Representation Wars:

Enacting an Armistice through Active Inference

**Authors**

Axel Constant1,2,3\*

Andy Clark4,5,6

Karl J. Friston3

**Institutions**

1. Charles Perkins Centre, The University of Sydney, John Hopkins, AU.

2. Culture, Mind, and Brain Program, McGill University, CA.

3. Wellcome Trust Centre for Human Neuroimaging, University College London, UK.

4. Department of Philosophy, The University of Sussex, UK.

5. Department of Informatics, The University of Sussex, UK.

6. Department of Philosophy, Macquarie University, AU.

**\*Corresponding author**

axel.constant.pruvost@gmail.com, Theory and Method in Biosciences, Level 6, Charles Perkins Centre D17, Johns Hopkins Drive (off Missenden Road), The University of Sydney, NSW 2006 Australia, phone: +610410756661

**Funding Acknowledgements**

Work on this article was supported by the Australian Laureate Fellowship project A Philosophy of Medicine for the 21st Century (Ref: FL170100160) (AConstant), by a Social Sciences and Humanities Research Council (SSHRC) doctoral fellowship (Ref: 752-2019-0065) (AConstant), by the European Research Council (ERC) Advanced Grant XSPECT - DLV-692739 (AClark), and by a Wellcome Trust Principal Research Fellowship (Ref: 088130/Z/09/Z) (KJF).

**General acknowledgements**

We thank Ian Robertson and Maxwell Ramstead for discussions that influenced this manuscript, and especially Paul Badcock for their comments on earlier versions of this manuscript.

**Abstract**

Over the last 30 years, intellectualist and dynamicist positions in the philosophy of cognitive science have been arguing over whether neurocognitive processes should be viewed as representational or not. Major scientific and technological developments over the years have furnished both parties with ever more sophisticated conceptual weaponry. In recent years, an enactive generalisation of predictive processing – known as active inference – has been proposed as a unifying theory of brain functions. Since then, active inference has fuelled intellectualist and dynamicist campaigns. However, we believe that when diving into the formal details of active inference, one should be able to find a solution to the war; if not a peace treaty, surely an armistice of sort. Based on an analysis of these formal details, this paper shows how both intellectualist and dynamicist sensibilities can peacefully coexist within the new territory of active inference.

**Keywords**

Representationalism, Embodiment, Active inference, Free energy principle, Philosophy of cognitive science

# 1 Introduction

This paper proposes a way to end the representation wars. Focusing on recent formal developments, we aim to show that the concept of generative models as applied to the brain under active inference accommodates a representationalist and a dynamicist (a.k.a. non-representational) view of cognition. More precisely, we show that the architecture or configuration of neuronal pathways under a (Markovian) generative model (for discrete state spaces) can – and generally speaking will – realise both representational and non-representational processes.

In section 2 of this paper, to help readers unfamiliar with the notion of representation in the philosophy of cognitive science, we present a heuristic overview of its history, focussing on salient moments of its war over the past 30 years. Although this history is complex and more nuanced than what we can present here, we believe that this discussion evinces some of the motivations behind the notion of representation and its contestations. In section 3, we present the architecture of generative models’ representationalist pathways. This allows us to segue into a discussion of dynamic pathways in section 4. Section 5 discusses some worries. We then conclude in section 6 with some brief remarks on good practice in the philosophy of cognitive science, when appealing to the mechanics of active inference.

Note that we do not engage with debates concerning active inference *per se*, nor do we venture into a philosophical justification of its use in cognitive neuroscience. Rather, we start from the premise that active inference is a suitable theory, as evidenced by the large literature that evidences, employs, argues for, and teaches its workings. For a comprehensive introduction and for a review of the formal fundaments and empirical evidence, we refer the reader to [(Buckley et al. 2017; Bogacz 2017; Beal 2003; Friston 2018; Keller and Mrsic-Flogel 2018; Parr et al., 2019)](https://paperpile.com/c/9SdDf0/6IVE+YQjE+DwUb+tZRZ+tj6H).

Also, note that the argument presented in this paper is of a different kind than those currently available in the literature on representationalism in predictive processing and active inference (e.g., Allen and Friston, 2016; Clark, 2015abc; Dolega, 2017; Gładziejewski, 2016; Kiefer & Hohwy, 2017), as it is not based on an intuitive conceptual analysis, but rather on a formal, analytic reading of the theory. The argument we present is simple. We show that if one agrees with the sufficient criteria for representationalism described in section 2, then one is compelled to agree with the claim made in this paper; namely, that formally, representational and non-representational cognitive processes can be implemented by the brain under active inference. Given that active inference is a formal theory of functional neuroanatomy, the debates on the representational nature of brain processes concerning active inference should come to an end. Thus, the upshot of this paper is to move forward the philosophical debates on representationalism in active inference and to enable practical debates about the varieties of possible implementations of representational and non-representational neuronal processes.

# 2 30 years of representation wars

## 2.1 80’s Connectionism

As the story goes, until the 90s, driven by advances in computer science, the philosophy of cognitive science was dominated by cognitivism and connectionism (e.g. Fodor, 1975; Churchland, 1989). Connectionism was presented as a first attack on cognitivism – cognitivism being an attempt at understanding the brain as a logical symbol manipulating system. For connectionists, the brain should not be studied as a symbol manipulating system, but rather, consistent with the brain’s actual neurophysiology, as a set of hierarchically deployed neural networks. The spirit of connectionism is still very much alive today, such as in deep learning research (for a review see LeCun, Bengio, and Hinton, 2015).

Both cognitivism and connectionism deal with a view of cognition as a problem-solving activity. And both paradigms have typically invoked some notion of representation, with the main difference being whether the representations had symbol-level content or something softer, something contentful yet ‘sub-symbolic’ (for extensive discussion, see Clark, 1989; 1993).

Although these are different types of representations, each involving different criteria, a cognitive process will – for the purposes of this paper – be deemed representational whenever that process can be said to fulfil the following sufficient conditions (See Hutto & Myin, 2013; Seigle, 2010):

(i)The cognitive process is about something else (a.k.a. aboutness).

(ii)The cognitive process has satisfaction conditions with respect the thing it is about.

## 2.2 90’s Dynamicism

The 90’s marked the rise of embodied views in cognitive science such as enactivism (Varela, Thompson, and Rosch, 1991) and radical embodied cognition (Chemero, 2009). Embodied approaches were motivated by developments in the field of dynamical system theory, which casts cognitive systems as coupled quantitative variables, mutually changing interdependently over time (Van Gelder, 1995; Thelen & Smith, 1996; Beer, 2000); one variable being the organism, the other being the environment. Dynamicism has been driven by two main criticisms of much previous work, (Thompson, 2007):

(i)Since the brain is embodied, we cannot abstract cognition from the body, and consequently from the environment;

(ii)Since representationalism posits the mediation of the world and cognition by the mental manipulation of representations, representationalism cannot genuinely acknowledge embodiment;

Therefore, for these kinds of dynamicists – see Clark (1997) for a more liberal approach – we should reject the representational view of cognition altogether. Instead, cognition should be viewed as a process of self-organisation among the components of the biological system performing the cognitive activity. These components include the brain (internal states) and the body and the environment (external states). On that view, cognition is a homeostatic and allostatic process of attunement to cope with environmental perturbations; a process of ‘coping, not computing’.

