



CrossMark
click for updates

Research

Cite this article: Barrett LF, Quigley KS, Hamilton P. 2016 An active inference theory of allostasis and interoception in depression. *Phil. Trans. R. Soc. B* **371**: 20160011. <http://dx.doi.org/10.1098/rstb.2016.0011>

Accepted: 19 August 2016

One contribution of 16 to a theme issue 'Interoception beyond homeostasis: affect, cognition and mental health'.

Subject Areas:

neuroscience

Keywords:

interoception, visceromotor, major depressive disorder, fMRI, prediction

Author for correspondence:

Lisa Feldman Barrett
e-mail: l.barrett@neu.edu

An active inference theory of allostasis and interoception in depression

Lisa Feldman Barrett^{1,2,3}, Karen S. Quigley¹ and Paul Hamilton⁴

¹Department of Psychology, Northeastern University, Boston, MA, USA

²Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital and Harvard Medical School, Charlestown, MA, USA

³Department of Psychiatry, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA

⁴Center for Social and Affective Neuroscience, Department of Clinical and Experimental Medicine, Linköping University, Linköping, Sweden

LFB, 0000-0003-4478-2051

In this paper, we integrate recent theoretical and empirical developments in predictive coding and active inference accounts of interoception (including the Embodied Predictive Interoception Coding model) with working hypotheses from the theory of constructed emotion to propose a biologically plausible unified theory of the mind that places metabolism and energy regulation (i.e. allostasis), as well as the sensory consequences of that regulation (i.e. interoception), at its core. We then consider the implications of this approach for understanding depression. We speculate that depression is a disorder of allostasis, whose myriad symptoms result from a 'locked in' brain that is relatively insensitive to its sensory context. We conclude with a brief discussion of the ways our approach might reveal new insights for the treatment of depression.

This article is part of the themed issue 'Interoception beyond homeostasis: affect, cognition and mental health'.

1. Introduction

Since ancient times, the human mind has been understood as a collection of mental faculties for thinking (cognitions), feeling (emotions) and volition (actions, and in more modern versions, perceptions). These categories come not from biology, but from the philosophical concerns about truth, beauty and ethics that anchor Western theories of human nature. But the human brain did not evolve to think or feel or see. Several decades of research points to a different hypothesis: the human brain has a common computational architecture that, first and foremost, supports the human body as it moves, grows, survives and reproduces [1]. As a consequence, *metabolism* and other forms of *energy regulation* may be at the core of the human mind, regardless of whether a person is thinking, feeling or perceiving. An emerging theoretical framework, centred on energy regulation, rests on three important insights. First, brains do not react to the world, but instead predict and then test their hypotheses against incoming sensory evidence. Their hypotheses constitute *internal models of the body in the world* that are constructed via Bayesian inferences constrained by sensory inputs, from which all perceptions and actions emerge (for a discussion of the computational details, see [2,3]). All animal brains, not just those found in a human body, host an internal model [1]. Collectively, these ideas are referred to as predictive coding, active inference or belief propagation accounts of brain function (e.g. [4–10]); informative discussions of these accounts are found in several other papers in this issue [2,11]. Second, a human brain's internal model (i.e. predictions) constructs all varieties of cognitions, emotions and perceptions, and guides actions, but the computational architecture for prediction did not evolve for these purposes. Predictions fundamentally serve to maintain energy balance, i.e. allostasis. Allostasis is not a condition or state of the body—it is how the brain efficiently maintains energy regulation in the body [1,12]. Allostasis is defined in terms of prediction: a brain maintains energy regulation by anticipating the body's needs

and preparing to satisfy those needs before they arise [1,12–14]. For example, allostasis describes the brain's capacity both to predict that to start running requires more oxygen in the body's striate muscles, as well as to mobilize the needed resources by increasing cardiac output, redistributing blood flow from organs that can spare oxygen (e.g. the stomach), etc. Third, these predictions cause changes in the body's internal systems (in humans, these include the immune, endocrine and autonomic nervous systems) and the sensations that arise from those changes are called *interoception* [15]. Interoceptive sensations are routinely experienced as lower dimensional feelings of affect [16,17]. If interoception plays a role in allostasis, and allostasis is at the core of the brain's computational architecture, then the properties of affect—valence and arousal [18,19]—are best thought of as basic features of consciousness [20–26], rather than properties of emotion *per se*. These insights offer the possibility of better understanding why affective dysregulation (and correspondingly, interoceptive difficulties) constitute a transdisorder vulnerability to mental illness [27], are key factors in mood disorders [28] and are common in physical illness [29]; for a discussion, see [11].

Recently, a powerful predictive coding account of interoception has emerged [2,11,30–33]. This account, when integrated with the insights above, creates a powerful framework for re-examining how the computational architecture for prediction can produce affective dysregulation (e.g. [5,17]). The purpose of this paper is to discuss these exciting and potentially transformative ideas with reference to the development and treatment of depression. We begin by considering predictive coding accounts of allostasis and interoception, briefly outlining our hypotheses about the brain's computational architecture with a focus on allostasis and interoception as core features. We then introduce the theory of constructed emotion, and present several novel hypotheses that unite the metabolic, mood and vegetative symptoms that occur in depression (and in other illnesses as well) within a common computational framework. We conclude by exploring the implications for treatment.

2. Energy regulation is at the core of the brain's computational architecture

There is a virtual revolution emerging in our understanding of brain structure and function, nicely illustrated by several papers in this special issue (i.e. [2,11]). It begins with a simple idea: maintaining a body is expensive. A body must be watered, fed and cared for, so that it can grow, thrive and ultimately, reproduce and care for its young. Growth, survival and reproduction (and therefore gene transmission) depend on the near continual intake of energy resources (metabolic and otherwise). Further, the physical movements necessary to move around in the world and acquire those resources in the first place (and protect against threats and dangers) require upfront energy expenditures which in mammals include spending resources such as glucose, water, oxygen, electrolytes, etc. To flourish, an animal must balance energy expenditures with deposits and see a return on its resource investments, not just in the quality and quantity of resources acquired, but also in having enough surplus energy to encode and consolidate the details of experience, making those experiences available within the brain's synaptic connections to guide future decisions about expenditures and deposits. From the brain's perspective, then, its body

and the world beyond are a system within which the body's overall metabolism and energy regulation must be managed. With these observations in mind, we can think of a brain as hosting an internal model of the world from the perspective of its body's physiological needs (following the well-known cybernetics principle that anything which regulates (i.e. acts on) a system must contain an *internal model* of that system [34]). An internal model is implemented by intrinsic brain activity (e.g. [35–38]) that, in humans, uses a whopping 20% of the total energy consumed [39].¹ The novel consideration here is that allostasis, and its sensory consequences (i.e. interoception), are core features of this model.

Allotasis is the capacity to vary physiological systems flexibly according to predicted energy demands [1]. *Efficiency* requires the ability to *anticipate* the body's needs and *prepare* to satisfy them before they arise [1,12].² Too much of a resource (e.g. obesity) or not enough (e.g. fatigue, [44]) is suboptimal, and inefficient. Prolonged imbalances can lead to illness (e.g. [45,46]) that remodels the brain, specifically causing atrophy in the regions that subserve allostasis [47,48] and causing increased arborization of the sympathetic nervous system, leading to enhanced sympathetic reactivity [49,50]. For a review, see [51]. This in turn makes physiological regulation even less efficient and therefore more metabolically expensive.

A broad range of evidence supports the hypothesis that the primary task of a brain is to implement allostasis in the service of efficient metabolism and energy regulation. First, there is the signal processing evidence across species that to enhance efficiency, brains send only the minimum data necessary, as efficiently as possible, and using the least possible neural wiring [1].³ Second, there is both structural and functional brain imaging evidence in humans that reveals a system which integrates allostasis and interoception at the core of the brain's intrinsic architecture [52,53] (figure 1). This system consists of two intrinsic networks, conventionally called the default mode network and the salience network, connected by an ensemble of the brain's 'rich club' hubs. These hubs belong to a community of densely interconnected regions that make up the structural core of the brain that serves as a high-capacity backbone for synchronizing information flow [54–60]. Rich clubs appear across species as diverse as roundworms, fruit flies and humans [61]. In primates, rich club regions have more complex pyramidal cell structure (larger dendritic branches, larger spines, etc.; [62]) and possess more excitatory (relative to inhibitory) chemoarchitecture [63]. Furthermore, older studies that applied strychnine to the exposed cortices of macaque monkeys (to block local GABA receptors and temporarily increase excitatory signals sent from the treated cortical area) reveal that rich club hubs have stronger connections that together have a net excitatory effect on other regions of cortex [64]. Similarly, higher intrinsic connectivity involving the pregenual anterior cingulate cortex (a rich club hub) is associated with higher glutamate and glutamine concentrations in that cortical site [65]. These findings are consistent with other brain imaging evidence showing that networks flexibly adjust their connectivity to one another via the rich club hubs to prepare for and adjust to changing task demands [66,67]. The default mode and salience networks are 'multi-use' or 'domain general' networks that are routinely engaged in a wide variety of tasks spanning almost every phenomenon and task domain within psychology [68,69]. Others have even argued that consciousness is enabled by

