Where There is Life There is Mind: In Support of a Strong Life-Mind Continuity Thesis

Michael D. Kirchhoff [1,*] and Tom Froese [2]

[1] Department of Philosophy, Faculty of Law, Humanities and the Arts, University of Wollongong, Wollongong 2500, Australia

[2] Department of Computer Science, Research Institute for Applied Mathematics and Systems, National Autonomous University of Mexico, 04510 Mexico City, Mexico; t.froese@gmail.com

[*] Correspondence: kirchhof@uow.edu.au; Tel.: +61-4221-5742

Academic Editors: Gordana Dodig-Crnkovic and Robert Lowe
Received: 22 February 2017; Accepted: 11 April 2017; Published: 14 April 2017

Abstract: This paper considers questions about continuity and discontinuity between life and mind. It begins by examining such questions from the perspective of the free energy principle (FEP). The FEP is becoming increasingly influential in neuroscience and cognitive science. It says that organisms act to maintain themselves in their expected biological and cognitive states, and that they can do so only by minimizing their free energy given that the long-term average of free energy is entropy. The paper then argues that there is no singular interpretation of the FEP for thinking about the relation between life and mind. Some FEP formulations express what we call an independence view of life and mind. One independence view is a cognitivist view of the FEP. It turns on information processing with semantic content, thus restricting the range of systems capable of exhibiting mentality. Other independence views exemplify what we call an overly generous non-cognitivist view of the FEP, and these appear to go in the opposite direction. That is, they imply that mentality is nearly everywhere. The paper proceeds to argue that non-cognitivist FEP, and its implications for thinking about the relation between life and mind, can be usefully constrained by key ideas in recent enactive approaches to cognitive science. We conclude that the most compelling account of the relationship between life and mind treats them as strongly continuous, and that this continuity is based on particular concepts of life (autopoiesis and adaptivity) and mind (basic and non-semantic).

Keywords: life-mind continuity; free energy principle; radical enactivism; autopoietic enactivism

1. Introduction

How are life and mind, respectively, characterized, and how are their relations to one another best conceived?

In this paper, we start by examining this question from the perspective of the free energy principle (FEP). The FEP is argued to deliver an overarching rationale for brain functioning; to give a unified theory of perception, cognition, and action (and all other psychological capacities); and to suggest a framework by which to understand the relation between life and mind [1–3]. It states that organisms act to maintain themselves in their expected biological and cognitive states, and that they can do so only by minimizing their free energy given that the long-term average of free energy is entropy [4,5] (By "state" we mean a state in a system's state space. One of the states that a system expects to find itself in is "to be alive". Thus, a system will seek to reduce the probability of finding itself in a non-anticipated state relative to its generative model. In other words, by minimizing free energy, on average and over time, the system will self-organize the parameters of its internal states

to occupy a limited number of states, on average and over time ([6], p. 180)). Hence, to minimize free energy is to reduce disorder, in the sense of uncertainty.

We then argue that there is no singular account of the FEP for thinking about the relation between life and mind (Or, minimally, there is no existing agreement on how best to interpret the properties of variational free energy for thinking about life and mind, and their relationship to one another). These different perspectives on the life-mind relation can be brought into view by considering the answer one would give to the following question: "Are mental phenomena restricted to living systems"?

Some free energy formulations answer this question negatively. We call this picture of the life-mind relation the *independence view of life and mind*. In the context of the FEP, it comes in at least two formulations, each of which has its own unique implications for thinking about life and mind, and their relation to one another:

- The Cognitivist Free Energy Principle [6–8].
- The Overly Generous Non-Cognitivist Free Energy Principle [9,10].

The cognitivist free energy principle (cognitivist FEP) treats the relationship between life and mind as a contingent one. It is sometimes referred to as the self-evidencing brain hypothesis [7] or simply the predictive mind [6]. An independence view of this kind might still hold that some cognitive systems are living systems, but it will treat this particular relation as purely contingent, for example by associating minds with computational processes with semantic (i.e., contentful) properties, or by allowing for the possibility that minds could be realized wholly independently from life given the right kind of artificially supporting system. Such minds might be said to be epistemically secluded from the world, comprised of powerful generative models, which cannot be "necessarily wedded to biological organs" ([8], p. 7).

The non-cognitivist free energy principle (non-cognitivist FEP) takes a very different starting point from the cognitivist FEP. The origins of the FEP were in thermodynamics, where non-equilibrium free energy theorems have been used to explain self-organizing dynamics in systems capable of remaining far from thermodynamic equilibrium [4]. In this sense the origins of the FEP have nothing intrinsically to do with life and mind, although it is now being directly applied to explain living and cognitive systems [2,9–11]. Overly generous non-cognitivist FEP is the view that all systems that maintain their variables within a limited range of values can be understood as having some form of mentality or proto-mentality given that the FEP casts any system that is able to maintain structural integrity in the face of a fluctuating environment as engaged in predicting its own future states. That is, retaining integrity rests upon processes the function of which is to maximize model evidence—i.e., these processes exhibit self-evidencing dynamics. Yet, this generous view of self-evidencing would appear to lead to some form of panpsychism.

Other free energy formulations answer the question of whether the mind is restricted by life positively. These formulations are part of a more general picture of the life-mind relation we refer to as the *dependence view of life and mind*. There are different accounts in the literature, and common in all of these is that they subscribe to a much less generous but still non-cognitivist view of the FEP. We dub the possible accounts:

- Non-Cognitivist FEP + Evolutionary Latecomer Views of Mind (e.g., [12])
- Non-Cognitivist FEP + Strong Life-Mind Continuity Views ([13,14] and *this paper*).

Evolutionary latecomer views of mind emphasize discontinuities between mere living systems and cognitive ones, such that the properties of mind can be thought of as complexifications of the properties of life (see [15] for discussion of such views, yet in a slightly different context). For example, mentality, but not life, requires the existence of sophisticated generative neural machinery that is not present in simple forms of life such as single-celled organisms [12]. Hence, on this view, it is possible to be alive and yet not (necessarily) cognitive (Despite his defense of an action-oriented representational theory of mind, we place Clark [12] in the non-cognitivist FEP camp given his advocacy of the complementarity of FEP and predictive processing schemes with work in embodied,

extended and enactive approaches to cognitive science. We will have more to say about Clark's [12] view in due course).

In this paper, we shall defend non-cognitivist FEP and a strong life-mind continuity view based on recent developments in embodied and enactive cognitive science. It is the classic premises of cognitivism, especially internalism and representationalism, which make straightforward applications of enactive approaches to the FEP problematic [1,16]. We are not alone in developing anti-cognitivist formulations of the FEP from the perspective of enactivism. Work by Bruineberg et al. [13] and Kirchhoff [14,17] set the stage. Here is a list of the points on which we agree with Bruineberg et al. [13]: (a) the Helmholtzian view of perception as unconscious inference is inherent to cognitivist formulations of the FEP; (b) there are good reasons to think that this Helmholtzian view of perceptual inference is incompatible with approximate Bayesian inference under a non-cognitivist formulation of the FEP; and (c) once viewed through the lens of enactivism, the FEP can address how life and mind share the same set of basic organizational properties. The main difference between Bruineberg et al. [13] and this paper is that while Bruineberg et al. ([13]; see also [18]) aim to establish that the function of generative models is to maintain a robust brain-body-niche system (see [17] for a metaphysical treatment), we directly target a strong life-mind continuity thesis, further developing recent arguments in [14].

We shall argue that non-cognitivist FEP, and its implications for thinking about the relation between life and mind, can be usefully constrained and augmented by key ideas in recent radical and autopoietic enactive approaches to cognitive science [14,17,19–22]. Our argument has two steps. The first addresses the nature of *basic minds* as selected by evolution for intentional directedness without semantic content [21]. The second step grounds the concept of basic minds in the concept of basic life cast in terms of autopoiesis and adaptivity [19,20]. Consequently, we arrive at a strong view of life-mind continuity, and we avoid the cognitivist position of no-continuity between life and mind, while, at the same time, remaining far removed from the kind of mental bloat associated with overly generous FEP interpretations of the place of mind in the natural world (For other related but different articulations of the strong life-mind continuity thesis, see [23–26], especially [27]).

