PERSPECTIVES

Interoceptive predictions in the brain

Lisa Feldman Barrett and W. Kyle Simmons

Abstract | Intuition suggests that perception follows sensation and therefore bodily feelings originate in the body. However, recent evidence goes against this logic: interoceptive experience may largely reflect limbic predictions about the expected state of the body that are constrained by ascending visceral sensations. In this Opinion article, we introduce the Embodied Predictive Interoception Coding model, which integrates an anatomical model of corticocortical connections with Bayesian active inference principles, to propose that agranular visceromotor cortices contribute to interoception by issuing interoceptive predictions. We then discuss how disruptions in interoceptive predictions could function as a common vulnerability for mental and physical illness.

For many decades, neuroscientists understood the brain as a 'stimulus–response' organ, consisting of individual neurons that lie dormant until stimulated. In this traditional model, learning and experience merely modulate neural activity that is driven by sensory events in the world. In recent years, scientists have come to realize that the brain probably does not work this way. Instead, research and theory are converging on the idea of the brain as an active inference generator that functions according to a Bayesian approach to probability: sensory inputs constrain estimates of prior probability (from past experience) to create the posterior probabilities that serve as beliefs about the causes of such inputs in the present[1,2].

According to this active inference account, the brain forms neural representations that are constructed from previous experience. These function as a generative model of how stimuli in the environment cause sensations. Rather than neurons simply lying dormant until information arrives via the external sensors of the body (that is, the eyes, ears and taste receptors, among others), the brain anticipates incoming sensory inputs, which it implements as predictions that cascade throughout the cortex. As predictions propagate across cortical regions (following their roughly centrifugal connections[3]), they modulate the firing of neurons within cortical columns in anticipation of these regions receiving actual sensory

input from the environment. In this way, predictions (with prior probabilities) function as hypotheses about the world that can be tested against sensory signals that arrive in the brain. The goal is to minimize the difference between the brain's prediction and incoming sensation (that is, the 'prediction error'). This can be achieved in any of three ways: first, by propagating the error back along cortical connections to modify the prediction; second, by moving the body to generate the predicted sensations; and third, by changing how the brain attends to or samples incoming sensory input. In this active inference framework, perception and action are tightly coupled, with both arising from the brain's hypotheses about the world and constrained by sensory inputs from the world. By this account, action drives perception to reduce prediction error.

At present, there is empirical evidence[4,5] that the visual and auditory processing systems operate according to the principles of active inference. However, from the perspective of the brain, the representation of the 'world' includes not only exteroceptive sensations from the external environment, but also interoceptive sensations from the internal milieu of the body (such as those relating to heart rate, glucose levels, build-up of carbon dioxide in the bloodstream, temperature, inflammation and so on). Although interoception is often studied in the context of emotion, it is a fundamental feature of the

human nervous system that has relevance for many biological, as well as psychological, phenomena[6–10], such as eating, craving and decision making.

In this Opinion article, we introduce the Embodied Predictive Interoception Coding (EPIC) model as an active inference account of interoception that is based on recent developments in the understanding of how predictions and prediction errors flow within the laminar architecture of corticocortical connections. To understand this flow, we use Barbas and colleagues' structural model of corticocortical connections[11,149]. Although other researchers have previously discussed the concept of interoceptive predictions[12–15], these accounts have focused primarily on particular brain structures, such as the anterior insula. Our integration of the structural model with the active inference account is highly consistent with early theoretical proposals for neurally implementing active inference schemes (such as that described by Mumford in 1991 (REF. 16)) that are now gaining increasing empirical support[17].

Our approach builds on existing discussions of interoceptive prediction in several distinct ways. First, the EPIC model incorporates a wider interoceptive system that has been verified in tracer studies of the macaque brain[18–26] (rather than focusing on individual brain regions). Second, the EPIC model implements active inference as the flow of prediction and prediction-error signals through the cortical lamination gradients within this interoceptive system using the structural model of corticocortical connections[11,149]. Third, the EPIC model provides a basis for further speculations on the broader role of interoceptive predictions in dynamic coordination of brain activity. Last, we discuss the implications of the EPIC model for the neuroscience of mental and physical illness.

Active inference in the cortex

A brief summary of laminar organization within the human cortex is presented in FIG. 1. The figure depicts elements of both the intracortical laminar architecture and the corticocortical connections that support the EPIC model. Barbas and colleagues' structural model describes how laminar organization
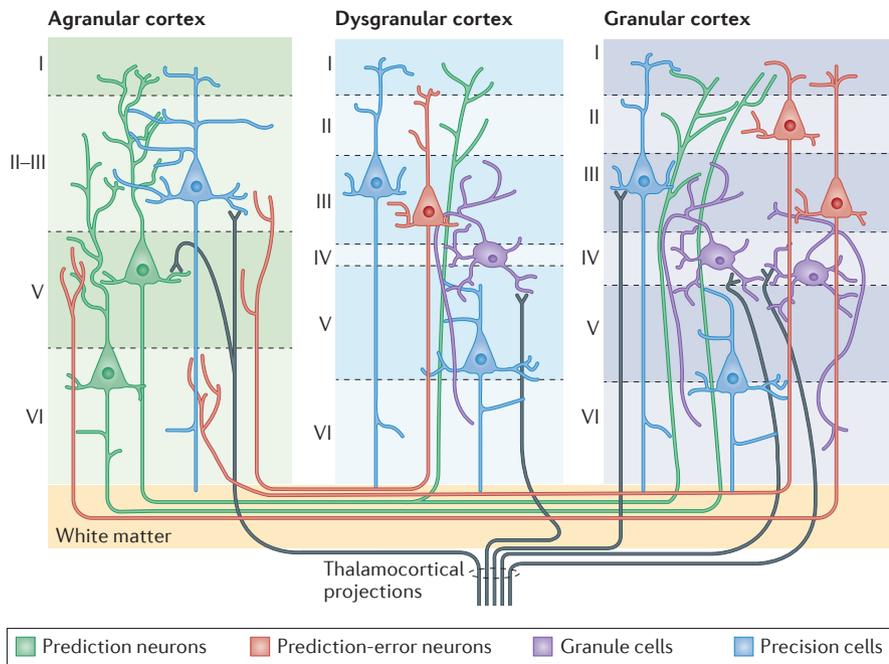
Figure 1 | **Intracortical architecture and intercortical connectivity for predictive coding.** Cortical columns are defined by different numbers of layers (also called laminae), with each layer having characteristic cell types and patterns of intracortical and intercortical connectivity. Granular cortex (right) is characterized by six differentiated laminae (layers I–VI), with layer IV containing granule cells, which are excitatory spiny stellate neurons (purple) that amplify and distribute thalamocortical inputs throughout the column. Granular cortex also contains many spiny pyramidal neurons throughout its infragranular and supragranular layers. Pyramidal neurons have: a triangular soma, from which basal dendrites project; an ascending apical dendrite, often with large dendritic tufts in layer I; and a single axon that descends and projects out of the cortical column (sometimes with multiple collaterals). By contrast, agranular cortex (left) does not have a fully expressed layer IV and has a poorly differentiated boundary between layer II and layer III. These upper laminae contain relatively fewer pyramidal neurons than does granular cortex. However, agranular cortex contains relatively greater numbers of large pyramidal neurons in layer V and layer VI than in its upper layers. Despite not having a defined layer IV with granule cells, agranular cortex still receives thalamic projections; however, the sensory information that enters agranular cortex is less amplified and less well redistributed throughout the column than in granular cortex. Dysgranular cortex is found in transition zones between granular and agranular regions and contains a small but defined layer IV and a distinctive (although rudimentary) layer II and layer III. This figure is not intended to be comprehensive but rather highlights laminar and cellular characteristics that are important for understanding the Embodied Predictive Interoception Coding (EPIC) model and its predictions. (For example, in the primate brain, cytoarchitecture varies from posterior to anterior, with posterior columns containing a greater number of cells in layer II and layer III than do anterior columns, and anterior columns containing relatively fewer neurons but a greater number of connections[147]; in addition, the cytoarchitectonic structure varies among different primate species, particularly in the density of connections within the anterior portions of the cortex[148].) According to the EPIC model, prediction neurons (depicted as green pyramidal neurons) in deep layers of agranular cortex drive active inference by sending sensory predictions via projections (green lines) to supragranular layers of dysgranular and granular sensory cortices. Prediction-error neurons (depicted as red pyramidal neurons) in the supragranular layers of granular cortex compute the difference between the predicted and received sensory signal, and send prediction-error signals via projections (red lines) back to the deep layers of agranular cortical regions. Precision cells (depicted as blue pyramidal neurons) tune the gain on predictions and prediction error dynamically, thereby giving these signals reduced (or, in some cases, greater) weight depending on the relative confidence in the descending predictions or the reliability of incoming sensory signals.

influences information flow between cortical regions[11,149], and this model also provides a plausible neural architecture for active inference throughout the brain (BOX 1; see Supplementary information S1 (box)). In our integration of the structural and active inference models, predictions originate in cortical columns with less laminar differentiation (such as in agranular cortex) and are propagated to areas with greater laminar differentiation (such as granular cortex). In the prototypical case, prediction signals originate in the deep layers (primarily layer V) of cortical regions in which there are many projection neurons, and terminate in the supragranular division of dysgranular and granular regions — principally on dendrites in layer I, as well as on neurons in layer II and layer III[11,27,28,149].

Predictions that arrive at a cortical column change the firing rates of neurons in layers I–III (but particularly in layer IIIb[28]) in anticipation of the thalamic input that arrives there; for example, in primary sensory regions, sensory input arrives — via core (class 1) thalamic projections — at layer IV[29], which contains neurons with dendrites that extend through layer III and layer V[30]. In this way, predictions from less-differentiated cortical columns modulate the ongoing pattern of activity in the more-differentiated cortical columns, thereby anticipating incoming sensory input to layer IV (as well as the lower portion of layer III and the upper portion of layer V). We speculate that this is the neural basis of what psychologists call 'perceptual inference' or 'simulation' (as reviewed by Barsalou[31]).

