PSYCHOLOGICAL INQUIRY
2018, VOL. 29, NO. 2, 67–73
https://doi.org/10.1080/1047840X.2018.1480693

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

Active Inference and Cognitive Consistency

Karl J. Friston

The Wellcome Centre for Human Neuroimaging, Institute of Neurology, London, United Kingdom

I know little about social psychology and less about *cognitive consistency theory*; however, I greatly enjoyed reading the deconstruction of this paradigm by Kruglanski et al. (this issue). I learned a lot and was struck repeatedly by the consilience between their treatment and complementary formulations in theoretical neurobiology. The underlying tenets of their argument emerge—in my world—from a Bayesian or variational treatment of perception and self-organized behavior. This is variously known as the *free energy principle*, *active inference*, *predictive processing*, or *self-evidencing* (Clark, 2013, Friston, 2010; Friston, FitzGerald, Rigoli, Schwartenbeck, & Pezzulo, 2017; Hohwy, 2016). Particular flavors of this formulation have dominated cognitive neuroscience and aspects of philosophy for the past decade. For example, the Bayesian brain hypothesis (Kersten, Mamassian, & Yuille, 2004; Knill & Pouget, 2004) and predictive processing (Barrett & Simmons, 2015; Michael & De Bruin, 2015; Rao & Ballard 1999; Seth, 2013) are now predominant paradigms in cognitive neuroscience. Active inference generalizes this approach to provide an enactive or embodied treatment of action and perception. This resulting treatment bears some remarkable similarities to the ideas reviewed in Kruglanski et al. This commentary tries to establish the formal links between *active inference* and constructs in *cognitive consistency theory* in the hope that there may be some useful cross-fertilisation. In what follows, I briefly overview active inference with a special focus on the quantities needed to describe affective responses to new information and then unpack some key dialectics that speak to cognitive consistency, epistemic and motivational value, specific and nonspecific closure, and so on.

Epistemic and Motivational Value in Active Inference

In brief, active inference explains everything we perceive and do in terms of one imperative, namely, the minimization of variational free energy (Friston, 2013). Variational free energy is an upper bound on surprise, where surprise (the negative log probability of sensory samples) corresponds to negative Bayesian log evidence. This means that everything we infer (and do) is in the service of minimizing surprise or maximizing the evidence for our internal or generative models of the world. By maximizing the evidence for our models of the world (i.e., self-evidencing), we come to infer hidden or latent states of the world "out there" that generate our sensations. This formulation of "perception as inference" has a long pedigree and was arguably best articulated by Helmholtz (Helmholtz, 1878/1971): "Objects are always imagined as being present in the field of vision as would have to be there in order to produce the same impression on the nervous mechanism."

Crucially, in active inference, we also have to consider beliefs about action or behavior. These (posterior) beliefs depend upon the degree to which a course of action will minimize expected free energy in the future. The notion of *expected free energy*, given a particular action, is key here. This follows because expected surprise corresponds to uncertainty. This means that everything we do is in the game of resolving uncertainty or ambiguity about states of affairs "out there," that is, beyond our sensory impressions. These states of affairs are often referred to as *hidden* or *latent* states because they cannot be observed directly, only inferred on the basis of observations or sensory evidence.

Technically, free energy can always be expressed as a *divergence* minus *log evidence*. This means that minimizing free energy reduces the divergence or difference between our current belief and the true posterior belief (because beliefs cannot change sensory evidence). In other words, our beliefs about states of the world are as close as possible to the true state of affairs, given the sensory evidence at hand.[1] Based on these beliefs, we can then form beliefs about "what to do" by choosing those actions that minimize expected free energy. The important move here is to separate the divergence and evidence parts of free energy and understand what their expected values mean. It turns out that they correspond to *epistemic* and *pragmatic* (i.e., motivational) value respectively. This is remarkable because exactly the same separation emerges from the treatment of cognitive

CONTACT Karl J. Friston k.friston@ucl.ac.uk The Wellcome Centre for Human Neuroimaging, Institute of Neurology, 12 Queen Square, London WC1N 3BG, United Kingdom.
Color versions of one or more of the figures in this article can be found online at www.tandfonline.com/HPLI.

[1]Technically, minimizing a variational free energy *bound* on log evidence is called variational Bayes or approximate Bayesian inference. It is also a nice metaphor for *bounded* rationality in behavioral economics. Variational free energy is also referred to as an evidence *bound* because the divergence cannot be less than zero.

consistency, namely, into epistemic and motivational value. Furthermore, these two components appear to underpin *nonspecific* and *specific closure*. In other words, the maximization of epistemic value offers a formal description of nonspecific closure, whereas the maximization of motivational (pragmatic) value corresponds to specific closure.

