Article
Sciences of Observation

Chris Fields [ID]

23 rue des Lavandières, 11160 Caunes Minervois, France; fieldsres@gmail.com

check for updates

Abstract: Multiple sciences have converged, in the past two decades, on a hitherto mostly unremarked question: what is observation? Here, I examine this evolution, focusing on three sciences: physics, especially quantum information theory, developmental biology, especially its molecular and "evo-devo" branches, and cognitive science, especially perceptual psychology and robotics. I trace the history of this question to the late 19th century, and through the conceptual revolutions of the 20th century. I show how the increasing interdisciplinary focus on the process of extracting information from an environment provides an opportunity for conceptual unification, and sketch an outline of what such a unification might look like.

Keywords: awareness; cognition; computation; cybernetics; differentiation; fitness; holographic encoding; memory; perception; quantum information; signal transduction; spatial representation; thermodynamics; unitarity

1. Introduction

Science is distinguished from speculation by its grounding in observation. What "observation" is, however, has been largely neglected. What does it mean to "ask a question of Nature" and receive a reply? How are the sought-after observational outcomes actually obtained? Even in quantum theory, where the "measurement problem" has occupied philosophers and physicists alike for nearly a century, the question of *how observations are made* is largely replaced by far more metaphysical-sounding questions of "wavefunction collapse" or the "quantum-to-classical transition" (see [1,2] for recent reviews). Once the world has been rendered "effectively classical", the thinking goes, observation becomes completely straightforward: one just has to look.

Here, I will advance two claims: (1) that substantial scientific effort is currently being devoted, across multiple disciplines, to understanding observation as a process, though seldom in these terms; and (2) that making this cross-disciplinary effort explicit offers an opportunity for conceptual cross-fertilization that leaves entirely aside troublesome issues of reductionism and disciplinary imperialism. Viewing a substantial fraction of current science as being fundamentally about observation, as opposed to some specialized domain or other, allows us to collapse a large amount of disparate ontology into a few very general concepts, and to see how these concepts inform science across the board. While suggestions along these lines have been made previously within the cybernetic and interdisciplinary traditions (e.g., [3–9]), developments in the past two decades in quantum information, cognitive science and the biology of signal transduction, among other areas, render them ever more compelling and productive. This multi-disciplinary conceptual landscape has recently been explored with somewhat different goals by Dodig Crnkovic [10,11]. As in [10,11], the focus here is on the formalized sciences, working "upwards" from the precise but relatively simple formal description of observation employed in physics toward the higher complexity needed for the life sciences. The alternative interdisciplinary tactic of working "downward" in complexity from studies of human language or culture toward the formalized sciences, as attempted in, e.g., cultural structuralism or general semiotics, is not considered here, though some points of contact are briefly mentioned.

Observation per se became a topic for investigation by physicists only in the late 19th century, with Boltzmann's realization that observers are characterized by uncertainty and must pay, in energetic currency, to reduce their uncertainty [12]. Energy, therefore, is an essential *resource* for observation. Any physically-implemented observer is limited to a finite quantity of this resource, so any such observer is limited to finite observations at finite resolution. Shannon, some 50 years later, recognized that while observational outcomes could be encoded in myriad ways, any finite sequence of finite-resolution outcomes could be encoded as a finite string of bits [13]. The third foundation of the classical, thermodynamic theory of observation was laid by Landauer, who emphasized that observational outcomes can be accessed and used for some further purpose only if they have been written in a thermodynamically-irreversible way on some physical medium [14,15]. This triad of energy, encoding and memory can be summarized by the claim that each bit of an irreversibly-recorded observational outcome costs ck_BT, where k_B is Boltzmann's constant, $T > 0$ is temperature, and $c > \ln 2$ is a measure of the observer's thermodynamic efficiency. Less formally, this classical theory defines "observation" as the exchange of energy for information.

As recognized by many and proved explicitly by Moore ([16] Theorem 2) in 1956, *finite* observations are necessarily ambiguous: no finite observation or sequence of observations can establish with certainty what system is being observed (see [17] for review). Hence any thermodynamically-allowed theory of observation is a theory of ontologically-ambiguous observation. Real observers are not "god's-eye" observers [18]. Bearing this in mind, we can ask several questions about observation:

- What kinds of systems exchange energy for information?
- How does this exchange work?
- What kinds of information do such systems exchange energy for?
- How is this information stored?
- What is it used for?

The sections that follow examine these questions in turn, from the perspectives of sciences from physics to psychology. We start with physics, both because physics defines observation precisely, albeit perhaps incompletely, and because it has historically been most concerned with how observation *works*, as opposed to the more philosophically- or linguistically-motivated question of what individual observations (or observations in general) *mean*. This latter question, which is related in at least its pragmatic sense to the final one above, is deferred to Section 6 below. How the answers to these various questions, however tentative, might be integrated into a theoretically-productive framework is then briefly considered.

2. What Is an Observer?

An "observer" as a perspective on the unfolding of events has been part of the lexicon of theoretical physics since Galileo's *Dialogue* of 1632. This Galilean observer is effectively a coordinate system; it has no active role in events and no effect on what transpires. An observer with an active role first appears in the mid-19th century, in the guise of Maxwell's demon. Macroscopic observers, in particular human scientists, gain an active role with the development of quantum theory, with Bohr's acknowledgement of the "free choice of experimental arrangement for which the mathematical structure of the quantum mechanical formalism offers the appropriate latitude" ([19] p. 71) and von Neumann's suggestion that a *conscious* observer is needed to collapse the wave function [20]. While observer-induced collapse has been largely superceded by the theory of decoherence [21–25], the observer's active role in freely choosing which observations to make remains a critical assumption of quantum theory [26,27].

In classical physics, the observer simply records information already present in the observed environment. This essentially passive view of observation is carried over into realist interpretations of quantum theory, in which collapse, branching, and/or decoherence are viewed as in some sense objective. Gell-Mann and Hartle [28], for example, describe the observer as an "information gathering and using system" (IGUS) that collects objective classical information generated by decoherence.

The decoherence-based quantum Darwinism of Zurek and colleagues [29–31] similarly requires an objectively redundant encoding of classical information in an environment shared by multiple observers. This realist assumption that classical information is objectively available to be "gathered" by observation was already challenged by Bohr in 1928: "an independent reality in the ordinary physical sense can neither be ascribed to the phenomena nor to the agencies of observation" ([32] p. 580). Bohr's and Heisenberg's view that observation was an active process, and that observational outcomes are well-defined only in the context of this process, finds recent expression in the "observer as participant" of Wheeler [33], Rovelli's [34] relational formulation of quantum theory, and Fuchs' [35] quantum Bayesianism (more recently, "QBism"). It is supported by the absence, within the quantum formalism, of any principled reason to decompose a state space into one collection of factors rather than another [36–44]. As loopholes in experimental demonstrations of Bell-inequality violations are progressively closed [45–47], physicists are increasingly forced to choose between giving up the objectivity of unobserved outcomes (i.e., "counterfactual definiteness" or just "realism") or giving up locality, including the idea that an experiment is a local operation on the world [48,49].

Describing an observer either as an IGUS or an outcome-generating participant raises an obvious question: what kinds of systems can play these roles? What, in other words, *counts as* an observer within a given interpretation or formulation of quantum theory? There are three common responses to this question. One is that any physical system can be an observer (e.g., [2,33,34]), a response difficult to reconcile with the assumption that observers can freely choose which observations to make. Another is to explicitly set the question outside of physics and possibly outside all of science (e.g., [35,50]), rendering a "theory of observation" impossible to construct. By far the most common response, however, is to ignore the question altogether. Prior to the development of quantum information theory, this reticence could be attributed to the general distaste for informal concepts memorably expressed by Bell ([51] p. 33):

> *Here are some words which, however legitimate and necessary in application, have no place in a formulation with any pretension to physical precision: system, apparatus, environment, microscopic, macroscopic, reversible, irreversible, observable, information, measurement.*

Since roughly 2000, however, quantum theory has increasingly been formulated in purely-informational terms [52–60]; see [61] for an informal review. In these formulations, quantum theory is itself a theory of observation. A physical interaction between A and B, represented by a Hamiltonian operator H_{AB}, is an information channel. If only quantum information passes through the channel, then A and B are entangled, i.e., are nominal components of a larger system $AB = A \otimes B$ in an entangled state $|AB\rangle$, by definition a state that cannot be factored into a product $|A\rangle|B\rangle$ of states of A and B individually (kets $|\cdot\rangle$ will be used to denote states, whether quantum or classical). In this case, no observation has occurred. If classical information flows through the channel, A and B must be separable, i.e., $|AB\rangle$ can be factored into $|A\rangle|B\rangle$, $|A\rangle$ encodes information about $|B\rangle$ and vice-versa, and observation can be considered to have occurred [17]. The only classical information associated with H_{AB} are its eigenvalues, so these are the only possible observational outcomes [62]. This picture, however, places no constraints at all on what counts as an observer. It does not, in particular, tell us the conditions under which an interaction transfers classical information. What more can be said? The next four sections pursue an indirect approach to this question, focusing first on the question of what is observed.

3. What Is Observed?

It is when the question, "what is observed?" is asked that one encounters the first serious disconnect between physics and the life sciences. At least in theoretical and foundational discussions, physicists describe observers interacting with *systems*, typically systems that have been defined a priori by specifying their state spaces. Biologists and psychologists, on the other hand, describe observers—organisms—interacting with the *world*. This world does not have an a priori specified state

space, at least not one known to or even knowable by any organism. It is the organism's job to figure out, by interacting with the world, what in it might be useful for the task of continuing to live.

The reliance of physics on predefined systems is sometimes explicitly acknowledged. Zurek, for example, remarks that "a compelling explanation of what the systems are - how to define them given, say, the overall Hamiltonian in some suitably large Hilbert space - would undoubtedly be most useful" ([63] p. 1818). Lacking such an explanation, however, he makes their existence axiomatic, assuming as "axiom(o)" of quantum mechanics that "(quantum) systems exist" ([50] p. 746; [64] p. 2) as objective entities. It is, as noted above, the objective, observer-independent existence of systems that allows the eigenvalues of their objective, observer-independent interactions with an objective, observer-independent environment to be encoded with objective, observer-independent redundancy in the theory of quantum Darwinism [29–31]. An even more extreme example of this assumption of given, a priori systems can be found in Tegmark's [65,66] description of decoherence (Figure 1). Here, the "system" S is defined as comprising *only* the "pointer" degrees of freedom of interest to the observer O, and O is defined as comprising *only* the degrees of freedom that record the observed pointer values. Everything else is considered part of the "environment" E and is traced over, i.e., treated as classical noise. The only information channel in this picture is the Hamiltonian H_{OS}, which is given a priori.
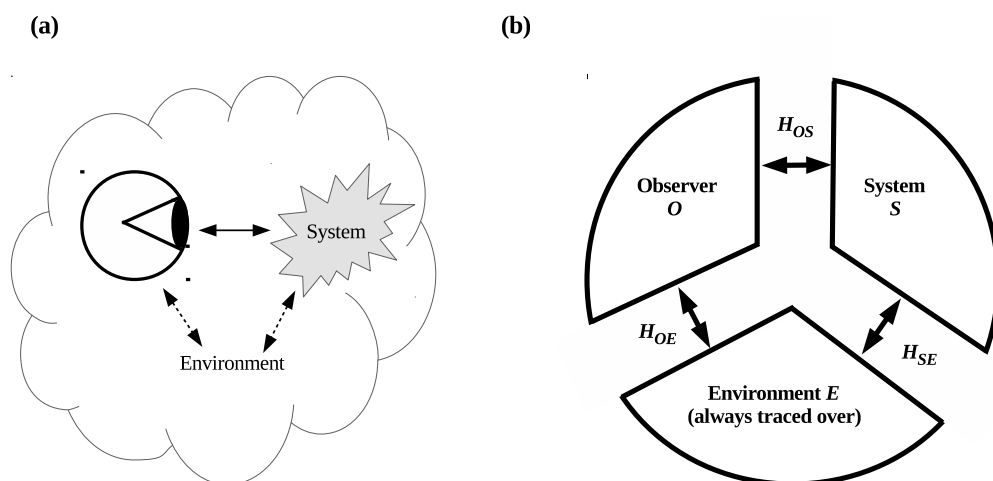
Figure 1. (**a**) The classical conception of observation, standardly carried over into quantum theory. The observer interacts with a system of interest; both are embedded in a surrounding environment. (**b**) Interactions between observer (O), system of interest (S) and environment (E) enabling environmental decoherence [1,2]. The Hamiltonian H_{OS} transfers outcome information from S to O; H_{SE} and H_{OE} decohere S and O respectively. Adapted from [66] Figure 1.