If the pattern of firing in a cortical column sufficiently anticipates the afferent thalamic input as it arrives at the column, then there will be little or no difference between the predicted signal and the sensory signal — that is, there will be little or no prediction error and thus no need to change the simulation. If some adjustment to firing among the upper and middle layers is required to accommodate the incoming thalamic input, the difference between the expected and received input is computed as a prediction error.

In the active inference framework, the goal is to minimize prediction error. In a normal functioning brain, prediction errors can be minimized in three ways[28,32]. First, the prediction error can be propagated back along the cortical connections to modify the prediction. Second, the incoming thalamic input can be modified by moving the body (via signals to the motor system) to actively change the incoming sensory stimulation. Third, the brain's intrinsic and widely distributed cognitive control networks[33–38] can change the focus of attention by biasing the influence of incoming sensory input. Physiologically, this third mechanism of error minimization corresponds to modulating the gain or excitability of neurons that represent different sensory features. Computationally, this represents the expected salience or precision (that is, reliability) of sensory signals[39].

When the active inference framework is integrated with the structural model[11,149], prediction-error signals are computed in cortical regions that have greater laminar

differentiation (such as granular cortices) and project to cortical regions with less laminar differentiation (such as agranular cortices). Prototypically, prediction-error signals originate in the superficial layers and project to the deep layers; this is consistent with prior suggestions (for example, by Shipp *et al.*[28]) that prediction-error signals are computed primarily in supragranular cortical layers. Some of the pyramidal neurons within a

cortical column, particularly within layer II and layer IIIa, function as precision units that dynamically modify the gain on neurons that compute prediction error. Precision units effectively reduce (or in some cases increase) the weight of incoming sensory input (and therefore prediction errors) on the basis of the relative confidence in the descending predictions or confidence in the reliability of incoming sensory signals[28].

Integrating the active inference approach[1] with the structural model of corticocortical connections[11,149] into one framework provides a flexible model for generating specific hypotheses about the transmission of predictions and prediction errors across the lamination gradients within the cortex. Importantly, granular cortices have a fully expressed layer IV that can amplify and redistribute thalamic input into the rest of their cortical columns. They also have many projection neurons in their upper layers that are optimized to compute and send prediction errors, and therefore modulate (but do not drive) perception and action. Conversely, agranular cortices, without a fully expressed layer IV and with fewer projection neurons in their upper layers but many in their deep layers, are optimized to drive (rather than modulate) perception and action. Although other authors have remarked that an active inference approach inverts the processing hierarchy of the brain[40], one of our unique contributions is incorporating the structural model into the EPIC model to consider the implications of this inversion for the role of agranular cortices and their interoceptive inferences within the brain.
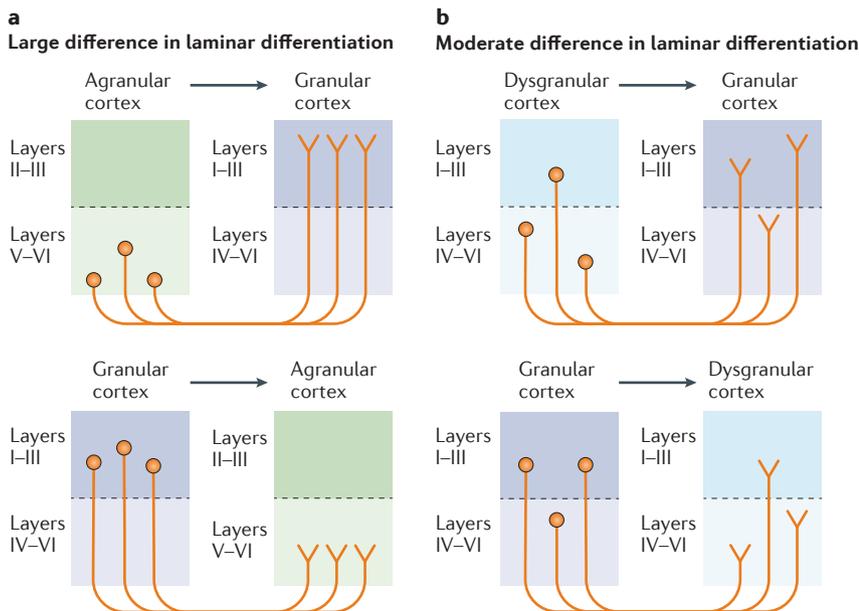
Active inference in primary motor cortex

Friston and colleagues[17,28,32] have recently proposed that proprioceptive and kinaesthetic perceptions of the body arise from predictions through the same computational principles of active inference as do visual and auditory perceptions of the external world, albeit with some notable differences. Following an analysis of the cytoarchitecture and function of the primary motor cortex (M1) by Shipp[41], Friston and colleagues suggest that M1 is agranular (that is, it lacks a fully expressed granular layer IV) and issues motor predictions to the spinal cord as deterministic commands to move the body. Simultaneously, M1 sends somatosensory predictions to stimulate neurons in the primary somatosensory cortex (S1) to represent the anticipated proprioceptive and kinaesthetic consequences of the movement. These somatosensory predictions are akin to 'efference copies' of corollary discharge that enable the brain to track self-initiated movements[42,43]. According to this view, M1 should be relatively 'immune' to prediction-error signals that ascend from the motor periphery; if M1 did receive such signals, it would infer that the intended movement had not been executed correctly[28]. Instead, according to Shipp *et al.*[28], rather than the ascending sensory input returning directly to M1, prediction error computed in the motor

---

Box 1 | Barbas and colleagues' structural model of corticocortical connections

In 1997, Barbas and Rempel-Clower[11] published a structural model of corticocortical connections by analysing projection patterns in macaque prefrontal and temporal cortices. Using anterograde and retrograde tracers, they showed that laminar structure is related to the distribution of projection neurons in cortical columns. Their resulting structural model has successfully predicted the flow of information in frontal, temporal, parietal and occipital cortices in experiments in macaques and cats[128–131] and outperforms other models of corticocortical connections[132]. For example, compared with the Felleman and van Essen model[133] — which only considers the differentiation in the termination site and suffers from many prediction violations[133] — the structural model considers the relative laminar differentiation in both the origin and termination of cortical connections. In addition, the structural model does not assume a strictly linear posterior–anterior lamination gradient across the brain to predict the flow of information, unlike Felleman and van Essen's model and Kennedy and colleagues' Distance Rule model[134]; this difference is important for our Embodied Predictive Interoception Coding (EPIC) model, as the regions that are most important for an active inference model of interoception do not show such a linear gradient[135] (see Supplementary information S1 (box)).

According to the structural model, projections from less-differentiated cortical regions (such as agranular cortex with undifferentiated layer II and layer III, and without a fully expressed layer IV) primarily originate in the deep layers (layer V and layer VI) and terminate in upper layers of regions with a more fully developed laminar organization (such as granular cortex; see upper half of part **a** in the figure). Conversely, more-differentiated cortical regions (for example, granular cortex with a fully expressed granular layer) have projections that originate mostly in the upper layers and that terminate predominantly in deep layers of regions with less-differentiated laminar architecture (for example, agranular cortex; see lower half of part **a** in the figure). In the case of lateral corticocortical connections for which there is little or no difference in laminar differentiation between sites of origin and termination (see part **b** in the figure), axons can originate and terminate in any layer, although they rarely originate in layer IV because the overwhelming majority of layer IV neurons are local circuit interneurons that do not project outside the column. Box figure is adapted from Barbas, H. & Rempel-Clower, N., Cortical structure predicts the pattern of corticocortical connections, *Cereb. Cortex*, 1997, **7**, 7, 635–646, by permission of Oxford University Press.

a Large difference in laminar differentiation

b Moderate difference in laminar differentiation

Glossary

Agranular cortex
An isocortical region with a relatively undifferentiated layer II and layer III and lacking a fully expressed layer IV.

Allostasis
The process of activating physiological systems (such as hormonal, autonomic or immune systems) with the aim of returning the body to homeostasis.

Bayesian approach to probability
Models for assessing the probability of an event (that is, the posterior probability) based on the prior likelihood of an event and the evidence currently available as to its existence.

Centrifugal
A proposed hierarchical organization whereby the agranular and heteromodal association cortices form a collection of hubs, from which connections to unimodal sensory systems can be depicted as concentric rings.

Corollary discharge
Signals generated by the motor cortex that influence or inhibit the sensory processing of self-generated motor actions. Such signals convey simultaneous 'efference copies' of motor commands to sensory regions.

Default mode network
(DMN). A collection of midline and parietal brain regions that show more activity when people are constructing representations of the past and the future, simulating the present or processing semantic and conceptual content.

Degeneracy
The capacity of a system to perform identical functions or yield identical outputs with structurally different sets of elements.

Deterministic models
Mathematical models in which, given initial conditions or parameter values, there is no variation in the outcome.

Dysgranular cortex
An isocortical region defined by the presence of a differentiated layer II and layer III, and a rudimentary layer IV that contains stellate granule cells receiving thalamocortical inputs.

Granular cortex
An isocortical region with six differentiated layers, including a well-defined layer IV that contains many stellate granule cells receiving thalamocortical inputs.

Homeostasis
A set of dynamic functions (not a single set point) that interact to maintain an optimal use of energy in the body across all conditions at all times.

Interoception
The perception and integration of autonomic, hormonal, visceral and immunological homeostatic signals that collectively describe the physiological state of the body.

Interoceptive sensations
Activity within the nervous system indexing the autonomic, hormonal, visceral and immunological homeostatic signals that collectively describe the physiological state of the body — for example, concerning vagal signals, levels of insulin or cortisol, heart rate, gastric distension or inflammatory cytokine levels.

