Cerebral hierarchies: predictive processing, precision and the pulvinar

Ryota Kanai[1,2], Yutaka Komura[1,2,3], Stewart Shipp[4] and Karl Friston[5]

[1]School of Psychology, Sackler Centre for Consciousness Science, University of Sussex, Brighton BN1 9QH, UK
[2]Precursory Research for Embryonic Science and Technology, Japan Science and Technology Agency, Saitama 332-0012, Japan
[3]Systems Neuroscience, Human Technology Research Institute, National Institute of Advanced Industrial Science and Technology, Tsukuba 305-8568, Japan
[4]UCL Institute of Ophthalmology, London EC1V 9EL, UK
[5]Wellcome Trust Centre for Neuroimaging, Institute of Neurology, UCL, London WC1 3BG, UK

This paper considers neuronal architectures from a computational perspective and asks what aspects of neuroanatomy and neurophysiology can be disclosed by the nature of neuronal computations? In particular, we extend current formulations of the brain as an organ of inference—based upon hierarchical predictive coding—and consider how these inferences are orchestrated. In other words, what would the brain require to dynamically coordinate and contextualize its message passing to optimize its computational goals? The answer that emerges rests on the delicate (modulatory) gain control of neuronal populations that select and coordinate (prediction error) signals that ascend cortical hierarchies. This is important because it speaks to a hierarchical anatomy of extrinsic (between region) connections that form two distinct classes, namely a class of driving (first-order) connections that are concerned with encoding the content of neuronal representations and a class of modulatory (second-order) connections that establish context—in the form of the salience or *precision* ascribed to content. We explore the implications of this distinction from a formal perspective (using simulations of feature–ground segregation) and consider the neurobiological substrates of the ensuing precision-engineered dynamics, with a special focus on the pulvinar and attention.

1. Introduction

There are many fascinating aspects of cerebral cartography that have been disclosed over the past centuries and—presumably—many more that have yet to reveal themselves. In this paper, we focus on one particular aspect, namely the imperatives for the coordination of message passing in the brain—and what those imperatives mandate in terms of cortical (and subcortical) anatomy. We base our treatment on the assumption that cerebral cartography is an anatomy of inference. In other words, the brain is a statistical organ predicting worldly states that generate its sensory inputs. In particular, we focus on predictive coding as a (biologically plausible) implementation of hierarchical inference in the brain and see how far this takes us in understanding the orchestration and contextualization of neuronal dynamics.

In what follows, we briefly review predictive coding with a special focus on how the brain encodes irreducible uncertainty inherent in the sensory evidence it samples [1,2]. It transpires that—under predictive coding—this uncertainty or relative confidence in sensory (and extrasensory) information can be succinctly encoded by the gain of certain neuronal populations that pass information from one hierarchal cortical level to the next [2,3]. This immediately brings us into the realm of cortical gain control and neuromodulation—that may be closely tied to synchronous gain and the (oscillatory) dynamics associated with binding, attention and dynamic coordination [4,5]. We then consider the computational anatomy implied by encoding the confidence or *precision* of ascending neuronal
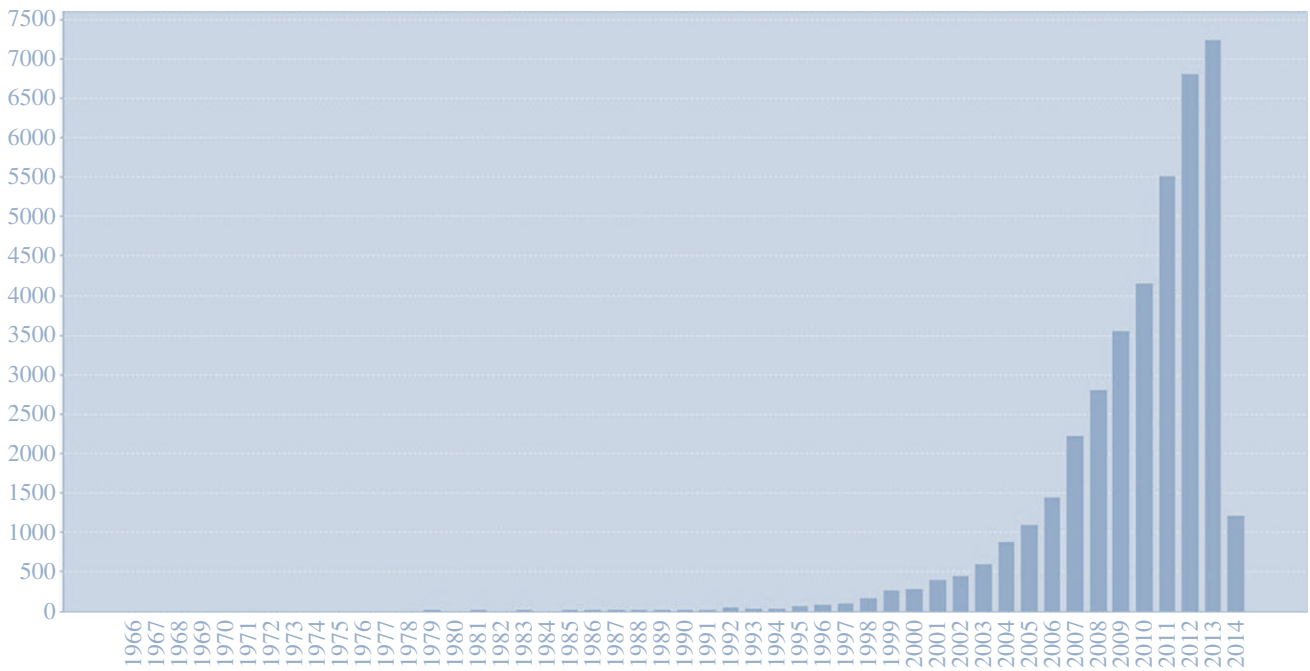
Figure 1. Citations per year, from 1966 to 2014, when searching for TOPIC: (Bayesian) AND TOPIC: (brain) in Web of Science. (Online version in colour.)

signals that is illustrated with a simple problem, namely figure–ground segregation in the visual hierarchy. We then turn to the neurobiology of cortical gain control, using the pulvinar as a prime example of a subcortical structure that has all the equipment necessary for contextualizing hierarchical inference in cortical hierarchies.

2. The Bayesian brain

Recent advances in theoretical neuroscience have inspired a paradigm shift in cognitive neuroscience (figure 1). This shift is away from the brain as a passive filter of sensations towards a view of the brain as a statistical organ that generates hypotheses or fantasies which are tested against sensory evidence [6]. In this formulation, the brain is, literally, a fantastic organ (fantastic: from Greek *phantastikos*, the ability to create mental images, from *phantazesthai*). This perspective can be traced back to Helmholtz and the notion of unconscious inference [7]. This notion has been generalized to cover deep or hierarchical Bayesian inference—about the causes of our sensations— and how these inferences induce beliefs, movement and behaviour [8–12].

(a) Predictive coding and the Bayesian brain

Modern formulations of Helmholtz's notion are now among the most popular explanations for neuronal message passing and are usually considered under the Bayesian brain hypothesis as predictive coding [12–15]. There is now considerable (circumstantial) anatomical and physiological evidence for predictive coding in the brain [15,16]. See Bastos *et al.* [17] for a review of canonical microcircuits and hierarchical predictive coding in perception and Adams *et al.* and Shipp *et al.* [18,19] for an equivalent treatment of the motor system. 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 (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 that 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 of a (hidden) cause or *expectation*. These causes are referred to as *hidden* because they have to be inferred from their sensory consequences.

In summary, predictive coding represents a biologically plausible scheme for updating beliefs (or expectations) about the world using sensory samples (figure 2). In this setting, neuroanatomy and neurophysiology can be regarded as a biological embodiment of how sensory signals are generated; for example, a smiling face generates luminance surfaces that generate textures and edges and so on, down to retinal input. This form of hierarchical inference explains a remarkable number of anatomical and physiological facts as documented elsewhere [15,17,18]. In brief, it explains the hierarchical nature of cortical cartography; the prevalence of backward connections and many of the functional and structural asymmetries in the extrinsic connections that link hierarchical levels. These asymmetries include the laminar specificity of 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 (prediction errors) and slower (e.g. beta) frequencies emerging as this evidence is accumulated in units that provide descending predictions [20–22].