The mathematical form of this separation can be expressed in terms beliefs Q about states of the world and choices or policies, where beliefs about states of the world under a particular policy minimize free energy F, whereas beliefs about policies minimize expected free energy G:

$$Q(s_\tau, \pi) = Q(s_\tau|\pi)Q(\pi)$$

$$Q(s_\tau|\pi) = \arg\min_Q F$$

$$F(\pi) = \underbrace{D[Q(s_\tau|\pi)||P(s_\tau|\tilde{o}, \pi)]}_{\textit{Divergence}} - \underbrace{\ln P(\tilde{o}|\pi)}_{\textit{log evidence}}$$

$$Q(\pi) = \exp\left[-\gamma \cdot \sum\nolimits_\tau G(\pi, \tau)\right]$$

$$G(\pi, \tau) = -\underbrace{D[Q(s_\tau|o_\tau, \pi)||\ln Q(s_\tau|\pi)]}_{\textit{Epistemic value}} - \underbrace{E[\ln P(o_\tau|\pi)]}_{\textit{Motivational value}} \quad (1)$$

Equation 1 provides a detailed description of these mathematical expressions for readers familiar with Bayesian or variational treatments. The variable γ is a constant of proportionality that will play an important role later in terms of affective responses to beliefs about action.

Pragmatic value acquires the semantic "motivational" because it is the expected log probability of prior beliefs about outcomes. In other words, if we treat beliefs about the consequences of behavior as prior preferences, then we will choose those behaviors that lead to goals or preferred outcomes. However, this is always in the context of reducing uncertainty (that can be also expressed in terms of *ambiguity* and *risk*; see Equation A.1 in the appendix). This is the epistemic part of expected free energy—variously known as Bayesian surprise, information gain, the value of information, intrinsic motivation, and so on (see Equation 1). In summary, the interaction between motivational and epistemic value appears to underpin the deconstruction of cognitive consistency and plays a central role in active inference. Note that pragmatic and epistemic values are log probabilities. This means that there is exactly the same sort of interaction in terms of probabilities per se suggested by Kruglanski et al. (this issue). In other words, our prior preferences contextualize attempts to resolve uncertainty. So how can this formalism used to understand "the affective response to new information" and, more specifically, how does it address "people's reactions to awkward news" (Kruglanski et al., this issue, p. 54)?

The important thing that active inference brings to the table is the distinction between beliefs about states of the world and beliefs about "what I am going to do." Expected free energy is a functional[2] of beliefs about states of the world and therefore determines (posterior) beliefs about "what to do" based on motivational (pragmatic) and epistemic value. So where does the affective bit come in? Generally, people describe affective or hedonic attributes to the beliefs about "what to do," as opposed to the "states of affairs." In particular, a positive valence is associated with precise, confident (low entropy) beliefs about action, and conversely, negative affect is associated with uncertainty and confusion about what to do next (Gu, Hof, Friston, & Fan, 2013; Seth & Friston, 2016). This distinction between beliefs about states and beliefs about action comfortably accommodates the key dialectics in the target article, as follows.

Trivial and Nontrivial Inconsistency

First, it distinguishes between *trivial* and *nontrivial* inconsistencies, in the sense that trivial inconsistencies can be associated with belief updates about latent states of the world and nontrivial inconsistencies change beliefs about behavior, "that is, relevant to individuals' goals in the situation" (Kruglanski et al., this issue, p. 54). Furthermore, a separation into beliefs about states and action offer a simple definition of consistency namely, the confidence or precision of beliefs about behavior. In other words, a consistent belief about what to do identifies one course of action unambiguously over all plausible alternatives. Conversely, inconsistency corresponds to an uncertainty about what to do that is generally associated with negative affect and—in the context of interoceptive inference—stress, anxiety, and negative emotion (Peters, McEwen, & Friston, 2017; Seth, 2013; Stephan et al., 2016).

The interesting thing about associating inconsistency with uncertainty about policies to pursue is that inconsistency becomes proportional to the free energy expected under the policies entertained. Mathematically, this can be expressed as follows:

$$\Im_\pi = H[Q(\pi)] = E_{Q(\pi)}[-\ln Q(\pi)] = \gamma \cdot E_{Q(\pi)}\left[\sum\nolimits_\tau G(\pi, \tau)\right] \quad (2)$$

This is remarkable because it is exactly the conclusion offered by Kruglanski et al. (this issue), namely, goal-sensitive or deliberative inconsistency is not a universal construct; it is composed of motivational and epistemic components that contextualize one another. Often, these imperatives can be in opposition. Indeed, a ubiquitous pattern of behavior—seen when simulating active inference—is that epistemic imperatives predominate initially, even at the expense of securing preferred outcomes. Only when the context has been explored sufficiently—and its epistemic value has been exhausted—do (synthetic) subjects exploit their epistemic foraging to pursue preferred outcomes (Friston et al., 2015; Moulin & Souchay 2015).