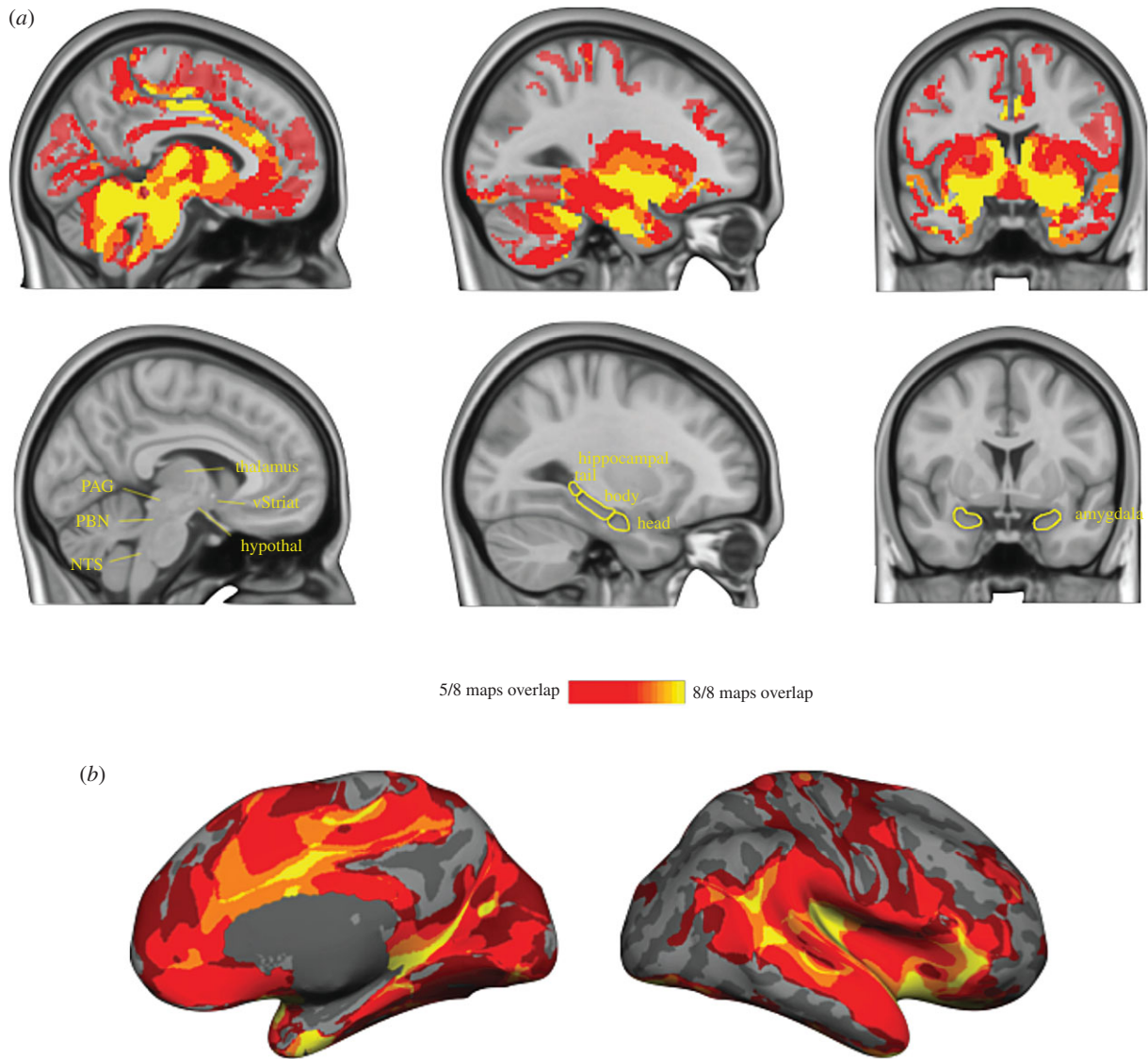


Figure 1. A large-scale system for allostasis in the human brain. We consulted anterograde and retrograde tract-tracing studies in macaque monkeys to select eight seed regions in limbic cortices with monosynaptic connections to midbrain and brainstem regions that are known to control the immune, endocrine and autonomic nervous systems in the service of allostasis (for details and coordinates, see [52]). For each seed region, we computed a ‘discovery map’ of voxels whose timecourse correlated with the seed region. (a) A conjunction of all eight maps presented in the volume to display subcortical regions. (b) A conjunction of maps depicted on the cortical surface. (c) Cluster analysis of the eight discovery maps revealed the system for allostasis was composed of two large-scale intrinsic networks (shown in red and blue) that share several hubs (shown in purple). Hubs belonging to the brain’s ‘rich club’ are labelled in yellow. Rich club hubs figure adapted with permission from [54]. Maps were constructed with resting state BOLD data from 280 participants binarized at $p < 10^{-5}$, and then replicated on a second sample of 270 participants. aMCC, anterior midcingulate cortex; dplns, dorsal posterior insula; IFG, inferior frontal gyrus; ITG, inferior temporal gyrus; valns, ventral anterior insula; MCC, midcingulate cortex; PHG, parahippocampal gyrus; pMCC, posterior midcingulate cortex; PostCG, postcentral gyrus; STS, superior temporal sulcus. (d) Reliable subcortical connections, thresholded $p < 0.05$ uncorrected. PAG, periaqueductal grey; hypothal, hypothalamus; PBN, parabrachial nucleus; vStriat, ventral striatum; NTS, nucleus of the solitary tract.

the sort of information integration provided by the hubs of the rich club [22].

Recently, we [30,70], along with others [2,11,31–33], have proposed a computational architecture for the brain that places allostasis and interoception at the core of its internal model. This assertion is based on integrating several different lines of research. First, decades of anatomical and functional research shows that the five exteroceptive sensory systems work via predictive inference (for a review of evidence, see [70]; for recent ultra-high field brain imaging evidence, see [71]), as does the motor system [72–76]. The brain predicts the timing and content of sensory events [77]. These findings strongly support the increasingly popular hypothesis that the brain constructs *embodied simulations* [78,79] that function as Bayesian filters [3,80] for incoming sensory input, *driving*

action and *constructing* perception. We integrated this view with a second line of research that has produced a well-validated neuroanatomical model for information flow within the brain [81–83]. Together, these research findings allowed us to propose that simulations, as ongoing, intrinsic activity, function as *prediction signals* (also known as ‘top-down’ or ‘feedback’ signals, and more recently as ‘forward’ models) that serve as plans for allostasis by continuously anticipating sensory events in the body and in the outer environment.⁴ Unanticipated information (prediction error) from both internal and external sensory domains modulates the predictions (also known as ‘bottom-up’ or, confusingly, ‘feedforward’ signals). Error signals track the difference between the predicted sensations and those that are incoming from the sensory world (including the body’s peripheral

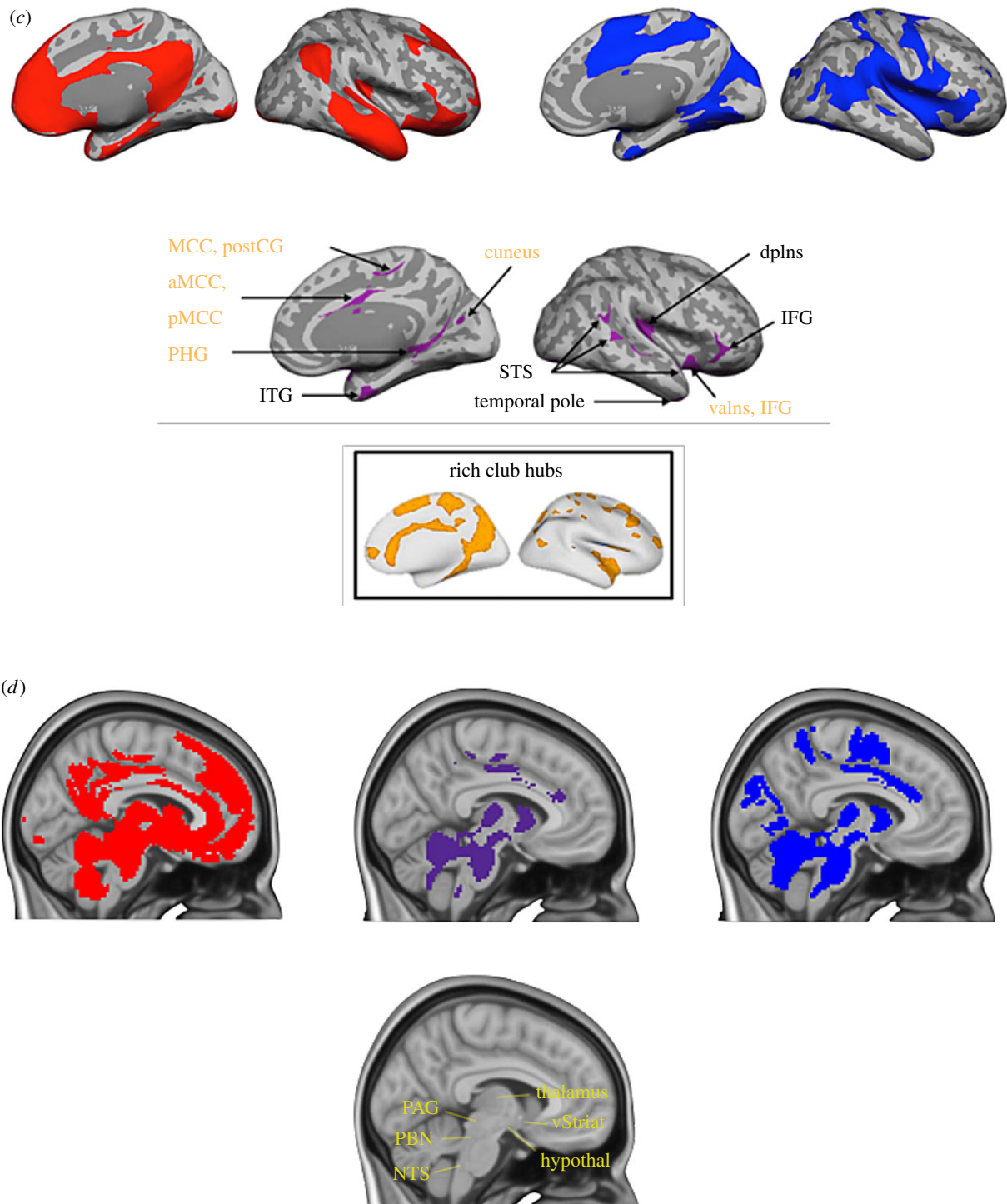


Figure 1. (Continued.)

physiological systems). The brain has various mechanisms for reducing prediction errors, and once they are sufficiently minimized, predictions (i.e. simulations) become the inferences about the causes of sensory events and plans for how to move (or not to move) the body to deal with them [7,91]. By modulating ongoing visceromotor actions (i.e. inner movements associated with the immune, endocrine and autonomic nervous systems) and motor movements to deal with upcoming sensory events, a brain infers their likely causes. In our view, then, sensory predictions arise from allostasis, and therefore allostasis (and interoception) guide all mental function. Sensory prediction errors (i.e. learning) are treated, at a very basic level, as information that guides a *predicted* allostatic plan. Whatever else the brain might be

doing—thinking, seeing, tasting—it is also predictively regulating the body's physiological systems in the service of allostasis.