It is useful to examine this assumption of a priori systems in a practical setting. Suppose you have a new graduate student, Alice, who has never set foot in your laboratory. You ask Alice to go to the laboratory and read the pointer value for some instrument S. What does Alice have to do when she enters the laboratory? The assumption that S is given a priori is, in this case, the assumption that S is given *to Alice* a priori. All she has to do is read the pointer value. *In practice*, however, Alice has to do much more than this. Before reading the pointer value, she has to *identify S*: she has to *find S* amongst the clutter of the laboratory, and *distinguish S* from the other stuff surrounding it. Doing this, obviously, requires observation. It requires observing not just S, but other things besides S—for example, the tables and chairs that she has to navigate around before she gets to S. These other things are part of the "world" W in which S is embedded; in the notation of Figure 1, $W = S \oplus E$. It is W that Alice has to interact with to identify S, which she must do before she can read its pointer

value (Figure 2). This *W* is, it bears emphasizing, *Alice's* world: it comprises everything in the universe except Alice.
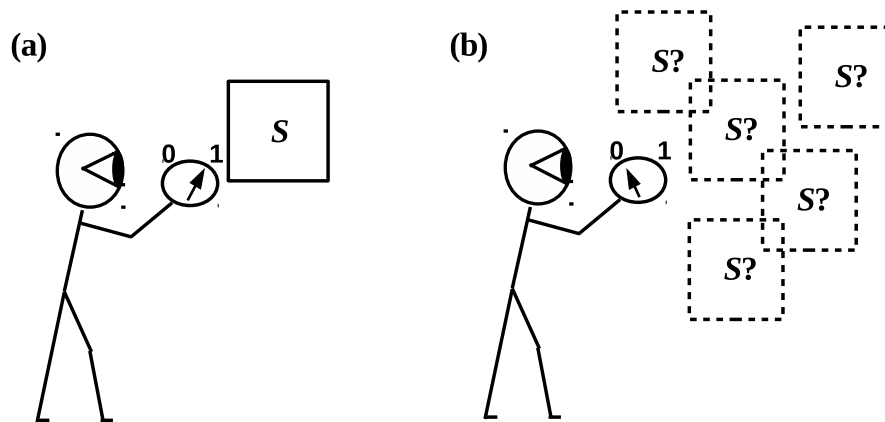

Figure 2. (**a**) Observers are standardly assumed to interact with a priori given systems, as in Figure 1. Here, the observer is equipped with an observable (e.g., a meter reading) with which to interact with the system *S*. Adapted from [35] Figure 1. (**b**) In practice, observers must *look for* the system of interest *S* by probing the "world" *W* in which it is embedded.

Classically, if I want Alice to observe the pointer of *S*, I need to give her a description of *S* that is good enough to pick it out from among the other objects in the laboratory. Such a description might specify the kind of instrument *S* is (e.g., a voltmeter), its size, shape, color, brand name, and possibly what it is sitting on or connected to. In quantum theory, these classical criteria are replaced with specified outcome values for some finite set of observables, which given Shannon's insight can be regarded as binary. It is important to emphasize that quantum theoretic observables are *operators* with which an observer *acts on the world*; the world then *acts on the observer* to deliver an outcome (e.g., [35]). This key idea of observation as *physical interaction*, formulated initially by Boltzmann and emphasized by Bohr and Heisenberg, is what is lost when systems are considered "given" and observation is regarded as a passive "gathering" of already-existing, observer-independent information. Recognizing that observers must *search for* systems of interest in the environments in which they are embedded brings this idea of observation as an activity to the fore.

Let us call the observables that identify some system *S reference observables* and their specified, criterial outcomes *reference outcomes* and denote them $\{M_i^{(R)}, x_i^{(R)}\}$. Given a set of reference observables and outcomes, Alice can then go into the laboratory, and measure everything she sees with the reference observables $\{M_i^{(R)}\}$. If every measurement yields its specified reference outcome $\{x_i^{(R)}\}$, Alice has found *S*; if not, she must keep searching. By giving Alice $\{M_i^{(R)}, x_i^{(R)}\}$, I am giving her what she needs, in practice, to identify *S* in the context of my laboratory. It is what she needs *in practice*, or as Bell [51] would have it "FAAP" due to Moore's [16] theorem noted above: such information is insufficient for objective, ontological precision. This set $\{M_i^{(R)}, x_i^{(R)}\}$ can be specified only because it is finite; hence, Alice needs only finite thermodynamic resources to employ it.

To read the pointer value of *S*, Alice also needs a finite set $\{M_j^{(P)}\}$ of *pointer* observables. These include not just the usual "meter readings" but also whatever is indicated by any adjustable control settings that serve to "prepare" the state of *S*. The outcome values $\{x_j^{(P)}\}$ of these observables are to be discovered, and in the case of control settings perhaps adjusted, so they are not specified in advance. For macroscopic systems, in practice, there are many more reference observables for any system than pointer observables: *identifying* a system against the inevitably cluttered background of the world requires more measurements, and hence more energy, than checking the variable parts of the state of the system once it has been identified (Figure 3).

$$\{M^{(R)}{}_i\} \qquad\qquad \{M^{(P)}{}_j\}$$
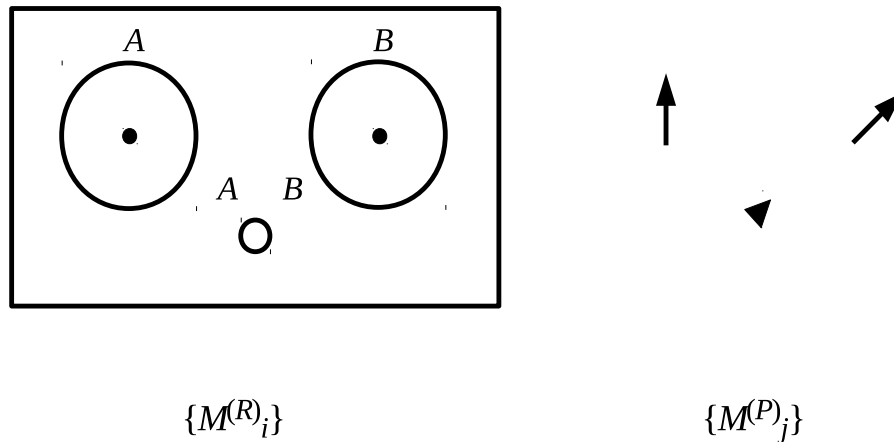
Figure 3. Identifying a system requires many more observables than are required to measure its pointer state, including the positions indicated by control knobs or other "preparation" settings. Adapted from [67] Figure 3.

The set $\{M_i^{(R)}, M_j^{(P)}, x_i^{(R)}\}$ of operators and expected values constitutes *semantic* information; it specifies a referent in W. It specifies, in particular, the desired system S. Moore's theorem [16], however, renders this referent intrinsically ambiguous. It can be thought of as the time-varying equivalence class of all components of W that satisfy $\{M_i^{(R)}, x_i^{(R)}\}$, i.e., all components of W that yield $\{x_i^{(R)}\}$ when measured with $\{M_i^{(R)}\}$, for all i, and can also be acted upon by $\{M_j^{(P)}\}$. This class can be represented, from the perspective of an observer deploying $\{M_i^{(R)}, M_j^{(P)}, x_i^{(R)}\}$, as a superposition of "systems" having dimension at least the cardinality of full set $\{M_i^{(R)}, M_j^{(P)}\}$ of operators [17]. Identifying S is, in fact, identifying such a superposition, as opposed to an observer- and observation-independent "thing" as assumed in classical physics.

As noted, the life sciences have treated observers as interacting not with pre-defined systems but with their worlds throughout their history; the continuing emphasis on understanding such interactions as methods shifted, over the 20th century, from strictly observational to experimental is reflected in continuing calls for "naturalism" and "ecological validity" (e.g., [68]). From a cognitive-science perspective, when Alice explores the laboratory in search of S, she is just doing what organisms do when exploring a novel environment. It is, moreover, clear from this perspective what the observables and values $\{M_i^{(R)}, M_j^{(P)}, x_i^{(R)}\}$ are: they specify a collection of semantically-coherent *categories* and a (partial) instance of this collection. The reference observables together specify what *type* of system S is, and their fixed values identify a particular instance or *token* of this type as S (for a general review of types and tokens, see [69]). These token-specifying values cannot change, or can change only very slowly and gradually, so long as S remains S. For Alice to search for a voltmeter with a specific size, shape, color and brand name, for example, she must know, in some sense to be determined empirically, what a voltmeter is, and that anything qualifying as a voltmeter has a specific size, shape, color and brand name, as well as some specific mass, surface texture, and layout of knobs or buttons, dials or digital displays, connectors, and so forth. She must, moreover, understand what voltmeters are used for, and that the pointer observables $\{M_j^{(P)}\}$ return the values of such properties as selector-switch positions and meter readings. Sets of observables and specified outcome values are therefore, from this perspective, *structured knowledge*; observers are systems—in this case, cognitive systems—capable of deploying such knowledge. Philosophers and, more recently, psychologists have expended considerable effort characterizing this categorical knowledge, investigating its implementation, and determining how it is used to identify and then

re-identify objects over time (for reviews, see [70–74]); we will return to the question of implementation in Section 5 below.

Cognitive science, therefore, tells us something important about the *physics* of observers: an observer needs sufficient degrees of freedom to represent both the category and the category-instance distinctions that are required to both identify and measure the pointer states of any systems that it can be regarded as observing [67]. Observers also need access to, and a means of acquiring and incorporating, the energy resources needed to register the outcome values of their observations, and they need a means of dissipating the waste heat. Finally, observers need a control structure that deploys their observables in an appropriate order. Observers cannot, in other words, be mere abstractions: they have physical, i.e., energetic, structure, and they must process energy to process information. Coordinate systems cannot be observers. The degrees of freedom that register the pointer state of a given meter on a given instrument cannot, by themselves, constitute an observer; additional degrees of freedom that encode categorical knowledge and others that manage control and energy input and output are mandatory. The former can be considered *memory* degrees of freedom; they are further characterized in Section 5.

Cognitive science also allows us to reformulate Zurek's question of where the systems come from to the question of where the categories that allow system identification and pointer-state measurements come from. This latter question has answers: evolutionary and developmental biology for observers that are organisms, and stipulation by organisms for observers that are artifacts. In the case of human observers, some of these categories (e.g., [face]) are apparently innate [75]; others (e.g., [chair]) are learned early in infancy [76], while still others (e.g., [voltmeter]) require formal education. For observers that are neither organisms nor artifacts, and ultimately for organisms and artifacts as well, the "where from" question demands an account of cosmogony. The relationship between physics and the life sciences is, in this case, not one of reduction but rather one of setting boundary conditions, as recognized by Polanyi [77] among others. It is, in other words, a historical as opposed to axiomatic relationship. The ontological consequences of this reformulation for a theory of observation are considered in Section 8 below.

4. What Information Is Collected?

The goal of observation is to get information about the *state* of what is observed. State changes are the "differences that make a difference" [78] for observers. In the traditional picture of Figure 1, the goal is to discover the pre-existing, observer-independent pointer state $|P\rangle$ of a given, pre-existing system S. We have seen above, however, that to measure $|P\rangle$, one must first measure the time-invariant reference state $|R\rangle = |x_1^{(R)}, x_2^{(R)}, ..., x_n^{(R)}\rangle$ by deploying the n reference observables $\{M_i^{(R)}\}$; hence measuring $|P\rangle$ requires measuring the entire state $|S\rangle = |R \oplus P\rangle$. As this identifying measurement is made *by the observer*, S cannot be considered given. The state $|S\rangle$ is, therefore, observation- and hence observer-dependent. To paraphrase Peres' [48] well-known aphorism, "unidentified systems have no states".

As in the case of Alice searching the laboratory, measuring $|S\rangle$ requires trying the $\{M_i^{(R)}\}$ out on many things besides S, some of which will yield some of the specified reference outcomes $\{x_i^{(R)}\}$ but not all of them. These extra measurements are *overhead*; they cost energy and time. As the complexity of W increases, this overhead expense increases with it.

This notion of overhead allows us to distinguish between two types of observers:

- "Context-free" observers that *waste* their observational overhead.
- "Context-sensitive" observers that *use* (at least some of) their observational overhead.

As an example context-free observer, consider an industrial robot that visually identifies a part to pick up and perform some operation on. The robot must identify the part regardless of, e.g., its orientation on a conveyor, but has no use for any "background" information that its sensors detect. Except for its assigned part, everything about its environment is noise. It can afford to be

context-free because it only has to deal with one, completely well-defined context: that of identifying a particular part on a conveyor and picking it up. Its control system is, in a sense, trivial: it needs to do only one thing, which it can do in the same way, up to minor variations, every time. It does not, in particular, have to worry about the frame problem [79], the problem of predicting what does *not* change as a consequence of an action or, in its more generalized (and controversial) form, the problem of relevance [80]. Nothing that happens outside of its context matters for *how* it performs its task; however, as we will see in Section 6 below, outside happenings do matter for *whether* it performs its task.

