6

The Mindful Filter

Free Energy and Action

Karl J. Friston

Abstract

This chapter frames key questions about embodied cognition and action in terms of *active inference*; namely, the premise that the brain is trying to infer the causes of its sensory input—and samples that input to minimize uncertainty about its inferences. This provides a process theory for embodied exchanges with the world that can be cast as a Bayesian filter—or more simply predictive coding—equipped with classical reflexes. The ensuing (embodied inference) perspective raises interesting questions about embodiment and enactivism: Can we ever truly observe worldly states? Are there such things as representations? Can we make inferences about other agents who are making inferences about us? These questions are unpacked in terms of the unobservability assumption, sensorimotor contingency theory, and the active inference perspective on mirror neurons and inferring the intention of others.

Active Inference and Predictive Coding

This chapter begins with a brief overview of active inference, with a particular focus on process theories and implementation in the embodied brain. Predictive coding will be used as a metaphor for neuronal message passing, as we consider how the basic imperatives of the Bayesian brain (predictive coding) can be met through active sampling of the sensorium. Having established the basic idea, I conclude with a statement of the three basic questions articulated above and the constraints afforded by active inference.

Recent advances in theoretical neuroscience have inspired a (Bayesian) paradigm shift in cognitive neuroscience. This shift is away from the brain as a passive filter of sensations—or an elaborate stimulus-response link—toward a view of the brain as a statistical organ that generates hypotheses or fantasies (fantastic: from Greek *phantastikos*, the ability to create mental images, from *phantazesthai*), which are tested against sensory evidence (Gregory 1968).

From "The Pragmatic Turn: Toward Action-Oriented Views in Cognitive Science,"
Andreas K. Engel, Karl J. Friston, and Danica Kragic, eds. 2016. Strüngmann Forum Reports, vol. 18,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03432-6.

This perspective dates back to the notion of unconscious inference (Helmholtz 1866/1962) and has been formalized in recent decades to cover deep or hierarchical Bayesian inference—about the causes of our sensations—and how these inferences induce beliefs, movement, and behavior (Dayan et al. 1995; Lee and Mumford 2003; Friston et al. 2006; Hohwy 2013; Clark 2013b).

Predictive Coding and the Bayesian Brain

Modern formulations of the Bayesian brain—such as predictive coding—are among the most popular explanations for neuronal message passing (Srinivasan et al. 1982; Rao and Ballard 1999; Friston 2008; Clark 2013b). Predictive coding is a biologically plausible process theory for which there is a considerable amount of anatomical and physiological evidence (Mumford 1992; Friston 2008). For a review of canonical microcircuits and hierarchical predictive coding in perception, see Bastos et al. (2012); for a treatment of the motor system, see Adams, Shipp et al. (2013) and Shipp et al. (2013). In these schemes, neuronal representations in higher levels of cortical hierarchies generate predictions of representations in lower levels. These top-down predictions are compared with representations at the lower level to form a prediction error (usually associated with the activity of superficial pyramidal cells). The ensuing mismatch signal is passed back up the hierarchy to update higher representations (associated with the activity of deep pyramidal cells). This recursive exchange of signals suppresses prediction error at each and every level to provide a hierarchical explanation for sensory inputs, which enter at the lowest (sensory) level. In computational terms, neuronal activity encodes beliefs or probability distributions over states in the world that cause sensations (e.g., my visual sensations are caused by a *face*). The simplest encoding corresponds to representing the belief with the expected value or *expectation* of a (hidden) cause. These causes are referred to as *hidden* because they have to be inferred from their sensory consequences. We will see later that these expectations are not limited to beliefs about the causes of sensations but include expectations about how those sensations are sampled through action. In summary, beliefs about hidden causes are engendered by ascending prediction errors, reporting sensory information that has yet to be explained. When these expectations are endowed with dynamics, descending predictions can preempt or anticipate sensory trajectories, thereby minimizing prediction errors as the sensorium unfolds.