Note that this construction necessarily implies that we will try to avoid bad news because we are inherently optimistic (Sharot, 2011; Sharot, Guitart-Masip, Korn, Chowdhury, & Dolan, 2012). This optimism is a necessary aspect of ideal Bayesian (active) inference, in which subjects make choices that are most likely given the sort of subject

[2]A function of a function; in this case, a function of a probability distribution

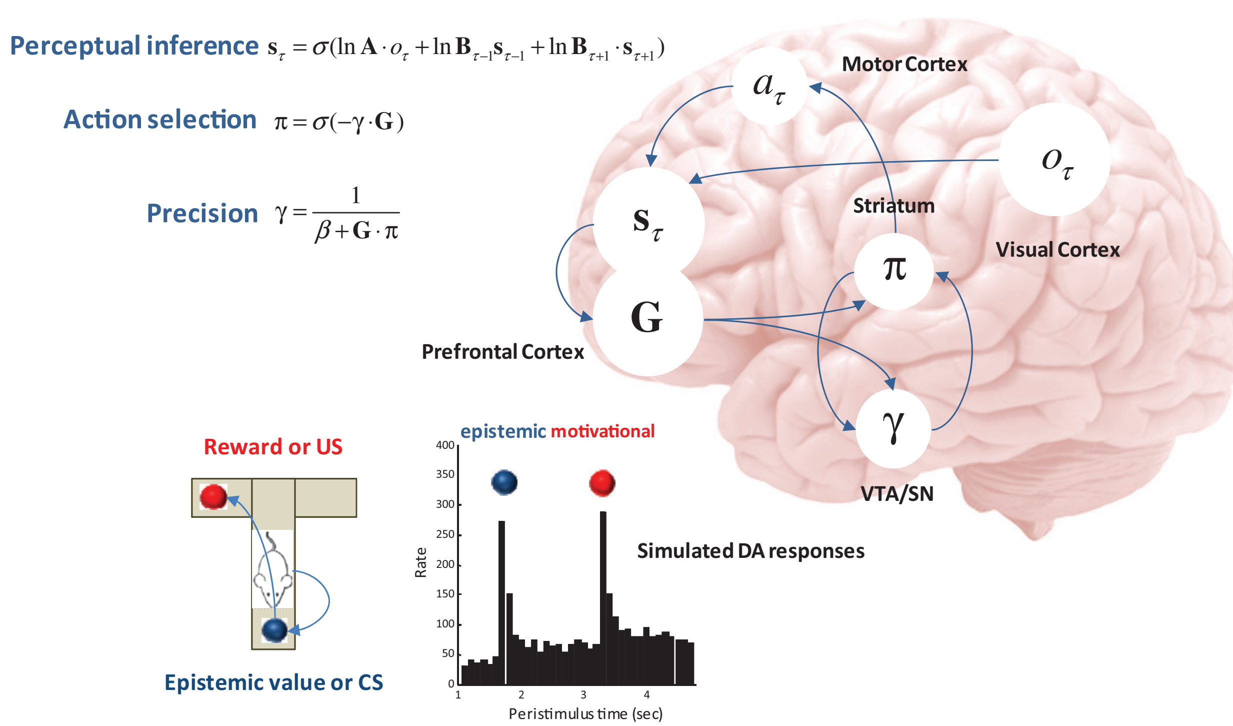

Figure 1. This figure illustrates the functional anatomy implied by a simple message passing scheme based on minimizing variational free energy; see Friston et al. (2014) for details. *Note.* The variables *o*, **s**, and *a* correspond to observations, expected states of the world, and action, respectively, where π and γ represent expected policies and their precision. **G** is the expected free energy of policies. The equations corresponds to (variational) Bayesian updates, where **A** and **B** are probability transition matrices, mapping hidden states to observations and hidden states to hidden states under different actions respectively. σ is a softmax function. Here, we have associated the Bayesian updates of hidden states of the world with perception, control states (policies) with action selection and expected precision with confidence. In this (purely iconic) schematic, we have associated perception (inference about the current state of the world) with the prefrontal cortex, while assigning action selection to the basal ganglia. Precision has been associated with dopaminergic projections from ventral tegmental area and substantia nigra. Lower panel: This shows the results of a simulation in terms of simulated dopamine discharges. The key thing to note is that the responses to an informative cue (CS) pre-empts subsequent responses to the reward (US). In this simulation, the agent was shown a cue that resolved uncertainty (i.e., had epistemic value) about where to find a reward (i.e., that had motivational value) in a simple T-maze (inset).

they are (Friston et al., 2015). Because the posterior probability of a choice rests upon prior beliefs, these prior beliefs bias everything we do. Any epistemic foraging (i.e., information seeking) that that leads to highly unlikely outcomes will therefore not be entertained; for example, we do not seek medical information that leads to a course of action that is so unlikely it is not worth contemplating, namely, dying.