Lamina I pathway
Small-diameter sensory fibres that carry ascending interoceptive sensory signals (about muscle contractions in blood vessels, temperature, pain, hormonal activity, immunological inflammation and other variables) in the lateral spinothalamic pathway.

Positive alliesthesia
Transformation of a sensation from aversive to pleasurable, depending on the homeostatic needs of the body.

Precision units
Pyramidal cells that represent prediction-error signals; these cells modulate the activity of other neurons within a cortical column according to confidence in the predictions or the reliability of the incoming sensory signals.

Vagus nerve
Cranial nerve X, which carries ascending interoceptive sensory information about internal organs and the enteric nervous system.

Viseromotor cortices
Agranular regions of isocortex and allocortex that modulate the regulation of the autonomic nervous system as well as of the hormonal and immune systems.

---

periphery would be relayed to S1, where it would update proprioceptive and kinematic representations of the body that are subsequently transmitted to M1 to ensure that motor actions are executed as planned.

Moreover, Friston and colleagues go on to suggest that within S1, attentional mechanisms reduce the gain of sensations that arise from self-made actions, thereby reducing the prediction error that is available to propagate to M1. In summary, according to Friston and colleagues' active inference account, M1 receives little direct (that is, ascending) sensory input with which it can compute prediction error and receives little prediction error from S1. In effect, the motor and sensorimotor predictions that are transmitted by M1 to

the spinal cord and the somatosensory cortex, respectively, are relatively insulated from correction by prediction errors, such that M1 adjusts its predictions primarily by issuing new predictions. As a consequence, motor predictions from M1 are thought to function more like deterministic models of actions that are to be executed. By contrast, sensory cortices generate more-probabilistic predictions about the causes of sensory input that are continuously updated by prediction errors.

This active inference account of M1 and its predictive capacities relies on several assumptions that may require further consideration. First, M1 does in fact receive ascending inputs from the thalamus, via the basal ganglia and the cerebellum[44,45]. These

inputs function as prediction errors to correct the motor predictions that descend from M1 to the spinal cord. As a consequence, motor predictions from M1 might function more like probabilistic (and less like deterministic) signals for action. Furthermore, Barbas and García-Cabezas[46] have identified a fully expressed granular layer IV in M1 of the macaque brain, suggesting that the relative paucity of granular cells in M1 makes this region distinct from other agranular regions (such as the anterior cingulate cortex (ACC) and the anterior insula, which lack such cells). (Interestingly, as the granular layer in M1 is less developed than the granular layer in S1 to which it projects, the structural model still suggests similar predictive dynamics to those proposed by Friston and colleagues[17].)

Nonetheless, Friston and colleagues' elegant analysis of agranular architecture in active inference[28] can be applied more generally to agranular cortices, notably to the agranular limbic regions that regulate visceral control of the body's internal milieu, (that is, to viseromotor cortices[21,22]). Extending Friston and colleagues' active inference account in the context of the structural model of cortical connectivity[11,149] enables us to propose new hypotheses about the role of the agranular limbic cortices in interoception as well as in general brain function. Specifically, our analysis suggests that agranular viseromotor cortices may be the drivers of active interoceptive perception within the brain.

Active interoceptive inference
Building on our integration of the active inference account with the structural model, we propose the EPIC model of interoception. According to the EPIC model, there is an interoceptive system in the brain in which agranular cortices send viseromotor predictions to the body and transmit interoceptive predictions about the viscerosensory consequences of those predictions (BOX 2). After describing this interoceptive system, we describe how various cytoarchitectural constraints might render agranular viseromotor cortices less sensitive than the skeletomotor system or sensory systems to prediction errors, and consider the functional consequences for interoceptive perceptions and broader brain dynamics.

The interoceptive system. On the basis of tract-tracing and cytoarchitectonic studies of the macaque brain[18–26], the EPIC model proposes that the human brain contains an 'interoceptive system' in the frontal cortex.

Box 2 | **Hypotheses for the EPIC model of active interoceptive inference**

Using the Embodied Predictive Interoception Coding (EPIC) model, we can formulate four main hypotheses about how active inference can be implemented in interoception. According to this model, we suggest that agranular visceromotor cortices — including the cingulate cortex (Brodmann area 24 (BA24), BA25 and BA32), the posterior ventral medial prefrontal cortex (BA14c), the posterior orbitofrontal cortex (BA13a) and the most ventral portions of the anterior insula — estimate the balance between the autonomic, metabolic and immunological resources that are available to the body, and the predicted requirements of the body, based on past experience.

Our first hypothesis is that these balance estimates enable these agranular visceromotor cortices to issue allostatic visceromotor predictions to the hypothalamus, brainstem and spinal cord nuclei to maintain homeostasis within the internal milieu.

Our second hypothesis is that agranular visceromotor cortices simultaneously issue predictions of the interoceptive signals that are expected to arise as consequences of those allostatic visceral changes to the primary interoceptive sensory cortex (see the figure). As agranular cortices have relatively few projection neurons in their relatively undifferentiated upper layers (layer II and layer III) and no fully expressed layer IV, they are poorly structured to compute and send prediction error. However, they do have many projection neurons in their deeper layers (FIG. 1) and are thus well equipped to send predictions.

Our third hypothesis is that the granular cortex in primary interoceptive sensory regions of the mid- and posterior insula[136] are architecturally well suited for computing and transmitting prediction error and for propagating prediction-error signals back to visceromotor regions to modify predictions. Granular cortices have a well-defined layer IV, such that incoming sensory input from the thalamus can be amplified and redistributed throughout the cortical column. They also have many projection neurons in layer II and layer III that can compare predictions to incoming thalamic input to compute and send

prediction error. Evidence consistent with our second and third hypotheses has been provided by studies[137,138] that report interoceptive-stimulus-related activity in the (dysgranular or granular) mid- and posterior insula, and anticipatory activity in the (agranular) anterior insula.

Our fourth hypothesis is that the agranular visceromotor regions that send visceromotor and interoceptive predictions may be relatively insensitive to prediction-error signals compared with regions that send exteroceptive predictions[64]. As a result, interoceptive predictions in the agranular visceromotor cortices may be fairly stable compared with the fluctuating state of the body and thus may alter predictions more slowly than homeostatic information accrues in the viscerosensory cortex and is forwarded to the visceromotor cortices. AC, anterior cingulate; PL, prelimbic cortex. Box figure adapted with permission from REF. 139, Elsevier.
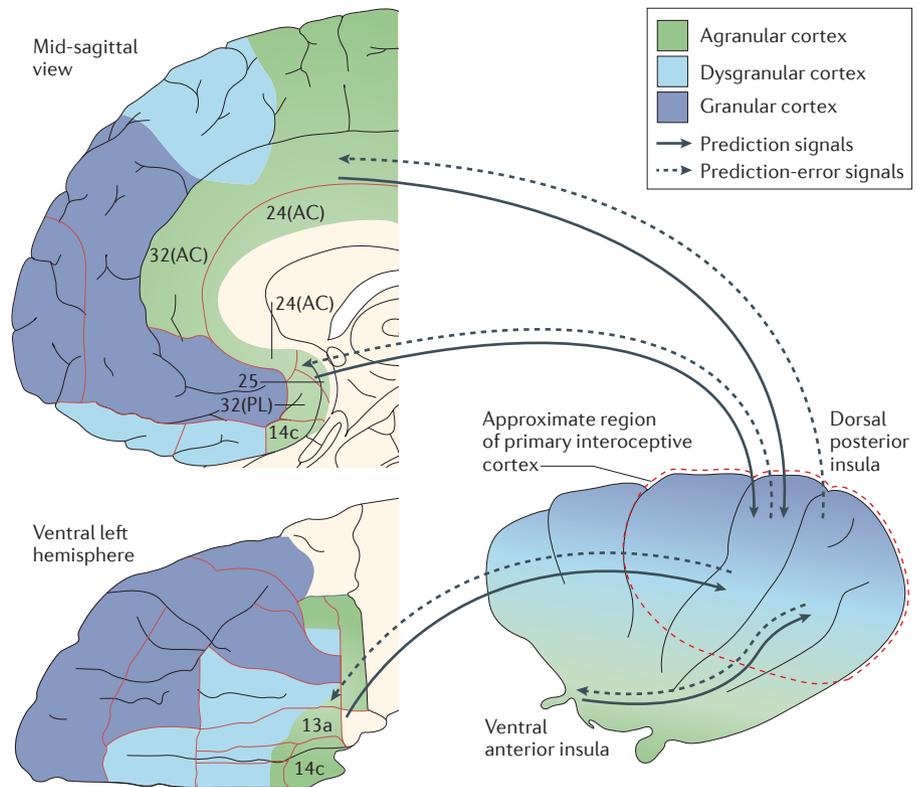
This system is composed of agranular visceromotor cortices that stretch along the medial wall of the cerebral cortex from the mid-cingulate cortex and the ACC (that is, Brodmann area 24 (BA24), BA25 and BA32) into the posterior ventromedial prefrontal cortex (posterior vmPFC; BA14c); laterally into the posterior orbitofrontal cortex (posterior BA13a); and finally posterolaterally into the anterior insula, the ventral extent of which is also agranular (see the figure in BOX 2). Collectively, these visceromotor regions project to the spinal cord via a set of connections that cascade through the amygdala, the ventral striatum, the hypothalamus and the periaqueductal grey[22,47–52]. Extending

Friston and colleagues' analysis of M1 (REFS 17,28,32) to agranular cortices more generally, we propose that the visceromotor cortices generate autonomic, hormonal and immunological predictions to adjust how the internal systems of the body deploy autonomic, metabolic and immunological resources to deal with the sensory world — not as it is right now, but as the brain predicts it will be in a moment from now. These visceromotor predictions underlie most of the body's allostatic and anticipatory responses to moment-by-moment physical movements and mental challenges.