Alice, walking around the laboratory, has different requirements and a different experience. She encounters many things roughly the right size and shape to be *S*, but that are not voltmeters, and perhaps some other voltmeters, but not the right brand or not connected to anything. Hence Alice, even before finding *S*, knows much more about the laboratory than when she entered. Her next pointer-value reporting task will be much easier to accomplish after this initial foray in search of *S*. For Alice, the search overhead is valuable. It is valuable in part because Alice *does* have to worry about the frame problem; every action she takes may have unintended and possibly unpredictable side effects relevant to her [81].

What is the difference between these two cases? Let us assume, for the sake of argument, that Alice and the robot have exactly the same input bandwidth: both record the same number of visual bits. The robot subjects these bit strings to a single analysis that returns 'yes' or 'no' for the presence of the target part. This analysis has multiple components, e.g., specification of a three-dimensional shape together with rotations and projections that accomplish "inverse optics" from the visual image (see e.g., [82] for an implementation of recognition of not one but several objects). These constitute the robot's reference measurements $\{M_i^{(R)}\}_{Robot}$; the pointer measurements then specify the position for grasping the part. Identifying the part requires all of the reference measurements to agree; all other scenes or scene components are "negatives" and are ignored. For this robot, objects of the "wrong" shape or size are task irrelevant and so effectively invisible.

Alice's analysis of her visual bits is superficially similar: she also employs a set of reference measurements $\{M_i^{(R)}\}_{Alice}$, agreement among all of which constitutes identifying *S*. Alice does not, however, ignore the negatives. They are not invisible to her; she has to see them to navigate around the room. This requirement for autonomous navigation already distinguishes Alice from the robot; Alice must deal with a different context, filled with different objects, every time she moves. She must, moreover, devote some of her observational overhead to observing *herself* in order to update her control system on where she is relative to whatever else she sees and how she is moving. The robot has no need for such self-observation (though again see [82] for a robot that must distinguish its own motions from those of another actor).

Alice can, moreover, classify some of the "negatives" she encounters as things that share some but not all of their characteristics with *S*. These similar-to-*S* things are not vague or undefined in whatever characteristics they do not share with *S*; Alice is not limited, as the robot is, to the "right" size, shape, or color and "other". Alice, unlike the robot, is capable of recognizing many different objects not just as "other" but as individuals; this tells us that she has a much larger set of deployable measurement operators, of which the $\{M_i^{(R)}\}_{Alice}$ that identify *S* are a tiny fraction. Her ability to group similar objects, e.g., to group voltmeters of different sizes, shapes and brands, tells us that her measurement operators are organized at least quasi-hierarchically. Whereas the robot needs only a single category, [my-target-part], Alice has an entire category network incorporating a large number of types, in many cases associated with one or more tokens. These types and tokens are, moreover, characterized by both abstraction and mereological relations (see e.g., [83,84] for reviews).

For human observers, categories are closely linked to, and commonly expressed in, language. Visual category learning is by far the best investigated; here, it is known that human infants can identify faces from birth [75], track moving objects by three months [85], and learn hundreds of initially-novel object categories by the onset of language use, approximately one year [76]. "Entry-level" categories

such as [dog], [person], [chair], or [house] are learned first and processed fastest [86]. Category learning accelerates with language use [87] and later, formal education, resulting in word repertoires in the tens of thousands (up to 100,000 in rich languages like English [88]) in adulthood and category repertoires somewhat smaller due to word redundancy. Multiple, quasi-independent systems contribute to learning distinct kinds of categories, e.g., perceptible objects versus abstrata [89]. New categories, like new words, can be constructed combinatorially, with no apparent in-principle limits beyond finite encoding. Human observers can, therefore, be viewed as encoding tens of thousands of distinct "observables" with sets of specified outcome values that identify subcategory members or individuals. The deployment of these categories on both input (i.e., recognition) and output (language production, bodily motion, etc.) sides is highly automated and hence fast [90] except for novel objects that must be categorized by examination or experimentation. While the categorization systems of non-human animals are not well studied and their lack of language suggests far smaller, niche-specific repretoires, their evident practical intelligence suggests robust and highly-automated categorization systems.

The ability to deploy a large number of different measurement operators, the results from which can contribute combinatorially to behavior is also ubiquitous in "simple" biological systems. Archaea and bacteria, for example, employ from a handful to well over a hundred different sensing and regulation systems, generally comprising just one [91] or two [92] proteins; the numbers of such systems roughly correlate with environmental niche complexity. These systems primarily regulate gene expression and hence metabolism, but also regulate motility [93] and aggregative and communal behaviors [94]. Eukaryotic signal transduction pathways are more complex and cross-modulating (e.g., Wnt [95] or MAPK [96]), typically forming "small-world" or "rich-club" patterns of connectivity [97,98] (Figure 4). Large numbers of distinct sensors—typically transmembrane proteins—are expressed constituitively, so the expressing cell is "looking" for signals that these sensors detect all the time. Especially in eukaryotic cells, the result of detecting one signal is in some cases to express other sensors for related signals, the cell-biology equivalent of "looking more closely" or "listening for what comes next".
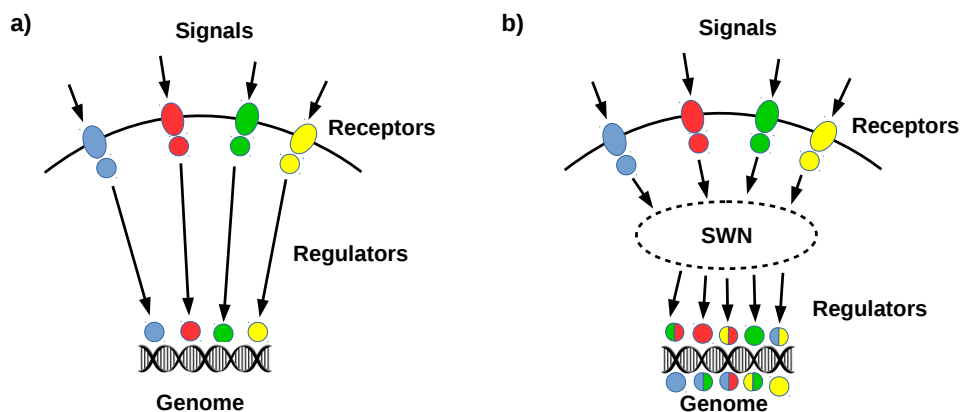
Figure 4. Simplified cartoon of sensing and regulation systems in (**a**) prokaryotes (archaea and bacteria) and (**b**) eukaryotes. Regulatory signals can interact and cross-modulate in eukaryotes more than in prokaryotes. Colors indicate communicated information. SWN = small-world network [99].

Context-sensitivity allows the response to one signal to be modulated by the response to another signal. Sets of signals identify systems; hence, context-sensitivity allows the response to one system to be modulated by responses to other systems. If Alice sees smoke coming out of a system to which *S* is connected, it will affect her report on the pointer state of *S*. To return to the language of physics, it will affect the probability distribution over her possible reports, by assigning high probabilities to reports that had low or zero probability before. Dzhafarov has termed this ubiquitous effect of context

on response probabilities "contextuality by default" [100,101]. Within this framework, the ideal of a fixed probability distribution on outcomes, impervious to the modulatory effects of other observations, becomes a minor special case. Contextuality is a hallmark of quantum systems, and the experimental recognition of ubiquitous contextuality in human cognition has motivated the adoption of quantum theoretic methods to psychological data (see [102–104] for reviews).

Most mammals, many birds [105] and at least some cephalopods [106] are not, however, just sensitive to context in real time; they are also able to notice and store currently task-irrelevant information that may be relevant to a different task at some time in the future. The combined use of landmarks, cognitive maps, and proprioception in wayfinding is a well-studied example [107,108]. Humans are particularly good at storing task-irrelevant information—for Alice, the other contents of the laboratory and their general layout—for possible future use. This ability underlies the human ability to discover Y while searching for X and is thus a key enabler of scientific practice.

5. What Is Memory?

As pointed out earlier, observers must, in general, devote some of their degrees of freedom to encoding the observables that they are able to deploy, as well as the reference outcomes used to identify systems. They must, moreover, devote degrees of freedom to the implementation of their control systems and to energy acquisition and management systems. Such dedicated degrees of freedom can be considered *memory* in the broadest sense of the term (but see [109] for an argument that this usage is too broad). Adventitious information, e.g., pointer-state outcomes not directly relevant now, collected by context-sensitive observers is also clearly of no use unless it can be remembered. Memory is often listed as an attribute of observers by physicists (e.g., [28,35,50]); clearly, any observer that prepares a system before observing it, or that re-identifies the same system for a repeat observation requires a memory [67].

Remembering a system S requires remembering the set $\{M_i^{(R)}, M_j^{(P)}, x_i^{(R)}\}$ that identifies it and enables measuring its pointer state. The $\{M_i^{(R)}, M_j^{(P)}\}$ are descriptors or, in quantum theory, operations: tests to which the world W is subjected. The $\{x_i^{(R)}\}$ are expectations about the outcomes of the reference tests that enable system identification. As noted earlier, the $\{M_i^{(R)}, M_j^{(P)}\}$ are effectively categories, specified by specifying their membership criteria.

The question of *how* such remembered information is encoded leads naturally to the ontological question of what kinds of "encodings" exist. Is there, for example, such a thing as a memory engram [110]? Newell [111] postulated that cognitive agents are "physical symbol systems": but are there such things as symbols? Most approaches to cognitive science postulate some form of representation [112]: do such things exist? The semiotic tradition postulates the existence of signs: are they real? Ecological realists postulate affordances encoded by an animal's environment [113]: do these exist? Is, indeed, *information* real in any sense? It is important to distinguish at least three formulations of these questions. One can ask whether the "existence" or "reality" of signs, symbols, representations or affordances is meant to be: (1) "objective" in the sense of observer- and observation-independent; (2) observer- and observation-relative; or (3) a convenience for the theorist, a "stance" [114] that aids description and prediction while making no ontological commitments. We defer the first of these questions to Section 8 below, here only noting its similarity to the question of whether "states" are observer- and observation-independent discussed above, and the second and third questions to Sections 6 and 7, respectively. Here, we focus on the initial "how" question, which is more tractable experimentally.

Influenced in part by J. J. Gibson's [113] idea of "direct uptake" of information from the environment, some theorists within the embodied/enactive tradition have concluded that "memories" can be stored entirely within the world; organisms, on this view, simply have to look to find out what they need to know (see [115–117] for reviews). Taken literally, this is another version of the hypothesis that systems and their states are "given" a priori. An observer required to identify systems of interest

must remember *how* to identify them. It must, in other words, remember the $\{M_i^{(R)}, M_j^{(P)}, x_i^{(R)}\}$. It cannot, on pain of circularity, store this "how" memory in W.

The alternative, "good old-fashioned AI" (GOFAI) view that cognitive systems store their memories as a collection of "beliefs" or as "knowledge" encoded in a "language of thought" (LoT) [118], including storing category networks as networks of connected "concepts" represented by either natural language or LoT words, has increasingly given way to a more implicit view of memory as a collection of *ways of processing* incoming signals. This shift to a "procedural" view of even "declarative" memories has been driven mainly by cognitive neuroscience, and comports well with "global workspace" models of neurocognitive architecture as comprising networks of networks, with small-world structure at every scale [119–124]. Functional imaging studies of pre- and perinatal human infants show that this architectural organization develops prenatally and is already functional at birth [125–127]. "Representations" in such models are network-scale activity patterns, which are reproducibly observable, specifically manipulable experimentally, and in some cases specifically disrupted either genetically or by pathology (for a review of such models and manipulations in the specific case of autism, see [128]). The only "symbolic" content, on this newer view, are the experienced outcomes themselves [129]. The best-supported current candidates for the implementation of such an implicit memory are multi-layer Bayesian predictive-coding networks, in which categories are effectively collections of revisable expectations about the perceptible structure of W [130–134]. Such models can be directly related to small- as well as large-scale neural architecture [135,136] and have achieved considerable predictive success in such areas as vision [137], motor control [138] and self-monitoring [139]. Cognitive systems powered by such networks are intrinsically exploratory observers, trading off category revision and hence learning to recognize new kinds of systems or states against behavioral changes that enable continuing interactions with familiar kinds of systems and states.