At a more macroscopic level, the implicit anatomy of inference also provides a simple explanation for functional segregation [23]; in the sense that hierarchically deeper representations or expectations (e.g. what and where) are
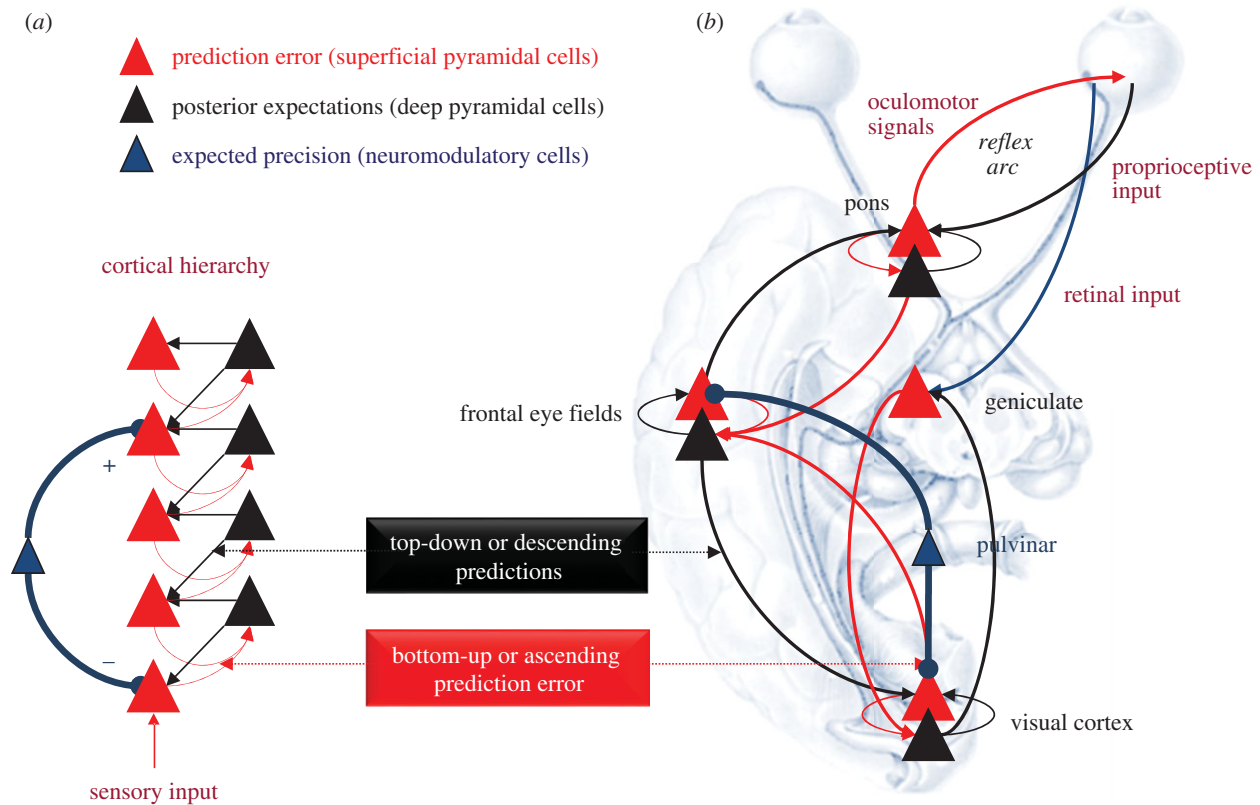
Figure 2. Summary of hierarchical message passing in predictive coding. Neuronal activity encodes expectations about the causes of sensory input, where these expectations minimize prediction error. Prediction error is the difference between (ascending) sensory input and (descending) predictions of that 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: see Bastos et al. [17] for a review of this evidence. (a) This schematic shows a simple cortical hierarchy with ascending prediction errors and descending predictions. We have included neuromodulatory gating or gain control (blue) of superficial pyramidal cells that determines their relative influence on deep pyramidal cells encoding expectations. (b) This provides a schematic example in the visual system: it shows the putative cells of origin of ascending or forward connections that convey prediction errors (red arrows) and descending or backward connections (black arrows) that construct predictions. The prediction errors are weighted by their expected precision— which we have associated with projections from the pulvinar. In this example, the frontal eye fields send predictions to primary visual cortex, which projects to the lateral geniculate body. However, the frontal eye fields also send proprioceptive predictions to pontine nuclei, which are passed to the oculomotor system to cause movement through classical reflexes. Every top-down prediction is reciprocated with a bottom-up prediction error to ensure predictions are constrained by sensory information.

statistically segregated and are combined in a nonlinear way to contextualize lower-level causes of sensory information (e.g. colour and form). Indeed, one could argue that the very existence of slender axonal connections in the brain speaks to the sparse causal dependencies or laws that endow our sensory world with statistical regularities (contrast the anatomy of the brain with the anatomy of the liver). Although hierarchical predictive coding appears to have broad explanatory power, there is something missing from this picture. In short, there appears to be no role for corticothalamic interactions or (recursive) coupling with other subcortical structures. So, what is lacking in the above description of predictive coding?

(b) Precision engineering and the encoding of context
One can regard ascending prediction errors as broadcasting 'newsworthy' information that has yet to be explained by descending predictions. However, the brain also 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. Computationally, this gain corresponds to the precision or confidence associated with ascending prediction errors; very much in the same way that we assess the statistical significance of an effect in relation to its standard error. However, to optimally select the prediction errors—that should be afforded greater influence—the brain has to estimate or encode their precision. Having done this, prediction errors can then be weighted by their precision, so that only precise information is accumulated and assimilated in high or deep hierarchical levels.

This broadcasting of precision-weighted prediction errors may rest on neuromodulatory gain control mechanisms at a synaptic level [24]. This gain control corresponds to a (Bayes-optimal) encoding of precision in terms of the excitability of neuronal populations reporting prediction errors [2,19]. This may explain why superficial pyramidal cells have so many synaptic gain control mechanisms such as N-methyl-D-aspartate (NMDA) receptors and classical neuromodulatory receptors like D1 dopamine receptors [25–28]. Furthermore, it places excitation–inhibition balance in a prime position to mediate precision-engineered message passing within and among hierarchical levels [29]. As noted above, the encoding of salience or precision can also be understood in terms of synchronous gain [30] and fast (oscillatory) dynamics associated with binding, attention and dynamic coordination [4,5].

4

rstb.royalsocietypublishing.org   Phil. Trans. R. Soc. B 370: 20140169

The dynamic and context-sensitive control of precision has been associated with attentional gain control in sensory processing [2,31] and has been discussed in terms of affordance and action selection [32–34]. 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 [35]. To illustrate the potential importance of precision—and implicit synaptic gain control—we will look at a particular problem from the point of view of predictive coding and see why neuromodulatory mechanisms are an integral part of its solution.

3. Predictive coding and figure–ground segregation

This section considers the figure–ground segregation problem where, crucially, a figure is defined texturally—in terms of its second-order statistics; in other words, a visual object is manifest in terms of its texture or spectral power in the spatial domain. This segregation problem precludes recourse to first-order attributes, such as differences in luminance or colour. In other words, the quantities causing visual impressions are only defined in terms of their precision (or inverse variance). This presents an interesting problem for predictive coding (and the brain) that we use to illustrate the importance of gain control in finessing the inference problem.

In statistics, this (inverse) problem is usually solved using some form of variance component estimation; for example, using covariance constraints in the electromagnetic source reconstruction problem. Here, we solve the same problem with predictive coding. In this setting, hidden causes in the generative model control the precision or variance of subordinate causes generating data. Expectations of these hierarchical causes are optimized with respect to variational free energy—using predictive coding. Here, variational free energy is a proxy for Bayesian model evidence and can be regarded as the sum of the (squared and precision-weighted) prediction error. The simulation used to illustrate this solution is trivially simple but sufficient to make our key point, namely that top-down predictions have very different effects on prediction error responses—depending upon whether they encode the first- or second-order statistical properties of a stimulus.

We simulated a contiguous object, whose texture was determined by the variance of random fluctuations in luminance, where this variance was modulated by (Gaussian) spatial basis functions of retinotopic space. The resulting signal was mixed with uniform Gaussian observation noise to produce sensory data. These data were then subjected to Bayesian inversion using (generalized) predictive coding to recover the object or figure. The implicit figure–ground segregation basically involves estimating the hidden causes modulating the spatial basis functions controlling textural features—in this case, the local variance of stimulus intensity over sensory channels.

Technically, predictive coding optimizes expectations of the hidden causes of data that, in this case, include the amplitude of radial basis functions controlling the precision (inverse variance) of retinotopic signals (see below). In brief, we see that the resulting figure–ground segregation rests on selectively attending to sensory input from the figure, relative to the background. However, this form of attention is distinct from simply boosting sensory precision (the precision of sensory prediction

errors) as in simulations of the Posner paradigm or biased competition [2]. Here, expectations of hidden causes are optimized in a way that renders them *less precise* and therefore more sensitive to ascending sensory (prediction error) input. This illustrates the importance of the *relative precision* of sensory and extrasensory prediction errors in modulating the influence of ascending sensory information (figure 2).