## 2.3 2000’s active inference Westphalia?

At the turn of the millennium, based on a Helmholtzian view of embodied perception, the theory of active inference was introduced as a realisation of the free energy principle [(Friston et al., 2006, Friston 2010; Friston et al. 2016)](https://paperpile.com/c/9SdDf0/TLip+O2kV). This enactive generalisation of predictive processing marked a paradigm shift in cognitive science: active inference became a potential candidate to meet the challenge of the grand unification of neurocognitive functions (Clark, 2013). Since then, many enthusiasts have leveraged active inference to attempt explanations of the underlying computational processes of biobehavioural functions such as action, perception, learning, attention, memory, decision making, emotions, planning and navigation, visual foraging, communication, social learning, and many more [(Friston et al. 2016; Feldman and Friston 2010; Joffily and Coricelli 2013; Kaplan and Friston 2018; Mirza et al. 2016; Constant et al. 2018; Friston and Frith 2015; Parr and Friston 2017; Badcock, Friston & Ramstead, 2019)](https://paperpile.com/c/9SdDf0/O2kV+QoVt+StgP+JZzS+5MaV+CpFD+WUtz+s7qa).

In line with much of Bayesian statistics, active inference claims that the brain is fundamentally in the business of finessing a generative model of the causes of its sensations; as if the brain was a scientist, trying to infer the causal architecture of its own relation to its world. Put another way, under active inference, the brain is a dynamical system that models the action-relevant causal structure of its coupling with the other dynamical system that embeds it – the body and the environment (i.e., the system generating its sensations). The mathematical formalism of active inference describes neuronal dynamics as a gradient flow that optimises the evidence for a generative model of the lived world. On this view, neuronal networks embodied by the brain form a set of nodes (modelling hidden states) and edges (modelling conditional dependencies) of a probabilistic (Bayesian) graphical model.

But active inference itself soon became contested ground too. The cognitivist campaign claimed that ‘brains as generative models’ were “rich and reconstructive, detached, truth-seeking inner representations” of the world (Hohwy 2013, 2016); others (such as Clark (2016)) resisted by claiming that generative models were in fact manifest as transient webs of neuronal coupling that are cost-efficient, and sometimes (though not always) freed from heavy-duty manipulation of internal representations – generating actions that exploited environmental opportunities by weaving themselves closely to the opportunities provided by body and the world (Clark 2013; 2016).

# 3 Representational pathways in active inference

Active inference assumes that the brain entails[[1]](#footnote-1) a causal model of the world (a.k.a. a generative model), whose structure represents the components involved in the cognitive function of interest, as well as the dynamics that realise that cognitive function. Formally, these components and dynamics are expressed as a Bayesian graphical model, with nodes and edges representing the dynamic relations among components, and the structure of which is assumed to map onto the neuroanatomy of neuronal systems realising any cognitive function. The cognitive function, then, is realised by these dynamics – that play the role of neuronal message passing in the service of belief updating (i.e., inference) that underwrites the cognitive function in question (Friston et al., 2017).

The representational interpretation of active inference is employed to study cognitive functions that rely on dynamics and components of the generative models that involve the internal manipulation of representational content (e.g., beliefs about hidden states of the world[[2]](#footnote-2), including one's body and physiology). (Hohwy, 2013; 2019). The motivation for appealing to representational generative models to explain perception and action in active inference stems from the inverse nature of the dual inference problems our brains solve (i.e., figuring ‘what causes what’ before inferring ‘what caused that’):

(i) Perceptual problem: the brain does not have direct access to causes of sensations, nor is there a stable one-to-one mapping between causes and sensations. For instance, a sensory input (e.g., red sensation) may be caused by multiple fluctuating causes (e.g., red jacket, red car, red traffic light). In the philosophical literature, this problem is sometimes referred to as the black box, seclusion or solipsism problem [(Clark 2013; Hohwy 2016; Metzinger and Wiese 2017)](https://paperpile.com/c/9SdDf0/uvAx+0H9b+aRQi).

(ii)Action planning problem: all the brain can work with are the sensory inputs it receives. If we are to engage adaptive action, we must not only infer the causes of our sensations (i.e., forming a sufficiently veridical perception – or conception – of the world in which we currently find ourselves), but we must also predict the consequences of engaging in this or that action in the future. In the philosophical literature, this problem is sometimes referred to as the problem of mere versus adaptive active inference [(Bruineberg, Kiverstein, and Rietveld 2016; Kirchhoff et al. 2018)](https://paperpile.com/c/9SdDf0/4iHG+Y4Jj), and requires action planning (c.f., planning as inference in machine learning).

This means that under active inference, agents like us must find a solution to infer, in an ill-posed setting, both the nature of the cause of our sensations (e.g., the jacket, the traffic light, or the car), and to infer what action will lead to outcomes that are consistent with our model of the lived world (e.g., being on the other side of the street vs. under the wheels of a car). Under active inference, perception and action are explained as solutions to these inverse problems – crucially, solutions that rest upon optimising exactly the same quantity, as we will see below.

## 3.1 Perception

Formally, the problem of indirect perception can be approached as follows. Consider a sensory outcome  generated by a hidden state of the world (). Taken together, these can be viewed as forming a joint probability distribution (). The only quantity to which the brain has access is a sensory consequence, not its cause. To perceive things, the brain must reconstruct the hidden state or cause (), or rather its posterior probability; i.e., the probability of the cause, after observing the sensory datum .

To infer this posterior probability, the brain learns the causal (i.e., generative) model of the manner in which the world caused the sensation. Learning here is a technical term. It refers to the optimisation of the parameters of a model – here the generative model. The brain learns the parameters of hidden states causing sensory outcomes; about which the brain may have prior beliefs. These prior beliefs are part of the generative model entailed by the brain. Hence, a generative model decomposes into prior beliefs about hidden states and a likelihood of these hidden states, given outcomes. One can easily follow this decomposition by visualising the graphical model that makes up the generative model in figure 1.



Fig.1. Elementary generative model for perception and the problem of indirect inference.

In an ideal scenario, the brain could use Bayes rule to infer the true probability of the cause, by using the probability of the data, known as model evidence or marginal likelihood:

 (4)

The marginal likelihood  refers to the probability of sensory data averaged – or marginalised – over all possible hidden states:

 (5)

To represent the marginal likelihood and perform exact inference (as in eq.4), the marginalisation that the brain would have to perform would be intractable, as there may be a near infinite number of causes with various probabilities for each sensory datum. This is at the core of the inverse problem of inference; direct calculation of the posterior probability of one’s beliefs given sensory data  is simply intractable. Thus, the problem of indirect inference may be restated as follows: the brain cannot access the true posterior probability over the causes of its sensations because this requires evaluating an intractable marginal likelihood. What the brain can do, however, is to perform ‘approximate Bayesian inference’ based on its prior beliefs and the sensory data it receives[[3]](#footnote-3). In active inference, the ‘manipulation of content’ rests on this method of inference known as approximate Bayesian inference [(Feynman 1972; Dayan et al. 1995; Beal 2003)](https://paperpile.com/c/9SdDf0/o6XH+2ROc+DwUb).