These various sources of evidence allowed us to propose a very specific computational architecture for implementing allostasis, guiding action and constructing perception [30,70], and we further develop this framework here (figure 2). Specifically, we proposed that prediction signals originate in cortical regions that have the least well-developed laminar structure, referred to as *agranular* cortices. The most agranular cortical sites are cytoarchitecturally arranged to send but not receive cortical prediction signals. Another name for agranular cortices is *limbic*. Limbic cortices, such as the cingulate cortices and the ventral portion of the anterior insula, as well as dysgranular

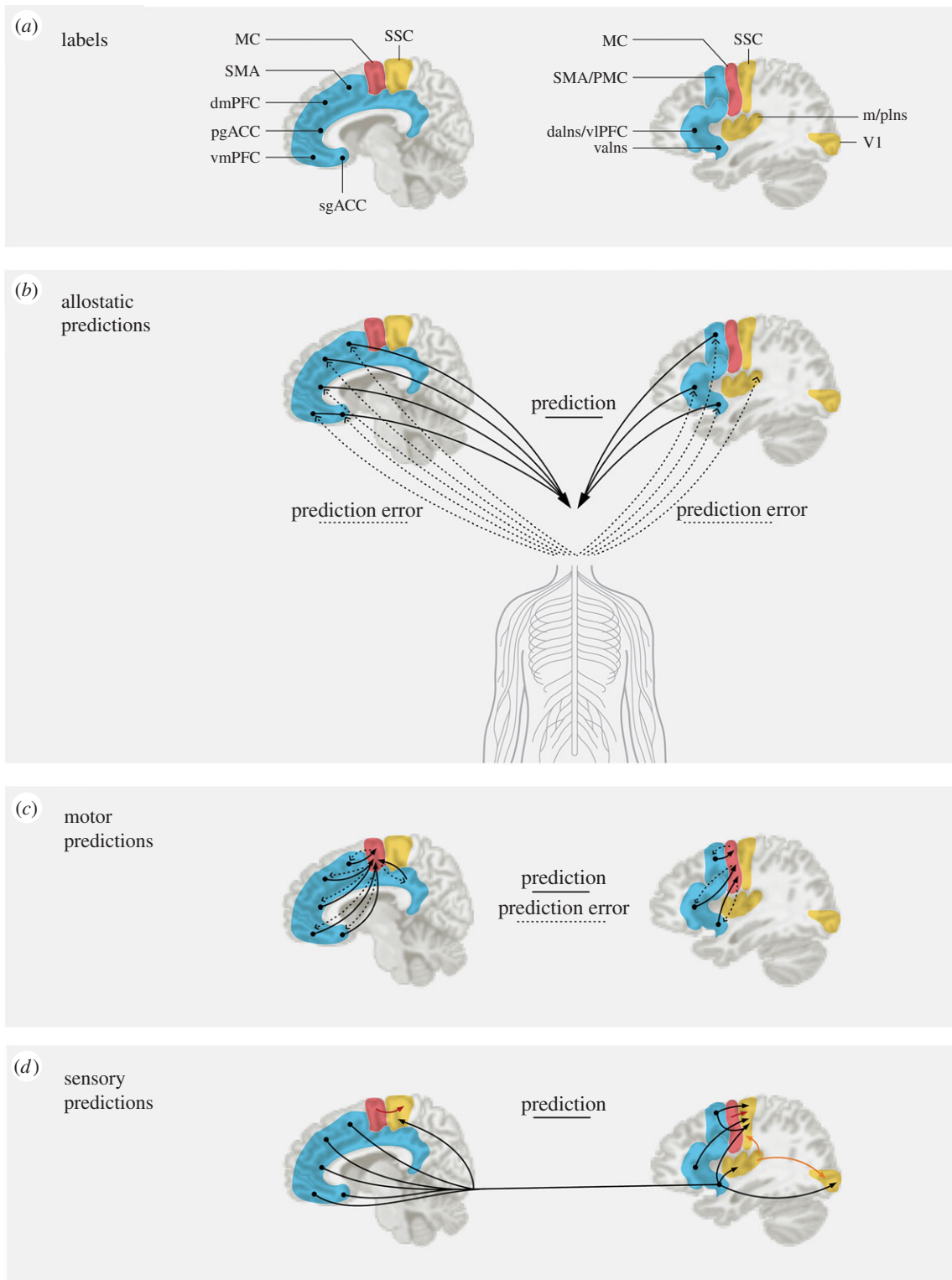


Figure 2. A depiction of predictive coding in the human brain. (a) We identified key limbic cortices (in blue) that provide cortical control of the body's internal milieu. Primary motor cortex is depicted in red, and primary sensory regions are in yellow. For simplicity, only primary visual, interoceptive and somatosensory cortices are shown; subcortical regions are not shown. (b) Limbic cortices initiate allostatic predictions to the hypothalamus and brainstem nuclei (e.g. periaqueductal grey, parabrachial nucleus, nucleus of the solitary tract) to regulate the autonomic, neuroendocrine and immune systems (solid lines). The incoming sensory inputs from the internal milieu of the body are carried along the vagus nerve and small diameter C and $A\delta$ fibres to limbic regions (dotted lines). Comparisons between prediction signals and ascending sensory input results in prediction error that is available to update the brain's internal model. In this way, prediction errors are learning signals and can adjust subsequent predictions. (c) Efferent copies of allostatic predictions are sent to motor cortex as motor predictions (solid lines) and prediction errors are sent from motor cortex to limbic cortices (dotted lines). (d) Sensory cortices receive sensory predictions from several sources. They receive efferent copies of allostatic predictions (black lines) and efferent copies of motor predictions (red lines). Sensory cortices with less well-developed lamination (e.g. primary interoceptive cortex) also send sensory predictions to sensory cortices that are more well developed (e.g. in this figures, somatosensory and primary visual cortices) (orange lines). For simplicity's sake, prediction errors are not depicted in panel (d). sgACC, subgenual anterior cingulate cortex; vmPFC, ventromedial prefrontal cortex; pgACC, pregenual anterior cingulate cortex; dmPFC, dorsomedial prefrontal cortex; MCC, midcingulate cortex, is ventral to dmPFC and SMA; valns, ventral anterior insula; dalns, dorsal anterior insula; vIPFC, ventrolateral prefrontal cortex; SMA, supplementary motor area; PMC, premotor cortex; m/plns, mid/posterior insula (primary interoceptive cortex); SSC, somatosensory cortex; V1, primary visual cortex; and MC, motor cortex (for relevant neuroanatomy references, see [52]).

cortical sites that project to subcortical regions controlling allostasis, such as medial prefrontal cortex, ventrolateral prefrontal cortex, and parts of temporal and parietal cortex, provide the substrate for allostasis (figure 2*a*). Descending allostatic prediction signals are relayed to the body's physiological systems via a collection of subcortical regions [52,53], including the central nucleus of the amygdala [92], the ventral and dorsal striatum, and the central pattern generators [93] of the hypothalamus, the parabrachial nucleus, periaqueductal grey and the solitary nucleus (figure 2*b*).⁵ The hippocampus very likely has a predictive role to play in allostasis as well (e.g. [95,96]).

Following the anatomic principles of information flow [81–83], we hypothesized that efferent copies of allostatic signals cascade to primary motor cortex as skeletomotor prediction signals, as well as to all primary sensory cortices as sensory prediction signals (see figure 2*c,d*, respectively; for a discussion, see [30,70,72,80]). Tract-tracing evidence indicates that prediction signals flow from deep layers of limbic cortices and terminate in the upper layers of cortical regions with more developed (i.e. more granular) structure, such as gustatory and olfactory cortex, primary motor cortex, primary interoceptive cortex, and the primary visual, auditory and somatosensory regions (which have the most developed laminar organization). Some prediction signals are conveyed directly to primary sensory regions via metabolically more expensive long-range connections, whereas others become progressively more detailed as they cascade through numerous synaptic relays (such that the posterior probabilities in the sending cortex become the priors in the receiving cortex in Bayesian terms).⁶

Because motor cortex has a laminar organization that is less well developed than primary visual, auditory, somatosensory and interoceptive regions [101], we hypothesize that motor cortex sends efferent copies to those sensory regions as sensory predictions (figure 2*d*, solid red paths; see [72]). Furthermore, because of their differential laminar organization, we hypothesize that primary interoceptive cortex in mid-to-posterior dorsal insula forwards sensory predictions to visual, auditory and somatosensory cortices which propagate across either a single or multiple synapses (figure 2*d*, solid orange paths). The skeletomotor prediction signals prepare the body for movement, whereas the interoceptive prediction signals are represented as a change in affect (i.e. the expected sensory consequences within the body) and the extrapersonal sensory prediction signals prepare upcoming perceptions. This ensemble of hypotheses is consistent not only with over three decades of tract-tracing studies in non-human animals, but also with engineering design principles (i.e. compute locally, and relay only the information that is needed to assemble a larger pattern; [1]).

As prediction signals cascade across the synapses within a brain, incoming sensory signals arriving to the brain (i.e. from the external environment and the internal periphery) simultaneously allow for computations of prediction error that are encoded to update the internal model (correcting visceromotor and motor action plans, as well as sensory representations; see figure 2, dotted paths). Sensory signals arise from changes within the body's physiological systems and ascend via vagal afferents and small diameter afferents in the dorsal horn of the spinal cord, through the nucleus of the solitary tract, the parabrachial nucleus, the periaqueductal grey and finally to the ventral posterior thalamus, before arriving in granular layer IV of the primary interoceptive insular cortex [15,102].

Prediction errors also arise within the amygdala, the basal ganglia and the cerebellum and are forwarded to the cortex to correct its internal model [52,103–105].

Within this framework, we hypothesize that information flowing from the amygdala to the cortex is not 'emotional' *per se*, but signals uncertainty [106] about the predicted sensory input (via the basolateral complex) and helps to adjust physiological functions in support of allostasis (via the central nucleus) as a result.⁷ The arousal signals that are associated with increases in amygdala activity (e.g. [108]) can be considered as learning signals [109]. Similarly, prediction errors from the ventral striatum to the cortex (referred to as 'reward prediction errors' [110]) convey information about sensory inputs that impact allostasis more than expected (i.e. indicating that this information should be encoded and consolidated in the cortex, and acted upon immediately). Dopamine is hypothesized to support vigorous action and learning that is necessary to secure the rewards that maintain efficient allostasis (or restore it in the event of disruption), rather than playing a necessary or sufficient role in rewards themselves [111,112]. Other neuromodulators, such as opioids, may be more intrinsically rewarding (e.g. [113]).