(a) Simulation details

Three Gaussian basis functions $B^{(2)} \in \mathbb{R}^{128 \times 3}$ of a one-dimensional retinotopic space (with a separation and standard deviation of eight channels) were modulated with three hidden causes $v^{(2)} = [8, 8, 0]$ to generate the log-precision of a visual signal over 128 visual channels. The resulting log precision vector $\pi_v^{(2)} \in \mathbb{R}^{128 \times 1}$ was used to modulate Gaussian fluctuations $\omega_v^{(2)}$ to generate textured signals by applying a Gaussian convolution matrix $B^{(1)} \in \mathbb{R}^{128 \times 128}$ (with a standard deviation of two channels); finally, uniform Gaussian noise $\omega_v^{(1)}$ with a precision of 16 was added to the signals to generate sensory data

$$s = B^{(1)} \cdot v^{(1)} + \tfrac{1}{4} \cdot \omega_v^{(1)},$$
$$v^{(1)} = \mathrm{diag}(\exp(-\tfrac{1}{2}\pi_v^{(2)})) \cdot \omega_v^{(2)}$$
$$\text{and} \quad \pi_v^{(2)} = 8 - B^{(2)} \cdot v^{(2)}. \tag{3.1}$$

This way of generating data rests upon a generative model of the following form, which is a special case of the generative models described in the appendix: it is a special case, because there are no dynamics or hidden states

$$\left. \begin{array}{c} s = g^{(1)}(v^{(1)}) + \exp(-\tfrac{1}{2}\pi_v^{(1)}(v^{(1)})) \cdot \omega_v^{(1)} \\ \text{and} \quad v^{(1)} = g^{(2)}(v^{(2)}) + \exp(-\tfrac{1}{2}\pi_v^{(2)}(v^{(2)})) \cdot \omega_v^{(i)}. \\ \vdots \end{array} \right\} \tag{3.2}$$

This generative model specifies the neuronal dynamics for posterior expectations about the hidden causes $\dot{\mu}_v^{(i)}$ that constitute predictive coding (see appendix for details and figure 3 for the particular equations of the current model), where $A \cdot B \triangleq A^{\mathrm{T}}B$:

$$\left. \begin{array}{c} \dot{\mu}_v^{(i)} = \left(\dfrac{\partial g^{(i)}}{\partial \mu_v^{(i)}} - \tfrac{1}{2}\tilde{\varepsilon}_v^{(i)}\dfrac{\partial \pi_v^{(i)}}{\partial \mu_v^{(i)}}\right) \cdot \xi_v^{(i)} + \dfrac{\partial tr(\pi_v^{(i)})}{\partial \mu_v^{(i)}} - \xi_v^{(i+1)}, \\ \xi_v^{(i)} = \exp(\pi_v^{(i)}) \cdot \varepsilon_v^{(i)} \\ \text{and} \quad \varepsilon_v^{(i)} = \mu_v^{(i-1)} - g^{(i)}(\mu_v^{(i)}) \end{array} \right\} \tag{3.3}$$

These equations provide a relatively simple update scheme, in which changes in posterior expectations are driven by a mixture of precision-weighted prediction errors—where prediction errors are defined by the equations of the generative model. Crucially, prediction errors are affected by descending predictions in one of two ways: expectations can either generate predictions of first-order effects, through the functions $g^{(i)}(\mu_v^{(i)})$. Alternatively, they can generate predictions of precision, through the functions $\pi_v^{(i)}(\mu_v^{(i)})$. These effects are formally distinct: the first-order predictions (of lower expectations) have a negative (driving) effect on the prediction errors, whereas the second-order predictions (of their precision) have a positive (modulatory) effect. We can see this separation clearly in the current example, because the
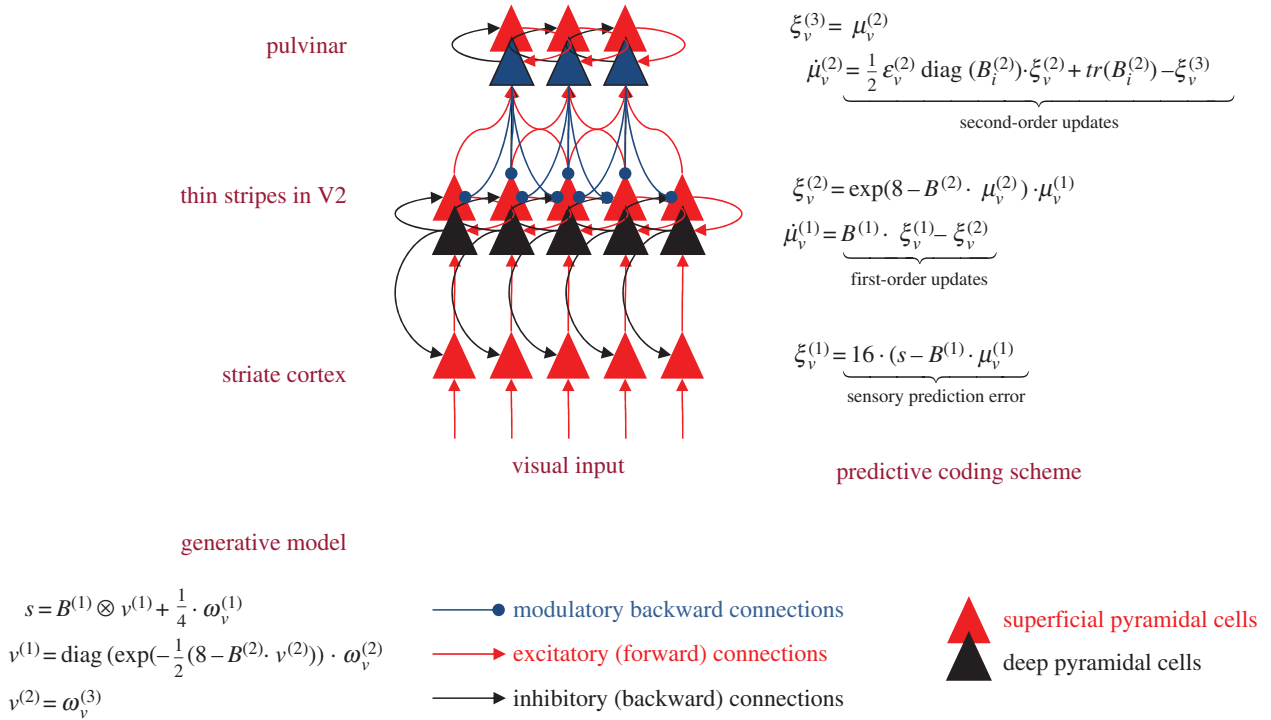
Figure 3. This schematic illustrates the message passing implicit in predictive coding based on the generative model described in the main text—and formulated mathematically on the lower left. In this scheme, sensory input is conveyed to visual cortex via ascending prediction errors from the lateral geniculate nucleus. Posterior expectations, encoded by the activity of deep pyramidal cells in primary visual cortex, are driven by ascending prediction errors while, at the same time, they are subjected to lateral interactions—with second-level prediction errors—that mediate (empirical) priors. Crucially, these constraints are modulated by top-down predictions of their precision (blue arrows). These predictions are based upon expectations about precision in the highest level that are effectively driven by the variance or power of prediction errors at the lower level. Heuristically, expectations about precision release posterior expectations from constraints in the vicinity of an inferred object and allow them to respond more sensitively to ascending geniculate input.

second-level hidden causes only predict second-order statistics (log precision), whereas the first-level hidden predict first-order statistics. This means equation (3.3) can be separated into first- and second-order updates

$$\dot{\mu}_v^{(1)} = \frac{\partial g^{(1)}}{\partial \mu_v^{(1)}} \cdot \xi_v^{(1)} - \xi_v^{(2)},$$

$$\xi_v^{(1)} = \underbrace{\exp\left(\pi_v^{(1)}\right) \cdot \left(s - g^{(1)}(\mu_v^{(1)})\right)}_{\text{first-order prediction errors}},$$

$$\dot{\mu}_v^{(2)} = \frac{\partial tr(\pi_v^{(2)})}{\partial \mu_v^{(2)}} - \frac{1}{2}\tilde{\varepsilon}_v^{(2)} \frac{\partial \pi_v^{(2)}}{\partial \mu_v^{(2)}} \cdot \xi_v^{(2)} - \mu_v^{(2)}$$

and $\quad \xi_v^{(2)} = \underbrace{\exp\left(\pi_v^{(2)}(\mu_v^{(2)})\right) \cdot \mu_v^{(1)}}_{\text{second-order prediction errors}}$. $\quad$ (3.4)

The key thing to take from these equations is the difference between first- and second-order message passing. The first-order expectations are driven by linear mixtures of first- and second-order prediction errors that play the role of the likelihood and (empirical) prior influences. Crucially, the second-order prediction errors (empirical priors) have more influence when they are more precise. Similarly, the first-order predictions enter the first-order prediction errors in a linear (subtractive) fashion. In contrast, the second-order expectations are driven by nonlinear (sum of squared) prediction errors and modulate the second-order prediction errors in a nonlinear fashion. It is this modulation we associate with precision-engineered message passing and the (attentional) contextualization of predictive coding. (See figure 3 for a

schematic of this message passing for the simple model considered here.)

Figure 4 shows the results of a typical simulation. The left columns show the results of predictive coding and the right columns show the true values generating sensory input. These inputs were inverted using the generative model that was used to produce them—but with unknown hidden causes at the first and second levels. The posterior expectations of these hidden causes are shown in the lower left panels—along with their 90% posterior confidence intervals (in grey). The upper left panel shows the predicted sensory input in blue, and the sensory prediction error in red.