Before turning to some of the implications in terms of affective responses to good and bad information, there is one interesting corollary that endorses the arguments in Kruglanski et al. (this issue)—and allows us to showcase the advantages of formal treatments, like active inference. In brief, the constant of proportionality γ above itself has a Bayes optimal (free energy minimizing) solution that turns out to be proportional to cognitive consistency [this can be written mathematically as follows by substituting Equation (2) into the update equations in Figure 1]:

$$\gamma = \frac{1}{\beta + \frac{1}{\gamma}\Im} = \frac{1-\Im}{\beta} \tag{3}$$

This is important because this constant of proportionality behaves in a way that has all the hallmarks of dopaminergic responses in real brains (Friston et al., 2014). This speaks to the neuromodulatory correlates of affective reactions to changes in cognitive consistency—and shows how formal treatments can link concepts in social psychology to neurobiological processes. In short, it suggests that whenever there is an increase in epistemic or motivational value during active inference, there will be an increase in consistency that will be reported by increases in dopamine activity. See Figure 1 for an example of phasic dopamine responses to epistemic and motivational value. Indeed, there is an emerging literature on the link between dopaminergic activity, in this setting, that may be usefully migrated to social psychology (Schwartenbeck, FitzGerald, Mathys, Dolan, & Friston, 2015). Although promising, this formulation of consistency—as the impact of new information on confidence about what to do—does not explain our responses to good and bad news (Kruglanski et al., this issue). Or does it?

Affective Responses and Closure

On the preceding arguments, a negative effective response to bad news would imply that new evidence induces posterior beliefs about the state of the world that induce uncertainty about the course of action. This seems sensible: For example, when opening a letter that contains exam results, one would imagine that this epistemic foraging is driven by the epistemic value (i.e., resolution of uncertainty) about your performance and the sort of person you are. If you have failed, under the belief that you are an overachiever, the overall consequences can be summarized as "What on Earth am I to do now?" In short, bad news almost invariably induces uncertainty about how to realize goals or prior

Table 1. The construct validity of cognitive consistency and constructs in active inference.

The social Psychology Paradigm	Mathematical Expression	The Free Energy Principle	Short Description
Cognitive inconsistency	$\Im = H[Q(s, \pi)]$	Posterior uncertainty about states and policies	Confidence or certainty about hidden states and policies generating outcomes
Trivial inconsistency	$\Im_s = H[Q(s)]$	Posterior uncertainty about states	A loss of confidence in beliefs about hidden states of the world
Nontrivial inconsistency	$\Im_\pi = H[Q(\pi)]$	Posterior uncertainty about policies	A loss of confidence about what to do (usually associated with stress and anxiety)
Motivational value	$E_{\tilde{Q}}[\ln P(o_\tau \| \pi)]$	Pragmatic value	The expected log likelihood (i.e. prior preferences) of outcomes in the future
Epistemic value	$E_{\tilde{Q}}[\ln Q(s_\tau \| o_\tau, \pi) - \ln Q(s_\tau \| \pi)]$	Epistemic value	The information gain, reduction of uncertainty, intrinsic motivation or epistemic affordance of a policy
Specific closure		Maximisation of pragmatic value	Selection of policies that optimise the motivational component of expected free energy
Nonspecific closure		Maximisation of epistemic value	Selection of policies that optimise the epistemic component of expected free energy
Self-verification	$Q(\pi) = \exp(-\gamma \cdot G(\pi))$	Self-evidencing	Selection of policies that maximise expected model evidence or minimise expected free energy
Affective reaction	$\gamma = \frac{1-\Im_\pi}{\beta}$	The precision of beliefs about policies	The consequences of Bayesian belief updating on the confidence placed in policies

preferences, whereas good news signifies a clear path forward, either in the short or long term. This seems to be a plausible explanation for our affective responses to new evidence that incorporates both the affective dimension and the interplay between epistemic affordance and prior preferences, both in resolving uncertainty and inducing it in terms of future prospects. There are more nuanced arguments (that entail the opportunity to resolve uncertainty in the future) that can be applied to the interesting phenomena of not wanting to know the end of a book (Kruglanski et al., this issue) or the final score of a football match before it is watched. See Friston, Lin, et al. (2017) for a discussion of novelty and curiosity in this context.

Conclusion

So is there a universal need for cognitive consistency? In the preceding view, one would argue that consistency is not the primary construct: It is an inevitable consequence of forming posterior beliefs about an active engagement with the world that has epistemic and motivational aspects. This dual-aspect construction has the latitude to dissociate (Bayesian) surprise from prior preference [see Equation (1)], providing an accommodating account of pleasant and unpleasant surprises. For example, the state of mind on opening a gift can place motivational and epistemic imperatives in opposition. Both will reduce uncertainty and afford Bayesian surprise (i.e., be informative); however, the outcome (i.e., the gift) may or may not be what you expected (i.e., preferred). Before closing, I wanted to briefly mention other prescient points of contact between active inference and the Kruglanski et al. (this issue) account.

It strikes me that “a reluctance to face ambiguity and uncertainty” (Kruglanski et al., this issue, p. 51) fits comfortably within the epistemic value of certain actions—and the overall propensity to minimize surprise (entailed by the free energy principle). The degree to which one person or another expresses a need for nonspecific closure may translate into the degree to which they, a priori, weight epistemic against pragmatic imperatives. This speaks to characterizing phenotypic traits in terms of priors, which is a prominent direction of research in things like computational psychiatry (Huys, Moutoussis, & Williams, 2011; Wang & Krystal, 2014).