At the same time as visceromotor predictions are issued to maintain homeostasis or to enable allostasis, the deep layers of agranular

visceromotor cortices send the same information via connections[23] to another part of the interoceptive system: the supragranular layers of the (granular) mid-to-posterior insular cortex, which serves as the primary interoceptive cortex[19,53]. Once there, these interoceptive predictions initiate a pattern of activity that represents expected interoceptive sensations. Interoceptive signals that result from changes in the viscera, muscles and skin ascend via the lamina I pathway and vagal afferents in the nucleus of the solitary tract, the parabrachial nucleus and the thalamus, before arriving at the granular layer (layer IV) of the primary interoceptive cortex[54]. Following our EPIC model, we hypothesize that these afferent interoceptive signals are amplified within

layer IV and expanded into the cortical columns of the mid- and posterior insula, which then computes the difference between the predicted interoceptive signal and the actual interoceptive signal as prediction error. The prediction-error signal is then propagated from the supragranular layers of the mid- and posterior insula back to the deep layers of the visceromotor regions, where the predictions originated.

Prediction error in agranular cortex. By analogy to Friston and colleagues' analysis, we suggest that agranular visceromotor cortices are less responsive to prediction-error signals than are more-granular cortices. First, without many projection neurons in layer II and layer III, and lacking a fully expressed layer IV, agranular visceromotor cortices lack the cytoarchitecture that computes and sends prediction errors efficiently (BOXES 1,2). Without a layer IV, thalamic information that reaches agranular visceromotor cortices will not be expanded and amplified into the rest of the column, limiting the capacity of the column to compute prediction error: fewer neurons in the upper layers of such columns are available to propagate whatever prediction error is computed. Furthermore, we would expect agranular visceromotor regions to be less sensitive to ascending spinal signals than is M1, because M1 receives ascending spinal afferents from the cerebellum, whereas the visceromotor cortices probably receive very light projections, although these have not been fully verified[44]. Visceromotor cortices do receive projections from basal ganglia structures via the thalamus, but there are far fewer thalamus–basal ganglia projections than there are projections going in the other direction (that is, from the basal ganglia to the thalamus) or other thalamocortical projections (that is, from the thalamus to the other parts of frontal cortex)[45,55,56]. (For example, the basal ganglia project heavily via the thalamus (and also, to a lesser extent, the amygdala) to the motor regions of the cingulate cortices and to the visceromotor subgenual ACC (BA25)[57–59], whereas they project sparsely to other visceromotor regions in the posterior vmPFC and in the posterior orbitofrontal cortex[60].) In addition to these structural considerations, visceral sensations are noisy[61] and ascend to the cortex via the largely unmyelinated vagus nerve[62], where precision-weighting units may reduce the transmission of ascending signals. As discussed above, the widely distributed control networks in the brain implement precision weighting of incoming sensory input by modulating the activity of precision units in the upper layers of cortical columns[33–38].

We propose that once prediction-error signals reach agranular visceromotor regions (either directly but minimally, or indirectly via the primary interoceptive cortex), a limited set of outcomes can occur. Visceromotor regions can modulate their outputs to the body by issuing new visceral predictions to the spinal cord that can actively generate the sensory input required to confirm the prediction. In addition, visceromotor cortices simultaneously transmit new interoceptive predictions to the mid- and posterior insula that will minimize the subsequent prediction errors. The anterior insula and the ACC — through their involvement in the intrinsic executive control and attention networks[33–38] — can further reduce the sampling or processing of the internal environment to reduce prediction error. Unlike most other cortical regions, the visceromotor circuitry can also direct sensory sampling to resolve interoceptive prediction errors by regulating the thalamic reticular nucleus[55] and thus gating incoming interoceptive input to the cortex in a highly specific manner. In this way, visceromotor cortices can effectively modulate the gain on corticothalamic and thalamocortical communications and thereby implement a potent form of interoceptive attention.

Interoceptive sensation is largely prediction. The active interoceptive inference perspective of the EPIC model posits that homeostatic–allostatic control and interoception are unified within an integrated neural architecture. Using expectations about the world that are based on past experience (that is, empirical prior probabilities), agranular circuitry estimates the body's upcoming autonomic, metabolic and immunological needs. These estimates not only function as predictions of which visceromotor changes are required in the next instance — that is, as visceromotor intentions — they are also interoceptive predictions. When prediction errors are small, interoceptive perceptions derive from simulated interoceptive sensations (that is, from interoceptive predictions or, in Bayesian terms, the posterior probability estimate); they are the brain's best hypothesis, based on past experiences, as to the cause of current sensations (and form the bases for interoceptive prior probability distributions that the brain will use to generate future predictions). The brain implements visceromotor intentions — thereby generating the interoceptive consequences that have been predicted — such that perceptions follow actions (rather than the other way around). Furthermore, by extending Friston and colleagues' analysis of M1, agranular visceromotor cortices are less sensitive to prediction error than are the granular cortices of exteroceptive sensory systems, such as those that process visual or auditory information (for the reasons discussed in the section 'Prediction error in agranular cortex'; also see FIG. 1). This means that interoceptive perception is largely a construction of beliefs that are kept in check by the actual state of the body (rather than vice versa). What you experience is in large part a reflection of what your brain predicts is going on inside your body, based on past experience. An example of how the process unfolds can be found in Supplementary information S2 (box).

If the EPIC account of interoception is correct, then interoceptive predictions in agranular cortices would be more stable than the ever-changing internal milieu of the body and would update predictions more slowly than the rate at which ascending sensory input accrues in primary viscerosensory cortex (BOX 2). This disparity may account for the long 'half-life' of the autonomic nervous system[63]. Indeed, evidence from a recent study provides support for this hypothesis: the (dysgranular or granular) dorsal mid-insula response to food images was attenuated by increases in circulating glucose levels (a homeostatic marker of energy availability), whereas the (agranular) anterior insula response to these images was unaffected by changes in glucose levels[64].

Visceromotor cortices, despite being relatively insensitive to interoceptive prediction error, are influenced (via connections to the vmPFC) by predictions from other agranular cortices, such as the hippocampus and other agranular regions of the default mode network (DMN). As a consequence, in a normally functioning brain, predictions from agranular cortices are less constrained by the current state of the body, enabling future-oriented visceromotor predictions to drive the body towards anticipated homeostasis. This allostatic control is constrained primarily by prediction errors that come from the extended interoceptive cortical system, as well as by the limbic regions from which the network receives connections.

Visceromotor prediction within the wider brain. It may be tempting to view the interoceptive system, as outlined in the EPIC model, as a modular system. However, the brain has a small-world architecture, whereby closely clustered nodes make up neighbourhoods that are connected into networks via long-range pathways[65,66]. This small-world architecture is augmented by 'rich-club' hubs (that is, highly connected nodes), which are the locations where intrinsic brain networks overlap with one another and which serve as the brain's 'backbone' for neural communication and synchrony[67,68]. Several agranular visceromotor regions — including the anterior insula and the cingulate cortices — are rich-club hubs[68], prompting the hypothesis that agranular visceromotor cortices send predictions to and receive prediction-error signals from cortices with greater laminar differentiation in an effort to create the kind of synchronized brain activity that is necessary for consciousness.

These visceromotor rich-club regions are part not only of the interoceptive system, but also of the intrinsic control and attention networks[33,36–38,69,70]. The anterior insula is a key hub in olfactory[71] and gustatory[72] networks, as well as in the brain's multimodal network, which integrates visual, auditory and somatosensory networks[73]. According to our EPIC account, agranular visceromotor cortices inform the rest of the brain of interoceptive changes by sending predictions to these intrinsic networks that are based on anticipated visceromotor consequences. Thus, the visceromotor cortices can modulate attentional, sensory and behavioural responses to stimuli that are homeostatically relevant (for instance, in the case of positive alliesthesia[74,75]). In this way, the anterior insula and the cingulate cortices help to create a multisensory representation of the body in the world so that what you see, hear and feel is influenced by your interoceptive predictions. The result is a multisensory representation of the world from the perspective of someone with a body[31,76,77]. The portions of the vmPFC that make up part of the visceromotor cortex belong to the DMN (which is also known as the 'mentalizing network') and could therefore transmit anticipated homeostatic consequences to the rest of that network as predictions during phenomena that robustly engage this network, such as semantic and conceptual processing, imagery, mind wandering, memory, empathy[15], person perception, mental-state inference, self-related processing, valuation and language processing, among others[78–82].

The fact that the agranular visceromotor cortices are key hubs in neural communication networks across the entire brain[67,68] ensures that interoceptive representations are a central part of every mental event. This view is wholly consistent with embodied accounts of perception, cognition and emotion[31,83,84] (see Supplementary information S3 (box)), and may explain why affective properties such as pleasure, displeasure and arousal — which are thought to be rooted in interoception — are fundamental properties of conscious experience[85]. The picture emerging here is one in which neural representations of the world that underlie perception and action are, in many cases, directed more by the homeostatic relevance of information than by the need for accuracy and completeness in representing the outside world. If the interoceptive system provides efference copies to multiple sensory systems[86] and thus forms the basis for unified conscious experience, then the EPIC model offers a new avenue for considering deficits in sensory integration (such as in autism spectrum disorders[86,87]). Similarly, aberrant predictions may also explain why pathophysiology within the regions that mediate interoception is associated with depression and other illnesses[88–90], as discussed below.