Figure 5 shows the results of the same simulation but using a two-dimensional visual input (and a grid of nine Gaussian radial basis functions). Here, the signal was an L-shaped feature (with anisotropic smoothness) in the lower left quadrant that has, effectively, attracted attention. This attentional 'spotlight' is nicely illustrated in terms of the expected variance (inverse precision) as shown in the middle panel. Although very simple, this example highlights the close relationship between attentional selection and figure–ground segregation based upon second-order statistics. Clearly, we could have used a much more elaborate generative model; for example, we could have considered anisotropic Gabor patches when mapping first level hidden causes to sensory (retinotopic) input (cf. classical receptive fields). One could also imagine having separate precision components for vertical and horizontal patches that themselves were constrained by higher hierarchical levels. We will illustrate these ideas in future papers. Here, we
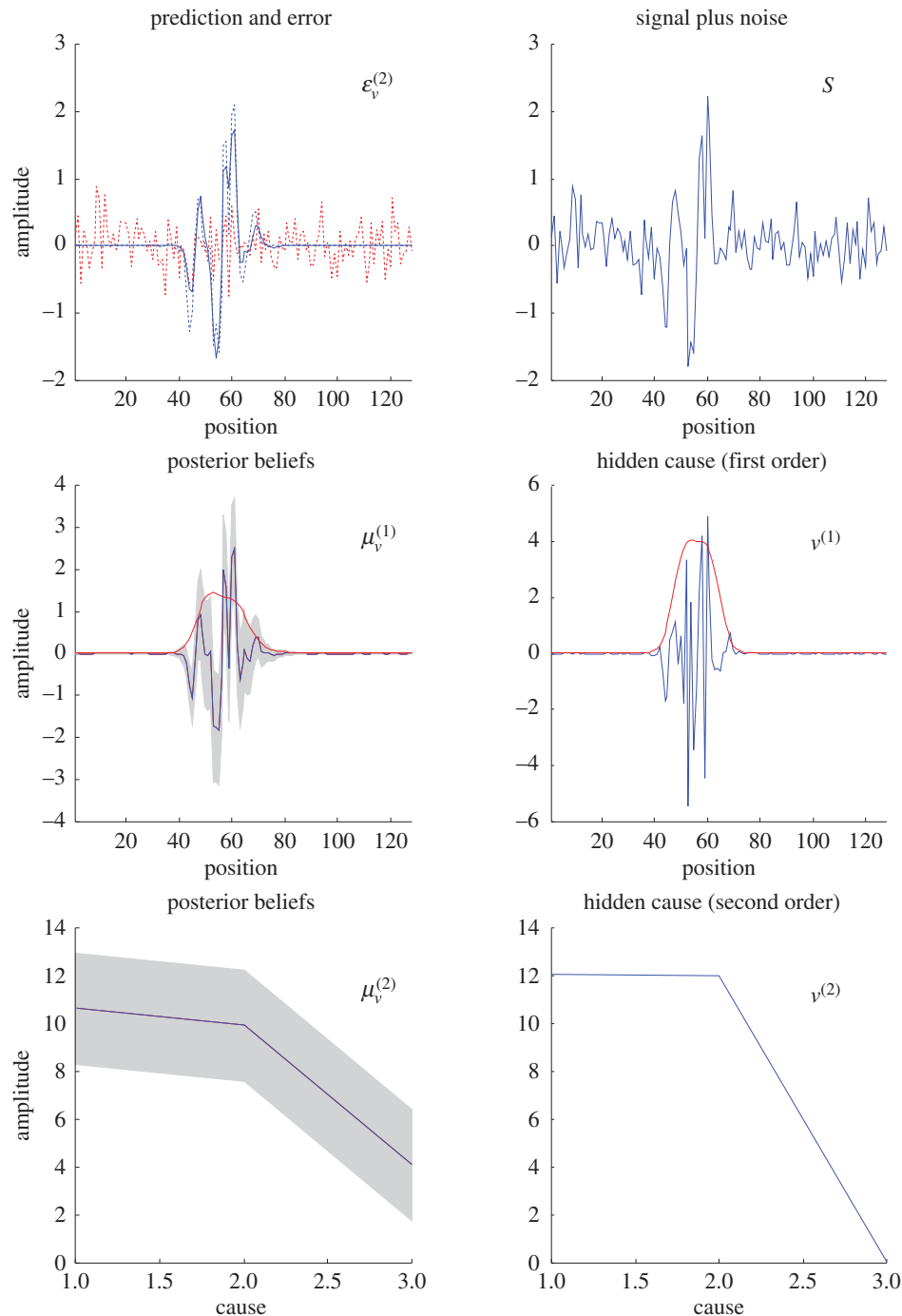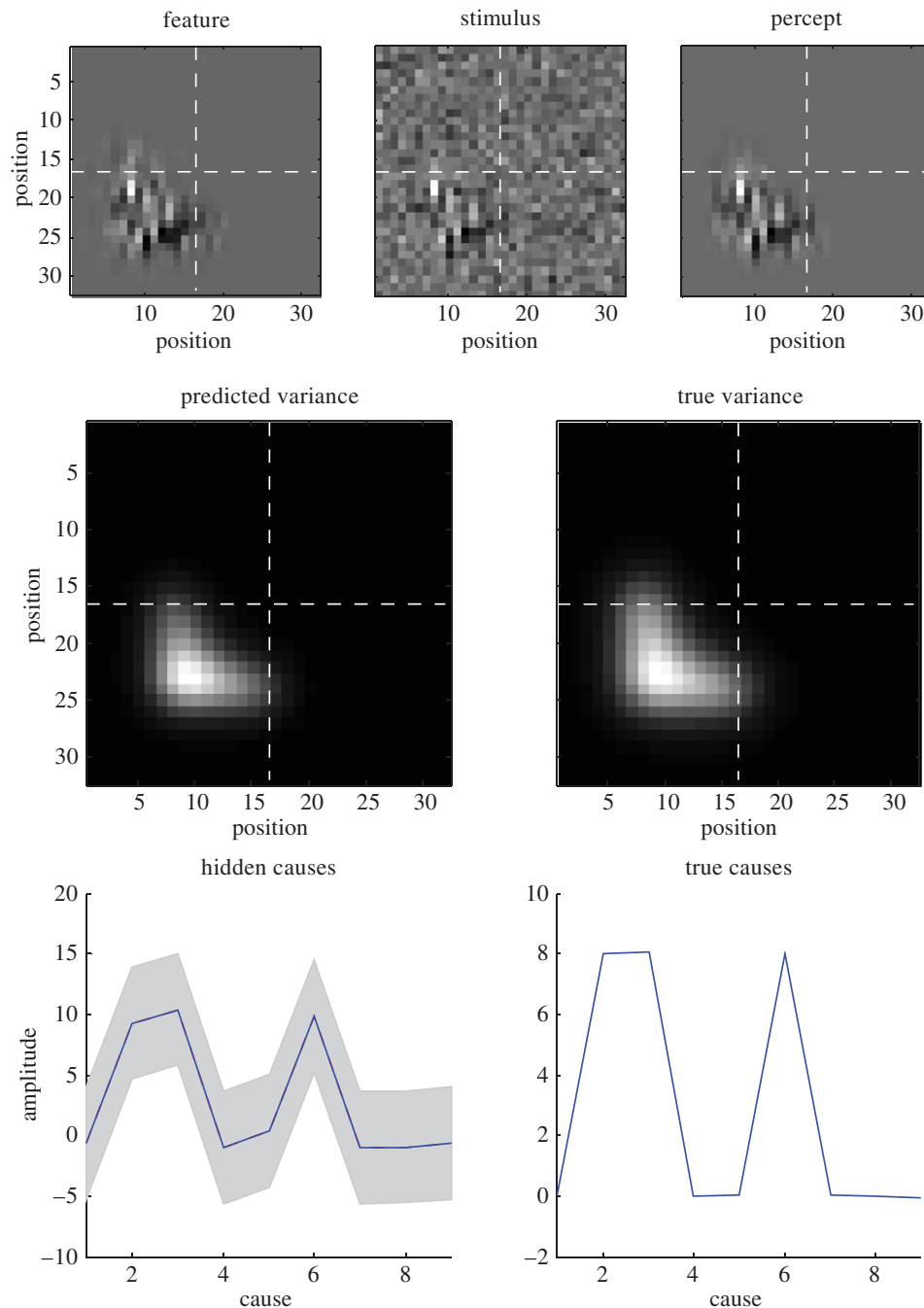
Figure 4. Results of generalized predictive coding. The key thing to take from these results is that the object has been segregated from the background, despite the high levels of sensory noise: predicted (solid blue line) and true (dotted blue line) sensory features are shown in the upper left panel with the associated prediction error (dotted red lines). Furthermore, the posterior confidence intervals (grey areas in the middle left panel) reveal the nature of precision-based predictive coding: note that the posterior confidence is reduced over the object or feature (whose location is shown on the middle right panel (red lines)—in terms of the expected standard deviation. This reduced confidence affords greater latitude for ascending sensory (prediction error) signals to influence posterior expectations. In contrast, over the radial basis functions in which signal was absent, the precision has increased, rendering these expectations insensitive to ascending prediction errors. This effectively means this (background) information is being ignored (or unattended).

focus on the basic computational anatomy implied by these schemes.