The cerebellum models sensory prediction errors from the periphery and relays them to cortex to rapidly modify motor predictions (i.e. it is hypothesized to predict the sensory consequences of a motor command much faster than actual sensory prediction errors can be received [74], and helps the cortex reduce the sensory consequences caused by one's own movements). The cerebellum may have the same role to play for allostatic predictions given the connectivity between the cerebellum and cingulate cortices, hypothalamus and the amygdala [104,114–116].⁸ This would give the cerebellum a major role in allostasis.

The limbic cortices that guide allostasis fall within the traditional territory of three intrinsic networks within the brain. We have discussed two of them already (figure 1). The first is the default mode network, which we hypothesize generatively uses prior experiences to construct the brain's internal model. This proposal is consistent with other proposals that the default mode network constructs mental models of the world from different points of view and different time points [36–38]. If a simulation is an embodied brain state, then the default mode network 'initiates' simulations and represents part of their pattern; its multimodal sensorimotor summaries become more detailed and particularized as they cascade out to primary sensory and motor regions.

We further hypothesize that the limbic cortices within the salience network send predictions that adjust the internal model to the conditions of the sensory periphery, again in the service of allostasis. This is consistent with the salience network's role in attention regulation; (e.g. [117–120]). We specifically propose that the salience network tunes the internal model by anticipating which prediction errors are likely to be allostatically relevant and therefore worth the metabolic cost of encoding and consolidation [121], and then modulating the gain on those errors accordingly.⁹ These predictions are called *precision signals* [123–126]. Precision signals optimize the sampling of the sensory periphery for allostasis. Via their core position in the brain's rich club, and their role in multisensory integration [57], the salience network's precision signals apply attention to every sensory system in the brain (this is sometimes called 'affective' attention). Precision signals directly alter the gain on neurons as

they compute prediction error from incoming sensory input (i.e. they apply attention to signal the degree of confidence in the reliability or quality of incoming sensory signals, and/or the predicted relevance for allostasis). Unexpected sensory inputs that are anticipated to have allostatic implications (because they are likely to impact survival, offering reward or threat, or are of uncertain value) will be treated as ‘signal’ and learned (i.e. encoded) to better predict energy needs in the future, with all other prediction error treated as ‘noise’ and safely ignored ([109]; for discussion, see [17]). Limbic regions within the salience network can also indirectly signal the precision of incoming sensory inputs via their modulation of the reticular nucleus that encircles that thalamus and controls the sensory input that reaches the cortex via thalamocortical pathways (for relevant anatomy, see [127–129]).

In a healthy brain, prediction error signals allow for learning to better tailor the brain’s internal model to the immediate circumstances and, thereby, to minimize future error and improve efficient allostasis. As we propose later in this paper, however, chronic prediction errors of a certain magnitude and frequency could constitute a transdisorder vulnerability to illness. In our view, the management of precision is a key computation that is relevant in the development and maintenance of depression.

Importantly, the salience network helps accomplish multimodal integration (its spatial topography strongly overlaps with the multimodal integration network as documented in [57]). Moreover, as we have documented, primary interoceptive cortex (in the dorsal mid to posterior insula) is a component of the salience network, ensuring that every mental event (not just emotions) is infused with interoception, which is made available to consciousness as affect. This state of affairs provides the recipe for affective realism, where people experience supposed facts about the world that are created in part by interoception and the associated affective feelings [17]. Food is ‘delicious’ or ‘distasteful’. People are ‘nice’ or ‘mean.’ Affective realism leads people to believe that objects and people in the world are inherently negative or positive.

Finally, we hypothesize that neurons within the frontoparietal control network sculpt and maintain simulations for longer than the several hundred milliseconds it takes to process imminent prediction errors. We hypothesize that they apply attention to adjust the degree of confidence in sensory predictions (i.e., adjusting priors) and they may also help to suppress or inhibit simulations whose priors are very low. It pays to be flexible, to construct, maintain and use patterns that extend over longer periods of time (different animals have different timescales that are relevant for their behavioural repertoire and ecological niche). It is also valuable to learn on a single trial, without the need for recurring statistical regularities in the world, particularly if you reside in a quickly changing environment or when the prediction error was large. As a prediction generator, the brain is constructing simulations across many different time-scales (i.e. it is integrating information across the few moments that constitute an event, but also across longer time frames at various scales; for similar ideas, see [96,130]). The frontoparietal control network (which contains key limbic rich-club hubs in the mid cingulate cortex and anterior insula) also may have a role to play in managing sensory prediction errors, by applying attention to select those body movements that will generate the expected sensory

inputs, presumably with help from cerebellar and striatal prediction errors. These movements then generate the sensory inputs that reduce prediction error and confirm an existing prediction. This dynamic may have importance for depression; as we will see, physical movements are often compromised in the most severe cases of depression, removing a mechanism for reducing prediction error.

3. The theory of constructed emotion

Thus far, we have proposed that the brain’s internal model consists of embodied, whole brain representations that predict what is about to happen in the external environment, the best course of action for dealing with these impending events, and their consequences for allostasis. The implication is that all perception (i.e. the ‘meaning’ of sensory events) contains interoceptive representations that are a consequence of allostasis. The processing of prediction errors is also guided by allostasis-relevant predictions (i.e. precision estimates, or ‘affective’ attention). These working hypotheses, and the anatomical and functional evidence that supports them, fill the computational and neural gaps in initial theoretical formulations in the conceptual act theory of emotion [131–135], transforming it into a unified theory of mind and brain that provides novel hypotheses about how a brain constructs emotional events [17].

Specifically, the unified theory proposes that the brain’s internal model runs on concepts, constructed as prediction signals. Traditionally, a *category* is a population of events or objects that are treated as similar because they all serve a particular goal in some context; a *concept* is the population of representations that correspond to those events or objects [136]. Evidence indicates that the brain prepares multiple competing predictions (with associated energy costs and potential rewards) before deciding between them and implementing one [137]. Because the brain regulates physiological systems to proactively provide the energy necessary for motor movements, our complimentary hypothesis is that it assembles a population of predictions, i.e. a concept, with a distribution of prior probabilities (what cognitive scientists refer to as an ‘*ad hoc*’ concept; [79,138,139]). In the language of the brain, a concept is a group of distributed *patterns* of neural spike trains, across a population of neurons, whose representations reflect the spatial and temporal scales that are most relevant to an animal’s behavioural repertoire. This hypothesis is generally consistent with brain imaging findings that the default mode network represents semantic concepts [140,141]. Predictions, as embodied brain states, emerge as default mode summaries cascade out to primary sensory and motor regions to become detailed and particularized (i.e. to modulate the spiking patterns of sensory and motor neurons [17]; for supporting evidence on embodied representations of concepts, see [142–145]).¹⁰

Using past experience as a guide, the brain is essentially asking, ‘what is this new sensory input most similar to?’ [147,148], where similarity is computed against the population of predictions and their associated energy costs and potential benefits for the body. That is, the conceptual representations (i.e. the prediction signals with some distribution of priors) are tested against the incoming sensory evidence to categorize incoming sensory signals (from outside the skull) according to past experience (producing a distribution of posterior probabilities). Incoming sensory

evidence, as prediction error, helps to select from or modify this distribution of predictions, because certain simulations will better fit the sensory array (i.e. they will have larger posteriors), such that incoming sensory events are categorized as similar to the past experiences with the highest posterior probabilities, thereby making them meaningful. The resulting categorization enables allostasis, allowing the brain to efficiently predict energy expenditures for motor actions, as well as the benefits that will result; this is how the brain proactively adjusts the body's physiological systems to satisfy those needs before they arise [12]. We hypothesize that, in this way, the brain manages physiological systems and motor actions to deal with upcoming sensory events, inferring their likely causes as they happen. The brain, via its trio of core networks, proactively anticipates demand across multiple body systems (e.g. need for glucose, oxygen, salt, etc.), evaluates the priorities (in terms of immediate and longer term needs, costs and likely pay-offs), and thereby implements allostasis. In neurotypical brains, completed predictions are categorizations that maintain efficient physiological regulation, guide appropriate action and construct perception. The meaning of every sensory event therefore includes a plan for allostasis during that event.

The unified theory also proposes that the processing of prediction error is equivalent to concept learning. Prediction errors (i.e. unanticipated sensory inputs) cascade in a feedforward cortical sweep, originating in the upper layers of cortices that have more-developed lamination and terminating in the deep layers of cortices with less well-developed lamination.¹¹ As information flows from primary sensory regions (whose upper layers contain many smaller pyramidal neurons with fewer connections) to limbic and other heteromodal regions in frontal cortex (whose upper layers contain fewer but larger pyramidal neurons with many more connections), it is compressed and reduced in dimensionality [149]. This dimension reduction efficiently represents a lot of information with a smaller number of neurons, reducing redundancy and saving metabolic cost, because smaller populations of neurons are summarizing statistical regularities in the spiking patterns in larger populations in the sensory and motor regions. Additional efficiency is achieved because conceptually similar representations utilize similar neural populations during simulation (e.g. [150]). As a result, different predictions are separable, but are not spatially separate (such that multimodal summaries are organized in a continuous neural territory that reflects their similarity to one another). In this way, the brain is condensing redundant firing patterns into more efficient (and cost-effective) multimodal summaries. This information is available for later use as limbic cortices generatively construct prediction signals, initiated as low-dimensional, multimodal summaries (i.e. 'abstractions'); as we noted earlier, these summaries, consolidated from prior encoding of prediction errors, become more detailed and particular as they propagate out to more architecturally granular sensory and motor regions. From this perspective, we hypothesize that interoceptive predictions are part of every concept that is learned and constructed, and categorization via concepts is the prime computation by which the emotion regulation process, cognitive reappraisal, takes place (the mechanisms for other regulation processes [151,152] can also be understood with our predictive inference account, but that is beyond the scope of this paper). This particular hypothesis is relevant to depression because,

like most disorders, depression is associated both with alexithymia, a condition defined by an impoverished conceptual understanding of emotion, and intense negative affect [153–155]. Interestingly, both alexithymia and depression are linked to diminished interoceptive awareness (e.g. [28,156]), which itself has recently been characterized as an inability to calibrate precision estimates for interoceptive prediction errors [11].

Within the unified theory of mind and brain, experiences of past, present and future are constructed within the same computational architecture via the same brain systems. In humans, it is well established that past experiences [157] and present experiences [158] contribute to experiences of the future. What is more striking, from our perspective, is that the present is, fundamentally, the *remembered* present [159]: the past becomes the present, corrected by the immediate future. What differs from moment to moment, context to context or even person to person is the extent to which the brain is prioritizing its own internal model versus accommodating that model to unexpected information from the sensory periphery (i.e. assimilation versus accommodation). These ideas provide us with the framework for examining the puzzle of depression.