Predictive coding represents a biologically plausible scheme for updating beliefs about states of the world using sensory samples (Figure 6.1). In this setting, cortical hierarchies become a neuroanatomical embodiment of how sensory signals are generated; for example, a face generates luminance surfaces which generate textures and edges and so on, down to retinal input. This form of hierarchical inference explains a large number of anatomical and physiological

From "The Pragmatic Turn: Toward Action-Oriented Views in Cognitive Science,"
Andreas K. Engel, Karl J. Friston, and Danica Kragic, eds. 2016. Strüngmann Forum Reports, vol. 18,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03432-6.

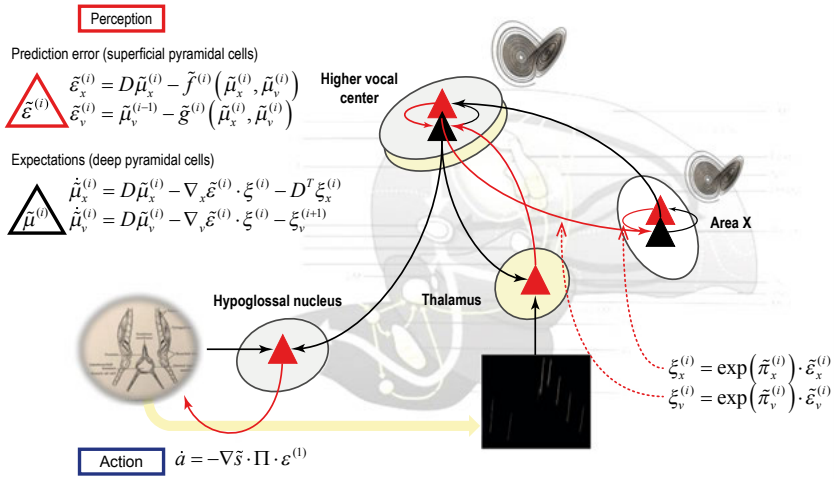

Figure 6.1 Hierarchical message passing in predictive coding using the (simplified) neuroanatomy of a songbird. Neuronal activity encodes expectations about the causes of sensory input, where these expectations minimize prediction error. Prediction error is the difference between (ascending) sensory input and (descending) predictions of that input. This minimization rests upon recurrent neuronal interactions among different levels of the cortical hierarchy. The available evidence suggests that superficial pyramidal cells (red triangles) compare the expectations (at each level) with top-down predictions from deep pyramidal cells (black triangles) of higher levels. Left: the equations represent the neuronal dynamics implicit in predictive coding. Prediction errors at the *i*-th level of the hierarchy are simply the difference between the expectations encoded at that level and top-down predictions of those expectations. The expectations per se are driven by prediction errors so that they perform a gradient ascent on the sum of squared (precision weighted) prediction error. [See Appendix for a detailed explanation of these (simplified) equations.] Right: scheme of the songbird's auditory system, showing the putative cells of origin of ascending or forward connections that convey (precision weighted) prediction errors (red arrows) and descending or backward connections (black arrows) that construct predictions. In this example, area X sends predictions to the higher vocal center, which projects to the auditory thalamus. However, the higher vocal center also sends proprioceptive predictions to the hypoglossal nucleus, which are passed to the syrinx to generate vocalization through classical reflexes. These predictions can be regarded as motor commands, while the descending predictions of auditory input correspond to corollary discharge. Note that every top-down prediction is reciprocated with a bottom-up prediction error to ensure predictions are constrained by sensory information. The Lorenz attractors associated with higher levels of the hierarchy indicate that generative models in the brain can possess autonomous, and possibly chaotic, dynamics with deep (hierarchical) structure (for details, see Kiebel et al. 2009; Friston and Kiebel 2009).