I was also taken with the notion of *self-verification*. This seems to be almost isomorphic to self-evidencing in its broadest sense (Hohwy, 2016). This becomes particularly interesting in relation to understanding certain psychiatric syndromes. For example, the explanation for depression as self-verification or self-evidencing—in the context of prior beliefs about adverse consequences of behavior—fits perfectly with current formulations of depression in terms of false inference (Stephan et al., 2016). In short, “individuals with such low self-esteem (e.g., depressed people) search for and prefer negative feedback consistent with their self-view” (Kruglanski et al., this issue, p. 51). This reduces uncertainty about their engagement with the (prosocial) world. There are clear metaphors for this in the behavioral psychology literature, for example, learned helplessness (Hammack, Cooper, & Lezak, 2012; Stephan et al., 2016).

In closing, I refer interested readers to Table 1 and Figure 2 for a summary of the similarity between the paradigm in the target article and active inference. Table 1 attempts to map the concepts and terms used in cognitive consistency theory with the corresponding constructs in active inference. The corresponding process theories are sketched out in Figure 2 using the same format as in Kruglanski et al. (this issue). As intimated in the closing comments of Kruglanski et al., these theoretical formulations will only prove themselves in terms of their predictive

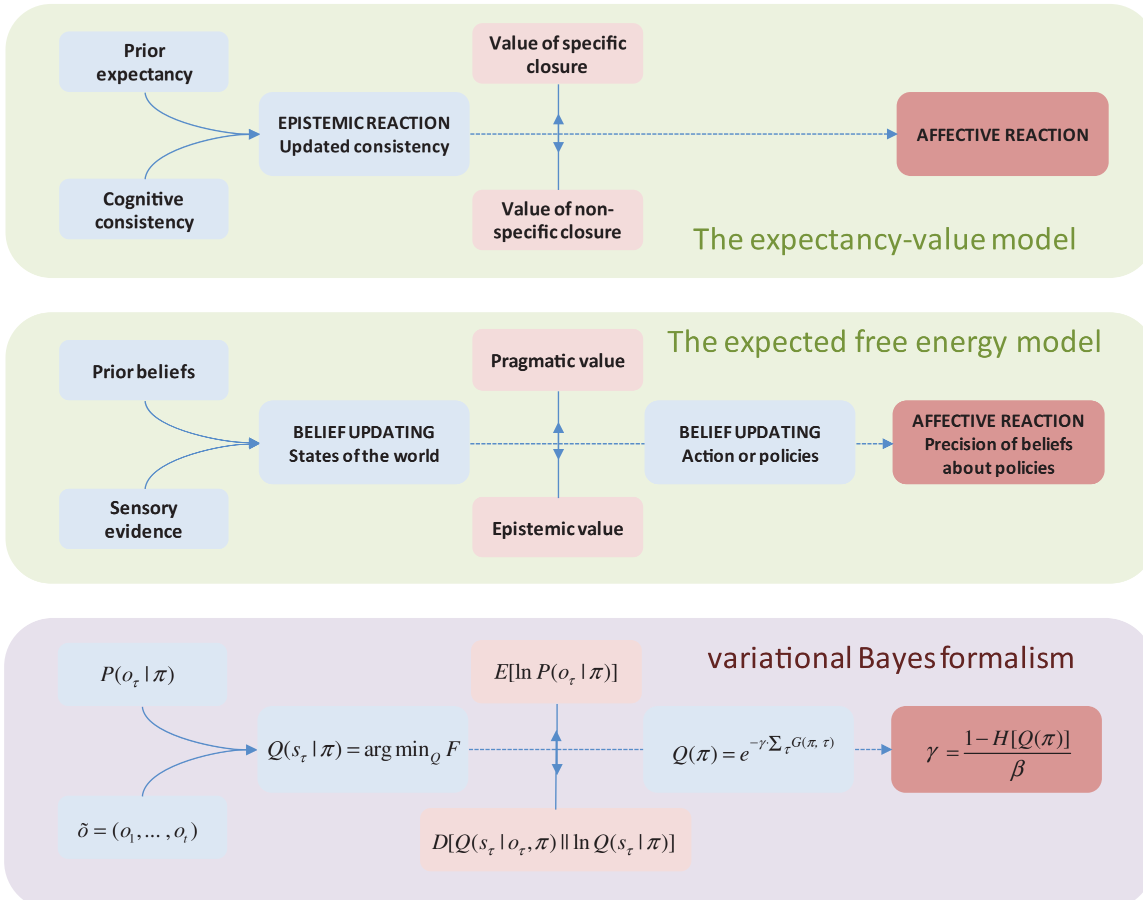

Figure 2. Schematic illustrating the architectural homology between the formulation of cognitive consistency in terms of closure and in terms of active inference (i.e., self-evidencing or minimisation of expected free energy). *Note.* The upper panel is based on Figure 1 in Kruglanski et al. (this issue).

validity. In this regard, there are some interesting paradigms that are emerging from the active inference literature that could be scrutinized from the perspective of social psychology. I have in mind here explicit tests of choice behavior when motivational and epistemic imperatives are set in opposition. One might even imagine the study of the neuronal correlates of belief updating—along the lines of the dopaminergic responses just described (Schwartenbeck, FitzGerald, Mathys, Dolan, & Friston, 2015).