Approximate Bayesian inference allows the inversion of the generative model to estimate the marginal likelihood via an approximation to the true posterior over sensory causes (i.e., what the brain would do using exact Bayesian inference if it had access to the marginal likelihood). Taking advantage of Jensen’s inequality, the method of approximate Bayesian inference involves the minimisation of an upper bound on (negative log) model evidence (a.k.a. surprisal), called variational free energy. This bound is constructed by using an arbitrary probability distribution[[4]](#footnote-4)  that is used to minimise the variational bound – and the generative model  :

 (6)

Eq. 6 says that the free energy of our approximate posterior (i.e., Bayesian) beliefs, given some sensory outcomes, is the Kullback-Leibler divergence () from the true posterior probability of external states, given the sensory input; minus the (negative log) marginal likelihood. Estimating the marginal likelihood can be achieved by minimising the free energy functional of (Bayesian) beliefs and sensations:

 (7)

The Kullback-Leibler (KL, or here) divergence represents the difference between the agent’s beliefs about external states  , and the true posterior probability over these states, given the sensory data . Any KL-divergence is always non-negative, which means that as the free energy gets smaller (i.e., as we minimise the functional) the divergence tends toward zero. This means that minimising free energy entails:

Marginal Likelihood Estimation (a.k.a. MLE, Beal, 2003) by making free energy a tight upper bound on the (negative log) marginal likelihood .

Perception (and learning) of external states by making the approximate posterior  a good approximation of the true posterior .

Perception (and learning), then, is simply the process whereby the approximate posterior  – parameterised or encoded by the internal states of the brain – are made ‘statistically consistent’ with the true posterior distribution over the external states of the world given sensory observations.

Note that there is some debate as to whether the reduction of the Kullback-Leibler divergence is a representational process (Kirchhoff and Robertson, 2018). Whether this process is representational or not, the probability distributions it manipulates are most certainly instances of representations (cf. Badcock, Friston & Ramstead, 2019). The divergence between two probability distributions can be said to be ‘right’ or ‘wrong’ with respect to some satisfaction conditions (i.e., a reducing divergence is better than an increasing divergence). Therefore, even if the process per se (i.e., reduction of the divergence or evidence bound) is non-representational, the components involved in this process make that process one of ‘manipulation’ of representations. A similar theme is seen in Bayesian decision theory, game theory and economics where the evidence bound can be interpreted as leading to bounded rationality (i.e., approximate Bayesian inference) (Friston et al., 2013). The rationality of decisions again speaks to an inherent representationalism that underwrites the ‘right’ sort of decisions.

Now, depending on the structure (i.e., entailed knowledge) in the generative model, approximate Bayesian inference not only optimises beliefs about the world ‘out there’ but also beliefs about the consequences of doing this or that. These beliefs yield inference to the best action to engage (see below). As we have seen, in the case of perception, approximate Bayesian inference involves minimising free energy, which is an upper bound on (negative log) marginal likelihood. We now turn to action planning as another instance of representational cognitive process.

## 3.2 Action planning

To account for action, one must start thinking about the manner in which states of the world change over time. This requires us to cast the generative model over multiple times steps , into the future and how an action policy  (i.e., possible sequence of actions) may influence these the trajectory of states when this or that policy is realised. Thus, our generative models will have the form  – to allow us to infer future hidden states and associated outcomes  relative to a policy  (see fig 2). Here, for the sake of simplicity, we will focus on a discrete formulation of the ensuing generative model for action.



Figure 2. Minimal discrete state space generative model for action. Open circles are random variables (hidden states and policies). Grey filled circles are observable outcomes. Squares are known variables, such as the model parameters.  refers to categorical distributions. The equations in the beige box (upper left) specify the architecture of the generative model (for a complete description see Friston, Parr & De Vries, 2017). The likelihood matrix  specifies the probability of outcomes for each combination hidden states . The novelty of this generative model rests on the addition of a policy  that determines state transitions, represented by the policy-dependent transition matrix . The initial state is specified by . The approximate posterior of the future hidden state  at time  relative to the policy  is found by evaluating the approximate posterior for the policy . This policy will be the one with the least expected free energy  that determines prior beliefs about the policy being pursued; namely  that can recovered from expected free energy, using the softmax operator  (for a complete description see Friston, FitzGerald, Rigoli, Schwartenbeck, & Pezzulo, 2017). Note that the edges that link the policy and the transition matrices are undirected. This is important, as it means that the evaluation of expected free energy requires a message from hidden states representations, thereby affording a representationalist pathway.

The structure of the graphical model in fig.2 allows us to work with a free-energy appropriate for outcomes that have yet to be observed. This is known as expected free energy (see Parr and Friston, 2017):

 (8)

In eq. 8, expected free energy of a policy at a given time  decomposes into a pragmatic or instrumental term and an epistemic term, also known as extrinsic and intrinsic values. The pragmatic term, or extrinsic value constitutes the goal seeking component of expected free energy (often referred to as expected value or utility in psychology and economics)(Kauder, 1953; Sutton and Barto, 1998). Extrinsic value is the expected value of a policy relative to preferred outcomes that will be encountered in the future . In turn, the epistemic term, or intrinsic value constitutes the information seeking component of expected free energy. Intrinsic value is the expected information gain relative to future states under a given policy (i.e., ‘what policy will best guarantee the minimisation of uncertainty in my beliefs about the causal structure of the world?’). In visual neurosciences, this is called salience and is a key determinant of epistemic foraging or exploratory behaviour (Itti and Baldi, 2009; Sun et al., 2011). As such, it is sometimes referred to as intrinsic motivation (Barto et al., 2013; Oudeyer and Kaplan, 2007; Ryan and Deci, 1985; Schmidhuber, 2010). Selecting the policy that affords the least expected free energy guarantees an adaptive action, that is, that first consolidates knowledge about the world, then optimises – i.e., works towards – preferred outcomes. For a complete discussion see (Friston et al., 2015).

## 3.3 Summary: The reason why perception and action planning rest on representational processes

In summary, under active inference, action selection is a process of manipulating representations about future states of the world to maximise one’s knowledge and secure desired (predicted) outcomes and sensory encounters. This inference or belief updating about ‘what I am doing’ rests on perceptual inference. Perception, in turn, is a process of updating mental representations of states of the world and their relationship to sensory consequences, so as to make these representations as consistent as possible with the true state of the world. Hence, more generally, perception and action planning, under active inference, are instances of representational processes. The statistical structure of the likelihood mapping tells me that the most likely cause of the sensory entry is the cause that my belief represents; and put bluntly, minimising uncertainty in beliefs is for the most part what ‘forming a percept’ is about. In turn, action selection is an inference process that relies on these optimised beliefs about sensory causes, and the consequences of future moves in a rich and reconstructive fashion. Action selection tells me that since I am a surprise or free energy minimising creature, I should selectively engage with the world to minimise expected surprise or uncertainty. This requires me to respond to epistemic affordances – to resolve uncertainty – while securing familiar (i.e. a priori preferred) sensory outcomes. This will minimise my uncertainty about future states and maximise the utility of my action.

In active inference, the need for rich, representations involving generative models stems directly from the problem of inverse inference about causes and adaptive actions to resolve uncertainty about those causes. The ill-posed nature of the inference problem we face forces us to first ‘figure out for ourselves’ ‘what causes what?’ before being able to zero-in on ‘what caused that’ (perception), and ‘I will cause that’ (i.e., action planning). This problem forces us to learn hierarchically (i.e., over multiple levels of prior beliefs) and temporally (i.e., over multiple time steps, such as in fig. 2) deep generative models (Friston, Rosch, Parr, Price, and Bowman, 2017).

