P.; Pezzulo, G. Active Inference: A Process Theory. *Neural Computation* **2016**, *29*, 1–49. doi:10.1162/NECO_a_00912.
163. Lehmann, A.; Masò, J.; Nativi, S.; Giuliani, G. Towards Integrated Essential Variables for Sustainability. *International Journal of Digital Earth* **2020**, *13*, 158–165. doi:10.1080/17538947.2019.1636490.
164. Allen, M.; Friston, K.J. From Cognitivism to Autopoiesis: Towards a Computational Framework for the Embodied Mind. *Synthese* **2018**, *195*, 2459–2482. doi:10.1007/s11229-016-1288-5.
165. Parr, T.; Friston, K.J. Generalised Free Energy and Active Inference. *Biological Cybernetics* **2019**, *113*, 495–513. doi:10.1007/s00422-019-00805-w.
166. Friston, K.; Rigoli, F.; Ognibene, D.; Mathys, C.; Fitzgerald, T.; Pezzulo, G. Active Inference and Epistemic Value. *Cognitive Neuroscience* **2015**, *6*, 187–214. doi:10.1080/17588928.2015.1020053.
167. Jaynes, E. Information Theory and Statistical Mechanics. *Physical Review* **1957**, *106*, 620–630.
168. Millidge, B. Deep Active Inference as Variational Policy Gradients. *arXiv:1907.03876 [cs]* **2019**, [arXiv:cs/1907.03876].
169. Friston, K.; Breakspear, M.; Deco, G. Perception and Self-Organized Instability. *Frontiers in Computational Neuroscience* **2012**, *6*. doi:10.3389/fncom.2012.00044.
170. Ramstead, M.J.D.; Kirchhoff, M.D.; Friston, K.J. A Tale of Two Densities: Active Inference Is Enactive Inference, 2019.
171. Friston, K. The History of the Future of the Bayesian Brain. *Neuroimage* **2012**, *62-248*, 1230–1233. doi:10.1016/j.neuroimage.2011.10.004.
172. Friston, K.J.; Lin, M.; Frith, C.D.; Pezzulo, G.; Hobson, J.A.; Ondobaka, S. Active Inference, Curiosity and Insight. *Neural Computation* **2017**, *29*, 2633–2683. doi:10.1162/neco_a_00999.
173. Goldberg, R.F.; Vandenberg, L.N. Distract, Delay, Disrupt: Examples of Manufactured Doubt from Five Industries. *Reviews on Environmental Health* **2019**, *34*, 349–363. doi:10.1515/revh-2019-0004.
174. Schuster, P. Increase in Complexity and Information through Molecular Evolution. *Entropy* **2016**, *18*, 397. doi:10.3390/e18110397.
175. Ma, W.; Hu, J. Computer Simulation on the Cooperation of Functional Molecules during the Early Stages of Evolution. *PLOS ONE* **2012**, *7*, e35454. doi:10.1371/journal.pone.0035454.
176. Levin, S.R.; West, S.A. The Evolution of Cooperation in Simple Molecular Replicators. *Proceedings of the Royal Society B: Biological Sciences* **2017**, *284*, 20171967. doi:10.1098/rspb.2017.1967.

177. Liu, B.; Pappas, C.G.; Ottel , J.; Schaeffer, G.; Jurissek, C.; Pieters, P.F.; Altay, M.; Mari , I.; Stuart, M.C.A.; Otto, S. Spontaneous Emergence of Self-Replicating Molecules Containing Nucleobases and Amino Acids. *Journal of the American Chemical Society* **2020**, *142*, 4184–4192. doi:10.1021/jacs.9b10796.
178. Monreal Santiago, G.; Liu, K.; Browne, W.R.; Otto, S. Emergence of Light-Driven Protometabolism on Recruitment of a Photocatalytic Cofactor by a Self-Replicator. *Nature Chemistry* **2020**, *12*, 603–607. doi:10.1038/s41557-020-0494-4.
179. Haken, H. What Can Synergetics Contribute to Embodied Aesthetics? *Behavioral Sciences* **2017**, *7*, 61. doi:10.3390/bs7030061.
180. Vasil, J.; Badcock, P.B.; Constant, A.; Friston, K.; Ramstead, M.J.D. A World Unto Itself: Human Communication as Active Inference. *Frontiers in Psychology* **2020**, *11*. doi:10.3389/fpsyg.2020.00417.
181. Boik, J.C. Optimality of Social Choice Systems: Complexity, Wisdom, and Wellbeing Centrality. *IDEAS [preprint]* **2016**.
182. Boik, J. Wellbeing Centrality: Toward an Economics and Politics of Meaning, 2017.