References

Attias, H. (2003). *Planning by probabilistic inference. Proceedings of the 9th International Workshop on Artificial Intelligence and Statistics.*

Barlow, H. (1961). Possible principles underlying the transformations of sensory messages. In W. Rosenblith (Ed.), *Sensory communication* (pp. 217–234). Cambridge, MA: MIT Press.

Barrett, L. F., & Simmons. W. K. (2015). Interoceptive predictions in the brain. *Nature Reviews Neuroscience*, *16*, 419–429.

Beal, M. J. (2003). *Variational algorithms for approximate Bayesian inference* (Unpublished doctoral thesis). University College London, London, England.

Botvinick, M., & Toussaint, M. (2012). Planning as inference. *Trends in Cognitive Science*, *16*, 485–488.

Braun, D. A., Ortega, P. A., Theodorou, E., & Schaal, S. (2011). *Path integral control and bounded rationality. Paper presented at the 2011 IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning*, Paris, France.

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

FitzGerald, T. H., Schwartenbeck, P., Moutoussis, M., Dolan, R. J., & Friston, K. (2015). Active inference, evidence accumulation, and the urn task. *Neural Computation*, *27*, 306–328.

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

Friston, K. (2013). Life as we know it. *Journal of the Royal Society Interface*, *10*, 20130475.

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

Friston, K. J., Lin, M., Frith, C. D., Pezzulo, G., Hobson, J. A., & Ondobaka, S. (2017). Active inference, curiosity and insight. *Neural Computation*, *29*, 2633–2683.

Friston, K., Rigoli, F., Ognibene, D., Mathys, C., Fitzgerald, T., & Pezzulo, G. (2015). Active inference and epistemic value. *Cognitive Neuroscience*, *6*, 187–214.

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., Schwartenbeck, P., FitzGerald, T., Moutoussis, M., Behrens, T., & Dolan, R. J. (2014). The anatomy of choice: Dopamine and decision-making. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *369*(1655).

Gu, X., Hof, P. R., Friston, K. J., & Fan, J. (2013). Anterior insular cortex and emotional awareness. *Journal of Comparative Neurology*, *521*, 3371–3388.

Hammack, S. E., Cooper, M. A., & Lezak, K. R. (2012). Overlapping neurobiology of learned helplessness and conditioned defeat:

implications for PTSD and mood disorders. *Neuropharmacology*, *62*, 565–575.
Helmholtz, H. (1971). The facts of perception. In R. Kahl (Ed.), *The selected writings of Hermann von Helmholtz* (p. 384). Middletown, CT: Wesleyan University Press. (Original work published 1878)
Hohwy, J. (2016). The self-evidencing brain. *Noûs*, *50*, 259–285.
Howard, R. (1966). Information value theory. *IEEE Transactions on Systems, Science and Cybernetics, SSC-2*(1), 22–26.
Huys, Q. J., Moutoussis, M., & Williams, J. (2011). Are computational models of any use to psychiatry? *Neural Networks*, *24*, 544–551.
Itti, L., & Baldi, P. (2009). Bayesian surprise attracts human attention. *Vision Research*, *49*, 1295–1306.
Kersten, D., Mamassian, P., & Yuille, A. (2004). Object perception as Bayesian inference. *Annual Review of Psychology*, *55*, 271–304.
Klyubin, A. S., Polani, D., & Nehaniv, C. L. (2005). Empowerment: A universal agent-centric measure of control. *Proceedings of the IEEE Congress on Evolutionary Computation*, *1*, 128–135.
Knill, D. C., & Pouget, A. (2004). The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends in Neurosciences*, *27*, 712–719.
Laughlin, S. B. (2001). Efficiency and complexity in neural coding. *Novartis Foundation Symposium*, *239*, 177–187.
Linsker, R. (1990). Perceptual neural organization: Some approaches based on network models and information theory. *Annual Review of Neuroscience*, *13*, 257–281.
Moulin, C., & Souchay, C. (2015). An active inference and epistemic value view of metacognition. *Cognitive Neuroscience*, *6*, 221–222.
Michael, J., & De Bruin, L. (2015). How direct is social perception? *Consciousness and Cognition*, *36*, 373–375.
Moutoussis, M., Trujillo-Barreto, N. J., El-Deredy, W., Dolan, R. J., & Friston, K. J. (2014). A formal model of interpersonal inference. *Frontiers in Human Neuroscience*, *8*, 160.
Olshausen, B. A., & Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, *381*, 607–609.
Ortega, P. A., & Braun, D. A. (2013). Thermodynamics as a theory of decision-making with information-processing costs. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, *469*, 2153.
Peters, A., McEwen, B. S., & Friston, K. (2017). Uncertainty and stress: Why it causes diseases and how it is mastered by the brain. *Progress in Neurobiology*, *156*, 164–188.
Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*, 79–87.
Schmidhuber, J. (1991). Curious model-building control systems. *Proceedings of the International Joint Conference on Neural Networks*, *2*, 1458–1463.
Schwartenbeck, P., FitzGerald, T. H., Mathys, C., Dolan, R., & Friston, K. (2015). The dopaminergic midbrain encodes the expected certainty about desired outcomes. *Cerebral Cortex*, *25*, 3434–3445.
Schwartenbeck, P., FitzGerald, T. H., Mathys, C., Dolan, R., Wurst, F., Kronbichler, M., & Friston, K. (2015). Optimal inference with suboptimal models: Addiction and active Bayesian inference. *Medical Hypotheses*, *84*, 109–117.
Seth, A. K. (2013). Interoceptive inference, emotion, and the embodied self. *Trends in Cognitive Sciences*, *17*, 565–573.
Seth, A. K., & Friston, K. J. (2016). Active interoceptive inference and the emotional brain. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *371*(1708).
Sharot, T. (2011). The optimism bias. *Current Biology*, *21*, R941–R945.
Sharot, T., Guitart-Masip, M., Korn, C. W., Chowdhury, R., & Dolan, R. J. (2012). How dopamine enhances an optimism bias in humans. *Current Biology*, *22*, 1477–1481.
Stephan, K. E., Manjaly, Z. M., Mathys, C. D., Weber, L. A. E., Paliwal, S., Gard, T., … Petzschner, F. H. (2016). Allostatic self-efficacy: A metacognitive theory of dyshomeostasis-induced fatigue and depression. *Frontiers in Human Neuroscience*, *10*, 550.
Still, S., & Precup, D. (2012). An information-theoretic approach to curiosity-driven reinforcement learning. *Theory in Biosciences*, *131*, 139–148.
Toussaint, M., & Storkey, A. (2006). *Probabilistic inference for solving discrete and continuous state Markov Decision Processes. Proceedings of the 23nd International Conference on Machine Learning*, pp. 954–952.
van den Broek, J. L., Wiegerinck, W. A. J. J., & Kappen, H. J. (2010). Risk-sensitive path integral control. *Proceedings of the Twenty-Sixth Conference on Uncertainty in Artificial Intelligence*, *6*, 1–8.
Wang, X. J., & Krystal, J. H. (2014). Computational psychiatry. *Neuron*, *84*, 638–654.