facts (Friston 2008; Adams, Shipp et al. 2013; Bastos et al. 2012). In brief, it explains the hierarchical nature of cortical connections, the prevalence of backward connections, as well as many of the functional and structural asymmetries in the extrinsic (between region) connections that link hierarchical levels (Zeki and Shipp 1988). These asymmetries include the laminar specificity of

From "The Pragmatic Turn: Toward Action-Oriented Views in Cognitive Science,"
Andreas K. Engel, Karl J. Friston, and Danica Kragic, eds. 2016. Strüngmann Forum Reports, vol. 18,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03432-6.

forward and backward connections, the prevalence of nonlinear or modulatory backward connections (that embody interactions and nonlinearities inherent in the generation of sensory signals), and their spectral characteristics—with fast (e.g., gamma) activity predominating in forward connections and slower (e.g., beta) frequencies that accumulate evidence (prediction errors) ascending from lower levels.

Precision and the Encoding of (Un)certainty

One can regard ascending prediction errors as broadcasting "newsworthy" information that has yet to be explained by descending predictions. However, the brain has to select the channels it listens to—by adjusting the volume or *gain* of prediction errors that compete to update expectations in higher levels (Clark 2013a). Computationally, this gain corresponds to the precision or confidence associated with ascending prediction errors. However, to select prediction errors, the brain has to estimate and encode their precision (i.e., inverse variance). Having done this, prediction errors can then be weighted by their precision so that only precise information is accumulated and assimilated at high or deep hierarchical levels. As with all expectations, expected precision maximizes Bayesian model evidence (see Appendix). In other words, not only does the brain have to infer the causes of sensations, it also is in the difficult game of inferring the context, in terms of the reliability of prediction errors at each level of the hierarchy and the implicit confidence in associated predictions.

The implicit broadcasting of precision-weighted prediction errors rests on synaptic gain control (Moran et al. 2013). This neuromodulatory gain control corresponds to a (Bayes-optimal) encoding of precision in terms of the excitability of neuronal populations reporting prediction errors (Feldman and Friston 2010; Shipp et al. 2013). This may explain why superficial pyramidal cells have so many synaptic gain control mechanisms, such as NMDA receptors and classical neuromodulatory receptors like D1 dopamine receptors (Goldman-Rakic et al. 1992; Braver et al. 1999; Doya 2008; Lidow et al. 1991). Furthermore, it places excitation-inhibition balance in a prime position to mediate precision-engineered message passing within and among hierarchical levels (Humphries et al. 2009). The dynamic and context-sensitive control of precision has been associated with attentional gain control in sensory processing (Feldman and Friston 2010; Jiang et al. 2013) and has been discussed in terms of affordance in active inference and action selection (Frank et al. 2007; Cisek 2007; Friston, Shiner et al. 2012). Crucially, the delicate balance of precision over different hierarchical levels has a profound effect on inference, and may also offer a formal understanding of false inference in psychopathology (Fletcher and Frith 2009; Adams, Stephan et al. 2013). We will see that it also plays a crucial role in sensory attenuation.

From "The Pragmatic Turn: Toward Action-Oriented Views in Cognitive Science,"
Andreas K. Engel, Karl J. Friston, and Danica Kragic, eds. 2016. Strüngmann Forum Reports, vol. 18,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03432-6.

Active Inference and Action

So far, we have only considered the role of predictive coding in perception or perceptual inference, through minimizing prediction errors. However, there is another way to minimize prediction errors; namely, by resampling sensory inputs so that they conform to predictions. This is known as active inference (Friston et al. 2011). In active inference, action is regarded as the fulfilment of descending proprioceptive predictions by classical reflex arcs. In more detail, the brain generates continuous proprioceptive predictions about the expected location of the limbs and eyes, and these predictions are hierarchically consistent with the inferred state of the world. In other words, we believe that we will execute a goal-directed movement, and this belief is unpacked hierarchically to provide proprioceptive and exteroceptive predictions. These predictions are then fulfilled automatically by minimizing proprioceptive prediction errors at the level of the spinal cord and cranial nerve nuclei (see Adams, Shipp et al. 2013 and Figure 6.1). Mechanistically, descending proprioceptive predictions provide a target or *set point* for peripheral reflex arcs, which respond by minimizing (proprioceptive) prediction errors.