The gradual rejection of explicit-memory models of cognition has been paralleled at the cellular and organismal levels by an increasing recognition that explicitly-encoded genetic memory is only one of many layers of biological memory [140–143] (Figure 5). This expanded view of memory is broadly consistent with thinking in the biosemiotic tradition, e.g., with the idea that "any biological system is a web of linked recognition processes" at multiple scales ([144] p. 15), particularly within the "physical" and "code" approaches to biosemiotics identified by Barbieri [145]. It frees the genome from the task of "remembering" how receptors are organized in the cell membrane or how pathways are organized in the cytoplasm; such information is maintained in the membrane and cytoplasm themselves, and is automatically passed on to offspring when these compartments are distributed between daughter cells by the process of cell division. Preliminary work suggests that this extended, supra-genomic biological memory can be productively modelled as implementing Bayesian predictive coding [146].

From this perspective, to *recall* a memory is simply to use it again: one recalls how to identify S when one identifies S. System identification and measurement are competences remembered as "know-how". Human observers can also formulate *descriptions* of these competences; in our running thought-experiment, the identification criteria for S given to Alice are such a description, as is the notation $\{M_i^{(R)}, M_j^{(P)}, x_i^{(R)}\}$. From the present perspective, these descriptions are themselves observational outcomes, and the process of generating them is a process of observation, implemented in humans by "metacognitive" measurement operators. This view accords well with the reconstructive nature and recall-context dependence of even episodic memories (see [147] for review) and with Chater's "flat" conception of the mind as a representation of outcomes only [129]. "Knowing-that" is knowing how to retrieve "that" by observation—by querying some memory, external or internal [148]—when needed.
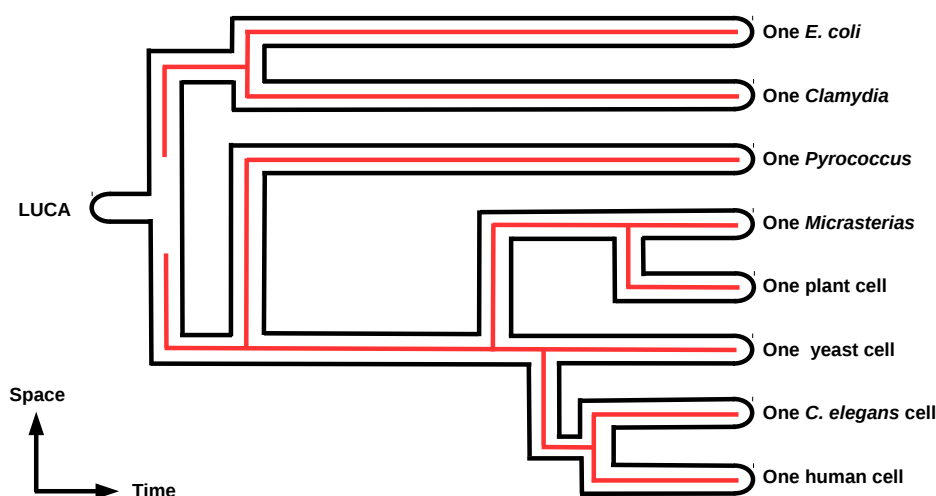
Figure 5. A phylogenetic tree showing the cell membrane (solid black lines) and cytoplasm (enclosed space) as well as the genome (red lines) from the last universal common ancestor (LUCA) to the present. Whether LUCA has a well-defined genome is left open. All compartments of the cell are continuous from LUCA to the present across all of phylogeny; hence, all cellular compartments can serve as memories.

6. What Is Observation for?

Evolutionary theory suggests that organisms make observations for one purpose: to survive and reproduce. The requirements of survival—obtaining and incorporating resources, self-maintenance and repair, an ability to detect and escape from threats—are seen by many as the key to making information "for" its recipient or user [116,149]. Roederer [150,151] similarly restricts "pragmatic information"—information that is *useful* for something, and hence can be considered to have causal effects—to organisms. This would suggest that outcomes can only be "for" organisms, and indeed that it only makes sense to consider organisms observers; writing from a biosemiotic perspective, Kull at al. make this explicit: "the semiosic/non-semiosic distinction is co-extensive with life/non-life distinction, i.e., with the domain of general biology" ([149] p. 27), here identifying observerhood with sign-use. It was this conclusion, in large part, that motivated Bell's rejection of observation and measurement, and hence information, as fundamental concepts in physics ([51] p. 34):

> *What exactly qualifies some physical systems to play the role of 'measurer'? Was the wavefunction of the world waiting to jump for thousands of millions of years until a single-celled living creature appeared? Or did it have to wait a little longer, for some better qualified system ... with a PhD?*

Wheeler famously took the opposite tack, making "observer-participancy" the bottom-level foundation of physics [33]; Wheeler's insistence on information and *active* observation ("participancy") as the sole ontological primitives of physics motivated not only the quantum-information revolution described in Section 2 above, but also the more recent drive to derive spacetime itself from processes of information exchange [152–158]. Observation, it is worth re-emphasizing, implies recordable observational outcomes and hence *classical* information. In practice, most physicists do not worry about where to put the "von Neumann cut" at which information is rendered classical, and hence regard ordinary laboratory apparatus as "registering outcomes". But are these outcomes *for* the apparatus? Should the apparatus be considered an observing "agent"? Do they have an effect on what the apparatus does next?

For organisms like *E. coli*, all observational outcomes are directly relevant to survival and reproduction. How its observational outcomes affect its behavior have, in many cases, e.g., flagellar

motility [93] or lactose metabolism [159], been worked out in exquisite detail. These functions have, however, been worked out by *us* from our perspective, using our capabilities as observers and theorists. *E. coli* itself has no ability to determine by observation how it changes direction or digests lactose; it has, as Dennett [160] puts it, competence without comprehension. It has no knowledge that its actions are "for survival and reproduction", though we can infer, using our understanding of the world in which both we and it live, that they are. Its observations and its current state together determine its actions, but it is the *world* that determines whether it will survive and reproduce. It has no ability to do the experiments that could reveal this causal connection. Saying that *E. coli*'s observational outcomes are "for" it is using a purely third-person sense of "for"; we could as well say that its observational outcomes have consequences, imposed by the world, that affect it.

Setting suicide and voluntary sterilization or celibacy aside, whether human observers survive and reproduce is also determined by the world. Our cognitive organization permits us to regard our observational outcomes as for us, but this ability can be lost, e.g., in psychosis [161], insular-cortex seizures [162], or Cotard's syndrome [163], and losing it does not prevent their consequences being "for" us in the third-person sense above. Unlike *E. coli*, we have some ability to explicitly associate observational outcomes with their consequences either pre- or postdictively, but this ability is limited and often prohibitively expensive. We are competent in many observational feats, from understanding natural languages to recognizing individual people decades after last seeing them, with little understanding of how we achieved or how we implement that competence [90,129,164,165]. Indeed, our understanding of *how* we do things tends to vary inversely with our competence; we can often explain cognitive abilities learned slowly and painfully through extensive practice, e.g., computer programming, but cannot explain abilities learned easily and automatically, e.g., grammatical sentence production in our native language. We are sometimes aware of our competence only by testing it, i.e., by performing further observations to determine whether our previous performances were competent. We often have *feelings* of competence, but they are unreliable and often spectacularly wrong (see [164] for various examples).

If an observational outcome is *for* an observer, it is natural to regard that observer as an *agent* that acts intentionally on the world to obtain an outcome. Here, again, Kull et al. make this explicit: a semiotic agent is "a unit system with the capacity to generate end-directed behaviors" ([149] p. 36), including the acquisition of information. It is, however, useful to ask whether a system is an agent from its own, first-person perspective—whether it is able to self-monitor its goals and agentive activity—or whether it is only an agent from our third-person theoretical perspective. Human observers have an essentially irresistible (i.e., highly automated) tendency to attribute agency to anything, animate or inanimate, exhibiting any but the simplest of motions, a tendency that develops in infancy and appears in every culture examined [166–169]. Humans self-monitor and hence experience their own agentive activity as agentive, but tend to over-attribute agency, in the sense of having reasons for actions, to themselves as well as to others (see [129] for examples). Hence, our third-person attributions of agency to other systems, whether Heider and Simmel's animated circles and triangles [166], *E. coli*, or each other, are of questionable reliability. Given our lack of access to the first-person perspective of other systems, however, we are left only with third-person attribution as a basis for theory construction. It is not, therefore, clear that characterizing something as an "agent" adds anything to its description beyond the claim that its observational outcomes have consequences for it, with "for" used in the third-person sense above.

Let us now consider the role of observational outcomes in influencing the survival and reproduction of artifacts like voltmeters or context-free industrial robots. When a voltmeter obtains an observational outcome, it "registers" it *for us* by displaying it on some output device, typically a meter or a digital display. If the voltmeter's behavior is erratic, we may attempt calibration or repair; if it remains erratic, we may discard or recycle it. These are consequences *for the voltmeter*; *it* ceases to exist as an organized entity, not something else. If we or fate destroy it, it has not survived. If, on the other hand, the voltmeter proves extremely reliable, we may buy another one like it. The voltmeter is

a "meme" in the broad sense defined by Blackmore [170]: a cultural artifact that can be reproduced, particularly one that can be reproduced by the accurate and efficient process of reproduction from recorded instructions, as opposed to by direct copying. The voltmeter's world, that which determines whether it will survive and reproduce, i.e., be reproduced, includes not just us but also numerous non-human actors, from curious housepets to earthquakes. As with any meme, the consequences of the voltmeter's actions for it (third-person "for") are different from the consequences of its actions for us (first- or third-person "for"). In this, the voltmeter is like *E. coli*, whose actions may be beneficial to it but detrimental to us or vice-versa, and indeed like most organisms.

For any observer, non-survival is not an observational outcome, but is rather the cessation of observational outcomes [62]. Survival is, therefore, the continuing of observational outcomes. Hence, we can reformulate the standard evolutionary goal of survival and reproduction as the general statement:

> *For any observer, the goal of observation is to continue to observe.*

This goal is a third-person theoretical attribution for almost all observers, including any humans who have never thought of it or do not believe it. If observer-participancy is the foundation of physics as Wheeler proposed, however, the goal of continuing to observe is the fundamental principle of cosmogony [33,171].

7. A Spectrum of Observers

What, then, is an observer? The physics of observation, as pointed out in Section 2, places no limits; any physical interaction can be considered observation. A proton moving through an accelerator samples the electromagnetic field at every point and behaves accordingly. Its observational outcomes—the field values, as detected by it—have consequences for its behavior. Indeed, they have consequences, in the context of human culture, for the production of more (unbound) protons and more accelerators; for us, the proton is a meme. The proton has no choice but to observe, and no choice beyond the freedom from local determinism granted by the Conway–Kochen "free will" theorem [172] of how to respond, but in this it differs little from *E. coli* observing and responding to an osmolarity gradient. Thinking of the proton as "computing" its trajectory is unhelpful [173], but is no more anthropomorphic than thinking of it as "obeying" a physical law. Thinking of the proton as observing and responding to its environment—as physicists naturally do when they describe it as "seeing" the electromagnetic field—is perhaps the least anthropomorphic approach.

Voltmeters and many other artifacts are designed by us to be observers. Their observations are far more fine-grained and accurate than we could achieve directly, and in many cases they probe phenomena to which we are otherwise insensitive. Such artifacts are context-free by design, and are likely to be discarded if they become irreparably context-sensitive. Like the proton, they have no choice beyond the Conway–Kochen prohibition of local determinism in what they do, but unlike the proton, it is often helpful (to us) to think of them as computing. It is worth noting, here, that while all artifacts are at least potentially and approximately reproducible and are hence memes, not all memes are observers. Abstract memes like coordinate systems, in particular, are not observers as noted earlier. Neither are words, symbols, representations, or other abstracta.

Some artifacts we want to be context-sensitive observers, and we expend considerable effort trying to equip them to be context-sensitive. We want them to recognize novelty unpredicted and possibly unpredictable by us, and to respond in ways unpredicted and possibly unpredictable by us. We want autonomous planetary-exploration rovers, for example, to take advantage of whatever circumstances they encounter, without the need for time-consuming communication with Earth. Explicitly evolutionary, developmental, and psychological thinking appears essential to the development of such artifacts [174–177].

While it has yet to become commonplace, treating cells, whether prokaryotic or eukaryotic and whether free-living or components of multicellular systems, not just as observers and actors but as

cognitive agents has become progressively more widespread and productive [178–181]. The risky and quite-literally shocking experience of action-potential generation by neurons has been proposed as the basis of awareness in organisms with brains [182]; this idea is easily extensible to the risky experiences of sudden osmolarity, membrane-voltage, or metabolic changes common in the unicellular world, or to the massively-risky experience of cell division. Canonical cognitive processes including learning [183], communication [184,185], memory and anticipation (priming) [186] are now often, though not yet routinely, used to characterize plants. They are, once again, used by *us* to make this characterization; we have no evidence that plants regard themselves, metacognitively, as learning, communicating, or remembering. It is interesting to note that less than 20 years ago, the idea that other mammals [187] and even human infants [188] were cognitive agents aware of their environments—observers, in other words—required vigorous defence. Not just human- or primate- but even animalocentrism about observation is fading, as is the notion that metacognitive awareness is required for awareness.