Appendix

Active Inference and Self-Evidencing

Active inference is based on the premise that creatures (like us) minimize variational free energy. This single premise leads to some surprisingly simple update rules for action, perception, policy selection, and the encoding of uncertainty or precision. The active inference scheme described here can be applied to any paradigm or choice behavior. It has been used to model waiting games (Friston et al., 2013), two-step maze tasks (Friston et al., 2015), the urn task and evidence accumulation (FitzGerald, Schwartenbeck, Moutoussis, Dolan, & Friston, 2015), trust games from behavioral economics (Moutoussis, Trujillo-Barreto, El-Deredy, Dolan, & Friston, 2014), addictive behavior (Schwartenbeck, FitzGerald, Mathys, Dolan, Wurst, et al., 2015), and engineering benchmarks such as the mountain car problem. It has also been used in the setting of computational functional magnetic resonance imaging (Schwartenbeck, FitzGerald, Mathys, Dolan, Wurst, et al., 2015).

Active Inference and Generative Models

Active inference rests on a generative model of observed outcomes. This model is used to infer the most likely causes of outcomes in terms of expectations about states of the world. These states are called *hidden states* because they can only be inferred indirectly through, possibly limited, sensory observations. Crucially, observations depend on action, which requires the generative model to entertain expectations under different policies or action sequences. Because the model generates the consequences of sequential action, it has explicit representations of the past and future; in other words, it is equipped with a working memory and expectations about future (counterfactual) states of the world under competing policies. These expectations are optimized by minimizing variational free energy, which renders them (approximately) the most likely (posterior) expectations about states of the world, given the current observations.

Expectations or beliefs about the most likely policy are based on the prior belief that policies are more likely if they pursue a trajectory or path that has the least free energy (or greatest model evidence). As we see next, this expected free energy can be expressed in terms of epistemic and motivational all pragmatic value, where epistemic value scores the information gain or reduction in uncertainty about states of the world and motivational value depends on prior beliefs about future outcomes. These prior preferences play the role of utility in economics and reinforcement learning.

Having evaluated the relative probability of different policies, expectations under each policy can then be averaged in proportion to their (posterior) probability. In statistics, this is known as *Bayesian model averaging*. The results of this averaging specify the next most likely outcome, which determines the next action. Once an action has been selected, it generates a new outcome and the (perception and action) cycle starts again. The resulting behavior is a principled interrogation and sampling of sensory cues that has both epistemic and motivated aspects. Generally, behavior in an ambiguous context is dominated by epistemic drives until there is no further uncertainty to resolve—and extrinsic value predominates. At this point, explorative behavior gives way to exploitative behavior.

In short, perception and action are subsumed under planning as inference (Attias, 2003; Botvinick & Toussaint, 2012; Toussaint & Storkey, 2006). Formally, this corresponds to inverting or fitting a generative model, given a sequence of outcomes. Technically, this is equivalent to optimizing the expected hidden states, policies, and precision with respect to variational free energy. These (posterior) estimates constitute posterior beliefs, usually denoted by the probability distribution $Q(x)$, where $x = \widetilde{s}, \pi, \gamma$ are the hidden or unknown variables.