4. Allostasis, interoception and depression

Depression is a devastating syndrome that is best characterized as abnormalities in neurologic (e.g. [160,161]), metabolic [162–165] and immunologic [44,166–170] systems, as well as aberrant hypothalamic–pituitary–adrenal (HPA)-axis function [171–173] and pervasive negative affect [174]. That is, depression is a disorder of allostasis.

Our unified theoretical framework, which is based on brain architecture and function supporting the efficient regulation of energy, may prove fruitful for understanding the dizzying variability in the pathophysiology of depression. Using the unified theory, we propose that the brain's internal model is adversely affected in the development and maintenance of depression (e.g. [30,175]). Specifically, the mood, motor, autonomic, immune, metabolic and circadian dysregulations all point to a central problem with inefficient energy regulation. We hypothesize that affective feelings characterized as pleasant or unpleasant may provide information about the moment-to-moment energy conditions of the body, whereas arousal might be a consequence of unresolved prediction error and indicate a need to learn [17]. Consistent with this, there is experimental evidence that momentary allostatic dysregulation is associated with momentary distress [176,177] and that arousal is a cue for novelty and learning [109,178].¹² Taken together, we propose that depression is the result of a relatively 'locked in' brain (i.e. relative insensitivity to prediction errors) coupled with inefficient energy regulation that is associated with intense suffering (i.e. negative affect), and difficulty engaging in vigorous mental or physical activity. The 'locked in' brain hypothesis is that internal models with certain characteristics result in inefficient energy regulation (either when they are insensitive to prediction errors and/or when they are subject to poorly calibrated precision estimates). Both problems of prediction error processing or precision estimation would lead to a failure of model updating. These can, in turn, prompt further inefficiency, producing a downward spiral.

Furthermore, we propose that there are many possible paths to depression (such many-to-one mappings are termed 'degeneracy'; [179]), not just because depression is a heterogeneous disorder, but also because there may be multiple pathways that shift the brain into a developmental trajectory towards inefficient metabolism and energy regulation. A clear example of a many-to-one mapping in depression is the reliable *and* discrepant functional finding of both increased [174,180] and decreased [181,182] resting metabolism in the subgenual portion of the anterior cingulate cortex (sgACC). In the light of the proposed role of sgACC in maintaining sympathetic and parasympathetic autonomic control of the viscera [183,184] and our conceptualization of depression as a state of inefficient metabolism and energy management, it becomes clear that *either* sgACC hyper-activation (fatigue, listlessness) or hypo-activation (agitation, irritation) can promote energy dysregulation leading to a depressed condition. A similar variable finding has been reported for anterior insula activity [185]. Not all pathways to depression require a pre-existing or longer lasting structural or functional nervous system dysfunction, however; many forms of depression are episodic, and some individuals have a single episode that never recurs, in part because there are many ways for allostasis to become costly and metabolically inefficient, not all of them permanent. For example, temporary changes in eating, sleeping or exercise behaviours can lead to transient changes in energy regulation [186–190] as can the loss of a loved one [191–194]. Using the unified theory, it is possible to hypothesize about the brain's computational architecture so as to better understand the various ways in which allostasis can become inefficient and resource management can be compromised. Based on figure 2, we hypothesize that major depressive disorder develops from at least one of three broad classes of inter-related problems (each of which can be caused by multiple, interacting neuropathological processes): the first class relating to the nature of the internal model (see figure 2, solid lines emanating from regions of the default mode network), the second relating to the nature of prediction errors (figure 2, dotted lines) and the third having to do with precision (figure 2, solid lines emanating from regions of the salience network). (Given space constraints, we will consider hypotheses related to the frontoparietal control network, cerebellar, striatal and thalamic portions of the theory in another venue.)

(a) A metabolically inefficient internal model

Using the logic of our unified theory, our first hypothesis is that a person becomes at risk for depression when his brain is inefficient at managing energy regulation for some relatively prolonged period of time (i.e. weeks to months). Inefficiency could result, for example, if the metabolic demand on the body was large in the past and the brain has not adjusted to its current context (e.g. the person was raised in an impoverished or adverse environment where rewards were rare, risks were frequent, and large investments of metabolic energy were routinely required). Consistent with these hypotheses, adverse childhood experiences such as traumatic events or neglect [195,196] are associated with later structural and functional abnormalities in the brain's core networks that predate the onset of depression [197–199] and may be associated with miscalibrated predictions. Metabolic efficiency may also be compromised by the loss of a loved one [191–194], as well as by the persistent presence of low-grade stressors (often present in adverse environments characterized by inconsistency

and uncertainty, or prolonged social evaluative stress) that cultivate sympathetic nervous system arborization; such arborization is known to enhance HPA axis reactivity [49,50]; for a review, see [51], producing increased reactivity and false alarms (i.e. the perception of threats where none exist). Even small but innumerable challenges (e.g. lack of sleep, poor nutrition) or physical illnesses (e.g. metabolic syndrome) can reduce metabolic efficiency because the brain has to work harder to achieve optimal energy regulation, even in the absence of structural or functional brain dysfunctions.

We also hypothesize that individuals may be vulnerable to depression because they have a narrower range of optimal energy regulation; in such people, the brain can efficiently implement allostasis, but they may be at greater risk for dysregulation in the face of varying environmental demands. By contrast, someone with a wider optimal range can more easily remain metabolically efficient in the face of broader variations in contextual demands. Consider, for example, a situation where you are reading quietly in your office and someone who has been critical of you in the past (say, a senior colleague) approaches your office door. If your colleague comes to your office on occasion, when you hear her footsteps, if the last (or a particularly salient) visit occurred when it was raining outside as it is on this day, etc., your brain will predict your colleague's arrival by constructing an embodied simulation. Part of this prediction will be how much glucose is needed to jump up and close the door, and how to mobilize this resource (by redirecting blood flow to the legs from other organs that need it less, by releasing cortisol, by reaching in your desk for some chocolate, etc.). Alternatively, you could be reading in your office and it is a beautiful day outside. The last time it was sunny and warm like this you heard a blue jay just outside your window, and your brain will predict how much glucose is needed to stand up, draw the curtains back, and open your window. When allostasis is working well, your brain has little difficulty efficiently predicting what your muscles need, what your heart and peripheral vasculature need, etc., to support any predicted action. We hypothesize that in the former case when your critical colleague approaches your door, the simulations created by the brain, as predictions, will involve negative affect associated primarily when there is metabolic inefficiency (and, if the inefficiency is of sufficiently long duration, an increase in visceral nociception may also result). The result may be what religion professor Wendy Farley calls 'tragic embodiment': discomfort in your body that is intense enough to draw your attention to yourself and away from the world.¹³ Pervasive negative affect may be a context in which the brain has difficulty processing prediction error. Indeed, depression is associated with inwardly focused attention [200]. By contrast, associative thinking has been linked, both conceptually and empirically, to positive affect [147,201].

Pervasive negative affect could also lead one to construct a profoundly negative internal model. The brain samples past experiences to create predictions of the immediate future and it is doing so in a current context of metabolic inefficiency. Feeling unpleasant could also lead to affective realism, trapping a person in a vicious cycle of negativity. In fact, there is abundant evidence that persistent distress plays a critical role in major depression. Diagnostically, sustained unpleasant mood, irritability and/or anhedonia are key symptoms of a major depressive episode. Furthermore, in the course of a

given episode, both prediction and prediction errors are biased towards the unpleasant (e.g. attention to and memory of information that will disrupt optimal energy regulation; e.g. [202]). Epidemiologically, sustained distress plays a significant role in depression, over the long course of risk, illness, remission and relapse as well as over the course of a given depressive episode. For example, prospective longitudinal studies have consistently shown that neuroticism—a personality trait characterized by excessive and sustained negative affect—predicts subsequent onset of depression [203]. In the context of the unified model, we might conceptualize such sustained negative affect as the phenomenological representation of metabolic inefficiency.

(b) Unreliable prediction errors

The nature of prediction errors might also contribute to a progressive isolation of the brain's internal working model from its sensory context. Accordingly, we hypothesize that unreliable prediction errors are a second possible computational ingredient to developing depression. For example, the increased arborization of the sympathetic nervous system leads to increased false alarms that can result in substantial differences between the prediction signals sent from limbic visceromotor cortices and the ascending interoceptive signals about the state of the body, resulting in increased prediction error. We hypothesize that increased interoceptive prediction error will be associated with inefficient allostasis, in much the same way that prediction error in the motor system is indicative of poor motor control [204]. The relationship between a motor prediction and a skeletomotor muscle movement is inherently noisy—commands produce noisy movements as both the body and environment change—and so we might expect something similar in the visceromotor domain. Moreover, the afferent sensory consequences of visceromotor changes are inherently noisy. (Recall that the autonomic, immune, and metabolic changes associated with allostatic dysregulation have sensory consequences that are communicated via the vagus nerve and the small diameter C and A δ fibres to primary interoceptive cortex in the dorsal mid- and posterior insula via relays in the brainstem and ventral posterior thalamic nuclei [205,206].) Many of these fibres are unmyelinated, meaning that they influence one another via a process known as ephaptic coupling as information is ascending (for discussion, see [102]).

We hypothesize that the resulting noisy afferent interoceptive inputs [207] would be progressively discounted more and more by precision predictions, reducing the output of prediction error signals from the cortical columns in primary interoceptive cortex, ultimately resulting in even greater energy dysregulation. Limbic regulation of the reticular nucleus of the thalamus is another avenue to modulate sampling of the sensory array, again reducing the availability of prediction error signals [127–129]. More speculatively, a third pathway for reducing sensory sampling might be the allostatic predictions descending through the brainstem to the spinal cord, which serve as a gain adjustment on ascending viscerosensory signals [17,208]. As is the case in the motor system (e.g. [209]), we expect that learning (i.e. correcting the internal model by modifying predictions) will increasingly diminish, because it is driven more by predictions of precision than by prediction errors themselves (a possible mechanism might be the precision-related prediction errors to agranular limbic regions of the default mode that are

being sent from the more developed dysgranular limbic regions within the salience portion of the allostatic/interoceptive system (e.g. the ventrolateral prefrontal cortex (vlPFC) and the midcingulate cortex (MCC) that is more associated with motor responses)).