2. The Free Energy Principle

What is called the free energy principle (FEP) is an imperative for self-organization in open dynamical systems. It specifies that for living systems to maintain their structural and functional integrity they must minimize "free energy" in the context of active inference: they must change their relationship to their niche in order preserve integrity [4,10,14]. The FEP is therefore the claim that all biological systems must actively resist a natural tendency for disorder [13,28].

Free energy was classically defined in terms of thermodynamic principles, but here we are only concerned with free energy cast as variational free energy that comes from probability theory and Bayesian statistics given that this is the conception of free energy involved in the FEP. Information theoretically, free energy is an upper bound on surprise (or formally, "surprisal"), where surprise is defined as the difference between an organism's predictions (or anticipations) about sensory input and the sensory input it actually receives. Thus, surprise is a measure of improbability, and should not be confused with the psychological notion of surprise (though the two sometimes converge). Organisms that succeed in remaining far from terminal phase boundaries (and therefore are able to remain alive), the FEP mandates, "do so by minimizing their tendency to enter into this special kind of surprising (that is, non-anticipated) states" ([5] p. 1).

The relationship between variational free energy and entropy should be understood in the following way. Free energy is an upper bound on surprise, and the long-term average of surprise is entropy. To see this more clearly, consider that a state can be said to have high surprise if it is deemed unlikely to occur relative to a generative model. The main idea is that organisms become (approximate) models of their local niche given that such systems, on average and over time, distill statistical regularities of their niche and thus come to embody such regularities in their gross-bodily form and internal global dynamics [10,11]. Were a state to be consistently high in surprise it would be a state with high entropy (with high statistical improbability). Alternatively, if a system is able to

predict the exteroceptive or interoceptive causes of its sensory input, it will be in a state with low entropy, and therefore low surprise. This is simply to say that expected states have a low entropy distribution. Conversely, the higher the average number of observations required to describe the dispersion of states for some random variable, the higher the entropy of that variable's probability distribution As a result, the FEP states that living systems can maintain themselves within entropic bounds by seeking to minimize their free energy.

Prima facie, at least, what the FEP makes possible is the generation of a deep and underlying unity connecting "processes of adaptation, mind, and life" ([5], p. 1) cast in terms of the information-theoretic notion of free energy. All this is to say that what enables living systems to survive is the same process that enables such systems to perceive, act, think, and so on. Thus, the FEP gives hope of providing a single framework by which to unify theorizing about life and mind, and it does so by appeal to a single imperative: free energy minimization [4,10,11]. Hence, the FEP provides one with reasons for linking processes of life with processes of mind via what is essentially an uncertainty minimization view of life and mind. However, in the context of the FEP, closer inspection reveals a tension between different conceptions of how best to understand the implications of free energy minimization and the central properties that free energy minimization involves.

2.1. Free Energy and Cognitivist Prediction Error Minimization

Some free energy formulations add what we call a cognitivist constraint, which has implications for how such a view understands the life-mind relation. By cognitivist constraint we mean a constraint about the nature of the information processing in question; that it be thought of as the processing of representations with semantic content. In addition to positing semantic mental representations, cognitivist FEP also conceives of free energy minimization through a particular epistemological lens; namely that the free energy theorem leads to global skepticism with respect to the mind-world relation ([8], p. 2). Although there is much to be said about the epistemological implications of cognitivist FEP, we shall restrict our attention to the issue of semantic mental representations.

Cognitivist FEP is usually framed almost exclusively within a particular understanding of approximate Bayesian inference cast in terms of "prediction error minimization" (PEM). Whereas the FEP takes its starting point in issues to do with self-organization in thermodynamic, non-equilibrium systems, and thus can be shown to apply to a wide range of different phenomena, prediction error minimization schemes have been more directly associated with brain functioning ([6]; see [1] for further references). Here the brain is depicted as a hierarchical generative model for minimizing a prediction error quantity reflecting the probability of sensorimotor input relative to an internal, knowledge-based, model. It is this knowledge-based formulation that adds the cognitivist assumption. We will unpack what we mean by this, since the FEP is generally taken to imply that internal states engage in Bayesian inference in which internal—comprising a system's generative model—as well as active states can be modeled as minimizing free energy inferentially [29]. So, it is not the issue of inference or the possession of a generative model as such that exemplifies the cognitivist assumption. Instead, it is the particular properties associated with such inferential processes that highlight a particular, cognitivist take on free energy minimization.

Cognitivist FEP gives special importance to internal, information processing with semantic (i.e., contentful) properties. This is a semantic view of the computational theory of mind, where internal representations are cast in terms of top-down probabilistic inferences on probability density distributions. Such a view of brain functioning is usually couched in the language of folk-psychology (the language of beliefs, desires, attention, and reasons), and proceeds from the premise that information processing with semantic properties is what constitutes cognition. The common reason for positing internal states with semantic properties is that unless there are such states they would not be cognitive, because purely physical systems would not be able to represent the world beyond their internal states. Since minds are generally assumed to manipulate representations and since most naturally occurring systems are not assumed to manipulate such things, it follows that cognitivist

FEP draws a hard line between mental and non-mental systems. For example, Hohwy [3,6,7] draws such cognitivist results from the FEP (but see also [30]).

Work of this particular kind motivates what we referred to as an independence view of life and mind, placing the origins of mind later than those of life and treating the relation between the living and the mental in a contingent fashion. However, given its functionalist framework, this approach threatens to introduce a hard to explain gap between more complex and cognitive forms of life and the rest of the living realm, thus denying the possibility of any kind of life-mind continuity altogether. On such a cognitivist view of mind only some living systems have evolved the neural machinery capable of realizing information processing involving semantic properties.

2.2. Non-Cognitivist Free Energy Minimization

Other less cognitivist or even anti-cognitivist variants draw conclusions to the effect that one can endorse the FEP without endorsing a cognitivist reading of approximate Bayesian inference in the context of prediction error minimization. Non-cognitivist FEP casts free energy minimization in physical systems in terms of Shannon "entropy" in information theory [5]. Hence, it is possible to posit the FEP as the unifying principle of life and mind and, at the same time, deny that the most basic features of life and mind involve probabilistic inferences with semantic content, even if the basis of life and mind involves probabilistic inference (see also [10,14,29]).

On some generous articulations of this view, free energy minimization is argued to occur not only in biological systems but to also take place in nonliving systems ranging from synchronization of clocks to the primordial soup and social networks [9,10]. Unlike cognitivist FEP, for which prediction error minimization is an evolutionary function of brains becoming hierarchical generative models, these "generous" views would seem to undermine the continuity and unity of life and mind. The reason for this is that their posited principle by which to unify life and mind applies to systems that are arguably non-living and non-mental. Yet, this presents some problems. The first is that if mentality is realized in processes of free energy minimization, and if free energy minimization applies to everything from human beings, pendulum clocks, to the primordial soup, then mentality may turn out to be nearly everywhere. Whether panpsychism of this form is correct is difficult to assess. But, without a clear way of separating mentality from non-mentality, or life from non-life, any overly generous version of the FEP becomes too general, thus losing its explanatory value for addressing the relationship between life and mind.

In the next two sections, we turn to develop these two different formulations of the FEP for the life-mind relation, while considering the key assumptions made in both accounts to drive them toward their respective conclusions.

3. From Cognitivist FEP to Life-Mind Discontinuity

Some psychologists, neuroscientists, and cognitive scientists adopt the view that the mind can be explained in terms of computation, and add that minds are computational. The main assumption of their view is that computation cannot be explained without an appeal to semantic content and mental representation. This is no different for the cognitivist FEP.