(b) Summary

These simulations demonstrate the nature of predictive coding, when sensory features generated by visual objects are textural or second-order in nature. They highlight the selective boosting or sensitivity to signals arising in the vicinity of a visual object or figure, suggesting a top-down augmentation of featural cues

that is consistent with attentional selection. The abovementioned examples illustrate a form of precision or gain control that is intrinsic to the cortical hierarchy and speaks to separate descending streams of prediction—that predict the first and second-order attributes of lower-level representations. This scheme suggests that modulatory or gain control mechanisms are restricted to backward connections (the forward connections convey prediction errors, which are linear and driving). Note also that the descending predictions are effectively inhibitory, which is consistent with the targeting of

Figure 5. Generalized predictive coding in two dimensions. This shows the same sort of results as figure 4; however, here, the input has been formulated in two dimensions (over 32 × 32 visual channels). The upper row shows the true signal feature (left panel), the visual input after sensory noise has been added (middle panel) and the prediction based upon posterior expectations (right panel). It can be seen that the background noise has been suppressed by rendering expectations of local luminance very precisely. Conversely, the regime of signal enjoys more relaxed constraints and is effectively attended. This second-order effect is summarized in the middle row that shows the estimated (left panel) and true (right panel) profile of single variance (inverse precision), which—in this example—is a small L-shaped feature in the lower left quadrant. The associated (second level) hidden causes are shown in the lower row in terms of their posterior expectations and the true values (as in the previous figure). (Online version in colour.)

inhibitory interneurons (particularly in layer one) by backward connections in the visual cortex [17]. In contrast, the predicted precision has an excitatory modulatory effect, consistent with mediation through voltage-dependent NMDA receptors in pyramidal cells of the superficial layers [19]. Later, we will also consider the important role of fast-spiking inhibitory interneurons and chandelier cells (that express NMDA receptors) in mediating synchronous gain.

Later, we consider general forms of descending precision control that have been associated with attentional processing. In this instance, the source of top-down gain control is not

necessarily an intrinsic part of the cortical hierarchy but may call upon the cortical systems that control precision throughout the cortical hierarchy. So, what are the cardinal features a neuronal system should possess to mediate such precision control? Precision is a function of hidden causes, which means that expected precision depends on expected hidden causes that—like all expectations—we presume are encoded neuronally. A universal feature of predictive coding is that connections to populations encoding expectations are from populations encoding prediction errors, and these connections are reciprocated. In the special case of hidden causes of precision, these projections

must show substantial (but possibly topographic) convergence and divergence: it can be seen from equation (3.4) (third equality) that the expected causes of precision gather information from each component or set of prediction errors that share the same covariance or precision. If the sum of (precision-weighted) squared prediction error is too large then expected precision falls and, conversely, rises when the sum of (precision-weighted) squared prediction error is too small. Furthermore, every prediction error unit contributing to the sum of squares receives reciprocal connections to modulate its gain or precision. Neuroanatomically, this suggests systems that encode and mediate expected precision must

— receive convergent projections from large (possibly topographically organized) regions of cortex, specifically from cells encoding prediction error (in supragranular layers);
— reciprocate divergent projections to the same regions;
— mediate some form of gain control over the cells encoding prediction error; and
— possess bilateral projections to cortical areas with cortico-cortical connections, to control the relative precision of their respective prediction errors.

In what follows, we consider corticothalamic systems—in which the thalamus (pulvinar) provides modulatory gain control—and what that implies for the cartography of attention and arousal.

4. Precision, attention and the pulvinar

There are two primary types of relay neurons in the thalamus, namely *core* cells and *matrix* cells [36]. Matrix cells are distributed widely over the nuclei of the dorsal thalamus and send thalamocortical axons that terminate principally in the superficial layers of the cortex [36,37]. Furthermore, thalamocortical signalling is primarily excitatory, but is largely mediated by inhibitory mechanisms that implement synchronous gain. Thalamocortical oscillations also provide modulatory inputs to the thalamus via GABAergic neurons that synapse in the reticular nucleus of the thalamus [38]. It seems natural therefore to consider (a subset of) the corticothalamic system as a candidate for precision control. In what follows, we review the evidence for such a role in the pulvinar—drawing on known neuroanatomy, neurophysiology and recent findings in cognitive neuroscience.

The pulvinar is the largest nucleus in the primate thalamus and has expanded in size during primate evolution—in parallel with other visual structures [39]. The pulvinar has long been thought to play a role in mediating visual attention [40,41] perhaps by registering the saliency of a visual scene [42,43]. Damage to the pulvinar in humans can result in visual hemineglect [44,45], deficits in feature binding [46] or focal attention [47]. Neurons in the pulvinar exhibit features of selective attention; as they respond more strongly to behaviourally relevant stimuli than to unattended stimuli [48], as such, they contribute to top-down suppression of distractors [49]. Human imaging studies report pulvinar activation that is consistent with the filtering of distractors [50–52], such that only information pertaining to the attended target can be decoded from patterns of activity [53].

Neuroanatomical observations of the pulvinar suggest that for every direct connection between two cortical regions, there is a parallel, indirect pathway that goes through the pulvinar. This is called the *replication principle* [41]. For example, consider the visual maps of ventral pulvinar [54,55] that receive retinotopically registered connections from the areas of the ventral visual pathway (V1, V2, V4, TEO and TE). These form a (diffuse) occipitotemporal gradient from V1 at one pole of the pulvinar map to area TE at the other. The relative overlap between the pulvinar fields of any given pair of areas roughly reflects their mutual level of cortical connectivity: for example, area V4 shares more pulvinar territory, and has stronger cortical connections with its neighbouring area TEO, than it does with the more distant area TE [41]. This neuroanatomical architecture of the cortico-pulvino-cortical pathway is therefore suited for concurrent precision estimation and to modulate the gain of reciprocal message passing between cortical hierarchical levels. In particular, the replication principle is entirely consistent with the control of the relative precision of prediction errors at different levels in the cortical hierarchy.

Based on these architectural properties—and the requirements of predictive coding—it seems reasonable to hypothesize that the functional role of the pulvinar is to optimize the gain of cortical prediction errors according to their expected precision. To fulfil this role, the pulvinar needs to encode expected precision and mediate gain modulation. Recent neurophysiological studies in behaving monkeys indicate that the pulvinar, indeed, performs these computational operations.

(a) Empirical evidence for precision engineering in the pulvinar

In terms of encoding precision, it has been recently reported that approximately 30% of neurons in the pulvinar are sensitive to the reliability of task-relevant sensory signals, representing the 'confidence' in perceptual decisions [56]. During a perceptual decision task, this subset of neurons does not selectively respond to the content of perception, but exhibits a higher firing rate when the monkeys behaved as though they were certain about their perceptual decision. When the monkeys were given a choice to opt-out, for a smaller reward, a lower firing rate of these neurons predicted the escape response of the monkeys—even when the signal-to-noise ratio in the stimulus was identical. The firing rate was lower for more difficult trials, and the deactivation of these neurons by GABA agonist (muscimol) injection enhanced escape responses—as though the monkeys lost confidence in their perceptual decision even though their objective task performance was unimpaired. These findings support the notion that neurons in the pulvinar encode expected precision or confidence in information used for perceptual decisions.

The pulvinar's contribution to gain control has been demonstrated in a compelling study of spike-field coherence [57]. By concurrently recording pulvinar spikes and local field potentials from V4 and TEO, the authors showed that the spike-field coherence between the pulvinar neurons and alpha oscillation in V4 and TEO was enhanced when attention was directed to the receptive field of the pulvinar neuron. Crucially, conditional Granger causality analysis across the three regions showed that the pulvinar neurons facilitated the transmission of information between V4 and TE by synchronizing the alpha oscillation in those cortical regions. This provides empirical evidence that the pulvinar serves as a gain control system for corticocortical interaction—via controlling neuronal synchronization. This synchronous gain control offers a neurobiological mechanism to adjust effective synaptic gain transiently across cortical regions [55,58]. Furthermore, it is

closely related to notions of communication through coherence (see below) and may reflect an important mechanism for precision engineering in attention [2].

These studies provide neurophysiological evidence that the pulvinar neurons encode expected precision, and modulate the gain of corticocortical communication. The notion of precision engineering in the pulvinar offers a coherent (computational) perspective on how seemingly disparate aspects of attention (gain modulation) and confidence (uncertainty estimation) are orchestrated. Although the concepts of salience, confidence and attention may appear distinct, their intimate relationship can be interpreted as an integral part of perceptual inference—reflecting the different faces of precision.

(b) Gain control mechanisms in the cortico-pulvino-cortical connectivity

There are multiple thalamocortical mechanisms that can modulate the gain of prediction error in superficial layers of the cortex. Here, we consider three possible mechanisms through which the pulvinar contributes to gain control: (i) induction of phase synchrony across presynaptic neurons within a cortical region, perhaps achieved by (ii) modulation of low-frequency (alpha) oscillation across cortical regions and (iii) modulation of prediction error units in the superficial layers via diffuse projections.

The first mechanism reflects the fact that synchronized presynaptic spikes generally make the postsynaptic impact stronger. Thus, controlling the degree of synchrony at presynaptic neurons can serve as a gain control mechanism [59,60]. This can be achieved by synchronous modulation of subthreshold membrane potentials at the gamma frequency [61]. While gamma oscillations can be generated by intracortical mechanisms, the thalamus plays a role in modulating gamma oscillations in sensory cortex. For example, it has been shown that stimulation of the posterior intralaminar nucleus modulates synchronous gamma oscillations in the auditory cortex [62]. Extending this notion to the visual cortex, one may speculate that the pulvinar could modulate subthreshold gamma oscillation in prediction error neurons in the cortex, thereby controlling the synchrony of spike timings of prediction error neurons.

The second mechanism is phase synchronization of distant cortical areas by the pulvinar [41], invoking the principle of 'communication through coherence'—that selective communication can be achieved through coherence between firing rate oscillation in the sending region and oscillatory gain modulation in the receiving region [63,64]. Corticothalamic connections play a prominent role in synchronizing oscillations [65], and the thalamus modulates phase relationships between cortical regions, thereby modulating the effective synaptic strengths. For example, Akam & Kullmann [66] demonstrate flexible signal routing in neural circuits, by exploiting sparsely synchronized network oscillations and temporal filtering by feed-forward inhibition.