Prediction errors might also remain uncorrected due to fatigue and reduced movements. For example, as the costs of energy dysregulation accrue, visceromotor regions slow the body down with feelings of fatigue [44] and the so-called sickness behaviours that conserve metabolic resources that are running low [210], including reduced interaction with the outside world (most notably with other people [211,212]). This solution reduces metabolic costs via reduced motor movements, and less exploration that usually requires processing prediction error and new encoding and consolidation, but it deprives the brain of one main way to reduce prediction error (i.e. move to generate the predicted sensory inputs).

In addition to fatigue, rumination may be another behavioural hallmark of a brain locked into the past, issuing predictions to explain incoming sensory events that remain uncorrected by sensory cues in the present. Indeed, the presence of a ruminative cognitive style (with a focus on potential causes and consequences of depression) predicts longer and more severe depressive episodes [200] and increases the risk of depressive relapse in remitted adults [213]. Consistent with the unified theory, rumination and repetitive thinking are associated with increased connectivity between the sgACC and the rest of the default mode network [214]. This increased connectivity allows the brain to idle in visceromotor and interoceptive predictions that are disruptive to efficient allostasis.

(c) Inaccurate precision signals

The third potential ingredient that can contribute to depression might be ineffective *precision signalling*, particularly when signalling a failure to predict resources that can improve metabolic efficiency (i.e. reward insensitivity; e.g. [215]). For example, low serotonin levels make it difficult to sustain effort when a reward is delayed [216–218]; for a review of neuromodulatory functions, see [219]. Low dopamine levels [220] in depression impair effortful movements and encoding of prediction errors. In addition, several meta-analyses demonstrate that nodes of the salience network are atrophied in depression [47,48] (as in other illnesses), and the major fibre pathways that link the salience network to other parts of the brain are compromised in depression [221]. Although these findings are broadly consistent with our hypotheses, stronger support still awaits; our hypotheses are, in effect, computational in nature, and so require functional testing on relevant timescales.

5. Implications for treatment

Our hypothesis is that a depressed brain is relatively 'locked in', and running a metabolically inefficient internal model of the body in the world, resulting in pervasive negative affect that is salient and difficult to modify. The fact that there are many potential sources of pathophysiology that will ultimately result in a relatively 'locked-in' depressed brain (i.e. 'degeneracy'; [179]) is probably one reason why major depressive disorder is so difficult to treat. Nonetheless, by suggesting that depression arises from a chronic energy inefficiency and altered interoceptive signalling through a

well-specified computational architecture, the unified theory of mind and brain suggests several targets of opportunity for intervention and treatment.

First, depression may be relieved by directly affecting the descending allostatic predictions that originate in agranular limbic cortices. For example, deep brain stimulation in the region of the subcallosal (agranular) anterior cingulate may achieve its antidepressant effects [222–224] via its ability to alter allostatic predictions to the body and brain. Indeed, the effectiveness of Mayberg's intracranial stimulation for untreatable depression could be due to the fact that it affects three white matter tracts (the cingulum bundle, the forceps minor and the uncinate fasciculus [225]), thereby modifying the connectivity within and between the default mode and salience networks.

Second, according to the unified theory, depression can arise from the effects of chronic, aberrant allostatic predictions leading to or resulting from HPA-axis dysregulation and inflammation. This suggests that interventions to address these systemic affects should help to slow the onset and intensity of depression. Indeed, recent efforts have explored the antidepressant effects of cortisol synthesis inhibitors [226,227] and anti-inflammatory medications [228–230] in depression. Greater aerobic physical activity may also help prevent the occurrence of chronic aberrant allostatic predictions, which is consistent with findings that those who are physically active have a prospectively lower risk for depression [231]. However, once the brain's internal model is altered and relatively insensitive to prediction errors, physical activity may be less impactful, consistent with findings that exercise has more modest positive effects in those who already have depression [232]. Alternatively, it may be that various methods for enhancing efficient energy regulation (such as exercising, and eating and sleeping just enough to maintain optimal energy balance) may need to be combined with other interventions (medication, cognitive behavioural therapy (CBT)) in depression to have a sufficiently potent effect to reset the brain's internal model.

Third, another approach to intervening on interoceptive neurocircuitry in depression is to reduce the 'noisy' afferent interoceptive prediction errors, thereby overcoming the gating effects of precision estimates. In principle, one way to accomplish this would be to greatly increase the signal-to-noise ratio (SNR) of afferent (primarily vagal) interoceptive signals. This may be one mechanism by which vagal nerve stimulation has its effects [233]. Additional novel approaches to increasing the SNR on interoceptive signals are also under development, including pharmacological interventions such as infusion regimens with the β -adrenoreceptor agonist isoproterenol that have been tested in other psychiatric illnesses [234,235] and therapy within sensory-attenuated environments that increase the salience of interoceptive signals [236].

Fourth, treatments that offer the opportunity for re-categorization provide yet another intervention pathway. For example, CBT may have its effects by helping a person construct new concepts that, as prediction signals, modify the gain on prediction errors via the salience network. Over time, this process may alter the sample of inputs that eventually become the 'empirical priors' that agranular limbic cortices use to initiate subsequent predictions. This dynamic may explain why CBT alters the activity of two key regions in the brain's rich club: the cingulate and anterior insular cortices [237,238]. Consistent with this idea, CBT is very effective in treating depression in individuals with low activity in the

anterior insula before treatment (presumably because CBT helps them to change their predictions, potentially by improving their processing of prediction errors and corresponding concept learning via salience network changes); alternatively, CBT is largely ineffective and medications are more effective in treating depression in individuals with high anterior insula activity before treatment [185,239], suggesting the hypothesis that activity in the anterior insula may be indicative of the extent to which multimodal precision estimates can be easily modified. Relatedly, the empirical priors themselves may be direct intervention targets, thereby altering the computation of interoceptive predictions. One example of this may be electroconvulsive therapy, which frequently interrupts memory consolidation and alters the activity of both visceromotor limbic regions and the hippocampus, with changes in both regions correlated with treatment response [240].

6. Conclusion

The size and complexity of human brains grant our species enormous energy range and behavioural flexibility, affording us the largest ecological niche of any mammal. By conceptualizing allostasis and interoception as unified processes within a predicting brain, the unified theory computationally recasts many of the 'mental' symptoms of affective distress, rumination and fatigue in metabolic terms. The theory suggests that understanding more about the role of metabolism in guiding basic perception and action will provide a richer, more powerful framework for studying major depressive illness. From this perspective, depression is an internal model, associated with distress, mental withdrawal from the world and sometimes a literal physical withdrawal from the world. Treating depression, then, will require providing the brain with the resources to modify its internal model of the body in the world and repair its energy regulation.

Data accessibility. No new data are reported in this article.

Authors' contributions. L.F.B. conceived of this paper. All authors contributed to the drafting and revising the article. All approve the final manuscript.

Competing interests. The authors have no competing interests.

Funding. This paper was prepared with support from the National Institutes on Aging (R01 AG030311), the National Cancer Institute (U01 CA193632), the National Science Foundation (1638234) and the US Army Research Institute for the Behavioral and Social Sciences (W911NF-15-1-0647 and W911NF-16-1-0191).

Acknowledgements. Many thanks to Kyle Simmons who contributed to the conception of this article, particularly the section on treatment, and who made invaluable comments on earlier versions of this manuscript, to Helen Mayberg for her invaluable insights on the nature of depression and to Jiahe Zhang for her help constructing figure 1.

Disclaimer. The views, opinions and/or findings contained in this paper are those of the authors and shall not be construed as an official Department of the Army position, policy or decision, unless so designated by other documents

Endnotes

¹Long-range neural connections, like those that form the human brain's broadly distributed intrinsic networks, are particularly metabolically expensive [1], with most of the energy costs going to signalling between neurons, particularly in postsynaptic processes [40–43].

²One aspect of allostasis involves the brain dynamically regulating resource allocations (i.e. diverting glucose, electrolytes, water, etc., from one system to another) to meet the body's spending needs

(e.g. in advance of standing up, the heart prepares to beat stronger and faster, blood vessels in the legs constrict and blood pressure goes up to ensure that the brain continues to receive the needed blood (and oxygen)). Allostasis also includes the brain's signalling the need for resources before a bodily resource becomes depleted (e.g. drinking before dehydration occurs) or preparing for the intake of resources in advance of their ingestion (e.g. saliva is pre-emptively secreted when the body is in need of glucose, even before anything is ingested, and a key component of saliva is alpha-amylase, an enzyme that breaks down glucose. Even just imagining eating food causes saliva secretion).

³All animal brains operate in the same manner (i.e. even insect brains coordinate visceral, immune and motor changes [1]).

⁴The term 'feedback' derives from a time when the brain was thought to be largely stimulus driven [84]. Nonetheless, the history of science is laced with the idea that the mind drives perception, e.g. in the eleventh century by Ibn al-Haytham (who helped to invent the scientific method), in the eighteenth century by Kant (1781) and in the nineteenth century by Helmholtz. In more modern times, see Craik's concept of internal models [85], Tolman's cognitive maps [86], Johnson-Laird's internal models [87,88] and for other relevant references, see [89,90]. The novelty in recent formulations can be found in (a) the hypothesis that predictions are embodied simulations of sensory motor experiences, (b) they are ultimately in the service of allostasis and therefore interoception is at their core, and, of course (c) the breadth of behavioural, functional and anatomic evidence supporting the hypothesis that the brain's internal model implements active inference as prediction signals, including (d) the specific computational hypotheses implementing our predictive coding account.

⁵Cortical regions with a dysgranular structure are referred to as limbic [82] or paralimbic [94].

⁶Prediction signals are conveyed as slow frequency oscillations (e.g. [97–99]) and, indeed, there is evidence that genes associated with slower rhythms are upregulated in limbic circuitry [100].

⁷Even more interestingly, there is some evidence to suggest that the cortical regions projecting to the brainstem nuclei which are the source nuclei for these neuromodulators (such as the locus coeruleus for norepinephrine) are largely entrained by limbic cortical regions via descending allostatic predictions that project directly from the cingulate cortices and medial prefrontal cortex, as well as

indirectly via projections from the central nucleus of the amygdala and the hypothalamus. The locus coeruleus also receives ascending interoceptive and nociceptive prediction errors [107]. This is yet another way that allostasis is linked to the modulating the gain or excitability of neurons that represent sensory and motor prediction errors.

⁸In a fly brain, the mushroom bodies may play an analogous role to the cerebellum [1].

⁹This allows for the encoding of statistical patterns of uncertain value that can later be reconstructed when they are of use [122].