Variational Free Energy and Inference

In variational Bayesian inference, model inversion entails minimizing variational free energy with respect to the sufficient statistics (i.e., expectations) of posterior beliefs:

$$
\begin{aligned}
&Q(x) = \operatorname{argmin}_{Q(x)} F \\
&F = E_Q\left[\ln Q(x) - \ln P(\widetilde{o}|x) - \ln P(x)\right] \\
&\quad = E_Q\left[\ln Q(x) - \ln P(x|\widetilde{o}) - \ln P(\widetilde{o})\right] \\
&= \underbrace{D\left[Q(x)||P(x|\widetilde{o})\right]}_{Divergence} - \underbrace{\ln P(\widetilde{o})}_{\log evidence} = \underbrace{D\left[Q(x)||P(x)\right]}_{Complexity} - \underbrace{E_Q[\ln P(\widetilde{o}|x)]}_{Accuracy}
\end{aligned}
\tag{A.1}
$$

where $\widetilde{o} = (o_1, \ldots, o_t)$ denotes observations up until the current time.

Because the Kullback-Leibler (KL) divergence cannot be less than zero, the penultimate equality means that free energy is minimized when the approximate posterior becomes the true posterior. At this point, the free energy becomes the negative log evidence for the generative model (Beal, 2003). This means minimizing free energy is equivalent to maximizing model evidence, which is equivalent to minimizing the complexity of accurate explanations for observed outcomes (last equality in B.1).

Minimizing free energy ensures expectations encode posterior beliefs, given observed outcomes. However, beliefs about policies rest on future outcomes. This means that policies should, a priori, minimize the free energy in the future. This can be formalized by making the log probability of a policy proportional to the free energy expected in the future (Friston et al., 2015):

$$
\begin{aligned}
Q(\pi) &= \exp(-\gamma \cdot G(\pi)) \\
G(\pi) &= \textstyle\sum_\tau G(\pi, \tau) \\
G(\pi, \tau) &= E_{\widetilde{Q}}[\ln Q(s_\tau|\pi) - \ln Q(s_\tau|o_\tau, \pi) - \ln P(o_\tau)] \\
&= -\underbrace{D[Q(s_\tau|o_\tau, \pi)||Q(s_\tau|\pi)]}_{Epistemic value} - \underbrace{E\widetilde{Q}\left[\ln P(o_\tau)\right]}_{Pragmatic\ value} \\
&= \underbrace{D\left[Q(s_\tau|\pi)||P(s_\tau)\right]}_{Risk} + \underbrace{E\widetilde{Q}[H[P(o_\tau|s_\tau)]]}_{Ambiguity}
\end{aligned}
\tag{A.2}
$$

where $\widetilde{Q} = Q(o_\tau, s_\tau|\pi) = P(o_\tau|s_\tau)Q(s_\tau|\pi)$.

The expected divergence now becomes mutual information or epistemic value, whereas the expected log-evidence becomes motivational value—if we associate the prior preferences with value or utility. The final expression shows how expected free energy can be evaluated relatively easily: It is just the divergence between the predicted and preferred outcomes, plus the ambiguity (i.e., entropy) expected under predicted states.

There are several helpful interpretations of expected free energy that appeal to (and contextualize) established constructs. For example, maximizing epistemic value is equivalent to maximizing (expected) Bayesian surprise (Itti & Baldi, 2009), where Bayesian surprise is the KL divergence between posterior and prior beliefs. This can also be interpreted in terms of the principle of maximum mutual information or minimum redundancy (Barlow, 1961; Laughlin, 2001; Linsker, 1990; Olshausen & Field, 1996). This is because epistemic value is the mutual information between hidden states and observations. In other words, it reports the reduction in uncertainty about hidden states afforded by observations. Because the information gain cannot be less than zero, it disappears when the (predictive) posterior ceases to be informed by new observations. Heuristically, this means that epistemic behavior will search out observations that resolve uncertainty about the state of the world (e.g., foraging to resolve uncertainty about the hidden location of prey or fixating on informative part of a face). However, when there is no posterior uncertainty—and the agent is confident about the state of the world—there can be no further information gain, and epistemic value will be the same for all policies, enabling extrinsic value to dominate. This resolution of uncertainty is closely related to satisfying artificial curiosity (Schmidhuber, 1991; Still & Precup, 2012) and speaks to the value of information (Howard, 1966).

The risk (i.e., expected complexity) is exactly the same quantity minimized in risk sensitive or KL control (Klyubin, Polani, & Nehaniv, 2005; van den Broek, Wiegerinck, & Kappen, 2010), and underpins related (free energy) formulations of bounded rationality based on complexity costs (Braun, Ortega, Theodorou, & Schaal, 2011; Ortega & Braun, 2013). In other words, minimizing expected complexity or cost renders behavior risk sensitive, and maximizing expected accuracy induces ambiguity-sensitive behavior.