The argument here is that the same inferential mechanisms underlie apparently diverse functions (e.g., exteroception in the visual cortex, interoception in the insula, and motor control in the motor cortex). Crucially, because these modality-specific systems are organized hierarchically, they are all contextualized by the same conceptual (amodal) expectations that generate descending predictions in multiple (exteroceptive, interoceptive and proprioceptive) modalities. In short, action and perception are facets of the same underlying imperative; namely, to minimize hierarchical prediction errors through selective sampling of sensory inputs. However, there is a potential problem:

Action and Sensory Attenuation

If proprioceptive prediction errors can be resolved by engaging classical reflexes (action) or changing expectations (perception), how does the brain adjudicate between these two options? The answer lies in the precision afforded to (proprioceptive) prediction errors and the consequences of movement sensed in other modalities. To engage classical reflexes, it is necessary to increase their gain through augmenting the precision of (efferent) proprioceptive prediction errors that drive neuromuscular junctions. However, to preclude (a veridical) inference that the movement has not yet occurred, it is necessary to attenuate the precision of (afferent) prediction errors that would otherwise update kinematic expectations or beliefs (Friston et al. 2011). Put simply, my prior belief that I am moving can be subverted by sensory evidence to the contrary; thereby precluding movement. In short, it is necessary to attenuate *all of the sensory consequences of moving*—leading to an active inference formulation

From "The Pragmatic Turn: Toward Action-Oriented Views in Cognitive Science,"
Andreas K. Engel, Karl J. Friston, and Danica Kragic, eds. 2016. Strüngmann Forum Reports, vol. 18,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03432-6.

of sensory attenuation. Sensory attenuation is the psychological phenomena where the magnitude of self-made sensations are perceived as less intense (Brown et al. 2013).

In Figure 6.1, the (afferent) proprioceptive prediction error was omitted from the hypoglossal nucleus (for discussion of this omission and the agranular nature of motor cortex, see Shipp et al. 2013). This renders descending proprioceptive predictions *motor commands*, where the accompanying exteroceptive predictions become *corollary discharge*. The ensuing motor control is effectively *open loop*. However, the hierarchical generation of proprioceptive predictions is contextualized by sensory input in other modalities, which register the sensory consequences of movement. These sensory consequences are transiently attenuated during movement. In summary, to act, one needs to temporarily suspend attention to the consequences of action in order to articulate descending predictions (Brown et al. 2013). With this conceptual framework in place, let us now consider some key questions it raises about the nature of representation and mindful action.

Some Key Questions

The first question raised by the above formalism is whether causes of sensations can ever be observed or perceived. Put simply, is perception the same as inference or are they somehow distinct? This argument becomes particularly acute when inferring the mental state of others, where one might call upon generative models of self-made acts to infer the intentions of others (Kilner et al. 2007; Teufel et al. 2010; Brown and Brüne 2012). For example, Bohl and Gangopadhyay (2013) distinguish among several versions of what they term unobservability assumption, as discussed by Michael and De Bruin (2015):

1. It may be understood as a phenomenological thesis: one does not experience oneself as perceiving others' mental states, that is, others' mental states do not (ever?) have the same kind of vividness or experiential presence as colors and shapes. Here, phenomenologists might urge that perceiving someone's sadness, for example, is a qualitatively different experience from merely believing that they are sad, and that this is a distinction to which we should do justice.
2. It may be interpreted as a metaphysical thesis: mental states cannot be perceived because they are immaterial entities.
3. It may be read as an epistemological thesis: perceptual experiences do not ground judgments about others' mental states, at least not without the help of inference and background knowledge.
4. It may be understood as a psychological thesis about the processes by which we come to ascribe mental states to others (i.e., these are not perceptual processes).