Multicellularity forcefully raises the question of how observers cooperate, compete, and possibly even coerce each other to form a larger-scale observer with its own capabilities, interests and boundary conditions [189–191]. Similar processes operate at every scale thus-far investigated, up to interactions between human population-culture combinations. A central finding of evolutionary developmental ("evo-devo") biology is that successful large-scale evolutionary transitions, e.g., from invertebrate to vertebrate body plans, involve the duplication and modification of modular packages of genetic instructions, typically instructions specifying where and when to express large sets of other genes [192,193]. While evo-devo thinking has entered psychology [194], what counts as a "module" and how they can be identified across species is considerably less clear. Whether duplication-with-modification mechanisms acting on knowledge "modules" can be demonstrated in, e.g., category learning, remains to be determined. The common use of analogy in abstract category learning [195–197] suggests that this may be common. The ubiquity of duplication-with-modification of instruction modules as a mechanism for producing progressively more complex memes is well recognized [160,170].

What happens above the scale of terrestrial populations and their cultures? Blackmore has argued that we are far more likely to encounter supra-terrestrial memes than organisms [170]; indeed, all SETI searches are searches for memes. Beyond familiar measures of non-randomness, however, we have few resources for meme-identification. If the Crab nebula were an artwork, could we identify it as such?

8. An Observer-Based Ontology

Every observer inhabits a perceived world, what Sellers [198] termed the "manifest image" for that observer (cf. [160]). The identifiable components of this world constitute the observer's "naive" ontology. This ontology contains everything that the particular observer it characterizes can detect, including objects, other organisms, environmental features, signs, words, affordances, memes, etc. It also contains whatever the particular observer can detect by self-monitoring, e.g., pains, pleasures, emotions, and feelings of believing, knowing, representing, owning, agency, or passivity. With this definition, the manifest image of any observer is clearly "personal" to it; it encompasses its possible first-person experiences. As described in Section 3 above, the observer identifies the components of its manifest image with finite sets $\{M_i^{(R)}, M_j^{(P)}, x_i^{(R)}\}$ of operators and expected values. Hence, specifying an observer O and the complete set $\{M_i^{(R)}, M_j^{(P)}, x_i^{(R)}\}$ of operators O can deploy and outcome values O can register specifies a (observer, world) pair.

As emphasized by von Foerster [5] among others, observation is an intrinsically symmetrical process. We can, therefore, ask how some observer O's world identifies O. As O's world can be considered an observer, it can be considered to deploy operators and expected values $\{Q_i^{(R)}, Q_j^{(P)}, q_i^{(R)}\}$ to identify O and characterize O's components and behavior. Specifying $\{M_i^{(R)}, M_j^{(P)}, x_i^{(R)}\}$ and $\{Q_i^{(R)}, Q_j^{(P)}, q_i^{(R)}\}$ thus specifies the "universe" comprising O and O's world entirely in terms of observations and outcomes, with no additional "systems" or "representations" of any kind needed.

The operators $\{M_i^{(R)}, M_j^{(P)}\}$ and $\{Q_i^{(R)}, Q_j^{(P)}\}$ specify the interaction between O and O's world, an interaction conceptually localizable to the "boundary" between them. The outcomes $\{x_i^{(R)}, x_j^{(P)}\}$ and $\{q_i^{(R)}, q_j^{(P)}\}$ obtained by O and O's world, respectively, are "written on" O's and O's world's, respectively, sides of this boundary [62]. Locating a *boundary on which outcomes are written* is, therefore, locating an observer; indeed, it is locating two interacting observers, each complementary to the other.

An observer may also be identified as a "system" by another observer, e.g., Alice may be identified by Bob. The outcomes of Bob's observation of Alice are, in this case, encoded on Bob's boundary with Bob's world, in which Bob sees Alice as embedded. Moore's theorem, as always, renders Bob's identification ambiguous: "apparent Alice" for Bob may be a different collection of degrees of freedom than "apparent Alice" for Charlie. Bob and Charlie cannot determine that they are observing the same Alice by increasing their measurement resolution [199]; to obtain reliable evidence of a shared system, they must contrive to violate a Bell inequality.

Conceptualizing observation in terms of operations *at* a boundary with outcomes encoded *on* that boundary makes explicit the fundamental epistemic position of any observer: observers can have no information about the *internal* structures of their worlds. This statement is familiar as the holographic principle, first stated by 't Hooft for black holes: "given any closed surface, we can represent all that happens inside it by degrees of freedom on this surface itself" ([200] p. 289). Hence, holography joins the principles of Boltzmann, Shannon, and Landauer as a fundamental principle of any theory of observation. The holographic principle reformulates and strengthens the ambiguity of system identification proved by Moore. As 't Hooft emphasizes, even *metric* information is inaccessible to observers outside of a closed system: "The inside metric could be so much curved that an entire universe could be squeezed inside our closed surface, regardless how small it is. Now we see that this possibility will not add to the number of allowed states at all." A boundary may encode *apparent* metric information, e.g., distances between systems "inside" the boundary, but both the apparent distances and the apparent systems are only observational outcomes encoded on the boundary itself.

A general account of perception explicitly compliant with the holographic principle has been developed by Hoffman and colleagues [201,202]. This "interface theory" of perception (ITP) postulates that percepts are compressed, iconic encodings of fitness information that are not homologous to, and encode no information about, structures in the world other than fitness. It assumes, in other words, that all information an observer is capable of obtaining is relevant to survival and reproduction, an assumption consonant with the significance of the frame problem for context-sensitive observers discussed in Section 4 above. The interface—the particular set of available icons and their behaviors—is species- and even individual-specific. ITP is supported by extensive evolutionary game theory simulations showing that agents sensitive only to fitness outcompete agents sensitive to world structures other than fitness [203] and by theorems showing that perception–action symmetries will induce apparent geometric and causal structures on worlds that lack such structures intrinsically [201,204].

As ITP implicitly assumes that perceiving agents are aware of what they perceive, Hoffman and Prakash [205] have proposed an ontology of "conscious agents" (CAs) that implement ITP. CAs comprise an interface through which they perceive and act on their world, together with a "decision" operator, modelled as a Markov kernel, that links perceptions to actions. The world with which any CA interacts is itself a CA or, equivalently, an arbitrarily-complex network of CAs; hence, the CA ontology comprises interfaces, each with an "internal" decision operator, linked bidirectionally by perception and action operators. It is, therefore, an ontology of boundaries and interactions as described above. Finite networks of CAs have the computational power of finite Turing machines, and networks of sufficient size can straightforwardly implement canonical cognitive processes including memory, categorization, attention, and planning [206].

The most fundamental problem that any theory of observation consistent with quantum theory faces is that of *unitarity*: the unitary dynamics required by quantum theory conserves net information,

just as it conserves net energy. Any *net* information present in the universe at any time must, therefore, be present as a boundary condition on the universe's initial state, and must be equally present as a boundary condition on its final state. This is the notorious "fine-tuning" problem typically addressed with some form of anthropic principle (e.g., [207]). Postulating observer-independent systems (as in Zurek's "axiom(o)"; Section 3), categories, or even observers themselves falls afoul of this problem; whatever information is required to specify the assumed entities, data structures, or operations must be included in past and future boundary conditions. The simplest solution to this problem, clearly, is for the universe to contain zero net information. Boundary or interface ontologies that comply with the holographic principle escape this problem provided two conditions are met: (1) information encoding on boundaries must be both signed and symmetric, so that the total information / entropy and energy transfers across any boundary are zero, and (2) every possible boundary must be allowed. The latter condition is consistent with Wheeler's postulate of observer-participancy: *every* characterizable system is an observer; the former enforces unitarity locally at every boundary. Both the formal and conceptual consequences of these conditions remain to be investigated; some initial considerations are discussed in [62].

9. Conclusions

When physical interaction is reconceptualized in informational terms, it becomes observation. The conflicts between this reconceptualization and both physical and psychological intuitions became obvious with the advent of quantum theory in the early 20th century, with the measurement problem and the explosion of competing interpretations of quantum theory as the result. Since the 1960s, the life sciences have investigated the implementation of observation by organisms at multiple scales with ever-increasing precision. Examining organisms tells us not only how observation works, but what it is for. Moreover, it reveals that observation is ubiquitous in nature and strongly linked to fitness. Its greatest contribution, however, is to emphasize and make obvious that observers must actively distinguish "systems" from the environments in which they are embedded. It is systems and their state changes that carry meaning for observers. Requiring that observers be capable of identifying systems removes the possibility of treating "the observer" as a mere abstraction or as a system of irrelevant structure. Indeed, the structure of any observer determines the measurement operators it can deploy, and the observational outcomes it can register.

As Dodig Crnkovic [11] also emphasizes, recognizing observation as a *relation* between observer and observed leads naturally to an observer-relative and observation-dependent "observed reality": an individual-specific *Umwelt* [208], "image" or interface. Moore's theorem renders this interface a boundary that cannot be looked behind. If interaction is observation, observational outcomes are holographically encoded on this boundary. An ontology of boundaries supporting interactions naturally emerges. "Systems" and "representations" are no longer necessary as ontological entities, though their utility in practice remains.

Little has been said, in the foregoing, about awareness or consciousness, terms I regard as synonyms. It is, however, difficult to conceive a meaning for "observation" that does not entail awareness. Strawson [209] has argued that any self-consistent physicalism entails panpsychism; it seems simpler to follow Wheeler [33] and treat awareness as an irreducible primitive characterizing the dynamics of the universe, whatever these may be. If this threatens the meaning of "physicalism," perhaps that meaning should be abandoned. "Materialism" in any strict sense has, after all, been dead for a century. Weiner famously insisted that "information is information, not matter or energy" ([210] p. 132), but this was in a vastly different cultural context, four decades before "it from bit" and the quantum-information revolution it provoked. Perhaps it is time to consider the possibility that our traditional distinctions between information, energy, and awareness no longer have value.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial intelligence
CA	Conscious agent
FAAP	For all practical purposes
GOFAI	Good old fashioned AI
IGUS	Information gathering and using system
ITP	Interface theory of perception
LoT	Language of Thought
QBism	Quantum Bayesianism
SETI	Search for extra-terrestrial intelligence

References

1. Landsman, N.P. Between classical and quantum. In *Handbook of the Philosophy of Science: Philosophy of Physics*; Butterfield, J., Earman, J., Eds.; Elsevier: Amsterdam, The Netherlands, 2007; pp. 417–553.
2. Schlosshauer, M. *Decoherence and the Quantum to Classical Transition*; Springer: Berlin, Germany, 2007.
3. Ashby, W.R. *Introduction to Cybernetics*; Chapman and Hall: London, UK, 1956.
4. Weiss, P. One plus one does not equal two. In *The Neurosciences: A Study Program*; Gardner, C., Quarton, G.C., Melnechuk, T., Schmitt, F.O., Eds.; Rockefeller University Press: New York, NY, USA, 1967; pp. 801–821.
5. Von Foerster, H. Objects: Tokens for (eigen-) behaviors. *ASC Cybern. Forum* **1976**, *8*, 91–96.
6. Rosen, R. On information and complexity. In *Complexity, Language, and Life: Mathematical Approaches*; Casti, J.L., Karlqvist, A., Eds.; Springer: Berlin, Germany, 1986; pp. 174–196.
7. Rössler, O.E. Endophysics. In *Real Brains—Artificial Minds*; Casti, J., Karlquist, A., Eds.; North-Holland: New York, NY, USA, 1987; pp. 25–46.
8. Kampis, G. Explicit epistemology. *Revue de la Pensee d'Aujourd'hui* **1996**, *24*, 264–275.
9. Pattee, H.H. The physics of symbols: Bridging the epistemic cut. *Biosystems* **2001**, *60*, 5–21. [CrossRef]
10. Dodig Crnkovic, G. Information, computation, cognition. Agency-based hierarchies of levels. In *Fundamental Issues of Artificial Intelligence*; Múller, V.C., Ed.; Springer: Berlin, Germany, 2016; pp. 141–159.
11. Dodig Crnkovic, G. Nature as a network of morphological infocomputational processes for cognitive agents. *Eur. Phys. J. Spec. Top.* **2017**, *226*, 181–195. [CrossRef]
12. Boltzmann, L. *Lectures on Gas Theory*; Dover Press: New York, NY, USA, 1995. First published in 1896.
13. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [CrossRef]
14. Landauer, R. Irreversibility and heat generation in the computing process. *IBM J. Res. Dev.* **1961**, *5*, 183–195. [CrossRef]
15. Landauer, R. Information is a physical entity. *Physical A* **1999**, *263*, 63–67. [CrossRef]
16. Moore, E.F. Gedankenexperiments on sequential machines. In *Autonoma Studies*; Shannon, C.W., McCarthy, J., Eds.; Princeton University Press: Princeton, NJ, USA, 1956; pp. 129–155.
17. Fields, C. Building the observer into the system: Toward a realistic description of human interaction with the world. *Systems* **2016**, *4*, 32. [CrossRef]
18. Koenderink, J. The all-seeing eye. *Perception* **2014**, *43*, 1–6. [CrossRef] [PubMed]
19. Bohr, N. *Atomic Physics and Human Knowledge*; Wiley: New York, NY, USA, 1958.
20. von Neumann, J. *The Mathematical Foundations of Quantum Mechanics*; Princeton University Press: Princeton, NJ, USA, 1955.
21. Zeh, D. On the interpretation of measurement in quantum theory. *Found. Phys.* **1970**, *1*, 69–76. [CrossRef]
22. Zeh, D. Toward a quantum theory of observation. *Found. Phys.* **1973**, *3*, 109–116. [CrossRef]