3.1. Concepts of Computation

Hohwy [3,6,7] draws explicitly on the FEP to develop a theory of the brain as engaged in processes of prediction error minimization with semantic content. As we have seen, the FEP states that for organisms to remain far removed from terminal phase boundaries, they must minimize their free energy. Hohwy claims that this free energy formulation "corresponds to the brain's job of minimizing prediction error, selective sampling of sensory data, optimizing expected precisions, and minimizing complexity of internal models" and that these job descriptions "map on to perception, action, attention, and model selection, respectively" ([3], p. 1).

PEM is a computational framework. Here the computational functions realized by the brain are cast in terms of minimizing prediction error, reflecting the probability of sensorimotor input (or data)

relative to a statistical model. On this view, the brain is conceived as realizing the functional roles required to encode a brain-wide, non-local, hierarchical Bayesian generative model [31]. A Bayesian model is generative given that its function is "to capture the statistical structure of some set of observed inputs by inferring a causal matrix able to give rise to that very structure" ([32], p. 21). Hence, a generative model is a statistical model mapping hidden causes with sensory consequences, and therefore encoding prior probabilistic beliefs about which sensorimotor effects have which worldly and/or bodily causes [33]. This means that at each level of the brain's cortical hierarchy, probability density functions are encoded of the prediction error signals arriving from the cortical level below [31].

PEM combines this computational picture of the brain with the view that information processing is inferential, and says that inference should be understood as approximate Bayesian inference akin to something like hypothesis testing. Statistical inference in this sense is a tool of science. Hohwy adopts this view of the brain, as he notes "scientists are in the business of minimizing the error in predictions generated by their hypotheses" ([7], p. 3). There are different ways by which to engage in scientific hypothesis testing. One might tweak certain parameters of one's model or one might intervene on the samples drawn in order to get a better match between the model and the samples one is sampling across. According to Hohwy, as "it is for statistical inference, so it is for the brain in perception […] and action" ([7], pp. 3–4). In perceptual inference, perception reduces prediction error in virtue of shaping prior predictions (i.e., probabilistic beliefs comprising the generative model). In active inference, action minimizes prediction error by working to change the sensory input itself through movement in the world [7]. This is perception and action, respectively.

Prediction error minimization can thus be shown to map onto a computational mechanism, the generative model realized by the brain, which functions to minimize error signals ([6], Chapter 2). On such a view, concrete computational systems are functional mechanisms of a special kind [34]. Although the jury is still out about the specific implementation details, PEM is portrayed as offering new insights into the computational mechanisms underlying perception, action, attention and other cognitive processes. Mechanistically, following Hohwy, "this is done by suppressing prediction error at multiple levels of the temporally ordered hierarchy" ([6], p. 85).

3.2. Concepts of Information: Toward a Semantic View of Computation

The FEP models internal states as engaging in approximate Bayesian inference, and where the latter is implemented in processes of prediction error minimization. This is how the FEP and PEM converge with one another. However, the cognitivist formulation of the FEP adds an explicitly defined *semantic* dimension to the computational profile of PEM.

To see what this involves, work is needed to arrive at what technical notion of "information" might arguably be involved in the computational processes of reducing prediction errors, on average and over time. It is possible to distinguish between different concepts of information. Here we consider three options.

First, one might define information thermodynamically. A closed thermodynamic system, say, is one that has contact with its surrounding environment only in virtue of work and heat exchange. Thermodynamically, information amounts to a difference between two distinguishable states (e.g., high and low gravitational potential). It is common practice to think of this discrepancy in terms of information, thereby linking information closely to the thermodynamic notions of entropy and free energy.

Second, one could also try to define information in information-theoretic terms. Here information is a measure of the mean of a probability distribution. Or, otherwise put, information is a measure of the average likelihood that some message is transmitted between a source and a receiver [35]. Friston [10] casts free energy minimization in terms of minimizing a surprisal quantity, and says that this concept of information should be understood in the precise sense of Shannon information [5]. Going further, consider what Godfrey-Smith and Sterelny say about Shannon information: "there is one kind of 'information' appealed to in biology, the kind originally described by Shannon, that is unproblematic […]. Information in this sense exists whenever there is contingency and correlation"

([36], p. 4). This gives us a picture where even simple life forms (e.g., bacteria) can be seen as tracking information and where information in this minimal sense can be shown to exist in complex systems such as brains and between organisms and their environment. Yet, it is important not to over-intellectualize this notion. As Godfrey-Smith and Sterelny go on to stress, if we say that simple life forms can be understood as instantiating Shannon information, or if we say that genes contain information about the proteins they make, then "we are saying no more than what we are saying when we say that there is an informational connection between smoke and fire, or between tree rings and a tree's age" ([36], p. 4). Indeed, according to Godfrey-Smith and Sterelny, Shannon information implies that "anything is a *source* of information if it has a range of possible states, and one variable *carries* information about another to the extent that their states are physically correlated" ([36], p. 1). Intuitively, the more one can infer about the state of one variable from the state of another, the more information the relation between the two variables conveys. This is also known as mutual information. Covariance is formally equivalent to mutual information in the following sense: if two variables stand in a relation of covariance, then information about one reduces "surprisal" about the other—which is just to say that covariance, just like mutual information, optimizes model evidence.

Both thermodynamic and Shannon information cannot underlie the cognitivist formulation of the FEP. Hohwy frames hierarchical predictive processing as providing us with a *representationalist* picture of the nature of mind or cognition [7] (See [30] for endorsement of the claim that PEM is a representational theory of mind). However this immediately prevents both thermodynamics and Shannon information from playing much of a role in minimizing surprise under the cognitivist free energy formulation. The reason is straightforward. Even if tree rings are a source of information with respect to other possible states (a tree's age), it does not follow that the rings represent anything at all about the age of a tree: the two states are informationally covariant; they are not in a representational relationship with one another.

Hohwy's representational gloss on minds is a more natural companion for a third concept of information, namely a semantic notion of information and thus a semantic view of computation. A semantic/representational account of computation imposes a semantic restriction. Piccinini puts this restriction as follows: "Only physical states that qualify as representations may be mapped onto computational descriptions, thereby qualifying as computational states." ([34], p. 9) Piccinini continues by saying that the "semantic account is probably the most popular in the philosophy of mind, because it appears to fit its specific needs better than other accounts. Since minds and digital computers are generally assumed to manipulate [...] representations, they turn out to compute" ([34], p. 9). So, cognitivist FEP posits that brains compute information, and where the relevant kind of information is the kind of information that we usually take to be important for epistemic purposes.

As it stands, the Bayesian generative model, realized in the brain's neural circuitry, is said to have a duplex representational architecture. This architecture is "one that at each level combines *quite traditional* representations of inputs with representations of error. According to the duplex proposal, what gets 'explained away' or cancelled out is the error signal, which (in these models) is depicted as computed by dedicated 'error units'. These are linked to, but distinct from, the so-called representation units meant to encode the causes of sensory signals" ([1], p. 187; italics added). So what cognitivist FEP presents us with is a view of the mind as engaged in computational processes of semantic prediction encoded in hierarchical representations of exteroceptive and interoceptive states of affairs.