# 4 Dynamic pathways in active inference

We turn now to the role of non-representational dynamics in active inference. There is a technical sense in which an austere, dynamicist reading of active inference is licensed in a fundamental way. This follows because the representational account above emerges from a certain kind of dynamics; namely, gradient flows on variational and expected free energy (cf. Ramstead, Kirchhoff, & Friston, K. J., 2019). In other words, the cognitivist functionality rests upon optimising free energy and this optimisation is a necessary consequence of neuronal dynamics that – not unlike a river flowing downhill – descend free energy gradients – to find free energy minima where the gradients are destroyed (Tschacher and Haken, 2007). Indeed, the back story to active inference shows that this kind of dynamical behaviour is a necessary aspect of any self-organisation to nonequilibrium steady-state in any random dynamical system that possesses a Markov blanket (Friston, 2013). On this view, any system that possesses some attracting states has dynamics that look ‘as if’ they are trying to minimise free energy and therefore acquire a representational and teleological interpretation. (cf. Ramstead, Kirchhoff, & Friston, K. J., 2019).

While there are interesting issues that attend the distinction between a purely dynamical formulation of active inference – and a representationalist reading in terms of dynamics and information geometry – we will consider non-representationalist formulations. These formulations speak to notions of extended and embedded optimisation, which call upon hierarchical dynamics that consider the dynamical exchange between an agent and its (physiological, evolutionary, and cultural) econiche. Accordingly, the dynamic pathways in generative models under active inference – examples appear below – do not appeal to the manipulation of representations of hidden states of the world to explain the cognitive processes underlying the behaviour they generate. Dynamic pathways can be exemplified by application to a specific class of unplanned action – more specifically enculturated action (more on that below) – that does not rest on the manipulation of rich webs of internal representations. Rather, heuristically, the dynamicist view is a view of action that only requires processing ‘something as doing’, such that the ‘doing’ (e.g., action sequences or realised policies) is directly conditioned upon the ‘something’ (e.g., sensory observation). In the recent literature on active inference, this sort of action has been coined ‘deontic’ action (Constant et al. 2019).

## 4.1 Deontic action

Deontic actions are actions for which the underlying policy has acquired a deontic value; namely, the shared, or socially admitted value of a policy (Constant et al. 2019). A deontic action is guided by the consideration of ‘what would a typical other do in my situation’. For instance, stopping at the red traffic light at 4am when no one is present may be viewed as such a deontically afforded action[[5]](#footnote-5).

Central to our agenda, deontic actions are processed through different mappings in the generative model. Technically, deontic value is the likelihood of a policy given an observation  that grounds posterior beliefs about policies[[6]](#footnote-6). This likelihood is an empirical prior which constitutes expected free energy. The deontic value  effectively supplements or supplants the likelihood of outcomes under different states  (see figure 3). From the point of view of the generative model, this means that if I am pursuing this policy then these outcomes are more likely (e.g., when I stop doing something, I am likely to see a stop sign). From the point of view of inference, this means that if I see these deontic outcomes, I will infer I am doing this (e.g., if I see a stop sign, I will stop). Put simply, a deontic action is an available (i.e., plausible) policy that is triggered by a sensory input, and which leads directly to an internally consistent action. Crucially, this means that deontic action selection bypasses representational beliefs about states of the world and associated sensory consequences.



Figure 3. Dynamic pathway. This graphical generative model is the same as the one presented in fig.2. However, it incorporates deontic value, which is read here as a dynamic pathway.

The computational architecture of deontic action is a clear candidate to implement a form of dynamicism under active inference. In effect, the information processing underlying deontic action eludes the two sufficient conditions of representationalism presented in section 2:

(i)Since they involve the inversion of a policy-outcome mapping, instead of state-outcomes mappings, deontic processes do not entail a propositional attitude involving the mediation of manipulations of one’s (Bayesian) beliefs standing for sensory causes in the world. Deontic processes do not have ‘aboutness’;

(ii) Success conditions in epistemology are about the way an agent’s act (e.g. an assertion, or another kind of speech act) needs to relate to the states of the world towards which it is directed for it to work. This means that usually, what the agent (or her brain) is seeking to optimise isn’t the issue. Under active inference, however, one ought to consider active brain processes and success conditions that do not relate to the external world *per se.*Brain processes can arbitrate between successful or unsuccessful alternatives with respect to the internal generative model *per se*, such as in the case of action planning, where action policies are compared in terms of the free energy – under the generative model – expected in the future. This means that success conditions can be given with respect to the generative model *per se*, not the world generating the observation (a.k.a. generative process). This is a subtle, yet crucial point, which becomes apparent when considering the probability distributions involved in various inference processes in the brain. For instance, when inferring hidden states (i.e., perceiving), the true posterior (*P*) approximated by the approximate posterior (*Q*) is the true posterior probability of the agent's beliefs, not of the cause of the agent's observations. 'Getting it right', in that case, again, is about getting it right with respect to one's own beliefs; e.g., successfully exploring the state space of one's own model of the world. This means that under active inference, there are two layers of success involved, one defined over the model, and one defined over the agent-world coupling (which corresponds to a more traditional epistemological point of view).

The second layer allows us to know when a generative model isn’t fit for purpose – e.g., in cases of mental disorders, where behavioural outcomes are maladaptive with respect to the individual’s environment. Those two layers of success are apparent in the fact that one can perform (Bayes) optimal inference while generating suboptimal behaviour because of suboptimal (prior) beliefs (Corlett & Fletcher, 2014). Deontic processes conform to success conditions of this second layer. They do not have success conditions *qua* brain processes, but rather have success condition *qua* agent-world coupling processes. They are simple observation-action loops; not rich and reconstructive policy selection loops. This is so because deontic processes circumnavigate the computation of expected free energy, which as we have seen, is used to compare different policies with respect to their ability to maximise instrumental and epistemic values. Under active inference, the maximisation of these values such as captured by expected free energy (*G*) is the success condition of non-deontic action selection in the brain. This construction means that inferences about states of the world – that admit a representationalist interpretation – are now replaced by direct action, without any intervening inference or representation of the consequences of action;

Having said this, the dual aspect architecture in Figure 3 means that the representationalist and dynamic pathways can happily live side-by-side, mutually informing each other – but both are sufficient for enactive engagement with the niche on their own. The distinction between the pathways – or routes to (subpersonal) action selection have some important implications. For example, deontic action circumnavigates expected free energy and therefore precludes planning as (active) inference. This means that under active inference, systems employing deontic strategies do not need to plan courses of action into the future. They simply act on the basis of the observation. Furthermore, in the absence of inference about hidden states, there could be no phenomenal opacity[[7]](#footnote-7). For instance, this speaks directly to the sort of actions experts perform (e.g., athletes), which are often complex, though, for which experts do not seem to plan ahead (i.e., ‘in the head’). Such skilled actions seem to yield very little phenomenal opacity (e.g., as when the athlete responds, ‘I don’t know, I just did it in the flow of action’, or ‘we simply executed the game plan’, when interviewed about her game winning shot).

## 4.2 Deontic action as a reflex?

The sort of ‘automatized’ deontic behaviour underwritten by dynamic pathways in the generative model might strike one as being conceptually close to the sort of cognitive processes underlying reflexes and other (homeostatic) functions processed through the autonomic nervous system. The computational pathway of deontic action indeed looks very much like a close control loop – secured by robust causal regularities in the world generating reliable sensory inputs – akin to a reflex processed at the brainstem and spinal cord level, but this time, processed in a ‘constructed local world’ (Constant et al., 2019, cf. Ramstead et al., 2016). Under active inference, motoric and autonomic reflexes are framed as an action that manages the sensory signal that comes from within the system that generated it; e.g., suppression of interoceptive prediction error (Pezzulo, Rigoli, and Friston, 2015).