The core cells in the pulvinar form a loop through layers 3 and 6 of extrastriate cortex [67]. This circuit could serve as an alpha generator for extrastriate cortex [68], through a mechanism analogous to the geniculocortical loop through layers 4 and 6 of V1 [69], and modulate the effectiveness of the output from one area to another distant cortical region. As discussed earlier, empirical evidence indicates that spikes from the pulvinar

generate alpha rhythms in the target cortical areas, and induce corticocortical synchrony in the alpha frequency that facilitates communication between the synchronized cortical regions [57]. Furthermore, there is evidence in this study [57] and others [70,71] for alpha–gamma cross-frequency coupling, thus forging a link between the two mechanisms considered so far.

The third mechanism considers gain control by projections from the pulvinar to the superficial layers 1–3 of a visual area (e.g. V1). This diffuse projection, originating from the matrix cells of the pulvinar, can modulate the activity of prediction error neurons in the target area—a functional analogy with the superficial component of backward cortical connections that we have previously attributed with a role in precision control [19]. Gain control via this pathway has been empirically demonstrated. Inactivation of the lateral pulvinar suppressed responses of superficial V1 neurons to visual input, whereas excitation of the pulvinar neurons increased the responsiveness of neurons in the superficial layers with overlapping receptive fields [72]. Given the organization of cortico-pulvino-cortical connections (noted above), we predict that the pulvinar neurons sending feedback to superficial layers of any given cortical area receive information about expected precision via the descending input from areas at both higher and lower levels in the cortical hierarchy. This may differ from the cortico-cortical transmission of precision that we have so far linked exclusively to backward connections [19].

These candidate mechanisms suggest that the pulvinar has multiple ways to control the gain in corticocortical communications. These neuronal implementations of gain control mechanisms are by no means comprehensive, and are not mutually exclusive. While all of the mechanisms discussed here have some empirical support, which mechanism plays a dominant role in the context of the predictive coding framework remains to be determined. Nevertheless, these examples collectively point to the pulvinar's role in gain control in corticocortical communication.

(c) Precision estimation in the corticothalamic network

The anatomy and laminar specificity of pulvinar projections to the cortex fits comfortably with the computational architecture implied by predictive coding (figure 6). We have outlined different roles for the core and matrix output neurons (in alpha generation for the core cells, projecting to the middle layers and in precision regulation for the matrix cells projecting superficially). The dual afferent projections to the pulvinar from the cortex derive from layers 5 and 6, thought to act as drivers and modulators, respectively [73,74]. It is these connections that should convey the (squared) prediction error to enable the pulvinar to estimate precision. However, if we consider the geniculocortical loop with striate cortex as a model for alpha generation [69] it is the layer 6 outputs to thalamus that serve this role, pointing to layer 5, perhaps, as the source of squared prediction error. Note that striate cortex output to pulvinar is not duplex, but arises exclusively from layer 5 [75]. All current analyses of canonical microcircuits place prediction error units in superficial layers [17–19]. Thus, we suppose that the principal cells reporting the squared prediction error (i.e. second-order forward connections) to the pulvinar are a secondary stream originating through the strong intrinsic connections from the superficial layers to layer 5. The particular arrangement that we arrive at—cortical drivers driving thalamic matrix cells, and cortical modulators modulating thalamic core cells—
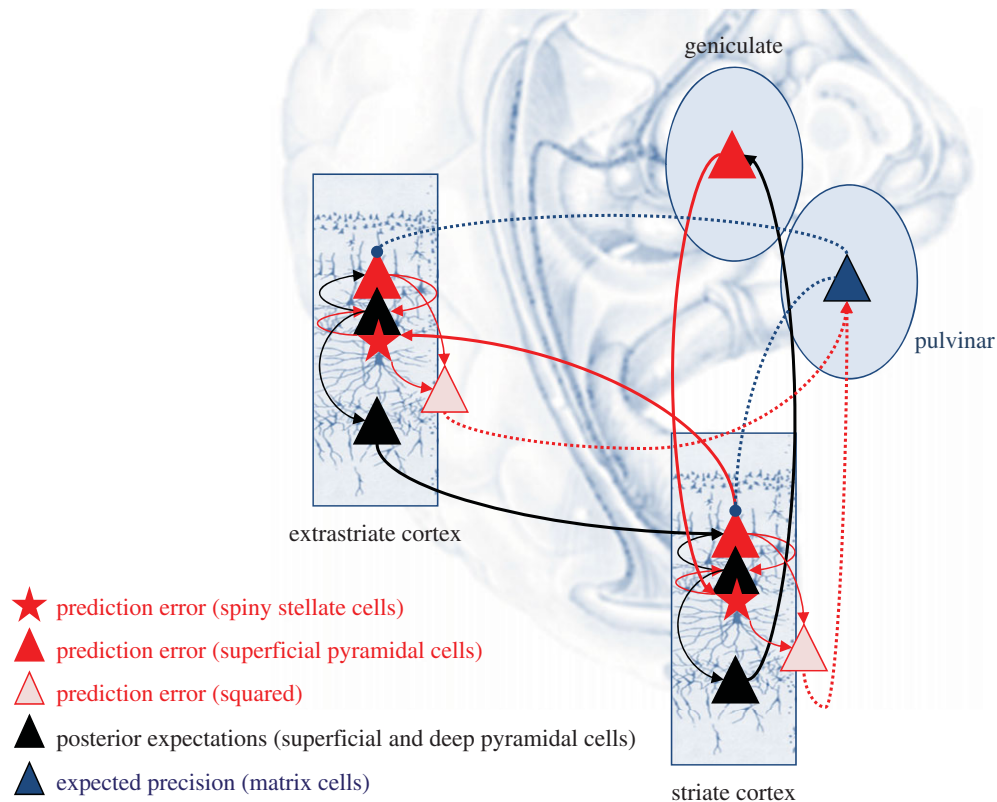
Figure 6. This schematic is a more detailed version of figure 2 that includes putative laminar-specific connections that are consistent with the precision-based predictive coding scheme in the main text. This architecture is based upon Bastos *et al.* [17] and Shipp *et al.* [19] and conforms roughly to the known neuroanatomy and physiology of canonical microcircuits and laminar specificity of extrinsic connections. The key aspect of this figure is the inclusion of deep pyramidal cells encoding the amplitude of prediction error (squared) that inform posterior expectations about precision in the (matrix cells) of the pulvinar. These cells reciprocate descending projections to modulate the gain of superficial pyramidal cells in cortex. Forward connections are in red and descending (backward) connections are in black. First-order streams are shown as full lines and second-order (precision-related) streams are shown as dashed lines.

matches the anatomy reported for connections between frontal cortex and the ventral anterior thalamic nucleus [76], but the specificity of contacts existing within pulvinar is unknown.

The message passing implied by predictive coding would require these layer 5 principal cells to respond, in a U-shaped fashion, to both high and low levels of prediction error firing in superficial layers; in other words, be both excited and disinhibited by (first-order) prediction errors. The notion of a (second-order) forward-type corticopulvinar stream fits comfortably with the notion that input from layer 5 is largely feed-forward and the hypothesis that layer 5 corticothalamic axons represent the afferent limb of a corticothalamocortical pathway.

(d) Summary

Here, the key requirements of a neuronal system that could coordinate precise corticocortical message passing among functioning segregated areas appears to be fulfilled by corticothalamic loops. A detailed consideration of the pulvinar, in relation to the computational anatomy of predictive coding, reveals a consistent picture at the architectural and microcircuit level—particularly with regard to the laminar specificity of intrinsic and extrinsic connections (and indeed suggests some new hypotheses about subpopulations and their connections). Furthermore, the emerging picture ties together a number of closely related themes, namely the distinction between driving and modulatory connections, cortical gain control, synchronous gain, communication through coherence and desynchronization of alpha rhythms. All of these physiological phenomena have been implicated

in attentional processing and the encoding of salience or confidence, which we associate with precision control.

5. Conclusion

In this paper, we have considered how inferences about first-order content and second-order context are orchestrated in hierarchical predictive coding, highlighting the importance of modulatory effects by second-order representations—such as precision and saliency—in optimizing inference. We have considered the neurobiological substrates of precision engineering in the brain, with a special focus on the pulvinar and attention. In this proposal, inference about the (first-order) content of perception was ascribed to corticocortical message passing, whereas parallel corticothalamocortical connections contextualize (second-order) corticocortical processing via precision-weighted gain control of ascending prediction errors. This proposal offers a formal understanding of attentional functions and the encoding of expected precision by the pulvinar.

More generally, the notion of hierarchical inference in the brain provides a potentially important link between structure and function: if the brain transcribes causal structure from the world, then this (hierarchical) structure should be embodied in cortical architectures. Predictive coding provides a particular process theory for this transcription and calls for an understanding of microscopic (laminar-specific) message passing in canonical microcircuits—that is consistent with macroscopic cartography defined by extrinsic connections. The particular contribution of this paper is to highlight the context-sensitive and dynamic aspects of functional anatomy—distinguishing

between the neuronal processing of (first-order) content and (second-order) context. The implications for the future of cerebral cartography are manifest at a number of levels, namely a fuller understanding of the asymmetries between forward and backward connections—and the distinction between streams responsible for perceptual synthesis *per se* and those (presumably more diffuse) streams that contextualize perceptual processing, enabling the selection and coordination of precise information. The formal constraints offered by schemes like predictive coding highlight the need to characterize cortical interactions at the level of cortical layers and the orchestration of cerebral processing through centrifugal exchanges with subcortical structures.