3.3. Problems of Integration

Cognitivist FEP raises the demands for information integration. In other words, sophisticated forms of computation cast in terms of semantic representational content would seem to place a high demand on the kind of onboard machinery required for the realization of such large-ranging and hierarchical generative models to get off the ground. It is machinery that many simple life forms and other organisms probably do not have, given that all the work of minimizing prediction error, selective sampling of data, precision optimization, and so on is taken to be "ordered hierarchically in the cortex" ([7], p. 15). Clark [12] draws precisely this conclusion. Considering very basic life forms

such as single-celled organisms capable of chemotaxis, Clark says: "Such a life-form may respond to environmental perturbations using a variety of tricks and ploys, *none of which require it to engage in a process* in which incoming sensory stimulations are met with attempts to generate the incoming signal 'from the top down'" ([12], p. 4). Furthermore, Clark claims that even if such bacteria are self-evidencing in the sense of optimizing a model of their environment, such basic life forms "need not rely upon top-down predictions to structure and inform their exchanges with the wider world. Predictive processing thus constitutes a biologically plausible process theory that may or may not be implemented in any given system" ([12], p. 5). One way to read Clark [12] is to take him to say that while single-celled organisms are alive they are not cognitive given that they lack the kind of architecture required for a system to instantiate a hierarchical generative model by which to engage in probabilistic inference. Indeed, if central nervous systems in general and cerebral cortex more particularly are required for Bayesian inference, and it is this kind of process that is required for mentality, then this shortens the continuity between life and mind, carving off entire branches from the tree of life from having mental properties.

This implies that some forms of life are without mind, while, at the same time, casting other (more neurally advanced) forms of life as wholly different in kind—being both alive and minded. Yet, this implies a deep separation between minds (with semantics) and the rest of the natural, living, and non-living world (without semantics), on the other. They are different in kind.

3.4. Problems of Meaning

The reliance of cognitivist FEP on semantic content raises a host of different issues about how to naturalistically explain semantic properties of mental representations, a problem to which we will return. Here we want to first draw attention to a different, yet related, issue. All physical systems can be understood in terms of Shannon information. However, on this conception information "is just information in terms of (meaningless) energy transfer and should not be confused with intentional or semantic content" ([13], p. 17). Yet, only some physical systems are generally assumed to manipulate representations with semantic content. A cognitivist formulation of the FEP along the lines of approximate Bayesian inference with semantic properties is therefore consistent with a picture that "keeps minds and computers in while leaving most everything else out, thereby vindicating the computational theory of cognition as a strong and nontrivial theory" ([37], p. 9).

Although such a view of mind is attractive to defenders of cognitivist FEP, it faces immediate difficulties. Hohwy [3] makes use of the FEP to arrive at a cognitivist version of mind. But starting from the FEP, which starts with thermodynamic considerations about self-organization in random dynamical systems, does not license an inference to a cognitivist formulation of free energy minimization. The reason is that the FEP casts biological self-organization in terms of Shannon information, viz., in terms of covariance, physical correlation, and generalized synchrony. What follows from this?

First, on a non-semantic formulation of the FEP, where information is understood in the sense of Shannon, the architecture of basic forms of cognition—e.g., adapting to and acting in a dynamic environment in more than a mere dispositional fashion—does not involve, necessarily, the acquisition and manipulation of internal states with semantic content [21]. Hence, cognitivist versions of the FEP place an unnecessarily high standard on what kinds of activities qualify as mental. Indeed, there are arguments that show (if correct) that information-as-covariance is not any kind of semantic content [21]. We elaborate on these points in Section 5.

Second, minds have evolved "for getting things done in the world in real time" ([38], p. 5). Hence, minds (and brains) have evolved primarily for action and only secondarily for thinking, as traditionally conceived. Cognitivist formulations of the FEP downplay the centrality of action in favor of its computational all-stars: semantic information and top-down representations realized by neurally encoded generative models. Yet consider that actual sensory systems are "not concerned with truth and accuracy as such but rather, with action and the need to maintain the functional stability of the organisms in which they are embedded" ([39], p. 90). This sits uneasily with the epistemology of cognitivist FEP that posits internal states with truth and/or accuracy conditions that

link such internal states to states of affairs in the world. Consider (again) the case of chemotaxis in single-celled organisms such as *E. coli*. Chemotaxis is an example of a control mechanism employed by organisms to maintain homeostasis and thereby prolong the probability of their survival. In this case, it is the search for food (or sugars) based on the concentration gradients of chemical repellents and attractants in a bacterium's fluid environment ([40], p. 311). Following Auletta [40], one can model chemotaxis as consisting in so-called homeostatic predictions or prior expectations endowed by natural selection, which approximate the external states that are preferred by the organisms—call these an organism's intrinsic preferences (we shall return to unpack this notion in the last section). According to Auletta, "genetically endowed prior beliefs, entailed by the internal states and configuration of the organism … specify what is innately surprising and enable action to counter unpredicted deviations from expected states" ([40], p. 315). On this account, the embodied generative model of *E. coli* comprise prior beliefs specifying that it should strive to find itself in states with high sugar gradients, as opposed to the opposite. So, *E. coli* can be modeled as searching out these states through active inference. This puts action or action-oriented priors at the basis of homeostasis and life. There is no reason to think that this process involves semantic content—maintaining homeostasis is not a matter of having internal states with content that can be true or false depending on how such contents correspond to the states of affairs in the external world. Nor is there any reason to think that life-enhancing chemotaxis is bereft of mentality given that nothing prevents one from treating the incoming nutrients as sensory data, the cell membrane as a Markov blanket, and the continuing running and tumbling behavior of bacteria as cases of active inference.

This difference in information-theoretic commitment—between semantic and non-semantic information—drives a wedge between life and mind. Life has a metabolic side. Metabolism is one of the key features of life [41]. Self-organizing systems can maintain their organization despite facing thermodynamic tendencies towards entropic disorder. What is required is that living systems are able to locate the right states in their overall state space that enables them to maintain themselves within viable bounds—as in the case of chemotaxis. For example, the "ideal temperature of a human is determined by its embodiment: at 37 °C the enzymes regulating metabolism perform optimally, while the metabolic costs of maintaining body temperature is still manageable within certain environmental conditions" ([13], p. 6). Higher temperature is equal to higher levels of surprisal, and vice-versa. The temperature gradient is a source of information. It is systematically related to an unknown probability distribution, in the sense that surprisal cannot be directly evaluated by an organism. However, this is simply to say that the information in question is one of covariation. It is not a representational, semantic state that carries contentful information about other states of affairs. Hence, basic forms of self-organizing processes in living systems have nothing to do with the requirements for mentality as posited by cognitivist FEP.

Hohwy [3] assumes that a cognitivist reading of prediction error minimization is co-extensive with the FEP. We have argued that the technical conception of information used in the FEP (Shannon information) differs substantially from semantic information. We further argued that the former does not imply the latter notion of information, which is part of the reason for why a cognitivist formulation of the FEP ends up severing life from mind. In the next section, we turn to consider non-cognitivist formulations of the FEP in more detail, and explore the implications for a life-mind continuity thesis on the basis of this non-cognitivist view of the FEP.

4. From Free Energy Minimization to an Overly Generous Life-Mind View

The FEP is an imperative for self-organization in dynamical systems. Given that life exists, the FEP states that living systems must exhibit the following properties. First, biological systems must have a model of how their sensorimotor input is generated and of the kind of states they expect to find themselves occupying. In other words, the FEP conceives of biological systems as close to optimal statistical models of their econiche. As Friston says: "the free-energy principle takes the existence of agents as its starting point and concludes that each phenotype or agent *embodies* an optimal model of its econiche. This optimality is achieved by minimizing free-energy, which bounds

the evidence for each agent (model), afforded by sensory interactions with the world" ([11], p. 89). The idea that a biological system embodies a model requires careful unpacking.

One way to read this claim is that evolution has set up biological systems such that they, on average and over time, distill (i.e., extract) statistical regularities of their econiche, and therefore come to embody these in their model. This may be what leads defenders of cognitivist FEP to say the following: "The mind can then be understood in internalist, solipsistic terms, throwing away the body, the world and other people" ([7], p. 7). For once the biological system has distilled the statistical regularities of its econiche, it can throw it away, and rely instead on its internal model of the world. This is the representationalist implication of cognitivist FEP.