Autonomic reflexes facilitate homeostatic regulation by engendering series of events necessary for the activity of the agent; e.g., salivation facilitates ingestion by easing the passing of the food. In this sense, they can be regarded as allostatic in nature. Similarly, one can think of sequences of deontic actions that facilitate social, affective and emotional regulation; e.g., the outcome generated by the red traffic light triggers a stop, which facilitates reaching in my pocket to grab my phone to check my notifications (which itself might trigger salivation). For some enculturated agents, such a sequence of ‘social reflexes’ may be necessary to pass through the day.

Now, the reader might worry that deontic action ends up being as unexciting as ‘digestive cognition’. But rest reassured, deontic action has been used to account for complex behavioural phenomena like social conformity: a.k.a., deference to the socially approved norm learnt through social influence or learning (Asch, 1955), and cooperative decision-making: a.k.a. decision-making under fairness psychology – as evidenced by the human tendency to zero in on fair decisions in economic games when compared to non-human animals (for a review see Henrich, 2015). Deontic action – as a social reflex – facilitates social interactions by easing the coordination among humans, if you will.

Deontic action is explained in terms of the circular causality between outsourcing decision-making to trusted others in the form of deontic cues (material or agential) – indicating the locally adaptive action – and learning the underlying cue-policy mappings. The ‘closed’ control loop, then, comprises the enculturated agent and regularities in her (social) environment. In effect, deontic cues are defined as such because they represent a reliable informational aggregate of ‘what would a creature like me would do in this situation’. These cues consolidate over development and through the modification of the environment by generations of other enculturated agents (i.e., creatures like me) (Constant et al., 2019)

Once the action afforded by these cues is learnt, there is no need for computing future states and associated outcomes; these are secured by the configuration of the cultural setting. For instance, in Canada, you can trust stopping or crossing, according to the deontic cue afforded by the traffic light – because the traffic light has come to represent what others typically do at an intersection – perhaps not in France though. And when faced with an uncertain outcome in an economic game (e.g., ‘if I don’t know what the opponent will do and my reward depends on her response, should I share or should I maximise my gain?’), you can trust that the fair option is the one the other is most likely to select since you’ve been socialised as a ‘typical other’, presumably, just as the other did (for a review see Veissière et al., 2019).

## 4.3 Summary: deontic actions rest on dynamic processes

In summary, for proponents of dynamicism, generative models are not rich and reconstructive internal models. Rather, they are fast and frugal. If internal representations play a role at all, that role is thin and simple. As we have seen above, a rich and reconstructive internal model is one in which multiple trajectories of hidden states (with different precisions – more on that below) would be entertained before selecting the action. The fast and frugal alternative is the one that underwrites deontic action. Hence, for enculturated, deontically constrained agents like us, “what may often be doing the work [in generative models] is a kind of perceptually-maintained motor-informational grip on the world: a low-cost perception-action routine that retrieves the right information just-in-time for use, and that is not in the business of building up a rich inner simulacrum” (Clark 2016, 11). This low-cost perception action routine corresponds to the web of deontic, or dynamic pathways learnt through enculturation (Constant et al., 2019).

# 5 Worries about rich settings for shallow strategies?

Although computationally viable under active inference, our description of fast and frugal dynamic pathways based on deontic value might still raise some conceptual worries. In this section, we provide a brief discussion of some such worries

## 5.1 First worry

One might worry that even deontic actions have to be selected through inferring the current context. The agent might need first to figure out if the context renders deontic action the most apt response. This worry raised by representationalists (e.g., Hohwy (2019)) might be a problem for the kind of account developed in this paper, and elsewhere (for example, Clark (2015b)). For even the selection of frugal dynamic strategies would require the on-going inference afforded by a rich inner model, able to determine when such strategies are warranted – and override them when necessary. In other words, the recruitment of the right transient webs of deontic activity, at the right time, is itself a high-grade cognitive achievement where the inner model plays, representationalists argue, a necessary and ongoing role. The upshot is a worry that truly ecumenical accounts may be hostage to “a potential tension…. between allowing and withholding a role for rich models” (Hohwy, 2019). For surely (so the argument goes) the active inference agent must repeatedly infer when she is in a situation where some low-cost deontic response is viable. In effect, according to representationalists, setting and learning the confidence of prior beliefs through perceptual processes – such as described in section 2 (a.k.a. precision, or gain control on sensory evidence, or prediction error) – needs to be a principled response, and that implicates the rich inner model even when the selected strategy is itself a frugal one.

In active inference, the mapping between causes (e.g., states and policies) and consequences (e.g., sensory outcomes) are parameterised in terms of probabilistic mappings that necessarily have a *precision*. In other words, the contingencies implicit in likelihood mappings can have different degrees of reliability, ambiguity, or uncertainty. For instance, if my child starts running towards the sea, as she gets further away (and closer to the water), my beliefs about whether she is in danger of drowning will become increasingly imprecise. Then, to disambiguate (hidden) states of affairs, I might plan an epistemic, representational strategy: running after my child to ensure she doesn’t go into the water without supervision. Had I known that my child would start running towards danger, I could have restrained her. After multiple visits at the beach, this might become my deontic, automatic, dynamic strategy (e.g., setting foot on the sand causes my arm to grab my child).

This means that the more ‘representationalist’ picture of the continuous rational influence of planning, we claim, is subtly mistaken. For example, suppose I am playing table-tennis well. My context sensitive ‘precision settings’ are all apt, no unexpected circumstances arise (alien invasions etc.). In such circumstances, I harvest a flow of expected kinds of prediction errors. These get resolved, in broadly predictable ways, as play unfolds without pushing far up the processing hierarchy. But if ‘unexpected surprises’ (for more on this distinction, see Yu & Dayan (2005)) occur, some errors are more fundamentally unresolved and get pushed higher. This provides the seed for re-organizing the precision of various likelihood mappings to lend more weight to different kinds of (internal, external, and action-involving) information. That, we suggest, is how we can remain constantly poised (e.g., Sutton’s compelling (2007) work on expert cricket) for nuance, even while behaving in the fast, fluent manner of a ‘habit machine’. In a deep sense, we exist in that moment as a habit machine – that is nonetheless constantly poised to become another transient machine should the need arise. This speaks to the coalition between representational and dynamic pathways illustrated in figure 3.

Put another way, wherever possible, simple ‘habit’ systems should guide behavior, dealing with expected prediction error fluently and fast. But where these fail, or where a change of context indicates the need, more and more knowledge-intensive resources (internal and external) can be assembled, via new waves of precision-weighting, to quash any outstanding prediction errors (i.e., free energy) – see Pezzulo and colleagues (2015) for a complete argument. Hence, we should not deny that there really is, in advanced minds, what representationalists correctly describes as “immense storage of causal knowledge” (Hohwy, 2019). But via moment by moment, self-organising, free energy minimising kinetics, we manifest as a succession of relatively special-purpose brain-body-world devices, strung together by those shifting but self-organising webs of precision-weighting. Importantly, it is self-organising around free energy that itself delivers the subsequent precision variations that recruit the ‘next machine’. There is no precision-master sitting atop this web, carefully deciding moment by moment just how to assign precision – there’s just the generative model itself.