23. Zurek, W.H. Pointer basis of the quantum apparatus: Into what mixture does the wave packet collapse? *Phys. Rev. D* **1981**, *24*, 1516–1525. [CrossRef]
24. Zurek, W.H. Environment-induced superselection rules. *Phys. Rev. D* **1982**, *26*, 1862–1880. [CrossRef]
25. Joos, E.; Zeh, D. The emergence of classical properties through interaction with the environment. *Z. Phys. B Condens. Matter* **1985**, *59*, 223–243. [CrossRef]
26. Gisin, N. Non-realism: Deep thought or a soft option? *Found. Phys.* **2012**, *42*, 80–85. [CrossRef]
27. Hofer-Szabó, G. How human and nature shake hands: The role of no-conspiracy in physical theories. *Stud. Hist. Philos. Mod. Phys.* **2017**, *57*, 89–97. [CrossRef]
28. Gell-Mann, M.; Hartle, J.B. Quantum mechanics in the light of quantum cosmology. In *Complexity, Entropy, and the Physics of Information*; Zurek, W., Ed.; CRC Press: Boca Raton, FL, USA, 1989; pp. 425–458.
29. Ollivier, H.; Poulin, D.; Zurek, W.H. Environment as a witness: Selective proliferation of information and emergence of objectivity in a quantum universe. *Phys. Rev. A* **2005**, *72*, 042113. [CrossRef]
30. Blume-Kohout, R.; Zurek, W.H. Quantum Darwinism: Entanglement, branches, and the emergent classicality of redundantly stored quantum information. *Phys. Rev. A* **2006**, *73*, 062310. [CrossRef]
31. Zurek, W.H. Quantum Darwinism. *Nat. Phys.* **2009**, *5*, 181–188. [CrossRef]
32. Bohr, N. The quantum postulate and the recent development of atomic theory. *Nature* **1928**, *121*, 580–590. [CrossRef]
33. Wheeler, J.A. Information, physics, quantum: The search for links. In *Complexity, Entropy, and the Physics of Information*; Zurek, W., Ed.; CRC Press: Boca Raton, FL, USA, 1989; pp. 3–28.
34. Rovelli, C. Relational quantum mechanics. *Int. J. Theory Phys.* **1996**, *35*, 1637–1678. [CrossRef]
35. Fuchs, C. QBism, the perimeter of Quantum Bayesianism. *arXiv* **2010**, arXiv:1003.5201v1.
36. Zanardi, P. Virtual quantum subsystems. *Phys. Rev. Lett.* **2001**, *87*, 077901. [CrossRef] [PubMed]
37. Zanardi, P.; Lidar, D.A.; Lloyd, S. Quantum tensor product structures are observable-induced. *Phys. Rev. Lett.* **2004**, *92*, 060402. [CrossRef] [PubMed]
38. Dugić, M.; Jeknić, J. What is "system": Some decoherence-theory arguments. *Int. J. Theory Phys.* **2006**, *45*, 2249–2259. [CrossRef]
39. Dugić, M.; Jeknić-Dugić, J. What is "system": The information-theoretic arguments. *Int. J. Theory Phys.* **2008**, *47*, 805–813. [CrossRef]
40. De la Torre, A.C.; Goyeneche, D.; Leitao, L. Entanglement for all quantum states. *Eur. J. Phys.* **2010**, *31*, 325–332. [CrossRef]
41. Harshman, N.L.; Ranade, K.S. Observables can be tailored to change the entanglement of any pure state. *Phys. Rev. A* **2011**, *84*, 012303. [CrossRef]
42. Thirring, W.; Bertlmann, R.A.; Köhler, P.; Narnhofer, H. Entanglement or separability: The choice of how to factorize the algebra of a density matrix. *Eur. Phys. J. D* **2011**, *64*, 181–196. [CrossRef]
43. Dugić, M.; Jeknić-Dugić, J. Parallel decoherence in composite quantum systems. *Pramana* **2012**, *79*, 199–209. [CrossRef]
44. Fields, C. A model-theoretic interpretation of environment-induced superselection. *Int. J. Gen. Syst.* **2012**, *41*, 847–859. [CrossRef]
45. Hensen, B.; Bernien, H.; Dreau, A.E.; Reiserer, A.; Kalb, N.; Blok, M.S.; Ruitenberg, J.; Vermeulen, R.F.L.; Schouten, R.N.; Abellán, C.; et al. Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres. *Nature* **2015**, *526*, 682–686. [CrossRef] [PubMed]
46. Giustina, M.; Versteegh, M.A.M.; Wengerowsky, S.; Handsteiner, J.; Hochrainer, A.; Phelan, K.; Steinlechner, F.; Kofler, J.; Larsson, J.-Å.; Abellán, C.; et al. A significant-loophole-free test of Bell's theorem with entangled photons. *Phys. Rev. Lett.* **2015**, *115*, 250401. [CrossRef] [PubMed]
47. Shalm, L.K.; Meyer-Scott, E.; Christensen, B.G.; Bierhorst, P.; Wayne, M.A.; Stevens, M.J.; Gerrits, T.; Glancy, S.; Hamel, D.R.; Allman, M.S.; et al. A strong loophole-free test of local realism. *Phys. Rev. Lett.* **2015**, *115*, 250402. [CrossRef] [PubMed]
48. Peres, A. Unperformed experiments have no results. *Am. J. Phys.* **1978**, *46*, 745–747. [CrossRef]
49. Jennings, D.; Leifer, M. No return to classical reality. *Contemp. Phys.* **2016**, *57*, 60–82. [CrossRef]
50. Zurek, W.H. Decoherence, einselection, and the quantum origins of the classical. *Rev. Mod. Phys.* **2003**, *75*, 715–775. [CrossRef]
51. Bell, J.S. Against measurement. *Phys. World* **1990**, *3*, 33–41. [CrossRef]

52. Clifton, R.; Bub, J.; Halvorson, H. Characterizing quantum theory in terms of information-theoretic constraints. *Found. Phys.* **2003**, *33*, 1561–1591. [CrossRef]

53. Bub, J. Why the quantum? *Stud. Hist. Philos. Mod. Phys.* **2004**, *35*, 241–266. [CrossRef]

54. Lee, J.-W. Quantum mechanics emerges from information theory applied to causal horizons. *Found. Phys.* **2011**, *41*, 744–753. [CrossRef]

55. Fuchs, C. Quantum mechanics as quantum information (and only a little more). *arXiv* **2002**, arXiv:quant-ph/0205039v1.

56. Chiribella, G.; D'Ariano, G.M.; Perinotti, P. Informational derivation of quantum theory. *Phys. Rev. A* **2011**, *84*, 012311, arXiv:1011.6451v3.

57. Chiribella, G.; D'Ariano, G.M.; Perinotti, P. Quantum theory, namely the pure and reversible theory of information. *Entropy* **2012**, *14*, 1877–1893. [CrossRef]

58. Rau, J. Measurement-based quantum foundations. *Found. Phys.* **2011**, *41*, 380–388. [CrossRef]

59. Leifer, M.S.; Spekkens, R.W. Formulating quantum theory as a causally neutral theory of Bayesian inference. *arXiv* **2011**, arXiv:1107.5849v1.

60. D'Ariano, G.M. Physics without physics: The power of information-theoretical principles. *Int. J. Theory Phys.* **2017**, *56*, 97–128. [CrossRef]

61. Goyal, P. Information physics—Towards a new conception of physical reality. *Information* **2012**, *3*, 567–594. [CrossRef]

62. Fields, C.; Hoffman, D.D.; Prakash, C.; Prentner, R. Eigenforms, interfaces and holographic encoding: Toward an evolutionary account of objects and spacetime. *Construct. Found.* **2017**, *12*, 265–291.

63. Zurek, W.H. Decoherence, einselection and the existential interpretation (the rough guide). *Philos. Trans. R. Soc. A* **1998**, *356*, 1793–1821.

64. Zurek, W.H. Probabilities from entanglement, Born's rule $p_k = |\psi_k|^2$ from envariance. *Phys. Rev. A* **2005**, *71*, 052105. [CrossRef]

65. Tegmark, M. Importance of quantum decoherence in brain processes. *Phys. Rev. E* **2000**, *61*, 4194–4206. [CrossRef]

66. Tegmark, M. How unitary cosmology generalizes thermodynamics and solves the inflationary entropy problem. *Phys. Rev. D* **2012**, *85*, 123517. [CrossRef]

67. Fields, C. If physics is an information science, what is an observer? *Information* **2012**, *3*, 92–123. [CrossRef]

68. Bronfenbrenner, U. Toward an experimental ecology of human development. *Am. Psychol.* **1977**, *32*, 513–531. [CrossRef]

69. Barwise, J.; Seligman, J. *Information Flow: The Logic of Distributed Systems*; (Cambridge Tracts in Theoretical Computer Science 44); Cambridge University Press: Cambridge, UK, 1997.

70. Scholl, B.J. Object persistence in philosophy and psychology. *Mind Lang.* **2007**, *22*, 563–591. [CrossRef]

71. Martin, A. The representation of object concepts in the brain. *Annu. Rev. Psychol.* **2007**, *58*, 25–45. [CrossRef] [PubMed]

72. Fields, C. The very same thing: Extending the object token concept to incorporate causal constraints on individual identity. *Adv. Cognit. Psychol.* **2012**, *8*, 234–247. [CrossRef]

73. Keifer, M.; Pulvermüller, F. Conceptual representations in mind and brain: Theoretical developments, current evidence, and future directions. *Cortex* **2012**, *48*, 805–825. [CrossRef] [PubMed]

74. Fields, C. Visual re-identification of individual objects: A core problem for organisms and AI. *Cognit. Process.* **2016**, *17*, 1–13. [CrossRef] [PubMed]

75. Baillargeon, R. Innate ideas revisited: For a principle of persistence in infants' physical reasoning. *Perspect. Psychol. Sci.* **2008**, *3*, 2–13. [CrossRef] [PubMed]

76. Rakison, D.H.; Yermoleva, Y. Infant categorization. *WIRES Cognit. Sci.* **2010**, *1*, 894–905. [CrossRef] [PubMed]

77. Polanyi, M. Life's irreducible structure. *Science* **1968**, *160*, 1308–1312. [CrossRef] [PubMed]

78. Bateson, G. *Steps to an Ecology of Mind: Collected Essays in Anthropology, Psychiatry, Evolution, and Epistemology*; Jason Aronson: Northvale, NJ, USA, 1972.

79. McCarthy, J.; Hayes, P.J. Some philosophical problems from the standpoint of artificial intelligence. In *Machine Intelligence*; Michie, D., Meltzer, B., Eds.; Edinburgh University Press: Edinburgh, UK, 1969; Volume 4, pp. 463–502.

80. Fodor, J.A. Modules, frames, fridgeons, sleeping dogs and the music of the spheres. In *The Robot's Dilemma: The Frame Problem in Artificial Intelligence*; Pylyshyn, Z., Ed.; Ablex: Norwood, NJ, USA, 1987; pp. 139–149.

81. Fields, C. How humans solve the frame problem. *J. Exp. Theory Artif. Intell.* **2013**, *25*, 441–456. [CrossRef]

82. Lyubova, N.; Ivaldi, S.; Filliat, D. From passive to interactive object learning and recognition through self-identification on a humanoid robot. *Auton. Robots* **2016**, *40*, 33–57. [CrossRef]

83. Sowa, J.F. Semantic networks. In *Encyclopedia of Cognitive Science*; Nadel, L., Ed.; John Wiley: New York, NY, USA, 2006.