Implications for illnesses

The EPIC model predicts that aberrant interoceptive predictions can lead to chronic physical burdens that result in various mental and physical illnesses — particularly disorders of mood, appetite and metabolism. Here, we describe one example: depression. Many of the key regions that have been implicated in the pathophysiology of depression, such as the subgenual and subcallosal ACC and the anterior insula[58,91,92], are agranular visceromotor limbic regions within the interoceptive system that is proposed here. It is well known that structural abnormalities and chronically hyperactive metabolism within agranular visceromotor regions precede the onset of depression (for example, see REFS 93,94). According to the EPIC model, the predictions that result from these structurally or functionally abnormal visceromotor regions may result in a break-down of body systems that are needed to maintain homeostasis in response to stressors or everyday events that are perceived as stressors. In the long term, this chronic imbalance — which is caused by constantly predicting the need for more metabolic energy to meet the demands of stressors[95] — can produce the well-known depression-related disruption and eventual downregulation of

hypothalamus–pituitary–adrenal (HPA)-axis negative-feedback loops, resulting in chronic hypercortisolaemia[96] (but see REF. 97). This in turn can promote a pro-inflammatory state that is associated with increased levels of cytokines and activated immune biochemical pathways[98].

The vagus nerve brings information about endocrine and immune functions from the periphery to the viscerosensory cortex in the mid- and posterior insula[99,100]. As a consequence of the peripheral endocrine and immune changes brought about by HPA-axis overactivity and pro-inflammatory states, afferent interoceptive input may become increasingly decoupled from the specific interoceptive predictions that are issued by the agranular visceromotor cortex, leading to increased prediction-error signals. This decoupling may present in the brain as "noisy afferent interoceptive inputs" (REF. 101). Active inference accounts typically posit that the brain deals with noisy prediction errors by decreasing the gain on cortical pyramidal neurons that function like precision units to regulate outputs from error-prediction computations, thereby reducing (or in some cases increasing) the influence of these outputs, on the basis of the relative confidence in the descending predictions[28] (reviewed in REF. 86). As a result — at least in the short term, when ascending sensory input is noisy or imprecise — the active inference framework hypothesizes that prediction errors will be resolved by maintaining predictions and reducing attention to the sensory input or by actively continuing to drive autonomic, metabolic and immune systems in the body to generate the predicted sensory input. This would set up a positive-feedback cycle whereby autonomic, endocrine[102] and immune changes lead to enhanced, but even noisier, ascending interoceptive sensory input. This prediction error may then be ignored, allowing further costs to the body's systems to accumulate. As the autonomic, endocrine and immune systems have a finite range of activity over which they can function, it is unlikely that these prediction-error signals can be ignored indefinitely.

We speculate that as the prodromal depressive illness progresses, the enhanced prediction-error signals may eventually overcome the inherent insensitivity of the agranular visceromotor cortex. The point at which this would occur would be related to individual differences in the presence and quality of precision units as well as in the connectivity between regions[103] that enable the propagation of noisy prediction-error signals up the active inference hierarchy.

When this happens, we hypothesize that, in a bid to ultimately reduce prediction error, limbic visceromotor cortices begin guiding the body towards a constellation of sickness behaviours[104] associated with fatigue[105] and negative affect[100] that are designed to reduce activity and energy expenditure. Collectively, these behaviours would be the initial behavioural symptoms of depression. When the attempts of the visceromotor regions to reduce interoceptive prediction error become extreme enough, the system tips into depression, anxiety or another of the syndromes that are associated with structural and processing abnormalities in the regions implicated in interoception[90,91,106].

If this account is correct, then two hypotheses should follow: first, depression should be associated with abnormal interoceptive activity in the mid- or posterior insula viscerosensory cortex; and second, the strength of the connections between this region and the agranular visceromotor cortices should be positively associated with the severity of an individual's depression. Indeed, we have recently demonstrated that both are the case[9]. When undergoing functional MRI (fMRI), unmedicated depressed adults had altered activity in the dorsal mid-insula while performing an interoceptive attention task, and the activity in this region was negatively correlated with their score on standard clinical measures of depression severity. In addition, the depression-severity scores of these same individuals were positively correlated with the intrinsic connectivity between the dorsal mid-insula and a host of regions that have previously been implicated in depression, including the visceromotor subgenual ACC and the anterior insula.

Our EPIC model may also inform treatment of depression and some anxiety disorders. For example, deep brain stimulation of the connections that project out of the subcallosal cingulate cortex in a region of the visceromotor system (particularly in BA25) is effective in remitting treatment-resistant depression[107,108]. This is consistent with the suggestion that the propagation of heightened interoceptive predictions from BA25 through the cingulum bundle and uncinate fasciculus to the salience network and the DMN contributes to depression[109], and with the finding that evoked activity in BA25 is reduced after treatment[107,108,110]. Likewise, cognitive behavioural therapy (CBT) decreases activity in agranular cingulate cortex in depression[111]. In addition, similar results with CBT have been observed in the anterior insula in anxiety disorders[112]. In a Bayesian sense, the effects of CBT may reflect changes in the way that precision-weighting pyramidal cells in the viscerosensory cortex adjust the weight of prediction-error signals that are communicated to agranular cortices, thus altering the sampling of inputs that become the 'empirical priors' in subsequent predictions. Interestingly, emerging evidence indicates that the activity within agranular visceromotor cortices predicts whether CBT or pharmacotherapy will be more effective as a treatment option[113,114].

Perhaps our most speculative but innovative hypothesis concerns the relationship between interoceptive predictions and certain physical illnesses that often co-occur with depression, such as diabetes, heart disease and cancer[115]. Aberrant interoceptive predictions and the compounding allostatic consequences that may result could help to explain the links among these disorders. For example, many of the same regions within the interoceptive system that show morphological changes in psychiatric illness[90] and chronic pain[116] also show morphological changes with accumulated stress across the lifespan[117] and leave individuals more vulnerable to these metabolic illnesses and with increased risk of mortality[118], particularly if this stress occurred in childhood[119]. All of these illnesses are also linked to homeostatic and inflammatory mechanisms.

Furthermore, if, as we speculate, interoceptive predictions are the basis for many normal anticipatory physiological phenomena (such as the cardiac and respiratory changes associated with stress, and the cephalic-phase insulin response, whereby insulin levels rise simply at the sight or smell of food[120]), then pathophysiology within the interoceptive system may result in aberrant endocrine and autonomic responses to stress and food cues. To date, studies of interoception in the context of diabetes, obesity and related metabolic disorders have primarily focused on the hypothalamus and the nucleus of the solitary tract for nutrient sensing, and on the locus coeruleus and the lateral tegmental area for the regulation of sympathetic nervous system activation. The fact that these subcortical structures project to regions in the interoceptive system or are closely connected to it — especially the insula and a visceromotor map in M1 and M2 (REF. 121) — strongly suggests that dysregulation within the interoceptive system may be involved in some physical illnesses, and thus these physical disorders may share a common neural substrate with mood disorders.

Conclusion

In this Opinion article, we have integrated the active inference account with the structural model of corticocortical connectivity to extend prior predictive-coding accounts of interoception. Our EPIC model represents the detailed anatomical implementation of the active inference framework, recognizing that lamination gradients within cortical architecture offer clues to the computations that occur within various regions during interoception (beyond merely the anterior insula; see BOX 3). The EPIC model also recognizes that homeostatically derived interoceptive predictions and prediction errors might have an important role in coordinating communication in brain networks, and that aberrant interoceptive predictions might account for some common chronic illnesses.

As with any novel approach, much research is required to verify the hypotheses we have proposed here. In particular, EPIC (as well as other predictive-coding accounts of interoception[12–14]) would benefit from a systematic attempt to verify the interoceptive system in the human brain, perhaps using resting-state connectivity analyses. The low-frequency fluctuations in blood-oxygenation-level-dependent (BOLD) signal that define resting-state connectivity MRI in humans are anatomically constrained by the structural connectivity of brain networks[122–125]. Such resting-state BOLD data would be particularly important for understanding the connectivity of the anterior insula, as macaques lack a homologue of the human anterior insula and so cannot be used to study this in great depth[6,24]. Moreover, such data may help to resolve an underspecified aspect of the model — namely, whether the anterior insula issues visceromotor predictions directly to the hypothalamus and brainstem[47], issues them indirectly via the subgenual ACC[58] or sends interoceptive predictions indirectly through the multimodal integration network[73].

Relatedly, recent developments in ultra-high-resolution laminar fMRI[126] may allow tests of the EPIC model's hypotheses about interoceptive prediction and prediction-error computations within specific laminae. Likewise, advanced electrocorticography techniques could be used to assess the temporal dynamics of information processing between visceromotor and viscerosensory cortices.

Finally, as viscerosensory signals probably undergo substantial signal conditioning in the brainstem, further work should specify the effects of the brainstem and

Box 3 | The insular cortex as a focus in models of interoception

For many years, interoception has been studied as a bottom-up, sensory-driven phenomenon that is linked to the neuroanatomy of the ascending homeostatic lamina I and vagal pathways that bring interoceptive sensations to the brain[54]. Other approaches to understanding interoception have focused on the structure and function of the anterior insula, because of its importance in interoceptive awareness and sensitivity (for example, see REFS 6,140). Both lines of research reflect an assumption that the brain simply reads out signals from the various interoceptive channels (such as those relating to the heart, respiration, digestion, pain, metabolism or immune function) and initiates action if and when those signals diverge too greatly from homeostatic set points. More recently, Craig[6] suggested that the insular cortex plays a part in transforming ascending interoceptive signals into a 'global emotional moment', with the anterior insula being a crucial area for human awareness.

The anterior insula also has a key role in recent active inference accounts of interoception[13,14], in which it is thought to compare predictions and prediction errors (rather than issue predictions and receive prediction errors that are computed within the more granular cortical regions; see BOX 2). In this approach, the anterior insula is the basis of conscious 'presence' and emotional awareness. In contrast to these approaches, our Embodied Predictive Interoception Coding (EPIC) model does not focus on the anterior insula as a necessary area for consciousness or emotional awareness per se. This is because, in this model, multiple pathways within the combined cortical interoceptive network and the ascending pathways can construct interoceptive perceptions — reflecting degeneracy in the system[141,142] — thereby making conscious presence and emotional awareness more resilient to brain injury. Indeed, two rare patients (known as Roger and Boswell) with bilateral destruction of the anterior insula (and most of their limbic circuitry) both show remarkably preserved affective experience, self-awareness and conscious 'presence' — that is, the psychological phenomena on which interoceptive predictions depend[143–145]. Our EPIC model also explains why bilateral insula lesions can result in profound viscerosensory interoceptive deficits[146], a finding that implies that the insula has a functional role that is more circumscribed than the concept of emotional awareness, which is probably broadly mediated by multiple networks distributed throughout the brain.

subcortical structures on the interoceptive system. For example, future research should clarify how information processing in the brainstem mediates the viscerosensory signals that are sent to the cortex, and the extent to which the brainstem and subcortical structures contribute directly to active inference — for instance, by computing prediction errors.