## 5.2 Second worry

At this point a new version of representationalist worry may arise. For it may seem that precision estimates – the roots of each episode of re-structuring – are cognitively expensive and purely inner-model-bound. But this too – or so we have been arguing – is subtly mistaken. If we shift perspectives and timescales, we can see the human-built cognitive niche as itself a prime reservoir, both of achieved precision estimations and of tools for cheaply estimating precisions on-the-fly. And once learnt, they allow non-representation involving deontic action pathway (e.g., positioning cheap cues in the world such as warning triangles around a broken-down vehicle). These otherwise arbitrary structures attract attention and act as local proxies for precision (e.g., Roepstorff et al. (2010), Hutchins (2014), Paton, Skewes, Frith and Hohwy (2013). Urgent fonts, food packaging, and priestly robes all provide handy shortcuts for our precision estimating brains. Squint just a little bit and much of the human-built world – including all those patterned social practices such as stopping at red traffic lights – can be seen as a bag of tricks for managing precision estimation and epistemic trust (Fonagy and Allison, 2014; Parr and Friston, 2017). And, as we behave in the present niche, we gradually alter it, ‘uploading’ (Constant et al., 2018) more and more of our individual and collective precision estimations into persisting (transmissible) material and social structures. These, in turn, alter the inner models that individuals need to command to negotiate their worlds.

## 5.3 Third worry

A final representationalist worry may be that fast and frugal, non-representational deontic action could simply not yield adaptive behaviour in a highly volatile world like ours and thus may lead to suboptimal, maladaptive decision making (e.g., decision making that fails to generate action that succeeds with respect to environmental challenges); especially, if our generative models of precision are not apt for a volatile world (Parr et al., 2018a; Parr et al., 2018b). Consequently, one should favour explanations based on rich reconstructing planning. This is a fair worry; a fair worry for humans in general, not for the dynamicists’ perspective, though. Indeed, humans learn to generate deontic actions that do not always lead to the ‘Machiavellian’, or perhaps ‘Darwinian’ utility maximising option relative to the current environment; we miss steps and fall down the stairs, forget to stop on the red, develop disorders such as PTSD that makes us misperceive threats, and generate many more maladaptive traits (Badcock et al., 2017; Cornwell et al., 2017; Peters et al., 2017). The tricks humans employ – to minimise the potential cost of normal maladaptive actions – is not to plan more ‘in the head’, but to plan more ‘in the world’; e.g., making sure that the synchronisation of the traffic lights is consistent with the traffic flow at different hours of the day. This ‘planning the world’ solution stabilises the environment to enable the acquisition (i.e., learning through representationalist processes) of cheap deontic action shared among ‘cultural’ conspecifics – ‘people enculturated like me, on a 9 to 5 schedule’ (Constant et al., 2018a; Constant et al., 2018b). Under that view, in certain situations, one can dispense with rich models that “stand-in for that world for the purposes of planning, reasoning, and the guidance of action” (Clark 2016, 6). In a word, for enculturated, deontically construed agents like us, the world is often ‘our shared’ best model.

# 6 Conclusions: bury the hatchet, or use it to carve a new path(way)

This paper offers a mathematically informed reading of generative models that could accommodate both richly representationalist and dynamicist views of cognition. Is cognition under active inference a richly representational or a dynamic, process reliant on simple cues and couplings? The answer is both. What remains unclear, however, is whether particular cognitive processes underlying certain behaviour are representational or not. To debate on that based on active inference, one ought to take the hatchet, and ask whether a new theoretical path(way) in generative models should be carved out. Indeed, any debate in the philosophy of cognitive science appealing to active inference (and its kin such as predictive processing, the Bayesian brain, and predictive coding) should clarify at the outset the manner in which the cognitive process of interest is implemented in the generative model, and what are the components of the graphical model involved in the process. Clarifying at the outset the architecture of the generative model of interest should be sufficient to settle the technical dimension of the debate.

Such good practice would allow researchers to save time and energy by simply showing the manner in which the cognitive process of interests may be already implemented by existing neurocomputational architecture. In effect, the name of the game with active inference is to show how cognitive processes can be expressed as rearrangements or decompositions of the free energy functional and the architecture it implements in the graphical model; i.e., to show the manner in which the dynamics of the process of interest are built in the free energy formalism, that is, the manner in which the formalism unifies the process of interest as a special case of free energy minimisation. Researchers could first explore the currently available generative models (relevant material is all freely available either in theoretical articles or as part of the Statistical Parametric Mapping 12 MATLAB toolbox). If the literature on the cognitive function of interest is not yet available, researchers could consider this a great opportunity for ‘getting their hands dirty’ and proposing novel architectures that could account for the cognitive process and the associated behaviour they want to characterise (Montague et al., 2012; Schwartenbeck and Friston, 2016). Ideally, these novel architectures should complement existing data on neuroanatomy and hierarchical neural dynamics (Friston et al., 2017).

Finally, a limitation of the current paper is that we do not know yet what existing neurophysiology implements the dynamicist pathway we describe in section 3. We have shown that dynamicism has a computational grip when implemented in the theory of active inference. However, one has yet to propose candidate neural correlates, which is a research enterprise for neuroscience made possible on the basis of implementable theory, such as the one discussed in this paper. Thus, despite the lack of empirical evidence, we consider settling the general active inference debate about representationalism a major development; since it is a first step towards scientifically informed debates on the representational nature of specific pathways, which could then feedback to further strengthen future philosophical discussions and inform research trajectories..

# References

Allen, M., & Friston, K. J. (2016). From cognitivism to autopoiesis: towards a computational framework for the embodied mind. *Synthese*, 1–24.

Asch, S. E. (1955). Opinions and Social Pressure. Scientific American, 193(5), 31–35.

Badcock, P. B., Friston, K. J., & Ramstead, M. J. D. (2019). The hierarchically mechanistic mind: A free-energy formulation of the human psyche. Physics of Life Reviews.

Badcock, P.B., Davey, C.G., Whittle, S., Allen, N.B., Friston, K.J., 2017. The Depressed Brain: An Evolutionary Systems Theory. Trends in Cognitive Sciences 21, 182-194.

Barto, A., Mirolli, M., Baldassarre, G., 2013. Novelty or Surprise? Frontiers in Psychology 4.

Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., & Friston, K. J. (2012). Canonical microcircuits for predictive coding. Neuron, 76(4), 695-711.).

Beal, M. J. (2003). Variational algorithms for approximate Bayesian inference. University of London London.

Beer, R. D. (2000). Dynamical approaches to cognitive science. Trends in cognitive sciences, 4(3), 91-99.

Bogacz, R. (2017). A tutorial on the free-energy framework for modelling perception and learning. Journal of Mathematical Psychology, 76(Pt B), 198–211.

Bruineberg, J., Kiverstein, J., & Rietveld, E. (2016). The anticipating brain is not a scientist: the free-energy principle from an ecological-enactive perspective. Synthese, 1–28.

Buckley, C. L., Kim, C. S., McGregor, S., & Seth, A. K. (2017). The free energy principle for action and perception: A mathematical review. Journal of Mathematical Psychology, 81(Supplement C), 55–79.

Chemero, A. (2009). Radical embodied cognition. Cambridge, MA: MIT Press.