From "The Pragmatic Turn: Toward Action-Oriented Views in Cognitive Science,"
Andreas K. Engel, Karl J. Friston, and Danica Kragic, eds. 2016. Strüngmann Forum Reports, vol. 18,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03432-6.

Mainstream approaches argue that even if we do perceive (some of) others' mental states, we still need to account for how this is achieved. Mainstream accounts of mind reading attempt to do this. Herschbach (2008), Michael et al. (2014), and Lavelle (2012) argue that mind-reading approaches offer accounts of the subpersonal inferential processes that underlie mental state ascriptions.

A first question, then, is: Just what version of unobservability assumption is affirmed or presupposed by which approach? The basic issue behind this question is whether there is something peculiar about theory of mind that distinguishes it from subpersonal inferences normally associated with perception (e.g., in the context of unconscious inference as proposed by Helmholtz 1866/1962). Clearly, one key difference between inferences about a world that contains other agents is that they have to be hierarchically deeper than the more elemental (possibly unconscious) inferences about, say, visual features. However, one may ask whether this is a sufficient distinction to dissolve questions posed by the unobservability assumption.

A second question is whether the brain entertains representations. In radical versions embodiment, representations per se are precluded. This is nicely articulated by Anil Seth in his treatment of predictive coding or processing (PP) (Seth 2015):

> The notion of counterfactual predictions connects PP with what at first glance seems its natural opponent: "enactive" theories of perception and cognition which explicitly reject internal models or representations (Hutto and Myin 2013; Thompson and Varela 2001). Central to the enactive approach are notions of "sensorimotor contingencies" and their "mastery" (O'Regan and Noë 2001), where a sensorimotor contingency refers to a rule governing how sensory signals change in response to action. Accordingly, the perceptual experience of, for example, redness is given by an implicit knowledge (mastery) of the way red things behave, given certain patterns of behavior. Mastery of sensorimotor contingencies is also said to underpin perceptual presence: the sense of subjective reality of the contents of perception (Noë 2006). From the perspective of PP, mastery of a sensorimotor contingency corresponds to the learning of a counterfactually equipped predictive model, connecting potential actions to expected sensory consequences. The resulting theory of Predictive Perception of Sensorimotor Contingencies provides a much needed reconciliation of enactive and predictive theories of perception and action.

Clearly, there is some consilience between active inference (predictive processing) and sensorimotor contingency theory. The inferences we make about our sensations are (at some level) amodal and entail necessarily predictions of a proprioceptive sort—consistent with our embodied interactions with objects causing sensations. This is true even at the level of visual searches (Wurtz et al. 2011). Do posterior beliefs, then, constitute representations? The active inference story suggests that the answer is yes—although the answer is nuanced

From "The Pragmatic Turn: Toward Action-Oriented Views in Cognitive Science,"
Andreas K. Engel, Karl J. Friston, and Danica Kragic, eds. 2016. Strüngmann Forum Reports, vol. 18,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03432-6.

to consider probabilistic representations. Does the encoding of expectations (sufficient statistics or posterior beliefs), therefore, violate the tenets of radical embodiment and, if so, what are the implications for this and other theoretical formulations?

A related issue here is whether inferences about hidden states are conscious or unconscious. At what point do unconscious inferences become conscious, and what is the relationship between inference and creature (or indeed state) consciousness? There are clearly some relatively simple arguments about reportability of (state) consciousness content that require inference at a sufficiently deep hierarchical level to generate proprioceptive (or interoceptive) predictions that engage action. It may be that the notion of # may help organize the formal distinctions between conscious inferences and unconscious (subpersonal) inferences.

Finally, we come to questions about theory of mind and the mirror neuron system, which are a necessary aspect of active inference. This is self-evident from the fact that the same dynamics (posterior expectations) of neuronal populations—that occupy a privileged position in the cortical hierarchy—encode predictions about sensations during perception and generate motor commands (proprioceptive predictions) during action. Does this provide a sufficient account of theory of mind, or does it subvert our understanding of agency in some fundamental way? Is it appropriate to associate the same processes underlying perceptual inference with inferences about others? What are the implications of subsuming inference about low-level perceptual features and the intentions of others within the same hierarchical framework?