Understanding the brain as issuing and sculpting interoceptive inferences genuinely inverts the traditional functional hierarchy, such that agranular (limbic) cortices are no longer assigned the function of reacting to stimulation from the world, but are instead anticipating it. Rather than interoceptive perceptions being solely the representation of afferent sensory input from the body, they can be thought of as inferences about the sensory consequences of homeostatic budgeting that are implemented as upcoming visceromotor commands; these inferences are constrained by error signals that result from the failure of previous predictions to accurately account for incoming interoceptive sensations[127]. Prediction errors have the capacity to feed back up the active inference hierarchy to sculpt future visceromotor outputs and predictions. In this way, representations that are built from previous experience drive neural activity and are modulated or constrained by actual sensory input from the internal milieu of the body. In the most general terms, interoceptive perceptions — that is, what is experienced — derive from the brain's best guess about the causes of events in the body, with incoming sensory inputs keeping those guesses in check. According to the EPIC model, not only has your past viscerosensory experience reached forward to create your present experience, but how your body feels now will again project forward to influence what you will feel in the future. It is an elegantly orchestrated self-fulfilling prophecy, embodied within the architecture of the nervous system.

Lisa Feldman Barrett is at Northeastern University, Department of Psychology, Boston, Massachusetts 02115, USA; and the Massachusetts General Hospital, Department of Psychiatry and the Martinos Center for Biomedical Imaging, Charlestown, Massachusetts 02129, USA.

W. Kyle Simmons is at the Laureate Institute for Brain Research, Tulsa, Oklahoma 74133, USA; and the Faculty of Community Medicine, University of Tulsa, Tulsa, Oklahoma 74104, USA.

L.F.B. and W.K.S. contributed equally to this work.

e-mails: l.barrett@neu.edu; wksimmons@laureateinstitute.org

1. Friston, K. The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* **11**, 127–138 (2010).
2. Clark, A. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.* **36**, 181–204 (2013).
3. Mesulam, M. M. From sensation to cognition. *Brain* **121**, 1013–1052 (1998).
4. Kok, P. & de Lange, F. P. Shape perception simultaneously up- and downregulates neural activity in the primary visual cortex. *Curr. Biol.* **24**, 1531–1535 (2014).
5. Chennu, S. *et al.* Expectation and attention in hierarchical auditory prediction. *J. Neurosci.* **33**, 11194–11205 (2013).
6. Craig, A. D. How do you feel — now? The anterior insula and human awareness. *Nat. Rev. Neurosci.* **10**, 59–70 (2009).
7. Paulus, M. P. & Stein, M. B. An insular view of anxiety. *Biol. Psychiatry* **60**, 383–387 (2006).
8. Paulus, M. P. & Stewart, J. L. Interoception and drug addiction. *Neuropharmacology* **76**, 342–350 (2014).
9. Avery, J. A. *et al.* Major depressive disorder is associated with abnormal interoceptive activity and functional connectivity in the insula. *Biol. Psychiatry* **76**, 258–266 (2014).
10. Simmons, W. K. *et al.* Keeping the body in mind: insula functional organization and functional connectivity integrate interoceptive, exteroceptive, and emotional awareness. *Hum. Brain Mapp.* **34**, 2944–2958 (2013).
11. Barbas, H. & Rempel-Clower, N. Cortical structure predicts the pattern of corticocortical connections. *Cereb. Cortex* **7**, 635–646 (1997).
12. Gu, X., Hof, P. R., Friston, K. J. & Fan, J. Anterior insular cortex and emotional awareness. *J. Comp. Neurol.* **521**, 3371–3388 (2013).
13. Seth, A. K. Interoceptive inference, emotion, and the embodied self. *Trends Cogn. Sci.* **17**, 565–573 (2013).
14. Seth, A. K., Suzuki, K. & Critchley, H. D. An interoceptive predictive coding model of conscious presence. *Front. Psychol.* **2**, 395 (2011).
15. Singer, T., Critchley, H. D. & Preuschoff, K. A common role of insula in feelings, empathy and uncertainty. *Trends Cogn. Sci.* **13**, 334–340 (2009).
16. Mumford, D. On the computational architecture of the neocortex. I. The role of the thalamo-cortical loop. *Biol. Cybern.* **65**, 135–145 (1991).
17. Bastos, A. M. *et al.* Canonical microcircuits for predictive coding. *Neuron* **76**, 695–711 (2012).
18. Carmichael, S. T. & Price, J. L. Connectional networks within the orbital and medial prefrontal cortex of macaque monkeys. *J. Comp. Neurol.* **371**, 179–207 (1996).
19. Evrard, H. C., Logothetis, N. K. & Craig, A. D. Modular architectonic organization of the insula in the macaque monkey. *J. Comp. Neurol.* **522**, 64–97 (2014).
20. Mufson, E. J. & Mesulam, M. M. Insula of the old world monkey. II: afferent cortical input and comments on the claustrum. *J. Comp. Neurol.* **212**, 23–37 (1982).
21. Öngür, D., Ferry, A. T. & Price, J. L. Architectonic subdivision of the human orbital and medial prefrontal cortex. *J. Comp. Neurol.* **460**, 425–449 (2003).
22. Öngür, D. & Price, J. L. The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cereb. Cortex* **10**, 206–219 (2000).
23. Saleem, K. S., Kondo, H. & Price, J. L. Complementary circuits connecting the orbital and medial prefrontal networks with the temporal, insular, and opercular cortex in the macaque monkey. *J. Comp. Neurol.* **506**, 659–693 (2008).
24. Semendeferi, K. & Damasio, H. The brain and its main anatomical subdivisions in living hominoids using magnetic resonance imaging. *J. Hum. Evol.* **38**, 317–332 (2000).
25. Vogt, B. A. & Pandya, D. N. Cingulate cortex of the rhesus monkey: II. Cortical afferents. *J. Comp. Neurol.* **262**, 271–289 (1987).
26. Vogt, B. A., Pandya, D. N. & Rosene, D. L. Cingulate cortex of the rhesus monkey: I. Cytoarchitecture and thalamic afferents. *J. Comp. Neurol.* **262**, 256–270 (1987).
27. Barbas, H. & Hilgetag, C. C. Rules relating connections to cortical structure in primate prefrontal cortex. *Neurocomputing* **44–46**, 301–308 (2002).
28. Shipp, S., Adams, R. A. & Friston, K. J. Reflections on agranular architecture: predictive coding in the motor cortex. *Trends Neurosci.* **36**, 706–716 (2013).