Ergodicity and the Markov Blanket

This is not how the concept of the model is understood in the FEP as formulated by Friston (Unless we explicitly add "cognitivist" to the FEP, we have in mind a non-cognitivist view of the FEP in what follows). The driving idea here is that a biological system *is* a model of its econiche [11]. As Friston puts it: a biological system "does not have a model of its world—it is a model. In other words, the form, structure, and states of our embodied brains do not contain a model of the sensorium—they are that model" ([10], p. 32). In this sense, each phenotype or living system "embodies an optimal model of its econiche" ([11], p. 89). Moreover, it is not only the biological system that can be said to embody the environment in which it is embedded, but "the environment embodies the agent" ([11], p. 89). Second, embodied models can thus be shown to rest on prior expectations about how environmental states unfold over time. This feature is also known as ergodicity. To say that a system is ergodic is simply to say that its states will tend to revisit a member of its attractive set over time, or that it will return—time and time again—to the same neighborhood in its state space during the span of its lifetime. In other words, ergodic systems will "occupy a small number of states with a high probability and avoid a large number of other states" [42].

The main points about how living systems endure over time apply as much to bacteria and plants as to organisms like us. This raises the questions about continuities and dis-continuities that we want to explore for the FEP. Answers to these questions are not well worked out in the literature on the FEP, and they tend to oscillate between radically different outcomes.

In some passages, Friston's formulation of the FEP seems to get us a deeply continuous view of the life-mind relation—what we called non-cognitivist FEP + strong life-mind continuity in the introduction. He emphasizes that there is more to life than mere self-organization. Living systems also "negotiate a changing […] environment in a way that allows them to endure over substantial periods of time" ([4], p. 422). In other words, they engage in adaptive behavior. Bruineberg et al. [13] characterize this adaptive ability as "an instance of what in philosophical phenomenology is described as being moved so as to tend towards an optimal grip" ([13], p. 10; see also [14,22] for similar remarks).

However, in other passages, Friston pedals what seems to us to be a different and overly generous view of the non-cognitivist FEP. Consider, for example, this quote from Friston: the FEP "applies to any … system that resists a tendency to disorder; from single-cell organisms to social networks" ([9], p. 293). Or, as Friston says here: "The motivation for minimizing free energy has hitherto used the following sort of argument: systems that do not minimize free energy cannot exist, because the entropy of their sensory states would not be bounded and would increase indefinitely— by the fluctuation theorem" ([10], p. 2). Strictly speaking, what Friston [10] says here is that for any system to exist it must work to minimize free energy. This commits Friston to one of the following three implications. First, if free energy minimization is sufficient for mentality, then every system has a mind, even if not all systems are alive. Second, if free energy minimization is enough for life and mind, then all systems that exist are both alive and mental. Finally, biological systems, like all other existing systems, need to work to minimize free energy. The last option states that free energy minimization is not a property of only living systems, and as such sets up one of the two following implications. Either (option one) the FEP places mentality in a class of systems that includes but is

not limited to living systems, and therefore veers toward some form of panpsychism. Or (option two) the FEP equates life-mind continuity with a view that sees life and mind nearly everywhere.

Some scientists and philosophers have embraced the first option for their theories of mind (e.g., [43,44]). However, barring such panpsychist or pancomputationalist implications, it follows that free energy minimization in and of itself cannot be sufficient for life and mentality, even if it is a necessary property of all living and cognitive systems as a means to place an upper limit on disorder. What else might be required? Ergodicity does not seem to be enough for either life or mind given than any random dynamical system will display ergodic behavior because such a system will evolve over time to what is called its random global attractor [45]. Consider that flipping a fair coin not once but several times is an example of an ergodic process given that the probability of observing "heads" will converge as a function of time to the same value.

In addition to ergodicity, Friston appeals to the idea that a Markov blanket bounds systems. A Markov blanket is comprised of sensory and active states, separating internal and external states (in the statistical sense of conditional independence). In a cell, say, the surface of a cell can be cast as constituting a Markov blanket, thus separating intracellular (internal) and extracellular (external) states. A Markov blanket can itself be further partitioned into perceptual/sensory states and active states [10,29]. Assuming that one can interpret internal states as parametrizing "some arbitrary (variational) density or Bayesian beliefs … about external states, then the dynamics of internal and active states can be described as a gradient descent on variational free energy" ([29], p. 3). In other words, one can describe the system in virtue of possessing a Markov blanket as maximizing model evidence via approximate Bayesian inference. Hence, the internal states of a Markov blanket can be said to infer the hidden causes of its sensory input either via perceptual or active inference.

It is worth mentioning that not all phenomena possess a Markov blanket because it shows why—from the perspective of the FEP—mere self-organization is enough neither for life nor for mind. A candle flame is a case in point. Although it can be shown to engage in self-maintaining processes, it does not do so recursively [46], and thus cannot ensure that the states comprising it and their interdependencies change slowly, which is required for a cell or an ergodic system to possess a Markov blanket [10]. In spite of these considerations about Markov blankets, the concerns we have raised here also apply to systems possessing a Markov blanket given that pendulum clocks can be modeled as comprising a Markov blanket, engaging in Bayesian inference, and to exhibit a degree of ergodicity [10]. In the case of pendulum clocks, Bruineberg et al. [13] have argued that the synchronization of chaotic systems can be explained more parsimoniously by appealing to generalized synchrony than by the idea that each clock infers the state of the other given that each clock can be understood as a model of the other's internal dynamics. According to Bruineberg et al. [13], while it is entirely unproblematic to interpret this case in terms of Markov blanketed dynamics, it is not necessary to insist on an inferential reading—an appeal to generalized synchrony will do. We do not have a stake in this particular discussion. Here we merely point out that in other publications, Friston formulates generalized synchrony as consistent with internal states engaging in approximate Bayesian inference [47]. Our point is: given that the core concepts of non-cognitivist FEP—approximate Bayesian inference, ergodicity, Markov blankets and so on—can be applied to living and cognitive systems, on the one hand, and seemingly non-living and non-cognitive systems, on the other, there is a clear danger of these concepts being over-broad in their application, resulting in either seeing life and mind nearly everywhere or in the FEP lacking explanatory power when having to address the nature of life and mind and their relation to one another (Say the left hand side of this disjunction could be defended, then the overly generous non-cognitivist FEP would not exemplify an independence view of life and mind but rather an extreme version of the dependence view of life and mind, given that life and mind would essentially be everywhere).

5. Restricting Non-Cognitivist FEP with REC: An Evolutionary Latecomer View of the Mind

Our strategy now will be to show that the implications for the overly generous view of non-cognitivist FEP can be restricted. Instead of placing mentality outside the domain of life, or even thinking that everything that exists is life-like and therefore mind-like, we shall argue that such

implications can be usefully constrained by work in radically embodied enactivism (REC) in naturalistic philosophy of cognition [21,48]. This restriction will help us in taking first steps toward developing a view of non-cognitivist FEP that avoids any kind of appeal to panpsychism, and where the presence of mental semantic content (i.e., internal states with correctness conditions) marks a transition within the realm of the mental—as opposed to highlighting a path from the non-mental to the mental.

The main commitments of REC can be articulated in two basic tenets. The first is a denial of the "usual" view in the cognitive sciences and its philosophy; namely, that cognition, in some fundamental sense, involves contentful mental representation. That is, according to REC, "the vast sea of what humans [and other organisms] do and experience is best understood by appealing to dynamically unfolding, situated embodied interactions and engagements with worldly offerings" ([48], p. 1). The second is the claim that these kinds of basic cognitive activities are realized in an organism's world-engaging, bodily activity [21]. We now turn to highlight two points of convergence between REC and non-cognitivist FEP.