Appendix A

This brief description of generalized predictive coding is based on Feldman & Friston [2]. A more technical description can be found in Friston *et al.* [77]. This scheme is based on three assumptions

— The brain minimizes a free energy of sensory inputs defined by a generative model.
— The generative model used by the brain is hierarchical, nonlinear and dynamic.
— 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 that 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.

$$\dot{\tilde{\mu}}(t) = D\tilde{\mu}(t) - \partial_{\tilde{\mu}} F(\tilde{s}, \tilde{\mu}). \tag{A 1}$$

These differential equations say that neuronal activity encoding posterior expectations about (generalized) hidden states of the world $\tilde{\mu} = (\mu, \mu', \mu'', \dots)$ reduce free energy—where free energy $F(\tilde{s}, \tilde{\mu})$ is a function of sensory inputs $\tilde{s} = (s, s', s'', \dots)$ 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 $D\tilde{\mu} = (\mu', \mu'', \mu''', \dots)$. The second (correction) term is usually expressed as a mixture of prediction errors that ensures 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 (A 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:

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

Equation (A 2) 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 ith 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 (3.2)) we can now write down the differential equations (equation (A 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):

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

$$\tag{A 3}$$

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.

In neural network terms, equation (A 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 ith 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 towards 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 [16,17].

Note that the precisions depend on the expected hidden causes and states. We have proposed that this dependency mediates attention [2]. Equation (A 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.

References

1. Hsu M, Bhatt M, Adolphs R, Tranel D, Camerer C. 2005 Neural systems responding to degrees of uncertainty in human decision-making. *Science* **310**, 1680–1683. (doi:10.1126/science.1115327)

2. Feldman H, Friston KJ. 2010 Attention, uncertainty, and free-energy. *Front. Hum. Neurosci.* **4**, 215. (doi:10.3389/fnhum.2010.00215)

3. Clark A. 2013 The many faces of precision. *Front. Psychol.* **4**, 270. (doi:10.3389/fpsyg.2013.00270)

4. Womelsdorf T, Fries P. 2006 Neuronal coherence during selective attentional processing and sensory-motor integration. *J. Physiol.* (*Paris*) **100**, 182–193. (doi:10.1016/j.jphysparis.2007.01.005)

5. Fries P, Womelsdorf T, Oostenveld R, Desimone R. 2008 The effects of visual stimulation and selective visual attention on rhythmic neuronal synchronization in macaque area V4. *J. Neurosci.* **28**, 4823–4835. (doi:10.1523/JNEUROSCI.4499-07.2008)

6. Gregory RL. 1968 Perceptual illusions and brain models. *Proc. R. Soc. Lond. B* **171**, 179–196. (doi:10.1098/rspb.1968.0071)

7. Helmholtz H. 1866 Concerning the perceptions in general. In *Treatise on physiological optics*, 3rd edn (Transl. by JPC Southall, 1962). New York, NY: Dover.

8. Dayan P, Hinton GE, Neal R. 1995 The Helmholtz machine. *Neural Comput.* **7**, 889–904. (doi:10.1162/neco.1995.7.5.889)

9. Lee TS, Mumford D. 2003 Hierarchical Bayesian inference in the visual cortex. *J. Opt. Soc. Am. Opt. Image Sci. Vis.* **20**, 1434–1448. (doi:10.1364/JOSAA.20.001434)

10. Friston K, Kilner J, Harrison L. 2006 A free energy principle for the brain. *J. Physiol.* (*Paris*) **100**, 70–87. (doi:10.1016/j.jphysparis.2006.10.001)

11. Hohwy J. 2013 *The predictive mind*. Oxford, UK: Oxford University Press.

12. Clark A. 2013 Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.* **36**, 181–204. (doi:10.1017/S0140525X12000477)

13. Srinivasan MV, Laughlin SB, Dubs A. 1982 Predictive coding: a fresh view of inhibition in the retina. *Proc. R. Soc. Lond. B* **216**, 427–459. (doi:10.1098/rspb.1982.0085)

14. Rao RP, Ballard DH. 1999 Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* **2**, 79–87. (doi:10.1038/4580)

15. Friston K. 2008 Hierarchical models in the brain. *PLoS Comput. Biol.* **4**, e1000211. (doi:10.1371/journal.pcbi.1000211)

16. Mumford D. 1992 On the computational architecture of the neocortex. II. *Biol. Cybern.* **66**, 241–251. (doi:10.1007/BF00198477)

17. Bastos AM, Usrey WM, Adams RA, Mangun GR, Fries P, Friston KJ. 2012 Canonical microcircuits for predictive coding. *Neuron* **76**, 695–711. (doi:10.1016/j.neuron.2012.10.038)

18. Adams RA, Shipp S, Friston KJ. 2012 Predictions not commands: active inference in the motor system. *Brain Struct. Funct.* **218**, 611–643. (doi:10.1007/S00429-012-0475-5)

19. Shipp S, Adams RA, Friston KJ. 2013 Reflections on agranular architecture: predictive coding in the motor cortex. *Trends Neurosci.* **36**, 706–716. (doi:10.1016/j.tins.2013.09.004)

20. Bosman C *et al.* 2012 Attentional stimulus selection through selective synchronization between monkey visual areas. *Neuron* **75**, 875–888. (doi:10.1016/j.neuron.2012.06.037)

21. Buffalo EA, Fries P, Landman R, Buschman TJ, Desimone R. 2011 Laminar differences in gamma and alpha coherence in the ventral stream. *Proc. Natl Acad. Sci. USA* **108**, 11 262–11 267. (doi:10.1073/pnas.1011284108)

22. van Kerkoerle T, Self MW, Dagnino B, Gariel-Mathis MA, Poort J, van der Togt C, Roelfsema PR. 2014 Alpha and gamma oscillations characterize feedback and feedforward processing in monkey visual cortex. *Proc. Natl Acad. Sci. USA* **111**, 14 332–14 341. (doi:10.1073/pnas.1402773111)

23. Zeki S, Shipp S. 1988 The functional logic of cortical connections. *Nature* **335**, 311–317. (doi:10.1038/335311a0)

24. Moran RJ, Campo P, Symmonds M, Stephan KE, Dolan RJ, Friston KJ. 2013 Free energy, precision and learning: the role of cholinergic neuromodulation. *J. Neurosci.* **33**, 8227–8236. (doi:10.1523/JNEUROSCI.4255-12.2013)

25. Goldman-Rakic PS, Lidow MS, Smiley JF, Williams MS. 1992 The anatomy of dopamine in monkey and human prefrontal cortex. *J. Neural Transm.* **36**, 163–177.

26. Braver TS, Barch DM, Cohen JD. 1999 Cognition and control in schizophrenia: a computational model of dopamine and prefrontal function. *Biol. Psychiatry* **46**, 312–328. (doi:10.1016/S0006-3223(99)00116-X)

27. Doya K. 2008 Modulators of decision making. *Nat. Neurosci.* **11**, 410–416. (doi:10.1038/nn2077)

28. Lidow MS, Goldman-Rakic PS, Gallager DW, Rakic P. 1991 Distribution of dopaminergic receptors in the primate cerebral cortex: quantitative autoradiographic analysis using [3H]raclopride, [3H]spiperone and [3H]SCH23390. *Neuroscience* **40**, 657–671. (doi:10.1016/0306-4522(91)90003-7)

29. Humphries MD, Wood R, Gurney K. 2009 Dopamine-modulated dynamic cell assemblies generated by the GABAergic striatal microcircuit. *Neural Netw.* **22**, 1174–1188. (doi:10.1016/j.neunet.2009.07.018)

30. Chawla D, Lumer ED, Friston KJ. 1999 The relationship between synchronization among neuronal populations and their mean activity levels. *Neural Comput.* **11**, 1389–1411. (doi:10.1162/089976699300016287)

31. Jiang J, Summerfield C, Egner T. 2013 Attention sharpens the distinction between expected and unexpected percepts in the visual brain. *J. Neurosci.* **33**, 18 438–18 447. (doi:10.1523/JNEUROSCI.3308-13.2013)

32. Frank MJ, Scheres A, Sherman SJ. 2007 Understanding decision-making deficits in

neurological conditions: insights from models of natural action selection. *Phil. Trans. R. Soc. B* **362**, 1641–1654. (doi:10.1098/rstb.2007.2058)

33. Cisek P. 2007 Cortical mechanisms of action selection: the affordance competition hypothesis. *Phil. Trans. R. Soc. B* **362**, 1585–1599. (doi:10.1098/rstb.2007.2054)

34. Friston KJ *et al*. 2012 Dopamine, affordance and active inference. *PLoS Comput. Biol.* **8**, e1002327. (doi:10.1371/journal.pcbi.1002327)

35. Adams RA, Stephan KE, Brown HR, Frith CD, Friston KJ. 2013 The computational anatomy of psychosis. *Front. Psychiatry* **4**, 47. (doi:10.3389/fpsyt.2013.00047)

36. Jones E. 2002 Thalamic circuitry and thalamocortical synchrony. *Phil. Trans. R. Soc. Lond. B* **357**, 1659–1673. (doi:10.1098/rstb.2002.1168)

37. Jones E. 1998 Viewpoint: the core and matrix of thalamic organization. *Neuroscience* **85**, 331–345. (doi:10.1016/S0306-4522(97)00581-2)

38. Gibbs J, Zhang Y, Shumate M, Coulter D. 1999 Regionally selective blockade of GABAergic inhibition by zinc in the thalamocortical system: functional significance. *J. Neurophysiol.* **83**, 1510–1521.