Sensory attenuation, in particular, requires us to either act or observe—but not do both at the same time. Is sensory attenuation, therefore, an integral part of agency and a sense of self? Can inferences about others be informed by predicting internal (as opposed to external) states of the world?

This brings us to another intriguing area: interoceptive inference and its putative role in inferring the emotional states of self, and possibly others (Seth 2013; Seth et al. 2011). Can one usefully transcribe the principles of active inference in the domain of motor control to emotional pragmatics and shared autonomic states? This is a fascinating area that speaks not just to our sense of self; it may also have had implications for psychopathology, when failing to attend or attenuate interoceptive prediction errors, as has been proposed in autism (Happe and Frith 2006; Lawson et al. 2014; Van de Cruys et al. 2014; Pellicano and Burr 2012).

In conclusion, deep questions arise from the challenges inherent in action and embodiment—questions that are particularly important for the notion of active inference. In subsequent chapters in this volume, various attempts are made to resolve, dissolve, or simply celebrate these questions.

From "The Pragmatic Turn: Toward Action-Oriented Views in Cognitive Science,"
Andreas K. Engel, Karl J. Friston, and Danica Kragic, eds. 2016. Strüngmann Forum Reports, vol. 18,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03432-6.

Appendix

This brief description of generalized predictive coding is based on Feldman and Friston (2010). A more technical description can be found in Friston, Stephan et al. (2010). This scheme is based on three assumptions:

1. The brain minimizes a free energy of sensory inputs defined by a generative model.
2. The generative model used by the brain is hierarchical, nonlinear, and dynamic.
3. Neuronal firing rates encode the expected state of the world under this model.

Free energy is a quantity from statistics that measures the quality of a model in terms of the probability that it could have generated observed outcomes. This means that minimizing free energy maximizes the Bayesian evidence for the generative model. The second assumption is motivated by noting that the world is both dynamic and nonlinear, and that hierarchical causal structure emerges inevitably from a separation of spatial and temporal scales. The final assumption is the Laplace assumption, which leads to the simplest and most flexible of all neural codes.

Given these assumptions, one can simulate a whole variety of neuronal processes by specifying the particular equations that constitute the brain's generative model. In brief, these simulations use differential equations that minimize the free energy of sensory input using a generalized gradient descent:

$$\dot{\tilde{\mu}}(t) = D\tilde{\mu}(t) - \partial_{\tilde{\mu}} F(\tilde{s}, \tilde{\mu}). \tag{6.1}$$

These differential equations say that neuronal activity encoding posterior expectations about (generalized) hidden states of the world $\tilde{\mu} = (\mu, \mu', \mu'', \ldots)$ reduce free energy, where free energy, $F(\tilde{s}, \tilde{\mu})$, is a function of sensory inputs, $\tilde{s} = (s, s', s'', \ldots)$, and neuronal activity. This is known as generalized predictive coding or Bayesian filtering. The first term is a prediction based upon a differential matrix operator, D, that returns the generalized motion of expected hidden states, $D\tilde{\mu} = (\mu', \mu'', \mu''', \ldots)$. The second (correction) term is usually expressed as a mixture of prediction errors that ensures the changes in posterior expectations are Bayes-optimal predictions about hidden states of the world. To perform neuronal simulations under this scheme, it is only necessary to integrate or solve Equation 6.1 to simulate the neuronal dynamics that encode posterior expectations. Posterior expectations depend upon the brain's generative model of the world, which we assume has the following hierarchical form:

From "The Pragmatic Turn: Toward Action-Oriented Views in Cognitive Science,"
Andreas K. Engel, Karl J. Friston, and Danica Kragic, eds. 2016. Strüngmann Forum Reports, vol. 18,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03432-6.