First, they both emphasize co-dependence between the internal and external. Recall that the FEP casts the biological system as a model of its econiche, and adds to this that the econiche is a model of the biological system. For example, the spider's morphology, possibility for action, and so on, are reflective of its niche, while the web and its wider embedding environment reflect the kind of organism that inhabits it. To be clear, the FEP does posit a separation of internal and external states—the Markov blanket implies such a separation [28]. However, unlike cognitivist FEP, it does not follow that any such separation is indicative of an *epistemic* separation. So, the mere existence of a Markov blanket does not entail that the internal states of an organism must represent states beyond the blanket in virtue of constructing internal states with semantic content about those external states. Rather, it suggests in a similar sense of the enactivist claim that "the organism and environment are bound together in a reciprocal specification and selection" ([49], p. 174; quoted in [32], p. 289). Allen and Friston emphasize this mutuality between the internal and external, as they say: "The point is that the boundary itself is constituted by ergodic dynamical interchange between 'internal' and 'external', rather than a cognitivist performance of internal processing" ([28], p. 16). REC casts this mutuality between the internal and the external in the form of covariance, which is an expression of how two or more variables change (or do not change) together. Covariance is formally equivalent to the emergence of generalized synchrony that results from active inference—a corollary of the FEP (see [47])—and expresses a dynamic coupling of two or more random dynamical systems. In this sense, both REC and the FEP formulate, from basic principles, why one should expect shrinkage in entropy given generalized synchrony or covariance in agent-environment couplings. Akin to active inference in the non-cognitivist FEP, REC understands the assembly and orchestration of such dynamic couplings to be a result of embodied activity.

Second, the non-cognitivist FEP and REC conceive of the kind of information available to a biological organism to be of the Shannon variety. However, where the FEP assumes this kind of information to provide an account of self-organization in biological systems [5], REC draws an additional implication that is intended to raise a serious problem for all cognitivist theories of mind, including cognitivist FEP [21]. Cognitivist FEP assumes that any kind of intelligent interaction with the world demands semantic content. Yet, this assumption runs afoul against what Hutto and Myin [21] dub the *hard problem of content*. As they say: "positing informational content is incompatible with explanatory naturalism. The root trouble is that Covariance doesn't Constitute Content" ([21], p. xv). The driving idea is that semantic content does not exist independently of certain socio-cultural practices. Hence, semantic content is not an inherent property of biological systems and therefore not of life ([21], p. xv).

We do not comment on the relation between content and social practices (but see [50]). Instead we shall hone in on the following implication: assuming that information-as-covariance cannot give rise to information-as-content, and assuming that responding adaptively to information-as-covariance is an essential property of living and cognitive systems, then mentality is not in its most basic forms a matter of processing any kind of content. It is precisely for this reason that we argued

that cognitivist FEP leads to a no-continuity or independence position with respect to the life-mind relation. We based this claim on the observation that cognitivist FEP cannot help itself to the kind of information at the base of the FEP as it applies to self-organization in biological systems, namely information-as-covariance, to account for processes involving semantic information, and that this thereby forces it to exclude most processes in biology from the domain of the mental.

REC goes further than just problematizing cognitivist theories of mind. It develops a positive picture of non-contentful cognitive activity. According to REC, if one denies that any kind of organismic interaction with the world must, necessarily, involve content, it does not follow that this kind of interaction is, necessarily, non-mental. On this view, it is possible for organisms to be *intentionally directed*—to be active, involved in world-directed engagements, and to be informationally responsive—without directedness of this kind being mediated by internal states with semantic content. As Hutto and Myin say: "The simplest life forms are capable of an intentionally directed responding" ([21], p. x), but this responding is not a matter of interpretation, understanding, or any other kind of sophisticated representational activity.

REC develops its account of non-contentful intentional directedness via a modified version of teleofunctionalism. Simply put, teleofunctionalism is the view that what makes something a mental activity has to do with the function it serves for an organism. The notion of function is cast in terms of *proper* (*biological*) *functions*, and such functions are commonly understood in light of the evolutionary and/or historical conditions under which the respective functions were selected for and/or acquired. Hence, teleofunctionalism explains why an organism has certain mental and biological capacities by appealing to the kind of operations that such functions were selected for in order to increase the probability of surviving and striving.

Many teleofunctionalist are representationalists about mentality. REC rejects this representational commitment of teleofunctionalism. According to Hutto and Myin [21], the intentional directedness of basic forms of cognitive activity—e.g., a frog flicking its tongue to catch a fly—are constituted by their natural history of selection. We made the same point as we discussed chemotaxis in single-celled organisms above. This implies that intentional directedness has a normative dimension. An intentionally directed form of activity (of the non-semantic kind) "aims at engendering certain types of organismic responses to certain things (or types of things) and not others" ([51], p. 142). A nice outcome of REC is that it implies a view of mind within which a transition from contentless to contentful activities does not mark a transition from no-mind to mind. REC is thus a transformative theory of mind within the mental realm.

Not all naturally occurring systems satisfy REC's condition for mentality. That is, not all things that exist exhibit intentional directedness. A thermometer, for example, does not. This is not because such a device lacks systemic functions of the kind that are defined by the role they play for the device. According to REC, such machinery lacks the kind of natural selection history associated with proper biological functions, which is required for a system to exhibit intentional directedness. Assuming that REC's proposal is along the right track, developing non-cognitivist FEP under the auspices of REC's teleofunctionalism would thus imply the following important constraint: that only a subset of free energy minimizing systems are intentionally directed at certain features of their local environment, and therefore only a subset of such systems are mental.

This issue requires a lot more discussion than we shall attempt here. Nonetheless, the distance between non-cognitivist versions of the FEP and REC is not very large—at least not once an appeal to a history of selection is recognized. Thus we read: "according to RPP [i.e., radical predictive processing based on the free energy principle] prediction error unfolds not only at ontogenetic but also phylogenetic timescales; if the brain (and body) constitute a generative model, then those [generative models] best suited to their environmental niche will be selected by evolution" ([28], p. 9). This is revealing. Allen and Friston [28] here recognize the need to develop a view of generative models, dynamic coupling, and bodily activity along the lines suggested by REC. By recognizing this overlap between non-cognitivist FEP and REC, we thus arrive at the idea that it is possible to develop a non-semantic view of free energy minimizing organisms, and to do so in a way that does not lead to unconstrained mental bloat.

Nevertheless, we will now argue that REC stops short—in the sense of undercutting—a strong version of the life-mind continuity thesis (Note that stopping short of a strong life-mind continuity thesis does not entail that REC faces a fatal dilemma when having to account for the origins of content in naturalistic terms (see [52] for discussion)). By extension, if the non-cognitivist FEP were to be developed within the confines of REC, it too will stop short of arriving at such a strongly unifying view of life and mind. The reason for this is that once REC is added to the FEP, this addition results in an evolutionary latecomer view of mind. REC places the origins of life prior to the emergence of mind, given its adoption of teleofunctionalism. It follows that REC can be characterized as accepting the following three claims: first, that there is life but no cognition; second, that there is life and that life and mind converge when living systems become capable of non-contentful forms of intentional directedness; and finally, that certain kinds of living systems (human beings, for example) are able to engage in cognitive activity with semantic content (such a writing and performing mathematics). We do not deal with this last implication of REC. Instead we focus and have focused on the second claim, which is the claim that for REC organisms come to exhibit mentality—defined as intentional directed actions—given a selection history. Thus, if added to the FEP, REC narrows the scope of non-cognitivist FEP, given that this appeal to a history of selection would then imply that the very first life forms do not have mental properties given that such life forms would not (at least not yet) have a history of natural selection.

We take issue with this view of the life-mind relation that REC entails, especially since it is somewhat unclear what explanatory work the appeal to selection is doing. Consider, e.g., when the first living system begins to divide. This process results in a number of clones of the first individual system. Yet, given environmental fluctuations, not all clones are able to survive. This raises a question: Have the survivors been "naturally selected"? If affirmative, then the survivors have intentional directedness, and the others do not. However, it is likely the case that there is no notable functional difference between the original individual and any of its clones—independent of them surviving or dying. The implication of this would thus be that physiologically identical organisms could differ in terms of mentality, raising a philosophical worry similar to the one evoked by the "swamp man" thought experiment or the notion of philosophical (free energetic) zombies.