39. Chalfin B, Cheung D, Muniz J, de Lima Silveira L, Finlay B. 2007 Scaling of neuron number and volume of the pulvinar complex in New World primates: comparisons with humans, other primates, and mammals. *J. Comp. Neurol.* **504**, 265–274. (doi:10.1002/cne.21406)

40. Chalupa L. 1977 A review of cat and monkey studies implicating the pulvinar in visual function. *Behav. Biol.* **20**, 149–167. (doi:10.1016/S0091-6773(77)90721-0)

41. Shipp S. 2004 The brain circuitry of attention. *Trends Cogn. Sci.* **8**, 223–230. (doi:10.1016/j.tics.2004.03.004)

42. Ungerleider L, Christensen C. 1979 Pulvinar lesions in monkeys produce abnormal scanning of a complex visual array. *Neuropsychologia* **17**, 493–501. (doi:10.1016/0028-3932(79)90056-3)

43. Robinson D, Petersen S. 1992 The pulvinar and visual salience. *Trends Neurosci.* **15**, 127–132. (doi:10.1016/0166-2236(92)90354-B)

44. Zihl J, von Cramon D. 1979 The contribution of the 'second' visual system to directed visual attention in man. *Brain* **102**, 835–856. (doi:10.1093/brain/102.4.835)

45. Karnath H, Himmelbach M, Rorden C. 2002 The subcortical anatomy of human spatial neglect: putamen, caudate nucleus and pulvinar. *Brain* **125**, 350–360. (doi:10.1093/brain/awf032)

46. Ward R, Danziger S, Owen V, Rafal R. 2002 Deficits in spatial coding and feature binding following damage to spatiotopic maps in the human pulvinar. *Nat. Neurosci.* **5**, 99–100. (doi:10.1038/nn794)

47. Snow J, Allen H, Rafal R, Humphreys G. 2009 Impaired attentional selection following lesions to human pulvinar: evidence for homology between human and monkey. *Proc. Natl Acad. Sci. USA* **106**, 4054–4059. (doi:10.1073/pnas.0810086106)

48. Benevento L, Port J. 1995 Single neurons with both form/color differential responses and saccade-related responses in the nonretinotopic pulvinar of the behaving macaque monkey. *Vis. Neurosci.* **12**, 523–544. (doi:10.1017/S0952523800008439)

49. Desimone R, Wessinger M, Thomas L, Schneider W. 1990 Attentional control of visual perception: cortical and subcortical mechanisms. *Cold Spring Harb. Symp. Quant. Biol.* **55**, 963–971. (doi:10.1101/SQB.1990.055.01.090)

50. LaBerge D, Buchsbaum M. 1990 Positron emission tomography of pulvinar activity during an attention task. *J. Neurosci.* **10**, 613–619.

51. Kastner S, O'Connor D, Fukui M, Fehd H, Herwig U, Pinsk M. 2004 Functional imaging of the human lateral geniculate nucleus and pulvinar. *J. Neurophysiol.* **91**, 438–448. (doi:10.1152/jn.00553.2003)

52. Strumpf H, Mangun G, Boehler C, Stoppel C, Schoenfeld M, Heinze H, Hopf J-M. 2012 The role of the pulvinar in distractor processing and visual search. *Hum. Brain Mapp.* **34**, 1115–1132. (doi:10.1002/hbm.21496)

53. Fischer J, Whitney D. 2012 Attention gates visual coding in the human pulvinar. *Nat. Commun.* **3**, 1051. (doi:10.1038/ncomms2054)

54. Bender D. 1981 Retinotopic organization of macaque pulvinar. *J. Neurophysiol.* **46**, 672–693.

55. Crick F. 1984 Function of the thalamic reticular complex: the search light hypothesis. *Proc. Natl Acad. Sci. USA* **81**, 4586–4590. (doi:10.1073/pnas.81.14.4586)

56. Komura Y, Nikkuni A, Hirashima N, Uetake T, Miyamoto A. 2013 Responses of pulvinar neurons reflect a subject's confidence in visual categorization. *Nat. Neurosci.* **16**, 749–755. (doi:10.1038/nn.3393)

57. Saalmann Y, Pinsk M, Wang L, Li X, Kastner S. 2012 The pulvinar regulates information transmission between cortical areas based on attention demands. *Science* **337**, 753–756. (doi:10.1126/science.1223082)

58. Saalmann Y, Kastner S. 2009 Gain control in the visual thalamus during perception and cognition. *Curr. Opin. Neurobiol.* **19**, 408–414. (doi:10.1016/j.conb.2009.05.007)

59. Sejnowski T, Paulsen O. 2006 Network oscillations: emerging computational principles. *J. Neurosci.* **26**, 1673–1676. (doi:10.1523/JNEUROSCI.3737-05d.2006)

60. Womelsdorf T, Schoffelen J, Oostenveld R, Singer W, Desimone R, Engel A, Fries P. 2007 Modulation of neuronal interactions through neuronal synchronization. *Science* **316**, 1609–1612. (doi:10.1126/science.1139597)

61. Fries P, Reynolds J, Rorie A, Desimone R. 2001 Modulation of oscillatory neuronal synchronization by selective visual attention. *Science* **291**, 1560–1563. (doi:10.1126/science.1055465)

62. Barth D, MacDonald K. 1996 Thalamic modulation of high-frequency oscillating potentials in the auditory cortex. *Nature* **383**, 78–81. (doi:10.1038/383078a0)

63. Fries P. 2005 A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends. Cogn. Sci.* **9**, 476–480. (doi:10.1016/j.tics.2005.08.011)

64. Akam T, Kullmann D. 2012 Efficient 'communication through coherence' requires oscillations structured to minimize interference between signals. *PLoS Comput. Biol.* **8**, e1002760. (doi:10.1371/journal.pcbi.1002760)

65. Contreras D, Destexhe A, Sejnowski M, Steriade M. 1996 Control of spatiotemporal coherence of a thalamic oscillation by corticothalamic feedback. *Science* **274**, 771–773. (doi:10.1126/science.274.5288.771)

66. Akam T, Kullmann DM. 2010 Oscillations and filtering networks support flexible routing of information. *Neuron* **67**, 308–320. (doi:10.1016/j.neuron.2010.06.019)

67. Shipp S. 2007 Structure and function of the cerebral cortex. *Curr. Biol.* **17**, 443–449. (doi:10.1016/j.cub.2007.03.044)

68. Liu Z, de Zwart J, Yao B, van Gelderen P, Kuo L, Duyn J. 2012 Finding thalamic BOLD correlates to posterior alpha EEG. *Neuroimage* **63**, 1060–1069. (doi:10.1016/j.neuroimage.2012.08.025)

69. Bollimunta A, Mo J, Schroeder C, Ding M. 2011 Neuronal mechanisms and attentional modulation of corticothalamic alpha oscillations. *J. Neurosci.* **31**, 4935–4943. (doi:10.1523/JNEUROSCI.5580-10.2011)

70. Spaak E, Bonnefond M, Maier A, Leopold D, Jensen O. 2012 Layer-specific entrainment of gamma-band neural activity by the alpha rhythm in monkey visual cortex. *Curr. Biol.* **22**, 2313–2318. (doi:10.1016/j.cub.2012.10.020)

71. Roux F, Wibral M, Singer W, Aru J, Uhlhaas P. 2013 The phase of thalamic alpha activity modulates cortical gamma-band activity: evidence from resting-state MEG recordings. *J. Neurosci.* **33**, 17 827–17 835. (doi:10.1523/JNEUROSCI.5778-12.2013)

72. Purushothaman G, Marion R, Li K, Casagrande V. 2012 Gating and control of primary visual cortex by pulvinar. *Nat. Neurosci.* **15**, 905–912. (doi:10.1038/nn.3106)

73. Sherman SM, Guillery RW. 1998 On the actions that one nerve cell can have on another: distinguishing 'drivers' from 'modulators'. *Proc. Natl Acad. Sci. USA* **95**, 7121–7126. (doi:10.1073/pnas.95.12.7121)

74. Sherman S, Guillery R. 2011 Distinct functions for direct and transthalamic corticocortical connection. *J. Neurophysiol.* **106**, 1068–1077. (doi:10.1152/jn.00429.2011)

75. Trojanowski J, Jacobson S. 1977 The morphology and laminar distribution of cortico-pulvinar neurons in the rhesus monkey. *Exp. Brain Res.* **28**, 51–62.

76. Zikopoulos B, Barbas H. 2007 Parallel driving and modulatory pathways link the prefrontal cortex and thalamus. *PLoS ONE* **2**, e848. (doi:10.1371/journal.pone.0000848)

77. Friston K, Stephan K, Li B, Daunizeau J. 2010 Generalised filtering. *Math. Problems Eng.* **2010**, 621670. (doi:10.1155/2010/621670)