$$
\begin{aligned}
s &= g^{(1)}\left(x^{(1)}, v^{(1)}\right) + \exp\left(-\tfrac{1}{2}\pi_v^{(1)}\left(x^{(1)}, v^{(1)}\right)\right)\cdot\omega_v^{(1)} \\
\dot{x}^{(1)} &= f^{(1)}\left(x^{(1)}, v^{(1)}\right) + \exp\left(-\tfrac{1}{2}\pi_x^{(1)}\left(x^{(1)}, v^{(1)}\right)\right)\cdot\omega_x^{(1)} \\
&\vdots \\
v^{(i-1)} &= g^{(i)}\left(x^{(i)}, v^{(i)}\right) + \exp\left(-\tfrac{1}{2}\pi_v^{(i)}\left(x^{(i)}, v^{(i)}\right)\right)\cdot\omega_v^{(i)} \\
\dot{x}^{(i)} &= f^{(i)}\left(x^{(i)}, v^{(i)}\right) + \exp\left(-\tfrac{1}{2}\pi_x^{(i)}\left(x^{(i)}, v^{(i)}\right)\right)\cdot\omega_x^{(i)} \\
&\vdots
\end{aligned}
\tag{6.2}
$$

This equation describes a probability density over the sensory and hidden states that generate sensory input. Here, the hidden states have been divided into hidden states and causes $(x^{(i)}, v^{(i)})$ at the i-th level within the hierarchical model. Hidden states and causes are abstract variables that the brain uses to explain or predict sensations—like the motion of an object in the field of view.

In these models, hidden causes link hierarchical levels, whereas hidden states link dynamics over time. Here, $(f^{(i)}, g^{(i)})$ are nonlinear functions of hidden states and causes that generate hidden causes for the level below and—at the lowest level—sensory inputs. Random fluctuations in the motion of hidden states and causes $(\omega_x^{(i)}, \omega_v^{(i)})$ enter each level of the hierarchy. Gaussian assumptions about these random fluctuations make the model probabilistic. They play the role of sensory noise at the first level and induce uncertainty at higher levels. The amplitudes of these random fluctuations are quantified by their precisions that may depend upon the hidden states or causes through their log-precisions $(\pi_x^{(i)}, \pi_v^{(i)})$.

Given the form of the generative model (Equation 6.2) we can now write down the differential equations (Equation 6.1) describing neuronal dynamics in terms of (precision-weighted) prediction errors. These errors represent the difference between posterior expectations and predicted values, under the generative model (using $A \cdot B \triangleq A^T B$ and omitting higher-order terms):

$$
\begin{aligned}
\dot{\tilde{\mu}}_x^{(i)} &= D\tilde{\mu}_x^{(i)} + \left(\frac{\partial \tilde{g}^{(i)}}{\partial \tilde{\mu}_x^{(i)}} - \tfrac{1}{2}\tilde{\varepsilon}_v^{(i)}\frac{\partial \tilde{\pi}_v^{(i)}}{\partial \tilde{\mu}_x^{(i)}}\right)\cdot\xi_v^{(i)} + \left(\frac{\partial \tilde{f}^{(i)}}{\partial \tilde{\mu}_x^{(i)}} - \tfrac{1}{2}\tilde{\varepsilon}_x^{(i)}\frac{\partial \tilde{\pi}_x^{(i)}}{\partial \tilde{\mu}_x^{(i)}}\right)\cdot\xi_x^{(i)} + \frac{\partial tr(\tilde{\pi}_v^{(i)} + \tilde{\pi}_x^{(i)})}{\partial \tilde{\mu}_x^{(i)}} - D^T\xi_x^{(i)} \\
\dot{\tilde{\mu}}_v^{(i)} &= D\tilde{\mu}_v^{(i)} + \left(\frac{\partial \tilde{g}^{(i)}}{\partial \tilde{\mu}_v^{(i)}} - \tfrac{1}{2}\tilde{\varepsilon}_v^{(i)}\frac{\partial \tilde{\pi}_v^{(i)}}{\partial \tilde{\mu}_v^{(i)}}\right)\cdot\xi_v^{(i)} + \left(\frac{\partial \tilde{f}^{(i)}}{\partial \tilde{\mu}_x^{(i)}} - \tfrac{1}{2}\tilde{\varepsilon}_x^{(i)}\frac{\partial \tilde{\pi}_x^{(i)}}{\partial \tilde{\mu}_v^{(i)}}\right)\cdot\xi_x^{(i)} + \frac{\partial tr(\tilde{\pi}_v^{(i)} + \tilde{\pi}_x^{(i)})}{\partial \tilde{\mu}_v^{(i)}} - \xi_v^{(i+1)}
\end{aligned}
$$