Furthermore, the appeal to "selection" raises additional issues. For example, at which point can we say that a function was "selected"? Is it the moment when the function for the very first time contributes to an organism's relatively higher rates of reproduction compared to its conspecifics? Or does it require several generations of individuals with above average fitness? And how can this particular function's contribution to fitness be non-trivially disentangled from the myriad of other contributing factors? To be sure, there is no doubt that over generations natural selection can change the particular way in which organisms are intentionally directed to their environment, but it is not clear whether this evolutionary process should be taken as constituting this very intentional directedness as such (and hence, basic mentality) in the first place.

6. From REC to AE and Non-Cognitivist FEP: Strong Life-Mind Continuity

We finally suggest a way by which to positively rehabilitate the above problem, and we do this by showing that non-cognitivist FEP is also in line with autopoietic enactivism, enabling us to weave the origins of life together with the origins of mind (This account is a further development of the line of argument pursued in [14]).

REC assumes as its starting point that there is a population of individuals that can be shaped by the forces of natural selection. However, what precisely is an individual? (We are not in a position to discuss this in depth here—a task for another occasion. However, for an initial detailed discussion see [53] and the rest of the papers in this special issue in *Biology and Philosophy*). If this concept is too broad, then it is hard to constrain the notion of intentional directedness appropriately, since many kinds of systems undergo histories of some kind of selection with regard to their functioning, like a thermostat. Moreover, it is undeniable that the thermostat's functioning can be described normatively, namely in terms of whether it correctly regulates temperature or not. However, its designers and users specify the conditions of "correctness" externally, and therefore its normativity is only a derived

kind of normativity. This is not so for the case of living systems. For example, our bodies are also continuously regulating their temperature so as to maintain it within specific boundaries. According to the FEP, this is achieved by free energy minimization, on average and over time. This function is also subject to conditions of success and failure, but here the conditions are intrinsic to the body rather than externally defined for the simple reason that they are determined by a range of viability, which ultimately is a question of existence, of life and death [28].

To be fair, even a malfunctioning thermostat may eventually disintegrate from overheating. However, this is not sufficient to attribute to it an intrinsic form of normativity. The key difference is that our body, or any living system for that matter, and in contrast to the thermostat, is a physically self-producing system (autopoiesis). Living is a process that maintains itself under far-from-thermodynamic-equilibrium conditions by investing work into preventing its own disintegration. In this manner it must define a boundary between what belongs to itself and what belongs to its environment. This boundary defines it as an autonomous individual. According to non-cognitivist FEP, such a system is an adaptive system capable of active inference with a Markov blanket, where the Markov blanket does not entail anything like an epistemic boundary demarcating mind from world. Whereas the mode of being of a thermostat is characterized by passive persistence, an organism is characterized by active existence (or, active inference). Additionally, because an organism's being is its own doing, its existence as an individual is inherently precarious, and yet living beings generally do all they can to maintain their way of life against all odds. That is, despite the possibility of simply letting themselves to succumb to the forces of decay and disintegration, they consistently regulate their boundaries to avoid this fate. In other words, living beings do not simply passively undergo perturbations like non-living systems, they respond selectively and the success of this response is normatively linked to the preservation of their way of life.

So, autopoietic enactivism (AE) provides us with the beginnings of an account of how the origin of life is also the origin of autonomous individuality and of intrinsic normativity (the origins of systems maintaining the internal states of their Markov blankets via active inference). However, at first this normativity is not specific to any particular function but rather concerns the individual as a whole: all activity that does not kill the individual has a positive value for the living system. Activity that kills the individual would have a negative value except that it can no longer be so for the system (which is gone). In order to be able to function in a more discerning manner it is necessary for the living system to be able to respond in a differential manner and particularly to detect and avoid potentially lethal interactions before they turn out to be lethal. It is unlikely that this kind of complexity can spontaneously emerge at the origin of life. Instead it requires a history of selection, conceived of in general terms as the outcome of interaction, development, and evolution. In other words, here is where AE connects with the concerns of REC and intentional directedness enters the picture as a way to adaptively respond to relevant features of the inner and outer milieu of the organism. While AE helps REC to appropriately constrain its appeal to selection history, REC in return can help AE to bridge the "cognitive gap" [26] between basic minds and non-basic minds.

Approaching intentional directedness in this manner has the advantage that it arrives in the context of an autonomous individual that is already actively distinguished from, as well as related to, what is other (its environment), and whose activity is already intrinsically characterized by a basic form of holistic normativity. Additionally, there is no need to worry about the mental status of artifacts with artificially selected functions, since they lack the criteria of basic life (autopoiesis and adaptivity, see [54]). Intentional directedness is, in this view, effectively a refinement and elaboration of a basic existential profile by means of selective history. The conditions of what REC considers a basic mind, namely an autonomous individual having intrinsic normativity, would already be in place. If so, then we may say that basic mind and basic life arose together at the origin of life, a momentous event, which then should be reconceived of as the origin of life-and-mind—as the origins of organisms with Markov blankets capable of maintaining their own internal states (or processes) through active inference (In this paper we have taken steps towards developing an approach to the strong life-mind continuity thesis that can bring together key insights in the free energy principle with the latest work in enactive cognitive science. However this task is far from finished. We suspect

that future work on this project will seek to evaluate the FEP not only in relation to the REC and AE frameworks but also in terms of other accounts of predictive and anticipatory dynamics. For work already started in this direction, see [55–57]).

7. Conclusions

If something like the FEP is in the right ballpark for addressing the life-mind relation, then life and mind co-emerge within the broader class of processes pertaining to the minimization of uncertainty. However, there is no singular account of this general class for thinking about the relation between life and mind. In this paper, we have explored two different FEP perspectives on the life-mind relation. First, we argued that cognitivist FEP implies that there is no continuity between life and mind. Essentially, cognitivism's appeal to semantic content as the mark of the mental places a very high bar that can only be passed by a small number of recent species and potentially by their artifacts. Second, we then argued that overly generous formulations of non-cognitivist FEP threaten to imply that mentality is nearly everywhere.

To avoid these implications, we then argued that a non-cognitivist version of the FEP could be helpfully constrained by integrating it with key ideas in enactive cognitive science about individuality and normativity. Finally, we concluded that the most compelling account of the relationship between life and mind treats them as strongly continuous, and that this continuity is based on particular concepts of basic life (autopoiesis and adaptivity) and basic mind (intentionally directed yet non-semantic) in the context of the free energy principle.

Author Contributions: Michael D. Kirchhoff and Tom Froese both contributed substantially to writing this paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Clark, A. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.* **2013**, *36*, 181–253.
2. Friston, K.J. The free-energy principle: A unified brain theory? *Nat. Rev. Neurosci.* **2010**, *11*, 127–138.
3. Hohwy, J. The Neural Organ Explains the Mind. Available online: http://open-mind.net/papers/the-neural-organ-explains-the-mind/at_download/paperPDF (accessed on 14 April 2017).
4. Friston, K.J.; Stephan, K.E. Free-energy and the brain. *Synthese* **2007**, *159*, 417–458.
5. Friston, K.; Thornton, C.; Clark, A. Free-energy minimization and the dark-room problem. *Front. Psychol.* **2012**, *3*, 130.
6. Hohwy, J. *The Predictive Mind*; Oxford University Press: Oxford, UK, 2013.
7. Howhy, J. The self-evidencing brain. *Noûs* **2014**, *50*, 259–285.
8. Hohwy, J. How to Entrain Your Evil Demon. Available online: http://predictive-mind.net/papers/how-to-entrain-your-evil-demon/at_download/paperPDF (accessed on 14 April 2017).
9. Friston, K.J. The free-energy principle: A rough guide to the brain? *Trends Cogn. Sci.* **2009**, *13*, 293–301.
10. Friston, K.J. Life as we know it. *J. R. Soc. Interface* **2013**, *10*, doi:10.1098/rsif.2013.0475.