Churchland, P. M. (1989). Some Reductive Strategies in Cognitive Neurobiology. In S. Silvers (Ed.), Rerepresentation: Readings in the Philosophy of Mental Representation (pp. 223–253). Dordrecht: Springer Netherlands. Clark, A. (2005). Intrinsic Content, Active Memory and the Extended Mind. Analysis, 65(1), 1–11.

Clark, A. (2013a). Whatever next? Predictive brains, situated agents, and the future of cognitive science. The Behavioral and Brain Sciences, 36(03), 181–204.

Clark, A. (2015b). Surfing uncertainty: prediction, action, and the embodied mind. New York, N.Y.: Oxford University Press.

Clark, A. (2015c). Predicting Peace: The End of the Representation Wars-A Reply to Michael Madary. In T. Metzinger & J. M. Windt (Eds). Open MIND: 7(R). Frankfurt am Main: MIND Group. doi: 10.15502/9783958570979

Clark, A. (2015). Radical Predictive Processing. *The Southern Journal of Philosophy*, *53*, 3–27.

Clark, A., & Chalmers, D. (1998). The Extended Mind. Analysis, 58(1), 7–19.

Clark, A. (1997). The Dynamical Challenge. *Cognitive Science*, *21*(4), 461–481.

Clark, A (1989) Microcognition: Philosophy, Cognitive Science and Parallel Distributed Processing, MIT Press/Bradford Books, Cambridge: MA)

Clark, A (1993) Associative Engines: Connectionism, Concepts And Representational Change (MIT Press, Bradford Books)

Constant, A., Clark, A., & J, K. F. (2019, June). Representation Wars: Enacting an Armistice through Active Inference. Retrieved from http://philsci-archive.pitt.edu/16125/

Constant, A., Ramstead, M. J. D., Veissière, S. P. L., Campbell, J. O., & Friston, K. J. (2018)a. A variational approach to niche construction. Journal of the Royal Society, Interface / the Royal Society, 15(141). https://doi.org/10.1098/rsif.2017.0685

Constant, A., Bervoets, J., Hens, K., & Van de Cruys, S. (2018)b. Precise Worlds for Certain Minds: An Ecological Perspective on the Relational Self in Autism. Topoi. An International Review of Philosophy. <https://doi.org/10.1007/s11245-018-9546-4>

Constant, A., Ramstead, M., Veissière, S., & Friston, K. J. (2019). Regimes of Expectations: An Active Inference Model of Social Conformity and Decision Making. Frontiers in Psychology. <https://doi.org/10.3389/fpsyg.2019.00679>

Corlett, P. R., & Fletcher, P. C. (2014). Computational psychiatry: a Rosetta Stone linking the brain to mental illness. The Lancet. Psychiatry, 1(5), 399–402.

Cornwell, B.R., Garrido, M.I., Overstreet, C., Pine, D.S., Grillon, C., 2017. The Unpredictive Brain Under Threat: A Neurocomputational Account of Anxious Hypervigilance. Biol Psychiatry 82, 447-454.

Dayan, P., Hinton, G. E., Neal, R. M., & Zemel, R. S. (1995). The Helmholtz machine. Neural Computation, 7(5), 889–904.

Dolega, K. (2017). *Moderate Predictive Processing*. https://doi.org/10.15502/9783958573116

Feldman, H., & Friston, K. J. (2010). Attention, uncertainty, and free-energy. Frontiers in Human Neuroscience, 4, 215.

Feynman, R. (1972). Statistical mechanics: a set of lectures. Reading, MA: Benjamin/Cummings Publishing.

Fodor, J. A. (1975). The language of thought (Vol. 5). Harvard University Press.

Fonagy, P., Allison, E., 2014. The role of mentalizing and epistemic trust in the therapeutic relationship. Psychotherapy (Chicago, Ill.) 51, 372-380.

Friston, K., 2013. Life as we know it. J R Soc Interface 10, 20130475.

Friston, K., Rigoli, F., Ognibene, D., Mathys, C., Fitzgerald, T., Pezzulo, G., 2015. Active inference and epistemic value. Cogn Neurosci 6, 187-214.

Friston, K.J., Parr, T., de Vries, B., 2017. The graphical brain: Belief propagation and active inference. Network neuroscience (Cambridge, Mass.) 1, 381-414.

Friston, K. (2018). Does predictive coding have a future? Nature Neuroscience, 21(8), 1019–1021.

Friston, K. J. (2010). The free-energy principle: a unified brain theory? Nature Reviews. Neuroscience, 11(2), 127–138.

Friston, K., 2013. Life as we know it. J R Soc Interface 10, 20130475.

Friston, K. J., FitzGerald, T., Rigoli, F., Schwartenbeck, P., O Doherty, J., & Pezzulo, G. (2016). Active inference and learning. Neuroscience and Biobehavioral Reviews, 68, 862–879.

Friston, K. J., FitzGerald, T., Rigoli, F., Schwartenbeck, P., & Pezzulo, G. (2017). Active inference: a process theory. Neural Computation, 29(1), 1–49.

Friston, K. J., & Frith, C. (2015). A Duet for one. Consciousness and Cognition, 36, 390–405.

Friston, K. J., Harrison, L., & Penny, W. (2003). Dynamic causal modelling. NeuroImage, 19(4), 1273–1302.

Friston, K. J., Rosch, R., Parr, T., Price, C., & Bowman, H. (2017). Deep temporal models and active inference. Neuroscience and Biobehavioral Reviews, 77, 388–402.

Friston, K., Schwartenbeck, P., Fitzgerald, T., Moutoussis, M., Behrens, T., & Dolan, R. J. (2013). The anatomy of choice: active inference and agency. Frontiers in Human Neuroscience, 7, 598.

Friston, K. J., Kilner, J., & Harrison, L. (2006). A free energy principle for the brain. *Journal of Physiology, Paris*, *100*(1-3), 70–87.

Gładziejewski, P. (2016). Predictive coding and representationalism. *Synthese*, *193*(2), 559–582.

Henrich, J. (2015). The secret of our success: how culture is driving human evolution, domesticating our species, and making us smarter. Princeton, NJ: Princeton University Press.

Hohwy, J. (2013). The predictive mind. Oxford: Oxford University Press.

Hohwy, J. (2016). The self‐evidencing brain. Noûs, 50(2), 259–285.

Hohwy, J. (2019). Quick’n'lean or Slow and Rich? Andy Clark on predictive processing and embodied cognition. In M. Colombo, E. Irvine, & M. Stapleton (Eds.), Andy Clark and His Critics (pp. 191–205). Oxford University Press.

Hutchins, E. (2014). The cultural ecosystem of human cognition. Philosophical Psychology, 27(1), 34–49.

Hutto, D. D., & Myin, E. (2013). Radicalizing enactivism: Basic minds without content. MIT Press.

Joffily, M., & Coricelli, G. (2013). Emotional valence and the free-energy principle. PLoS Computational Biology, 9(6),e1003094.

Kaplan, R., & Friston, K. J. (2018). Planning and navigation as active inference. Biological Cybernetics. https://doi.org/10.1007/s00422-018-0753-2

Keller, G. B., & Mrsic-Flogel, T. D. (2018). Predictive Processing: A Canonical Cortical Computation. Neuron, 100(2), 424–435.

Kiefer, A., & Hohwy, J. (2017). Content and misrepresentation in hierarchical generative models. Synthese, 1–29.