$$
\begin{aligned}
\xi_x^{(i)} &= \exp(\tilde{\pi}_x^{(i)})\cdot\tilde{\varepsilon}_x^{(i)} \\
\xi_v^{(i)} &= \exp(\tilde{\pi}_v^{(i)})\cdot\tilde{\varepsilon}_v^{(i)}
\end{aligned}
\tag{6.3}
$$

$$
\begin{aligned}
\tilde{\varepsilon}_x^{(i)} &= D\tilde{\mu}_x^{(i)} - \tilde{f}^{(i)}(\tilde{\mu}_x^{(i)}, \tilde{\mu}_v^{(i)}) \\
\tilde{\varepsilon}_v^{(i)} &= \tilde{\mu}_v^{(i-1)} - \tilde{g}^{(i)}(\tilde{\mu}_x^{(i)}, \tilde{\mu}_v^{(i)}).
\end{aligned}
$$

This produces a relatively simple update scheme, in which posterior expectations $\tilde{\mu}^{(i)}$ are driven by a mixture of prediction errors $\tilde{\varepsilon}^{(i)}$ that are defined by the equations of the generative model.

From "The Pragmatic Turn: Toward Action-Oriented Views in Cognitive Science," Andreas K. Engel, Karl J. Friston, and Danica Kragic, eds. 2016. Strüngmann Forum Reports, vol. 18, series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03432-6.

In neural network terms, Equation 6.3 says that error-units compute the difference between expectations at one level and predictions from the level above (where $\xi^{(i)}$ are precision weighted prediction errors at the *i*-th level of the hierarchy). Conversely, posterior expectations are driven by prediction errors from the same level and the level below. These constitute bottom-up and lateral messages that drive posterior expectations toward a better prediction to reduce the prediction error in the level below. In neurobiological implementations of this scheme, the sources of bottom-up prediction errors are generally thought to be superficial pyramidal cells, because they send forward (ascending) connections to higher cortical areas. Conversely, predictions are thought to be conveyed from deep pyramidal cells by backward (descending) connections, to target the superficial pyramidal cells encoding prediction error (Mumford 1992; Bastos et al. 2012).

Note that the precisions depend on the expected hidden causes and states. We have proposed that this dependency mediates attention (Feldman and Friston 2010). Equation 6.3 tells us that the (state-dependent) precisions modulate the responses of prediction error units to their presynaptic inputs. This suggests something intuitive—attention is mediated by activity-dependent modulation of the synaptic gain of principal cells that convey sensory information (prediction error) from one cortical level to the next. This translates into a top-down control of synaptic gain in principal (superficial pyramidal) cells and fits comfortably with the modulatory effects of top-down connections in cortical hierarchies that have been associated with attention and action selection.

Acknowledgments

KJF is funded by the Wellcome Trust.

From "The Pragmatic Turn: Toward Action-Oriented Views in Cognitive Science,"
Andreas K. Engel, Karl J. Friston, and Danica Kragic, eds. 2016. Strüngmann Forum Reports, vol. 18,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-03432-6.