29. Larkum, M. A cellular mechanism for cortical associations: an organizing principle for the cerebral cortex. *Trends Neurosci.* **36**, 141–151 (2013).
30. Sherman, S. M. The thalamus is more than just a relay. *Curr. Opin. Neurobiol.* **17**, 417–422 (2007).
31. Barsalou, L. W. Grounded cognition. *Annu. Rev. Psychol.* **59**, 617–645 (2008).
32. Adams, R. A., Shipp, S. & Friston, K. J. Predictions not commands: active inference in the motor system. *Brain Struct. Funct.* **218**, 611–643 (2013).
33. Corbetta, M. & Shulman, G. L. Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* **3**, 201–215 (2002).
34. Fox, M. D., Corbetta, M., Snyder, A. Z., Vincent, J. L. & Raichle, M. E. Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems. *Proc. Natl Acad. Sci. USA* **103**, 10046–10051 (2006).
35. Fox, M. D. *et al.* The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc. Natl Acad. Sci. USA* **102**, 9673–9678 (2005).
36. Seeley, W. W. *et al.* Dissociable intrinsic connectivity networks for salience processing and executive control. *J. Neurosci.* **27**, 2349–2356 (2007).
37. Vincent, J. L., Kahn, I., Snyder, A. Z., Raichle, M. E. & Buckner, R. L. Evidence for a frontoparietal control system revealed by intrinsic functional connectivity. *J. Neurophysiol.* **100**, 3328–3342 (2008).
38. Yeo, B. T. *et al.* The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J. Neurophysiol.* **106**, 1125–1165 (2011).
39. Feldman, H. & Friston, K. J. Attention, uncertainty, and free-energy. *Front. Hum. Neurosci.* **4**, 215 (2010).
40. Friston, K. The free-energy principle: a rough guide to the brain? *Trends Cogn. Sci.* **13**, 293–301 (2009).
41. Shipp, S. The importance of being agranular: a comparative account of visual and motor cortex. *Phil. Trans. R. Soc. B* **360**, 797–814 (2005).
42. Wolpert, D. M. & Kawato, M. Multiple paired forward and inverse models for motor control. *Neural Netw.* **11**, 1317–1329 (1998).
43. Crapse, T. B. & Sommer, M. A. Corollary discharge across the animal kingdom. *Nat. Rev. Neurosci.* **9**, 587–600 (2008).
44. Barbas, H., García-Cabezas, M. Á. & Zikopoulos, B. Frontal–thalamic circuits associated with language. *Brain Lang.* **126**, 49–61 (2013).
45. McFarland, N. R. & Haber, S. N. Thalamic relay nuclei of the basal ganglia form both reciprocal and nonreciprocal cortical connections, linking multiple frontal cortical areas. *J. Neurosci.* **22**, 8117–8132 (2002).
46. Barbas, H. & García-Cabezas, M. Á. Motor cortex layer 4: less is more. *Trends Neurosci.* **38**, 259–261 (2015).
47. Carmichael, S. T. & Price, J. L. Limbic connections of the orbital and medial prefrontal cortex in macaque monkeys. *J. Comp. Neurol.* **363**, 615–641 (1995).
48. Mesulam, M. M. & Mufson, E. J. Insula of the old world monkey. III: efferent cortical output and comments on function. *J. Comp. Neurol.* **212**, 38–52 (1982).
49. Freedman, L. J., Insel, T. R. & Smith, Y. Subcortical projections of area 25 (subgenual cortex) of the macaque monkey. *J. Comp. Neurol.* **421**, 172–188 (2000).
50. Vogt, B. A. Pain and emotion interactions in subregions of the cingulate gyrus. *Nat. Rev. Neurosci.* **6**, 533–544 (2005).
51. Hsu, D. T. & Price, J. L. Midline and intralaminar thalamic connections with the orbital and medial prefrontal networks in macaque monkeys. *J. Comp. Neurol.* **504**, 89–111 (2007).
52. Chiba, T., Kayahara, T. & Nakano, K. Efferent projections of infralimbic and prelimbic areas of the medial prefrontal cortex in the Japanese monkey, *Macaca fuscata*. *Brain Res.* **888**, 83–101 (2001).
53. Mesulam, M. M. & Mufson, E. J. Insula of the old world monkey. I: Architectonics in the insulo–orbito–temporal component of the paralimbic brain. *J. Comp. Neurol.* **212**, 1–22 (1982).
54. Craig, A. D. How do you feel? Interoception: the sense of the physiological condition of the body. *Nat. Rev. Neurosci.* **33**, 655–666 (2002).
55. Zikopoulos, B. & Barbas, H. Prefrontal projections to the thalamic reticular nucleus form a unique circuit for attentional mechanisms. *J. Neurosci.* **26**, 7348–7361 (2006).
56. Haber, S. N. & Behrens, T. E. The neural network underlying incentive-based learning: implications for interpreting circuit disruptions in psychiatric disorders. *Neuron* **83**, 1019–1039 (2014).
57. Ghashghaei, H. T., Hilgetag, C. C. & Barbas, H. Sequence of information processing for emotions based on the anatomic dialogue between prefrontal cortex and amygdala. *Neuroimage* **34**, 905–923 (2007).
58. Hamani, C. *et al.* The subcallosal cingulate gyrus in the context of major depression. *Biol. Psychiatry* **69**, 301–308 (2011).
59. Timbie, C. & Barbas, H. Specialized pathways from the primate amygdala to posterior orbitofrontal cortex. *J. Neurosci.* **34**, 8106–8118 (2014).
60. Zikopoulos, B. & Barbas, H. Parallel driving and modulatory pathways link the prefrontal cortex and thalamus. *PLoS ONE* **2**, e848 (2007).
61. Damasio, A. & Carvalho, G. B. The nature of feelings: evolutionary and neurobiological origins. *Nat. Rev. Neurosci.* **14**, 143–152 (2013).
62. Hoffman, H. H. & Schnitzlein, H. N. The numbers of nerve fibers in the vagus nerve of man. *Anat. Rec.* **139**, 429–435 (1961).
63. Sapolsky, R. M. *Monkeyluv: And Other Essays on Our Lives as Animals* (Scribner, 2006).
64. Simmons, W. K. *et al.* Category-specific integration of homeostatic signals in caudal but not rostral human insula. *Nat. Neurosci.* **16**, 1551–1552 (2013).
65. van den Heuvel, M. P., Stam, C. J., Boersma, M. & Hulshoff Pol, H. E. Small-world and scale-free organization of voxel-based resting-state functional connectivity in the human brain. *Neuroimage* **43**, 528–539 (2008).
66. Zippo, A. G. *et al.* Small-world networks in neuronal populations: a computational perspective. *Neural Netw.* **44**, 143–156 (2013).
67. van den Heuvel, M. P. & Sporns, O. An anatomical substrate for integration among functional networks in human cortex. *J. Neurosci.* **33**, 14489–14500 (2013).
68. van den Heuvel, M. P. & Sporns, O. Rich-club organization of the human connectome. *J. Neurosci.* **31**, 15775–15786 (2011).
69. Dosenbach, N. U. *et al.* Distinct brain networks for adaptive and stable task control in humans. *Proc. Natl Acad. Sci. USA* **104**, 11073–11078 (2007).
70. Dosenbach, N. U., Fair, D. A., Cohen, A. L., Schlaggar, B. L. & Petersen, S. E. A dual-networks architecture of top-down control. *Trends Cogn. Sci.* **12**, 99–105 (2008).
71. Seubert, J., Freiherr, J., Djordjevic, J. & Lundström, J. N. Statistical localization of human olfactory cortex. *Neuroimage* **66**, 333–342 (2013).
72. Veldhuizen, M. G. *et al.* Identification of human gustatory cortex by activation likelihood estimation. *Hum. Brain Mapp.* **32**, 2256–2266 (2011).
73. Sepulcre, J., Sabuncu, M. R., Yeo, T. B., Liu, H. & Johnson, K. A. Stepwise connectivity of the modal cortex reveals the multimodal organization of the human brain. *J. Neurosci.* **32**, 10649–10661 (2012).
74. Paulus, M. P., Tapert, S. F. & Schulteis, G. The role of interoception and alliesthesia in addiction. *Pharmacol. Biochem. Behav.* **94**, 1–7 (2009).
75. van Elk, M., Lenggenhager, B., Heydrich, L. & Blanke, O. Suppression of the auditory N1-component for heartbeat-related sounds reflects interoceptive predictive coding. *Biol. Psychol.* **99**, 172–182 (2014).
76. Oosterwijk, S. *et al.* States of mind: emotions, body feelings, and thoughts share distributed neural networks. *Neuroimage* **62**, 2110–2128 (2012).
77. Blanke, O. Multisensory brain mechanisms of bodily self-consciousness. *Nat. Rev. Neurosci.* **13**, 556–571 (2012).
78. Andrews-Hanna, J. R., Smallwood, J. & Spreng, R. N. The default network and self-generated thought: component processes, dynamic control, and clinical relevance. *Ann. NY Acad. Sci.* **1316**, 29–52 (2014).
79. Barrett, L. F. & Satpute, A. B. Large-scale brain networks in affective and social neuroscience: towards an integrative functional architecture of the brain. *Curr. Opin. Neurobiol.* **23**, 361–372 (2013).
80. Binder, J. R., Desai, R. H., Graves, W. W. & Conant, L. L. Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb. Cortex* **19**, 2767–2796 (2009).
81. Clithero, J. A. & Rangel, A. Informatic parcellation of the network involved in the computation of subjective value. *Soc. Cogn. Affect. Neurosci.* **9**, 1289–1302 (2014).
82. Roy, M., Shohamy, D. & Wager, T. D. Ventromedial prefrontal–subcortical systems and the generation of affective meaning. *Trends Cogn. Sci.* **16**, 147–156 (2012).
83. Wilson-Mendenhall, C. D., Barrett, L. F. & Barsalou, L. W. Situating emotional experience. *Front. Hum. Neurosci.* **7**, 764 (2013).
84. Barrett, L. F. Solving the emotion paradox: categorization and the experience of emotion. *Pers. Soc. Psychol. Rev.* **10**, 20–46 (2006).
85. Barrett, L. F. & Bliss-Moreau, E. Affect as a psychological primitive. *Adv. Exp. Soc. Psychol.* **41**, 167–218 (2009).
86. Quattrocki, E. & Friston, K. Autism, oxytocin and interoception. *Neurosci. Biobehav. Rev.* **47**, 410–430 (2014).
87. Sinha, P. *et al.* Autism as a disorder of prediction. *Proc. Natl Acad. Sci. USA* **111**, 15220–15225 (2014).
88. Menon, V. Large-scale brain networks and psychopathology: a unifying triple network model. *Trends Cogn. Sci.* **15**, 483–506 (2011).
89. Crossley, N. A. *et al.* The hubs of the human connectome are generally implicated in the anatomy of brain disorders. *Brain* **137**, 2382–2395 (2014).
90. Goodkind, M. *et al.* Identification of a common neurobiological substrate for mental illness. *JAMA Psychiatry* **72**, 305–315 (2015).
91. Price, J. L. & Drevets, W. C. Neurocircuitry of mood disorders. *Neuropsychopharmacology* **35**, 192–216 (2010).
92. Drevets, W. C. *et al.* Subgenual prefrontal cortex abnormalities in mood disorders. *Nature* **386**, 824–827 (1997).
93. Boes, A. D., McCormick, L. M., Coryell, W. H. & Nopoulos, P. Rostral anterior cingulate cortex volume correlates with depressed mood in normal healthy children. *Biol. Psychiatry* **63**, 391–397 (2008).
94. Levesque, M. L. *et al.* Altered patterns of brain activity during transient sadness in children at familial risk for major depression. *J. Affect. Disord.* **135**, 410–413 (2011).
95. Nieuwenhuizen, A. G. & Rutters, F. The hypothalamic–pituitary–adrenal-axis in the regulation of energy balance. *Physiol. Behav.* **94**, 169–177 (2008).
96. Gold, P. W. & Chrousos, G. P. Organization of the stress system and its dysregulation in melancholic and atypical depression: high versus low CRH/NE states. *Mol. Psychiatry* **7**, 254–275 (2002).
97. Carroll, B. J. *et al.* Pathophysiology of hypercortisolism in depression: pituitary and adrenal responses to low glucocorticoid feedback. *Acta Psychiatr. Scand.* **125**, 478–491 (2012).
98. Dantzer, R., O'Connor, J. C., Freund, G. G., Johnson, R. W. & Kelley, K. W. From inflammation to sickness and depression: when the immune system subjugates the brain. *Nat. Rev. Neurosci.* **9**, 46–56 (2008).
99. Critchley, H. D. & Harrison, N. A. Visceral influences on brain and behavior. *Neuron* **77**, 624–638 (2013).
100. Harrison, N. A. *et al.* Neural origins of human sickness in interoceptive responses to inflammation. *Biol. Psychiatry* **66**, 415–422 (2009).
101. Paulus, M. P. & Stein, M. B. Interoception in anxiety and depression. *Brain Struct. Funct.* **214**, 451–463 (2010).
102. Ryan, J. P., Sheu, L. K., Critchley, H. D. & Gianaros, P. J. A neural circuitry linking insulin resistance to depressed mood. *Psychosom. Med.* **74**, 476–482 (2012).
103. Buckholtz, J. W. *et al.* Genetic variation in MAOA modulates ventromedial prefrontal circuitry mediating individual differences in human personality. *Mol. Psychiatry* **13**, 313–324 (2008).
104. Raison, C. L. & Miller, A. H. Malaise, melancholia and madness: the evolutionary legacy of an inflammatory bias. *Brain Behav. Immun.* **31**, 1–8 (2013).
105. Dantzer, R., Heijnen, C. J., Kavelaars, A., Laye, S. & Capuron, L. The neuroimmune basis of fatigue. *Trends Neurosci.* **37**, 39–46 (2014).
106. Etkin, A. & Wager, T. D. Functional neuroimaging of anxiety: a meta-analysis of emotional processing in PTSD, social anxiety disorder, and specific phobia. *Am. J. Psychiatry* **164**, 1476–1488 (2007).
107. Mayberg, H. S. *et al.* Deep brain stimulation for treatment-resistant depression. *Neuron* **45**, 651–660 (2005).
108. Kennedy, S. H. *et al.* Deep brain stimulation for treatment-resistant depression: follow-up after 3 to 6 years. *Am. J. Psychiatry* **168**, 502–510 (2011).
109. Riva-Posse, P. *et al.* Defining critical white matter pathways mediating successful subcallosal cingulate deep brain stimulation for treatment-resistant depression. *Biol. Psychiatry* **76**, 963–969 (2014).