¹⁰Note that this hypothesis is species-general: rats have a default mode network and are not able to engage in mental time travel as far as we can tell [146], but this in no way disconfirms the hypothesis that the network is running an internal model of the animal's world in the service of allostasis.

¹¹They also arise from ascending inputs to terminate in the deep layers of cortex (e.g. [92]), but these can be thought of as errors related to predictions of precision.

¹²The human nervous system is not wired to represent interoceptive sensations in high-dimensional detail. Some of the ascending viscerosensory inputs that reach the brain are not labelled lines sending precise, modality-specific information. Furthermore, interoceptive predictions only traverse one or two synapses from limbic visceromotor regulation regions to primary interoceptive cortex, which does not leave much opportunity for them to become elaborated with high-dimensional detail (compared to the sensory predictions reaching primary exteroceptive cortices, which usually traverse more synapses and therefore can be more elaborated with details). Relative low dimensionality is probably a good thing, because if we could detect the ongoing dynamic sensory changes in the body in a lot of detail, we would never pay attention to anything else (think about the last time you had abdominal cramps).

¹³To paraphrase an example Farley once used during an address: you could be listening to the radio about a terrible plane crash that took the lives of 200 people while opening the mail, but if you give yourself a paper cut on an envelope, your attention will be completely diverted from the disaster while you nurse your minor discomfort. It is not your moral position that a paper cut deserves more attention than the death of 200 people, but your discomfort demands your attention, even if just for a moment. So imagine your distraction if you felt intensely unpleasant much of the time.

References

1. Sterling P, Laughlin SB. 2015 *Principles of neural design*. Cambridge, MA: MIT Press.
2. Seth AK, Friston KJ. 2016 Active interoceptive inference and the emotional brain. *Phil. Trans. R. Soc. B* **371**, 20160007. (doi:10.1098/rstb.2016.0007)
3. Deneve S. 2008 Bayesian spiking neurons I: Inference. *Neural Comput.* **20**, 91–117. (doi:10.1162/neco.2008.20.1.91)
4. Clark A. 2013 Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.* **36**, 281–253. (doi:10.1017/S0140525X12002919)
5. Deneve S, Jardri R. 2016 Circular inference: mistaken belief, misplaced trust. *Curr. Opin. Behav. Sci.* **11**, 40–48. (doi:10.1016/j.cobeha.2016.04.001)
6. Friston K. 2010 The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* **11**, 127–138. (doi:10.1038/nrn2787)
7. Hohwy J. 2013 *The predictive mind*. Oxford, UK: Oxford University Press.
8. Larkum M. 2013 A cellular mechanism for cortical associations: an organizing principle for the cerebral cortex. *Trends Neurosci.* **36**, 141–151. (doi:10.1016/j.tins.2012.11.006)
9. Mumford D. 1991 On the computational architecture of the neocortex. I. The role of the thalamo-cortical loop. *Biol. Cybern.* **65**, 135–145. (doi:10.1007/BF00202389)
10. Mumford D. 1992 On the computational architecture of the neocortex. *Biol. Cybern.* **66**, 241–251. (doi:10.1007/BF00198477)
11. Ainley V, Apps MAJ, Fotopoulou A, Tsakiris M. 2016 'Bodily precision': a predictive coding account of individual differences in interoceptive accuracy. *Phil. Trans. R. Soc. B* **371**, 20160003. (doi:10.1098/rstb.2016.0003)
12. Sterling P. 2012 Allostasis: a model of predictive regulation. *Physiol. Behav.* **106**, 5–15. (doi:10.1016/j.physbeh.2011.06.004)
13. Franklin DW, Wolpert DM. 2011 Computational mechanisms of sensorimotor control. *Neuron* **72**, 425–442. (doi:10.1016/j.neuron.2011.10.006)
14. McEwen BS, Wingfield JC. 2010 What's in a name? Integrating homeostasis, allostasis and stress. *Horm. Behav.* **57**, 105. (doi:10.1016/j.yhbeh.2009.09.011)
15. Craig AD. 2015 *How do you feel? An interoceptive moment with your neurobiological self*. Princeton, NJ: Princeton University Press.
16. Barrett LF, Bliss-Moreau E. 2009 Affect as a psychological primitive. *Adv. Exp. Soc. Psychol.* **41**, 167–218. (doi:10.1016/S0065-2601(08)00404-8)
17. Barrett LF. 2017 *How emotions are made: the secret life the brain*. New York, NY: Houghton-Mifflin-Harcourt.
18. Barrett LF, Russell JA. 1999 Structure of current affect: controversies and emerging consensus. *Curr. Dir. Psychol. Sci.* **8**, 10–14. (doi:10.1111/1467-8721.00003)
19. Kuppens P, Tuerlinckx F, Russell JA, Barrett LF. 2013 The relation between valence and arousal in subjective experience. *Psychol. Bull.* **139**, 917–940. (doi:10.1037/a0030811)
20. Damasio AR. 1999 *The feeling of what happens: body and emotion in the making of consciousness*. Boston, MA: Houghton Mifflin Harcourt.
21. Dreyfus G, Thompson E. 2007 Asian perspectives: Indian theories of mind. In *The Cambridge handbook of consciousness* (eds PD Zelazo, M