84. Hampton, J.A. Categories, prototypes and exemplars. In *Routledge Handbook of Semantics*; Reimer, N., Ed.; Routledge: Abington, UK, 2016; Chapter 7.

85. Johnson, S.P.; Hannon, E.E. Perceptual development. In *Handbook of Child Psychology and Developmental Science, Volume 2: Cognitive Processes*; Liben, L.S., Müller, U., Eds.; Wiley: Hoboken, NJ, USA, 2016; pp. 63–112.

86. Clarke, A.; Tyler, L.K. Understanding what we see: How we derive meaning from vision. *Trends Cognit. Sci.* **2015**, *19*, 677–687. [CrossRef] [PubMed]

87. Waxman, S.R.; Leddon, E.M. Early word learning and conceptual development: Everything had a name, and each name gave birth to a new thought. In *The Wiley-Blackwell Handbook of Childhood Cognitive Development*; Goswami, U., Ed.; Wiley-Blackwell: Malden, MA, USA, 2002; pp. 180–208.

88. Pinker, S. Language as an adaptation to the cognitive niche. In *Language Evolution*; Christiansen, M.H., Kirby, S., Eds; Oxford University Press: Oxford, UK, 2003; pp. 16–37.

89. Ashby, F.G.; Maddox, W.T. Human category learning 2.0. *Ann. N. Y. Acad. Sci.* **2011**, *1224*, 147–161. [CrossRef] [PubMed]

90. Bargh, J.A.; Ferguson, M.J. Beyond behaviorism: On the automaticity of higher mental processes. *Psychol. Bull.* **2000**, *126*, 925–945. [CrossRef] [PubMed]

91. Ulrich, L.E.; Koonin, E.V.; Zhulin, I.B. One-component systems dominate signal transduction in prokaryotes. *Trends Microbiol.* **2005**, *13*, 52–56. [CrossRef] [PubMed]

92. Ashby, M.K. Survey of the number of two-component response regulator genes in the complete and annotated genome sequences of prokaryotes. *FEMS Microbiol. Lett.* **2004**, *231*, 277–281. [CrossRef]

93. Wadhams, G.H.; Armitage, J.P. Making sense of it all: Bacterial chemotaxis. *Nat. Rev. Mol. Cell Biol.* **2004**, *5*, 1024–1037. [CrossRef] [PubMed]

94. Schuster, M.; Sexton, D.J.; Diggle, S.P.; Greenberg, E.P. Acyl-Homoserine Lactone quorum sensing: From evolution to application. *Annu. Rev. Microbiol.* **2013**, *67*, 43–63. [CrossRef] [PubMed]

95. Logan, C.Y.; Nuss, R. The Wnt signalling pathway in development and disease. *Annu. Rev. Cell. Dev. Biol.* **2004**, *20*, 781–810. [CrossRef] [PubMed]

96. Wagner, E.F.; Nebreda, A.R. Signal integration by JNK and p38 MAPK pathways in cancer development. *Nat. Rev. Cancer* **2009**, *9*, 537–549. [CrossRef] [PubMed]

97. Agrawal, H. Extreme self-organization in networks constructed from gene expression data. *Phys. Rev. Lett.* **2002**, *89*, 268702. [CrossRef] [PubMed]

98. Barabási, A.-L.; Oltvai, Z.N. Network biology: Understanding the cell's functional organization. *Nat. Rev. Genet.* **2004**, *5*, 101–114. [CrossRef] [PubMed]

99. Watts, D.J.; Strogatz, S.H. Collective dynamics of 'small-world' networks. *Nature* **1998**, *393*, 440–442. [CrossRef] [PubMed]

100. Dzhafarov, E.N.; Kujala, J.V. Context-content systems of random variables: The Contextuality-by-Default theory. *J. Math. Psychol.* **2016**, *74*, 11–33. [CrossRef]

101. Dzhafarov, E.N.; Kon, M. On universality of classical probability with contextually labeled random variables. *J. Math. Psychol.* **2016**, *85*, 17–24. [CrossRef]

102. Pothos, E.M.; Busemeyer, J.R. Can quantum probability provide a new direction for cognitive modeling? *Behav. Brain Sci.* **2013**, *36*, 255–327. [CrossRef] [PubMed]

103. Aerts, D.; Sozzo, S.; Veloz, T. Quantum structure in cognition and the foundations of human reasoning. *Int. J. Theory Phys.* **2015**, *54*, 4557–4569. [CrossRef]

104. Bruza, P.D.; Kitto, K.; Ramm, B.J.; Sitbon, L. A probabilistic framework for analysing the compositionality ofconceptual combinations. *J. Math. Psychol.* **2015**, *67*, 26–38. [CrossRef]

105. Emery, N.J. Cognitive ornithology: The evolution of avian intelligence. *Philos. Trans. R. Soc. B* **2006**, *361*, 23–43. [CrossRef] [PubMed]

106. Godfrey-Smith, P. *Other Minds: The Octopus, the Sea, and the Deep Origins of Consciousness*; Farrar, Straus and Giroux: New York, NY, USA, 2016.

107. Etienne, A.S.; Jeffery, K.J. Path integration in mammals. *Hippocampus* **2004**, *14*, 180–192. [CrossRef] [PubMed]

108. Wiltschko, R.; Wiltschko, W. Avian navigation: From historical to modern concepts. *Anim. Behav.* **2003**, *65*, 257–272. [CrossRef]

109. Klein, S.B. What memory is. *WIRES Interdiscip. Rev. Cognit. Sci.* **2015**, *1*, 1–38.

110. Eichenbaum, H. Still searching for the engram. *Learn. Behav.* **2016**, *44*, 209–222. [CrossRef] [PubMed]

111. Newell, A. Physical Symbol Systems. *Cognit. Sci.* **1980**, *4*, 135–183. [CrossRef]

112. Markman, A.B.; Dietrich, E. In defense of representation. *Cognit. Psychol.* **2000**, *40*, 138–171. [CrossRef] [PubMed]

113. Gibson, J.J. *The Ecological Approach to Visual Perception*; Houghton Mifflin: Boston, MA, USA, 1979.

114. Dennett, D. Intentional systems. *J. Philos.* **1971**, *68*, 87–106. [CrossRef]

115. Anderson, M.L. Embodied cognition: A field guide. *Artif. Intell.* **2003**, *149*, 91–130. [CrossRef]

116. Froese, T.; Ziemke, T. Enactive artificial intelligence: Investigating the systemic organization of life and mind. *Artif. Intell.* **2009**, *173*, 466–500. [CrossRef]

117. Chemero, A. Radical embodied cognitive science. *Rev. Gen. Psychol.* **2013**, *17*, 145–150. [CrossRef]

118. Fodor, J. *The Language of Thought*; Harvard University Press: Cambridge, MA, USA, 1975.

119. Dehaene, S.; Naccache, L. Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework. *Cognition* **2001**, *79*, 1–37. [CrossRef]

120. Baars, B.J. Global workspace theory of consciousness: Toward a cognitive neuroscience of human experience. *Prog. Brain Res.* **2005**, *150*, 45–53. [PubMed]

121. Sporns, O.; Honey, C.J. Small worlds inside big brains. *Proc. Natl. Acad. Sci. USA* **2006**, *111*, 15220–15225. [CrossRef] [PubMed]

122. Shanahan, M. The brain's connective core and its role in animal cognition. *Philos. Trans. R. Soc. B* **2012**, *367*, 2704–2714. [CrossRef] [PubMed]

123. Baars, B.J.; Franklin, S.; Ramsoy, T.Z. Global workspace dynamics: Cortical "binding and propagation" enables conscious contents. *Front. Psychol.* **2013**, *4*, 200. [CrossRef] [PubMed]

124. Dehaene, S.; Charles, L.; King, J.-R.; Marti, S. Toward a computational theory of conscious processing. *Curr. Opin. Neurobiol.* **2014**, *25*, 76–84. [CrossRef] [PubMed]

125. Ball, G.; Alijabar, P.; Zebari, S.; Tusor, N.; Arichi, T.; Merchant, N.; Robinson, E.C.; Ogundipe, E.; Ruekert, D.; Edwards, A.D.; et al. Rich-club organization of the newborn human brain. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 7456–7461. [CrossRef] [PubMed]

126. Gao, W.; Alcauter, S.; Smith, J.K.; Gilmore, J.H.; Lin, W. Development of human brain cortical network architecture during infancy. *Brain Struct. Funct.* **2015**, *220*, 1173–1186. [CrossRef] [PubMed]

127. Huang, H.; Shu, N.; Mishra, V.; Jeon, T.; Chalak, L.; Wang, Z.J.; Rollins, N.; Gong, G.; Cheng, H.; Peng, Y.; et al. Development of human brain structural networks through infancy and childhood. *Cereb. Cortex* **2015**, *25*, 1389–1404. [CrossRef] [PubMed]

128. Fields, C.; Glazebrook, J.F. Disrupted development and imbalanced function in the global neuronal workspace: A positive-feedback mechanism for the emergence of ASD in early infancy. *Cognit. Neurodyn.* **2017**, *11*, 1–21. [CrossRef] [PubMed]

129. Chater, N. *The Mind Is Flat*; Allen lane: London, UK, 2018.

130. Friston, K.J.; Kiebel, S. Predictive coding under the free-energy principle. *Philos. Trans. R. Soc. B* **2009**, *364*, 1211–1221. [CrossRef] [PubMed]

131. Friston, K.J. The free-energy principle: A unified brain theory? *Nat. Rev. Neurosci.* **2010**, *11*, 127–138. [CrossRef] [PubMed]

132. Clark, A. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.* **2013**, *36*, 181–204. [PubMed]

133. Hohwy, J. The self-evidencing brain. *Noûs* **2016**, *50*, 259–285. [CrossRef]

134. Spratling, M.W. Predictive coding as a model of cognition. *Cognit. Process.* **2016**, *17*, 279–305. [CrossRef] [PubMed]

135. Bastos, A.M.; Usrey, W.M.; Adams, R.A.; Mangun, G.R.; Fries, P.; Friston, K.J. Canonical microcircuits for predictive coding. *Neuron* **2012**, *76*, 695–711. [CrossRef] [PubMed]

136. Adams, R.A.; Friston, K.J.; Bastos, A.M. Active inference, predictive coding and cortical architecture. In *Recent Advances in the Modular Organization of the Cortex*; Casanova, M.F., Opris, I., Eds.; Springer: Berlin, Germany, 2015; pp. 97–121.