Kirchhoff, M., Parr, T., Palacios, E., Friston, K. J., & Kiverstein, J. (2018). The Markov blankets of life: autonomy, active inference and the free energy principle. Journal of the Royal Society, Interface / the Royal Society, 15(138). https://doi.org/10.1098/rsif.2017.0792

Kirchhoff, M. D., & Robertson, I. (2018). Enactivism and Predictive Processing: A Non-Representational View. Philosophical Explorations: An International Journal for the Philosophy of Mind and Action, 21(2), 264–281.

Itti, L., Baldi, P., 2009. Bayesian Surprise Attracts Human Attention. Vision Res. 49, 1295-1306.

Kauder, E., 1953. Genesis of the Marginal Utility Theory: From Aristotle to the End of the Eighteenth Century. The Economic Journal 63, 638-650.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444.

Limanowski, J., Friston, K., 2018. ‘Seeing the Dark’: Grounding Phenomenal Transparency and Opacity in Precision Estimation for Active Inference. Frontiers in Psychology 9, 643.

Metzinger, T., & Wiese, W. (2017). Philosophy and predictive processing. MIND Group: Frankfurt Am Main.

Mirza, M. B., Adams, R. A., Mathys, C. D., & Friston, K. J. (2016). Scene Construction, Visual Foraging, and Active Inference.Frontiers in Computational Neuroscience, 10, 56.

Parr, T., & Friston, K. J. (2017). Working memory, attention, and salience in active inference. Scientific Reports, 7(1), 14678.

Montague, P.R., Dolan, R.J., Friston, K.J., Dayan, P., 2012. Computational psychiatry. Trends Cogn Sci 16, 72-80.

Oudeyer, P.-Y., Kaplan, F., 2007. What is intrinsic motivation? a typology of computational approaches. Frontiers in Neurorobotics 1, 6.

Parr, T., Benrimoh, D., Vincent, P., Friston, K., 2018a. Precision and false perceptual inference. Front. Integr. Neurosci.

Parr, T., Friston, K.J., 2017. Uncertainty, epistemics and active inference. Journal of the Royal Society Interface 14, 20170376.

Parr, T., Rees, G., Friston, K.J., 2018b. Computational Neuropsychology and Bayesian Inference. Front Hum Neurosci 12, 61.

Peters, A., McEwen, B.S., Friston, K., 2017. Uncertainty and stress: Why it causes diseases and how it is mastered by the brain. Prog Neurobiol.

Parr, T., Markovic, D., Kiebel, S. J., & Friston, K. J. (2019). Neuronal message passing using Mean-field, Bethe, and Marginal approximations. *Scientific Reports*, *9*(1), 1889.

Pezzulo, G., Rigoli, F., & Friston, K. (2015). Active Inference, homeostatic regulation and adaptive behavioural control. Progress in Neurobiology, 134, 17–35.

Ryan, R., Deci, E., 1985. Intrinsic motivation and self-determination in human behavior. Plenum, New York.

Ramstead, M. J. D., Veissière, S. P. L., & Kirmayer, L. J. (2016). Cultural affordances: scaffolding local worlds through shared intentionality and regimes of attention. Frontiers in Psychology, 7, 1090.

Ramstead, M. J. D., Badcock, P. B., & Friston, K. J. (2017). Answering Schrödinger’s question: A free-energy formulation. Physics of Life Reviews, 24, 1–16.

Ramstead, M. J. D., Kirchhoff, M., & Friston, K. J. (2019). A tale of two densities: Active inference is enactive inference. Manuscript submitted for publication.

Roepstorff, A, Niewöhner, J, & Beck, S. (2010). Enculturating brains through patterned practices. Neural Networks, 23, 1051 1059.

Siegel, S. (2010). The Contents of Visual Experience. Oxford University Press.

Schmidhuber, J., 2010. Formal Theory of Creativity, Fun, and Intrinsic Motivation (1990-2010). Ieee Transactions on Autonomous Mental Development 2, 230-247.

Schwartenbeck, P., Friston, K., 2016. Computational Phenotyping in Psychiatry: A Worked Example. eNeuro 3.

Sun, Y., Gomez, F., Schmidhuber, J., 2011. Planning to be surprised: optimal Bayesian exploration in dynamic environments, Proceedings of the 4th international conference on Artificial general intelligence. Springer-Verlag, Mountain View, CA, pp. 41-51.

Sutton, John (2007) 'Batting, Habit and Memory: The Embodied Mind and the Nature of Skill', Sport in Society, 10:5, 763 –786

Sutton, R.S., Barto, A.G., 1998. Reinforcement Learning: An Introduction. MIT Press, Cambridge, MA.

Tschacher, W., Haken, H., 2007. Intentionality in non-equilibrium systems? The functional aspects of self-organised pattern formation. New Ideas in Psychology 25, 1-15.

Thelen, E., & Smith, L. B. (1996). A dynamic systems approach to the development of cognition and action. MIT press.

Thompson, E. (2007). Mind in life: biology, phenomenology, and the sciences of mind. Cambridge, MA: Harvard University Press.

Van Gelder, T. (1995). What might cognition be if not computation? Journal of Philosophy, 91, 345–381

Varela, F. J., Thompson, E., & Rosch, E. (1991). The embodied mind: Cognitive science and human experience, 308.

Veissière, S. P. L., Constant, A., Ramstead, M. J. D., Friston, K. J., & Kirmayer, L. J. (2019). Thinking Through Other Minds: A Variational Approach to Cognition and Culture. The Behavioral and Brain Sciences, 1–97.

Williams, D. (2018). Predictive Processing and the Representation Wars. Minds and Machines, 28(1), 141–172.

1. The word ‘entails’ has a particular meaning here because, strictly speaking, the brain is viewed as performing inference under a generative model. Technically, the generative model is just a probability distribution over the (unobserved) causes of (observed) sensations. As such, the generative model only exists to the extent that neuronal dynamics maximise the evidence for that model (for a discussion see Ramstead, Badcock, and Friston, 2017). [↑](#footnote-ref-1)
2. Hidden states are sometimes called latent states and refer to variables that cannot be observed. There are effectively hidden behind observations and have to be inferred as random variables. [↑](#footnote-ref-2)
3. Approximate Bayesian inference should not be read as inference that is approximate; rather, it should be read as inference that can be realised. Indeed, as we will see later, exact Bayesian inference is a special case of approximate Bayesian inference when certain conditions are met. [↑](#footnote-ref-3)
4. This arbitrary probability distribution is variously called an approximate posterior, a recognition distribution or more simply a Bayesian belief. [↑](#footnote-ref-4)
5. Note that not every deontic action need be socially sanctioned; or rather, we humans have the ability to socially sanction ourselves. For instance, one might come up with a solo practice; a habit that is all mine, such as putting a sock on the doorknob to stop me leaving without my keys. After some time, I might forget why I put a sock on my doorknob, until the day where for some reason my habit fails me. Coming back home, realising that I forgot my key, and explaining the situation to myself, I will realise that I forgot my key because I should have put the sock on the doorknob. I might think, “this is what someone ‘like me’ would and should have done!” [↑](#footnote-ref-5)
6. E.g., when I see the sock on my night table, I put it in my pocket. [↑](#footnote-ref-6)
7. In the sense that there would be no opportunity to optimise the precision of the likelihood mapping between hidden states and outcomes that, in active inference, is usually associated with (covert) mental action and attention (Limanowski et al., 2018). [↑](#footnote-ref-7)