11. Friston, K.J. Embodied inference: Or 'I Think Therefore I Am, If I Am What I Think'. In *The Implications of Embodiment (Cognition and Communication)*; Tschacher, W., Bergomi, C., Eds.; Imprint Academic: Exeter, UK, 2011; pp. 89–125.

12. Clark, A. How to Knit Your Own Markov Blanket: Resisting the Second Law with Metamorphic Minds. Available online: http://www.x-spect.org/uploads/9/8/1/5/98154170/knittingmarkov8.pdf (accessed on 14 April 2017).

13. Bruineberg, J.; Kiverstein, J.; Rietveld, E. The anticipating brain is not a scientist: The free-energy principle from an ecological-enactive perspective. *Synthese* **2016**, doi:10.1007/s11229-016-1239-1.

14. Kirchhoff, M.D. Autopoiesis, free energy, and the life-mind continuity thesis. *Synthese* **2016**, doi:10.1007/s11229-016-1100-6.

15. Godfrey-Smith, P. Mind, matter, and metabolism. *J. Philos*. **2017**, submitted.

16. Froese, T.; Ikegami, T. The brain is not an isolated "black box," nor is its goal to become one. *Behav. Brain Sci.* **2013**, *36*, 213–214.

17. Kirchhoff, M.D. Species of realization and the free energy principle. *Australas. J. Philos.* **2015**, *93*, 706–723.

18. Bruineberg, J.; Rietveld, E. Self-organization, free energy minimization, and optimal grip on a field of affordances. *Front. Hum. Neurosci.* **2014**, doi:10.3389/fnhum.2014.00599.

19. Di Paolo, E. Extended life. *Topoi* **2009**, *28*, 9–21.

20. Froese, T.; di Paolo, E.A. The enactive approach: Theoretical sketches from cell to society. *Pragmat. Cogn.* **2011**, *19*, 1–36.

21. Hutto, D.D.; Myin, E. *Radicalizing Enactivism*; The MIT Press: Cambridge, MA, USA, 2013.

22. Kirchhoff, M.D. Experiential fantasies, prediction, and enactive minds. *J. Conscious. Stud.* **2014**, *22*, 68–92.

23. Stewart, J. Life = Cognition: The Epistemological and Ontological Significance of Artificial Life. In *Toward a Practice of Autonomous Systems: Proceedings of the First European Conference on Artificial Life*; Varela, F.J., Bourgine, P., Eds.; MIT Press: Cambridge, MA, USA, 1992; pp. 475–483.

24. Stewart, J. Cognition = life: Implications for higher-level cognition. *Behav. Process.* **1996**, *35*, 311–326.

25. Bitbol, M.; Luisi, P.L. Autopoiesis with or without cognition: Defining life at its edge. *J. R. Soc. Interface* **2004**, *1*, 99–107.

26. Froese, T.; di Paolo, E. Sociality and the life-mind continuity thesis. *Phenomenol. Cogn. Sci.* **2009**, *8*, 439–463.

27. Thompson, E. *Mind in Life*; MIT Press: Cambridge, MA, USA, 2007.

28. Allen, M.; Friston, K.J. From cognitivism to autopoiesis: Towards a computational framework for the embodied mind. *Synthese* **2016**, doi:10.1007/s112 29-016-1288-5.

29. Friston, K.J.; Levin, M.; Sengupta, B.; Pezzulo, G. Knowing one's place: A free energy approach to pattern regulation. *J. R. Soc. Interface* **2015**, *12*, 20141383.

30. Gladziejewski, P. Predictive coding and representationalism. *Synthese* **2015**, doi:10.1007/s11229-015-0762-9.

31. Hohwy, J. Attention and conscious perception in the hypothesis testing brain. *Front. Psychol.* **2012**, *3*, 74–87.

32. Clark, A. *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*; Oxford University Press: Oxford, UK, 2016.

33. Colombo, M.; Wright, C. Explanatory pluralism: An unrewarding prediction error for free energy theorists. *Brain Cogn.* **2016**, doi:org/10.1016/j.bandc.2016.02.003.

34. Piccinini, G. Computation in Physical Systems. Available online: https://plato.stanford.edu/entries/computation-physicalsystems (accessed on 14 April 2017).

35. Shannon, C. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 279–423.

36. Godfrey-Smith, P.; Sterelny, K. Biological Information. Available online: https://plato.stanford.edu/entries/information-biological/ (accessed on 14 April 2017).

37. Piccinini, G. *Physical Computation: A Mechanistic Account*; Oxford University Press: Oxford, UK, 2015.

38. Wilson, R.; Foglia, L. Embodied Cognition. Available online: https://plato.stanford.edu/entries/embodied-cognition/ (accessed on 14 April 2017).

39. Shani, I. Intentional directedness. *Cybern. Hum. Know.* **2006**, *13*, 87–110.

40. Auletta, G. Information and metabolism in bacterial chemotaxis. *Entropy* **2013**, *15*, 311–326.

41. Boden, M.A. Is metabolism necessary? *Br. J. Philos. Sci.* **1999**, *50*, 231–248.

42. Rabinovich, M.I.; Friston, K.J.; Varona, P. (Eds.) *Principles of Brain Dynamics: Global State Interactions*; The MIT Press: Cambridge, MA, USA, 2012.

43. Chalmers, D. Panpsychism and panprotopsychism. In *Consciousness in the Physical World: Perspectives on Russellian Monism*; Alter, T., Nagasawa, Y., Eds.; Oxford University Press: Oxford, UK, 2015; pp. 246–276.

44.   Tononi, G.; Koch, C. Consciousness: Here, there and everywhere? *Philos. Trans. R. Soc. B* **2015**, *370*, doi:10.1098/rstb.2014.0167.

45.   Crauel, H. Global random attractors are uniquely determined by attracting deterministic compact sets. *Annali di Matematica Pura ed Applicata* **1999**, *4*, 57–72.

46.   Bickhard, M.H. Interactivism: A manifesto. *New Ideas Psychol.* **2008**, *27*, 85–95.

47.   Friston, K.J.; Frith, C. A duet for one. *Conscious. Cogn*. **2015**, *36*, 390–405.

48.   Hutto, D.D.; Kirchhoff, M.D.; Myin, E. Extensive enactivism: Why keep it all in? *Front. Hum. Neurosci.* **2014**, doi:10.3389/fnhum.2014.00706.

49.   Varela, F.; Thompson, E.; Rosch, E. *The Embodied Mind*; MIT Press: Cambridge, MA, USA, 1991.

50.   Hutto, D.D. *Folk-Psychological Narratives*; The MIT Press: Cambridge, MA, USA, 2008.

51.   Hutto, D.D. Against passive intellectualism: Reply to Crane. In *Radical Enactivism*; Menary, R., Ed.; John Benjamins Publishing Company: London, UK, 2006.

52.   Hutto, D.D.; Satne, G. Continuity scepticism in doubt: A radically enactive take. In *Embodiment, Enaction, and Culture*; Durt, C., Fuchs, T., Tewes, C., Eds.; The MIT Press: Cambridge, MA, USA, 2016.

53.   Pradeu, T. The many faces of biological individuality. *Biol. Philos.* **2016**, *31*, 761–773.

54.   Froese, T.; Ziemke, T. Enactive artificial intelligence: Investigating the systemic organization of life and mind. *Artif. Intell.* **2009**, *173*, 366–500.

55.   Bickhard, M.H. The anticipatory brain: Two approaches. In *Fundamental Issues of Artificial Intelligence*; Müller, V.C., Ed.; Springer: Berlin/Heidelberg, Germany, 2016; pp. 259–281

56.   Rosen, R. *Anticipatory Systems: Philosophical, Mathematical, and Methodological Foundations*; Springer: Berlin/Heidelberg, Germany, 2012.

57.   Nasuto, S.J.; Hayashi, Y. Anticipation: Beyond synthetic biology and cognitive robotics. *BioSystems* **2016**, *148*, 22–31.