137. Spratling, M.W. A hierarchical predictive coding model of object recognition in natural images. *Cognit. Comput.* **2017**, *9*, 151–167. [CrossRef] [PubMed]
138. Shipp, S.; Adams, R.A.; Friston, K.J. Reflections on agranular architecture: Predictive coding in the motor cortex. *Trends Neurosci.* **2013**, *36*, 706–716. [CrossRef] [PubMed]
139. Seth, A.K. Interoceptive inference, emotion, and the embodied self. *Trends Cognit. Sci.* **2013**, *17*, 565–573. [CrossRef] [PubMed]
140. Harold, F.M. Molecules into cells: Specifying spatial architecture. *Microbiol. Mol. Biol. Rev.* **2005**, *69*, 544–564. [CrossRef] [PubMed]
141. Laland, K.N.; Uller, T.; Feldman, M.W.; Sterelny, K.; Müller, G.B.; Moczek, A.; Jablonka, E.; Odling-Smee, J. The extended evolutionary synthesis: Its structure, assumptions and predictions. *Proc. R. Soc. B* **2015**, *282*, 20151019. [CrossRef] [PubMed]
142. Pezzulo, G.; Levin, M. Top-down models in biology: Explanation and control of complex living systems above the molecular level. *J. R. Soc. Interface* **2016**, *13*, 20160555. [CrossRef] [PubMed]
143. Fields, C.; Levin, M. Multiscale memory and bioelectric error correction in the cytoplasm- cytoskeleton-membrane system. *WIRES Syst. Biol. Med.* **2018**, *10*, e1410. [CrossRef] [PubMed]
144. Kull, K.; Emmeche, C.; Hoffmeyer, J. Why biosemiotics? An introduction to our view on the biology of life itself. In *Towards a Semiotic Biology: Life Is the Action of Signs*; Emmeche, C., Kull, K., Eds.; Imperial College Press: London, UK, 2011; pp. 1–21.
145. Barbieri, M. A short history of biosemiotics. *Biosemiotics* **2009**, *2*, 221–245. [CrossRef]
146. Friston, K.; Levin, M.; Sengupta, B.; Pezzulo, G. Knowing one's place: A free-energy approach to pattern regulation. *J. R. Soc. Interface* **2015**, *12*, 20141383. [CrossRef] [PubMed]
147. Schwabe, L.; Nader, K.; Pruessner, J.C. Reconsolidation of human memory: Brain mechanisms and clinical relevance. *Biol. Psychiatry* **2014**, *76*, 274–280. [CrossRef] [PubMed]
148. Clark, A.; Chalmers, D. The extended mind. *Analysis* **1998**, *58*, 7–19. [CrossRef]
149. Kull, K.; Deacon, T.; Emmeche, C.; Hoffmeyer, J.; Stjernfelt, F. Theses on biosemiotics: Prolegomena to a theoretical biology. In *Towards a Semiotic Biology: Life Is the Action of Signs*; Emmeche, C., Kull, K., Eds.; Imperial College Press: London, UK, 2011; pp. 25–41.
150. Roederer, J. *Information and Its Role in Nature*; Springer: Berlin, Germany, 2005.
151. Roederer, J. Pragmatic information in biology and physics. *Philos. Trans. R. Soc. A* **2016**, *374*, 20150152. [CrossRef] [PubMed]
152. Gu, Z.-C.; Wen, X.-G. Emergence of helicity ± 2 modes (gravitons) from qubit models. *Nucl. Phys. B* **2012**, *863*, 90–129. [CrossRef]
153. Swingle B. Entanglement renormalization and holography. *Phys. Rev. D* **2012**, *86*, 065007. [CrossRef]
154. Dowker, F. Introduction to causal sets and their phenomenology. *Class. Quantum Gravity* **2013**, *45*, 1651–1667. [CrossRef]
155. Pastawski, F.; Yoshida, B.; Harlow, D.; Preskill, J. Holographic quantum error-correcting codes: Toy models for the bulk/boundary correspondence. *J. High Energy Phys.* **2015**, *2015*, 149. [CrossRef]
156. Tegmark, M. Consciousness as a state of matter. *Chaos Solitons Fractals* **2015**, *76*, 238–270. [CrossRef]
157. Hooft, G. *The Cellular Automaton Interpretation of Quantum Mechanics*; Springer: Heidelberg, Germany, 2016.
158. Han, M.; Hung, L.-Y. Loop quantum gravity, exact holographic mapping, and holographic entanglement entropy. *Phys. Rev. D* **2017**, *95*, 024011. [CrossRef]
159. Setty, Y.; Mayo, A.E.; Surette, M.G.; Alon, U. Detailed map of a cis-regulatory input function. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 7702–7707. [CrossRef] [PubMed]
160. Dennett, D. *From Bacteria to Bach and Back*; Norton: New York, NY, USA, 2017.
161. Simons, J.S.; Henson, R.N.A.; Gilbert, S.J.; Fletcher, P.C. Separable forms of reality monitoring supported by anterior prefrontal cortex. *J. Cognit. Neurosci.* **2008**, *20*, 447–457. [CrossRef] [PubMed]
162. Picard, F. Ecstatic epileptic seizures: A glimpse into the multiple roles of the insula. *Front. Behav. Neurosci.* **2016**, *10*, 21.
163. Debruyne, H.; Portzky, M.; Van den Eynde, F.; Audenaert, K. Cotard's syndrome: A review. *Curr. Psychiatry Rep.* **2009**, *11*, 197–202. [CrossRef] [PubMed]
164. Kahneman, D. *Thinking, Fast and Slow*; Farrar, Straus and Giroux: New York, NY, USA, 2011.
165. Bargh, J.A.; Schwader, K.L.; Hailey, S.E.; Dyer, R.L.; Boothby, E.J. Automaticity in social-cognitive processes. *Trends Cognit. Sci.* **2012**, *16*, 593–605. [CrossRef] [PubMed]

166. Heider, F.; Simmel, M. An experimental study of apparent behavior. *Am. J. Psychol.* **1944**, *57*, 243–259. [CrossRef]

167. Scholl, B.J.; Tremoulet, P. Perceptual causality and animacy. *Trends Cognit. Sci.* **2000**, *4*, 299–309. [CrossRef]

168. Scholl, B.J.; Gao, T. Perceiving animacy and intentionality: Visual processing or higher-level judgment? In *Social Perception: Detection and Interpretation of Animacy, Agency and Intention*; Rutherford, M.D., Kuhlmeier, V.A., Eds.; MIT Press: Cambridge, MA, USA, 2013; pp. 197–230.

169. Fields, C. Motion, identity and the bias toward agency. *Front. Hum. Neurosci.* **2014**, *8*, 597. [CrossRef] [PubMed]

170. Blackmore, S. Dangerous memes; or, What the Pandorans let loose. In *Cosmos & Culture: Culture Evolution in a Cosmic Context*; Dick, S.J., Lupisdella, M.L., Eds.; NASA: Washington, DC, USA, 2009; pp. 297–317.

171. Patton, C.M.; Wheeler, J.A. Is physics legislated by cosmogony? In *Quantum Gravity*; Isham, C.J., Penrose, R., Sciama, D.W., Eds.; Clarendon: Oxford, UK, 1975; pp. 538–605.

172. Conway, J.; Kochen, S. The free will theorem. *Found. Phys.* **2006**, *36*, 1441–1473. [CrossRef]

173. Horsman, C.; Stepney, S.; Wagner, R.C.; Kendon, V. When does a physical system compute? *Proc. R. Soc. A* **2014**, *470*, 20140182. [CrossRef] [PubMed]

174. Oudeyer, P.-Y.; Kaplan, F. What is intrinsic motivation? A typology of computational approaches. *Front. Neurorobot.* **2007**, *1*, 6. [CrossRef] [PubMed]

175. Baldassarre, G.; Mirolli, M. (Eds.) *Intrinsically Motivated Learning in Natural and Artificial Systems*; Springer: Berlin, Germany, 2013.

176. Cangelosi, A.; Schlesinger, M. *Developmental Robotics: From Babies to Robots*; MIT Press: Cambridge, MA, USA, 2015.

177. Adams, A.; Zenil, H.; Davies, P.C.W.; Walker, S.I. Formal definitions of unbounded evolution and innovation reveal universal mechanisms for open-ended evolution in dynamical systems. *Nat. Rep.* **2017**, *7*, 997. [CrossRef] [PubMed]

178. Pattee, H.H. Cell psychology. *Cognit. Brain Theory* **1982**, *5*, 325–341.

179. Stewart, J. Cognition = Life: Implications for higher-level cognition. *Behav. Process.* **1996**, *35*, 311–326. [CrossRef]

180. di Primio, F.; Müller, B.F.; Lengeler, J.W. Minimal cognition in unicellular organisms. In *From Animals to Animats*; Meyer, J.A., Berthoz, A., Floreano, D., Roitblat, H.L., Wilson, S.W., Eds.; International Society for Adaptive Behavior: Honolulu, HI, USA, 2000; pp. 3–12.

181. Baluška, F.; Levin, M. On having no head: Cognition throughout biological systems. *Front. Psychol.* **2016**, *7*, 902. [CrossRef] [PubMed]

182. Cook, N.D. The neuron-level phenomena underlying cognition and consciousness: Synaptic activity and the action potential. *Neuroscience* **2008**, *153*, 556–570. [CrossRef] [PubMed]

183. Gagliano, M.; Renton, M.; Depczynski, M.; Mancuso, S. Experience teaches plants to learn faster and forget slower in environments where it matters. *Oecologia* **2014**, *175*, 63–72. [CrossRef] [PubMed]

184. Heil, M.; Karban, R. Explaining evolution of plant communication by airborne signals. *Trends Ecol. Evolut.* **2010**, *25*, 137–144. [CrossRef] [PubMed]

185. Rasmann, S.; Turlings, T.C.J. Root signals that mediate mutualistic interactions in the rhizosphere. *Curr. Opin. Plant Biol.* **2016**, *32*, 62–68. [CrossRef] [PubMed]

186. Hilker, M.; Schwachtje, J.; Baier, M.; Balazadeh, S.; Bäurle, I.; Geiselhardt, S.; Hincha, D.K.; Kunze, R.; Mueller-Roeber, B.; Rillig, M.C.; et al. Priming and memory of stress responses in organisms lacking a nervous system. *Biol. Rev.* **2016**, *91*, 1118–1133. [CrossRef] [PubMed]

187. Panksepp, J. Affective consciousness: Core emotional feelings in animals and humans. *Conscious. Cognit.* **2005**, *14*, 30–80. [CrossRef] [PubMed]

188. Rochat, P. Primordial sense of embodied self-unity. In *Early Development of Body Representations*; Slaughter, V., Brownell, C.A., Eds.; Cambridge University Press: Cambridge, UK, 2012; pp. 3–18.

189. Queller, D.C.; Strassmann, J.E. Beyond society: The evolution of organismality. *Philos. Trans. R. Soc. B* **2009**, *364*, 3143–3155. [CrossRef] [PubMed]

190. Strassmann, J.E.; Queller, D.C. The social organism: Congresses, parties and committees. *Evolution* **2010**, *64*, 605–616. [CrossRef] [PubMed]

191. Fields, C.; Levin, M. Are planaria individuals? What regenerative biology is telling us about the nature of multicellularity. *Evolut. Biol.* **2018**, *45*, 237–247. [CrossRef]

192. Müller, G.B. Evo-devo: Extending the evolutionary synthesis. *Nat. Rev. Genet.* **2007**, *8*, 943–949. [CrossRef] [PubMed]

193. Carroll, S.B. Evo-devo and an expanding evolutionary synthesis: A genetic theory of morphological evolution. *Cell* **2008**, *134*, 25–36. [CrossRef] [PubMed]

194. Bjorklund, D.F.; Ellis, B.J. Children, childhood, and development in evolutionary perspective. *Dev. Rev.* **2014**, *34*, 225–264. [CrossRef]

195. Gentner, D. Why we're so smart. In *Language and Mind: Advances in the Study of Language and Thought*; Gentner, D., Goldin-Meadow, S., Eds.; MIT Press: Cambridge, MA, USA, 2003; pp. 195–235.

196. Dietrich, E.S. Analogical insight: Toward unifying categorization and analogy. *Cognit. Process.* **2010**, *11*, 331–345. [CrossRef] [PubMed]

197. Fields, C. Metaphorical motion in mathematical reasoning: Further evidence for pre-motor implementation of structure mapping in abstract domains. *Cognit. Process.* **2013**, *14*, 217–229. [CrossRef] [PubMed]

198. Sellars, W. Philosophy and the scientific image of man. In *Frontiers of Science and Philosophy*; Colodny, R., Ed.; University of Pittsburgh Press: Pittsburgh, PA, USA, 1962; pp. 35–78.

199. Fields, C. Quantum Darwinism requires an extra-theoretical assumption of encoding redundancy. *Int. J. Theory Phys.* **2010**, *49*, 2523–2527. [CrossRef]

200. Hooft, G. Dimensional reduction in quantum gravity. In *Salamfestschrift*; Ali, A., Ellis, J., Randjbar-Daemi, S., Eds.; World Scientific: Singapore, 1993; pp. 284–296.

201. Hoffman, D.D.; Singh, M.; Prakash, C. The interface theory of perception. *Psychon. Bull. Rev.* **2015**, *22*, 1480–1506. [CrossRef] [PubMed]

202. Hoffman, D.D. The interface theory of perception. *Curr. Dir. Psychol. Sci.* **2016**, *25*, 157–161. [CrossRef]

203. Mark, J.T.; Marion, B.B.; Hoffman, D.D. Natural selection and veridical perceptions. *J. Theory Biol.* **2010**, *266*, 504–515. [CrossRef] [PubMed]

204. Prakash, C. On invention of structure in the world: Interfaces and conscious agents. *Found. Sci.* **2018**, in press.

205. Hoffman, D.D.; Prakash, C. Objects of consciousness. *Front. Psychol.* **2014**, *5*, 577. [CrossRef] [PubMed]

206. Fields, C.; Hoffman, D.D.; Prakash, C.; Singh, M. Conscious agent networks: Formal analysis and application to cognition. *Cognit. Syst. Res.* **2018**, *47*, 186–213. [CrossRef]

207. Carter, B. Anthropic principle in cosmology. In *Current Issues in Cosmology*; Pecker, J.-C., Narlikar, J.V., Eds.; Cambridge University Press: Cambridge, UK, 2006; pp. 173–179.

208. von Uexküll, J. A Stroll through the Worlds of Animals and Men. In *Instinctive Behavior*; Schiller, C., Ed.; International Universities Press: New York, NY, USA, 1957; pp. 5–80.

209. Strawson, G. *ConsciOusness and Its Place in Nature: Does Physicalism Entail Panpsychism?* Imprint Academic: Exeter, UK, 2006.

210. Weiner, N. *Cybernetics*; MIT Press: Cambridge, MA, USA, 1948.