110. Hamani, C. *et al.* Deep brain stimulation of the subcallosal cingulate gyrus for depression: anatomical location of active contacts in clinical responders and a suggested guideline for targeting. *J. Neurosurg.* **111**, 1209–1215 (2009).

111. Goldapple, K. *et al.* Modulation of cortical–limbic pathways in major depression: treatment-specific effects of cognitive behavior therapy. *Arch. Gen. Psychiatry* **61**, 34–41 (2004).

112. Aupperle, R. L. *et al.* Neural responses during emotional processing before and after cognitive trauma therapy for battered women. *Psychiatry Res.* **214**, 48–55 (2013).

113. McGrath, C. L. *et al.* Toward a neuroimaging treatment selection biomarker for major depressive disorder. *JAMA Psychiatry* **70**, 821–829 (2013).

114. McGrath, C. L. *et al.* Pretreatment brain states identify likely nonresponse to standard treatments for depression. *Biol. Psychiatry* **76**, 527–535 (2014).

115. Moussavi, S. *et al.* Depression, chronic diseases, and decrements in health: results from the World Health Surveys. *Lancet* **370**, 851–858 (2007).

116. Smallwood, R. F. *et al.* Structural brain anomalies and chronic pain: a quantitative meta-analysis of gray matter volume. *J. Pain* **14**, 663–675 (2013).

117. Ganzel, B. L., Morris, P. A. & Wethington, E. Allostasis and the human brain: integrating models of stress from the social and life sciences. *Psychol. Rev.* **117**, 134–174 (2010).

118. Piazza, J. R., Charles, S. T., Sliwinski, M. J., Mogle, J. & Almeida, D. M. Affective reactivity to daily stressors and long-term risk of reporting a chronic physical health condition. *Ann. Behav. Med.* **45**, 110–120 (2013).

119. Dannlowski, U. *et al.* Limbic scars: long-term consequences of childhood maltreatment revealed by functional and structural magnetic resonance imaging. *Biol. Psychiatry* **71**, 286–293 (2012).

120. Teff, K. L. How neural mediation of anticipatory and compensatory insulin release helps us tolerate food. *Physiol. Behav.* **103**, 44–50 (2011).

121. Levinthal, D. J. & Strick, P. L. The motor cortex communicates with the kidney. *J. Neurosci.* **32**, 6726–6731 (2012).

122. Deco, G., Jirsa, V. K. & McIntosh, A. R. Emerging concepts for the dynamical organization of resting-state activity in the brain. *Nat. Rev. Neurosci.* **12**, 43–56 (2011).

123. Hermundstad, A. M. *et al.* Structural foundations of resting-state and task-based functional connectivity in the human brain. *Proc. Natl Acad. Sci. USA* **110**, 6169–6174 (2013).

124. Pernice, V., Staude, B., Cardanobile, S. & Rotter, S. How structure determines correlations in neuronal networks. *PLoS Comput. Biol.* **7**, e1002059 (2011).

125. van den Heuvel, M. P., Mandl, R. C., Kahn, R. S. & Hulshoff Pol, H. E. Functionally linked resting-state networks reflect the underlying structural connectivity architecture of the human brain. *Hum. Brain Mapp.* **30**, 3127–3141 (2009).

126. Huber, L. *et al.* Cortical lamina-dependent blood volume changes in human brain at 7T. *Neuroimage* **107**, 23–33 (2015).

127. de-Wit, L., Machilsen, B. & Putzeys, T. Predictive coding and the neural response to predictable stimuli. *J. Neurosci.* **30**, 8702–8703 (2010).

128. Rempel-Clower, N. L. & Barbas, H. The laminar pattern of connections between prefrontal and anterior temporal cortices in the rhesus monkey is related to cortical structure and function. *Cereb. Cortex* **10**, 851–865 (2000).

129. Medalla, M. & Barbas, H. Specialized prefrontal "auditory fields": organization of primate prefrontal–temporal pathways. *Front. Neurosci.* **8**, 77 (2014).

130. Medalla, M. & Barbas, H. Diversity of laminar connections linking periarcuate and lateral intraparietal areas depends on cortical structure. *Eur. J. Neurosci.* **23**, 161–179 (2006).

131. Hilgetag, C. C. & Grant, S. Cytoarchitectural differences are a key determinant of laminar projection origins in the visual cortex. *Neuroimage* **51**, 1006–1017 (2010).

132. Goulas, A., Uylings, H. B. & Stiers, P. Mapping the hierarchical layout of the structural network of the macaque prefrontal cortex. *Cereb. Cortex* **24**, 1178–1194 (2014).

133. Felleman, D. J. & Vas Essen, D. C. Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* **1**, 1–47 (1991).

134. Markov, N. T. *et al.* Cortical high-density counterstream architectures. *Science* **342**, 1238406 (2013).

135. Sanides, F. in *Advances in Primatology* (eds Noback, C. H. & Montagna, W.) 137–208 (Appleton-Century-Crofts, 1970).

136. Nieuwenhuys, R. The insular cortex: a review. *Prog. Brain Res.* **195**, 123–163 (2012).

137. Lovero, K. L., Simmons, A. N., Aron, J. L. & Paulus, M. P. Anterior insular cortex anticipates impending stimulus significance. *Neuroimage* **45**, 976–983 (2009).

138. Holtz, K., Pane-Farre, C. A., Wendt, J., Lotze, M. & Hamm, A. O. Brain activation during anticipation of interoceptive threat. *Neuroimage* **61**, 857–865 (2012).

139. Wise, S. P. Forward frontal fields: phylogeny and fundamental function. *Trends Neurosci.* **31**, 599–608 (2008).

140. Terasawa, Y., Shibata, M., Moriguchi, Y. & Umeda, S. Anterior insular cortex mediates bodily sensibility and social anxiety. *Soc. Cogn. Affect. Neurosci.* **8**, 259–266 (2013).

141. Price, C. J. & Friston, K. J. Degeneracy and cognitive anatomy. *Trends Cogn. Sci.* **6**, 416–421 (2002).

142. Edelman, G. M. & Gally, J. A. Degeneracy and complexity in biological systems. *Proc. Natl Acad. Sci. USA* **98**, 13763–13768 (2001).

143. Feinstein, J. S. *et al.* Preserved emotional awareness of pain in a patient with extensive bilateral damage to the insula, anterior cingulate, and amygdala. *Brain Struct. Funct.* http://dx.doi.org/10.1007/s00429-014-0986-3 (2015).

144. Philippi, C. L. *et al.* Preserved self-awareness following extensive bilateral brain damage to the insula, anterior cingulate, and medial prefrontal cortices. *PLoS ONE* **7**, e38413 (2012).

145. Damasio, A., Damasio, H. & Tranel, D. Persistence of feelings and sentience after bilateral damage of the insula. *Cereb. Cortex* **23**, 833–846 (2013).

146. Khalsa, S. S., Rudrauf, D., Feinstein, J. S. & Tranel, D. The pathways of interoceptive awareness. *Nat. Neurosci.* **12**, 1494–1496 (2009).

147. Markov, N. T. *et al.* A weighted and directed interareal connectivity matrix for macaque cerebral cortex. *Cereb. Cortex* **24**, 17–36 (2014).

148. Finlay, B. L. & Uchiyama, R. Developmental mechanisms channeling cortical evolution. *Trends Neurosci.* **38**, 69–76 (2015).

149. Barbas, H. General cortical and special prefrontal connections: principles from structure to function. *Ann. Rev. Neurosci.* http://dx.doi.org/ 10.1146/ annurev-neuro-071714-033936 (2015).

Competing interests statement
The authors declare no competing interests.

SUPPLEMENTARY INFORMATION
See online article: S1 (box) | S2 (box) | S3 (box)
ALL LINKS ARE ACTIVE IN THE ONLINE PDF