- Moscovitch, E Thompson), pp. 89–114. Cambridge, UK: Cambridge University Press.
22. Edelman GM, Tononi G. 2000 A universe of consciousness: how matter becomes imagination. New York, NY: Basic books.
 23. James W. 1890/2007 *The principles of psychology*. New York, NY: Dover.
 24. Searle JR. 1992 *The rediscovery of the mind*. Cambridge, MA: MIT Press.
 25. Searle JR. 2004 *Mind: a brief introduction*. Oxford, UK: Oxford University Press.
 26. Wundt W. 1897 *Outlines of psychology*. Leipzig, Germany: Wilhelm Engelmann.
 27. Kring AM. 2008 Emotion disturbances as transdiagnostic processes in psychopathology. In *Handbook of emotion* (eds M Lewis, JM Haviland-Jones, LF Barrett), pp. 691–705, 3rd edn. New York, NY: Guilford Press.
 28. Harshaw C. 2015 Interoceptive dysfunction: toward an integrated framework for understanding somatic and affective disturbance in depression. *Psychol. Bull.* **141**, 311–363. (doi:10.1037/a0038101)
 29. Balon R. 2006 Mood, anxiety, and physical illness: body and mind, or mind and body? *Depress. Anxiety* **23**, 377–387. (doi:10.1002/da.20217)
 30. Barrett LF, Simmons WK. 2015 Interoceptive predictions in the brain. *Nat. Rev. Neurosci.* **16**, 419–429. (doi:10.1038/nrn3950)
 31. Pezzulo G, Rigoli F, Friston K. 2015 Active inference, homeostatic regulation and adaptive behavioural control. *Progress Neurobiol.* **134**, 17–35. (doi:10.1016/j.pneurobio.2015.09.001)
 32. Seth AK. 2013 Interoceptive inference, emotion, and the embodied self. *Trends Cogn. Sci.* **17**, 565–573. (doi:10.1016/j.tics.2013.09.007)
 33. Seth AK, Suzuki K, Critchley HD. 2011 An interoceptive predictive coding model of conscious presence. *Front. Psychol.* **2**, 395. (doi:10.3389/fpsyg.2011.00395)
 34. Conant RC, Ashby WR. 1970 Every good regulator of a system must be a model of that system. *Int. J. Syst. Sci.* **1**, 89–97. (doi:10.1080/00207727008920220)
 35. Berkes P, Orbán G, Lengyel M, Fiser J. 2011 Spontaneous cortical activity reveals hallmarks of an optimal internal model of the environment. *Science* **331**, 83–87. (doi:10.1126/science.1195870)
 36. Buckner RL. 2012 The serendipitous discovery of the brain's default network. *Neuroimage* **62**, 1137–1145. (doi:10.1016/j.neuroimage.2011.10.035)
 37. Hassabis D, Maguire EA. 2009 The construction system of the brain. *Phil. Trans. R. Soc. B* **364**, 1263–1271. (doi:10.1098/rstb.2008.0296)
 38. Mesulam MM. 2002 The human frontal lobes: transcending the default mode through contingent encoding. In *Principles of frontal lobe function* (eds DT Stuss, RT Knight), pp. 8–30. New York, NY: Oxford University Press.
 39. Raichle ME. 2010 Two views of brain function. *Trends Cogn. Sci.* **14**, 180–190. (doi:10.1016/j.tics.2010.01.008)
 40. Attwell D, Iadecola C. 2002 The neural basis of functional brain imaging signals. *Trends Neurosci.* **25**, 621–625. (doi:10.1016/S0166-2236(02)02264-6)
 41. Attwell D, Laughlin SB. 2001 An energy budget for signaling in the grey matter of the brain. *J. Cereb. Blood Flow Metab.* **21**, 1133–1145. (doi:10.1097/00004647-200110000-00001)
 42. Alle H, Roth A, Geiger JR. 2009 Energy-efficient action potentials in hippocampal mossy fibers. *Science* **325**, 1405–1408. (doi:10.1126/science.1174331)
 43. Harris JJ, Jolivet R, Attwell D. 2012 Synaptic energy use and supply. *Neuron* **75**, 762–777. (doi:10.1016/j.neuron.2012.08.019)
 44. Dantzer R, Heijnen CJ, Kavelaars A, Laye S, Capuron L. 2014 The neuroimmune basis of fatigue. *Trends Neurosci.* **37**, 39–46. (doi:10.1016/j.tins.2013.10.003)
 45. McEwen BS, Bowles NP, Gray JD, Hill MN, Hunter RG, Karatsoreos IN, Nasca C. 2015 Mechanisms of stress in the brain. *Nat. Neurosci.* **18**, 1353–1363. (doi:10.1038/nn.4086)
 46. Hunter RG, McEwen BS. 2013 Stress and anxiety across the lifespan: structural plasticity and epigenetic regulation. *Epigenomics* **5**, 177–194. (doi:10.2217/epi.13.8)
 47. Crossley NA, Mechelli A, Scott J, Carletti F, Fox PT, McGuire P, Bullmore ET. 2014 The hubs of the human connectome are generally implicated in the anatomy of brain disorders. *Brain* **137**, 2382–2395. (doi:10.1093/brain/awu132)
 48. Goodkind M *et al.* 2015 Identification of a common neurobiological substrate for mental illness. *JAMA Psychiatry* **72**, 305–315. (doi:10.1001/jamapsychiatry.2014.2206)
 49. Capitanio JP, Cole SW. 2015 Social instability and immunity in rhesus monkeys: the role of the sympathetic nervous system. *Phil. Trans. R. Soc. B* **370**, 20140104. (doi:10.1098/rstb.2014.0104)
 50. Sloan EK, Capitanio JP, Tarara RP, Mendoza SP, Mason WA, Cole SW. 2007 Social stress enhances sympathetic innervation of primate lymph nodes: mechanisms and implications for viral pathogenesis. *J. Neurosci.* **27**, 8857–8865. (doi:10.1523/JNEUROSCI.1247-07.2007)
 51. Sloan EK, Capitanio JP, Cole SW. 2008 Stress-induced remodeling of lymphoid innervation. *Brain Behav. Immunity* **22**, 15–21. (doi:10.1016/j.bbi.2007.06.011)
 52. Kleckner IR, Zhang J, Touroutoglou A, Chanes L, Xia C, Simmons WK, Quigley KS, Dickerson BC, Barrett LF. Submitted. Evidence for an intrinsic brain system supporting allostasis and interoception in humans.
 53. Bar K-J, de la Cruz F, Schumann A, Koehler S, Sauer H, Critchley H, Wagner G. 2016 Functional connectivity and network analysis of midbrain and brainstem nuclei. *Neuroimage* **134**, 53–63. (doi:10.1016/j.neuroimage.2016.03.071)
 54. van den Heuvel MP, Sporns O. 2013 An anatomical substrate for integration among functional networks in human cortex. *J. Neurosci.* **33**, 14 489–14 500. (doi:10.1523/JNEUROSCI.2128-13.2013)
 55. de Pasquale F *et al.* 2010 Temporal dynamics of spontaneous MEG activity in brain networks. *Proc. Natl Acad. Sci. USA* **107**, 6040–6045. (doi:10.1073/pnas.0913863107)
 56. Honey CJ, Thivierge J-P, Sporns O. 2010 Can structure predict function in the human brain? *Neuroimage* **52**, 766–776. (doi:10.1016/j.neuroimage.2010.01.071)
 57. Sepulcre J, Sabuncu MR, Yeo TB, Liu H, Johnson KA. 2012 Stepwise connectivity of the modal cortex reveals the multimodal organization of the human brain. *J. Neurosci.* **32**, 10 649–10 661. (doi:10.1523/JNEUROSCI.0759-12.2012)
 58. van den Heuvel MP, Kahn RS, Goni J, Sporns O. 2012 High-cost, high-capacity backbone for global brain communication. *Proc. Natl Acad. Sci. USA* **109**, 11 372–11 377. (doi:10.1073/pnas.1203593109)
 59. van den Heuvel MP, Sporns O. 2011 Rich-club organization of the human connectome. *J. Neurosci.* **31**, 15 775–15 786. (doi:10.1523/JNEUROSCI.3539-11.2011)
 60. van den Heuvel MP, Sporns O. 2013 Network hubs in the human brain. *Trends Cogn. Sci.* **17**, 683–696. (doi:10.1016/j.tics.2013.09.012)
 61. van den Heuvel MP, Bullmore ET, Sporns O. 2016 Comparative connectomics. *Trends Cogn. Sci.* **20**, 345–361. (doi:10.1016/j.tics.2016.03.001)
 62. Scholtens LH, Schmidt R, de Reus MA, van den Heuvel MP. 2014 Linking macroscale graph analytical organization to microscale neuroarchitectonics in the macaque connectome. *J. Neurosci.* **34**, 12 192–12 205. (doi:10.1523/JNEUROSCI.0752-14.2014)
 63. van den Heuvel MP, Scholtens LH, Turk E, Mantini D, Vanduffel W, Barrett LF. 2016 Multimodal analysis of cortical chemoarchitecture and macroscale fMRI resting-state functional connectivity. *Hum. Brain Mapping* **37**, 3103–3113. (doi:10.1002/hbm.23229)
 64. Turk E, Scholtens LH, van den Heuvel MP. 2016 Cortical chemoarchitecture shapes macroscale effective functional connectivity patterns in macaque cerebral cortex. *Hum. Brain Mapping* **37**, 1856–1865. (doi:10.1002/hbm.23141)
 65. Enzi B, Duncan NW, Kaufmann J, Tempelmann C, Wiebking C, Northoff G. 2012 Glutamate modulates resting state activity in the perigenual anterior cingulate cortex—a combined fMRI–MRS study. *Neuroscience* **227**, 102–109. (doi:10.1016/j.neuroscience.2012.09.039)
 66. Ekman M, Derrfuss J, Tittgemeyer M, Fiebach CJ. 2012 Predicting errors from reconfiguration patterns in human brain networks. *Proc. Natl Acad. Sci. USA* **109**, 16 714–16 719. (doi:10.1073/pnas.1207523109)
 67. Kitzbichler MG, Henson RN, Smith ML, Nathan PJ, Bullmore ET. 2011 Cognitive effort drives workspace configuration of human brain functional networks. *J. Neurosci.* **31**, 8259–8270. (doi:10.1523/JNEUROSCI.0440-11.2011)
 68. Barrett LF, Satpute AB. 2013 Large-scale brain networks in affective and social neuroscience: towards an integrative functional architecture of the brain. *Curr. Opin. Neurobiol.* **23**, 361–372. (doi:10.1016/j.conb.2012.12.012)

69. Yeo BTT, Krienen FM, Eickhoff SB, Yaakub SN, Fox PT, Buckner RL, Asplund CL, Chee MW. 2014 Functional specialization and flexibility in human association cortex. *Cereb. Cortex* **25**, 3654–3672. (doi:10.1093/cercor/bhu217)
70. Chanes L, Barrett LF. 2016 Redefining the role of limbic areas in cortical processing. *Trends Cogn. Sci.* **20**, 96–106. (doi:10.1016/j.tics.2015.11.005)
71. Kok P, Bains LJ, van Mourik T, Norris DG, de Lange FP. 2016 Selective activation of the deep layers of the human primary visual cortex by top-down feedback. *Curr. Biol.* **26**, 371–376. (doi:10.1016/j.cub.2015.12.038)
72. Adams RA, Shipp S, Friston KJ. 2013 Predictions not commands: active inference in the motor system. *Brain Struct. Funct.* **218**, 611–643. (doi:10.1007/s00429-012-0475-5)
73. Ito M. 2008 Control of mental activities by internal models in the cerebellum. *Nat. Rev. Neurosci.* **9**, 304–313. (doi:10.1038/nrn2332)
74. Shadmehr R, Smith MA, Krakauer JW. 2010 Error correction, sensory prediction, and adaptation in motor control. *Annu. Rev. Neurosci.* **33**, 89–108. (doi:10.1146/annurev-neuro-060909-153135)
75. Wolpert DM, Kawato M. 1998 Multiple paired forward and inverse models for motor control. *Neural Netw.* **11**, 1317–1329. (doi:10.1016/S0893-6080(98)00066-5)
76. Wolpert DM, Miall RC, Kawato M. 1998 Internal models in the cerebellum. *Trends Cogn. Sci.* **2**, 338–347. (doi:10.1016/S1364-6613(98)01221-2)
77. O'Reilly JX, Mesulam MM, Nobre AC. 2008 The cerebellum predicts the timing of perceptual events. *J. Neurosci.* **28**, 2252–2260. (doi:10.1523/JNEUROSCI.2742-07.2008)
78. Barsalou LW. 2008 Grounded cognition. *Annu. Rev. Psychol.* **59**, 617–645. (doi:10.1146/annurev.psych.59.103006.093639)
79. Barsalou LW, Simmons WK, Barbey AK, Wilson CD. 2003 Grounding conceptual knowledge in modality-specific systems. *Trends Cogn. Sci.* **7**, 84–91. (doi:10.1016/S1364-6613(02)00029-3)
80. 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)
81. Barbas H, Rempel-Clower N. 1997 Cortical structure predicts the pattern of corticocortical connections. *Cereb. Cortex* **7**, 635–646. (doi:10.1093/cercor/7.7.635)
82. Barbas H. 2015 General cortical and special prefrontal connections: principles from structure to function. *Annu. Rev. Neurosci.* **38**, 269–289. (doi:10.1146/annurev-neuro-071714-033936)
83. Beul SF, Barbas H, Hilgetag CC. 2015 A predictive structural model of the primate connectome. (<http://arxiv.org/abs/151107222>).
84. Lamme VA, Roelfsema PR. 2000 The distinct modes of vision offered by feedforward and recurrent processing. *Trends Neurosci.* **23**, 571–579. (doi:10.1016/S0166-2236(00)01657-X)
85. Craik K. 1943 *The nature of exploration*. Cambridge, UK: Cambridge University Press.
86. Tolman EC. 1948 Cognitive maps in rats and men. *Psychol. Rev.* **55**, 189–208. (doi:10.1037/h0061626)
87. Johnson-Laird P. 2004 The history of mental models. In *Psychology of reasoning: theoretical and historical perspectives* (eds K Manktelow, MC Chung), pp. 179–212. New York, NY: Psychology Press.
88. Johnson-Laird PN. 2006 *How we reason*. New York, NY: Oxford University Press.
89. Neisser U. 1967 *Cognitive psychology*. New York, NY: Appleton-Century-Crofts.
90. Gregory RL. 1980 Perceptions as hypotheses. *Phil. Trans. R. Soc. Lond. B* **290**, 181–197. (doi:10.1098/rstb.1980.0090)
91. Lochmann T, Deneve S. 2011 Neural processing as causal inference. *Curr. Opin. Neurobiol.* **21**, 774–781. (doi:10.1016/j.conb.2011.05.018)
92. Ghashghaie H, Hilgetag C, Barbas H. 2007 Sequence of information processing for emotions based on the anatomic dialogue between prefrontal cortex and amygdala. *Neuroimage* **34**, 905–923. (doi:10.1016/j.neuroimage.2006.09.046)
93. Swanson LW. 2012 *Brain architecture: understanding the basic plan*. Oxford, UK: Oxford University Press.
94. Mesulam MM. 1998 From sensation to cognition. *Brain* **121**, 1013–1052. (doi:10.1093/brain/121.6.1013)
95. Davachi L, DuBrow S. 2015 How the hippocampus preserves order: the role of prediction and context. *Trends Cogn. Sci.* **19**, 92–99. (doi:10.1016/j.tics.2014.12.004)
96. Hasson U, Chen J, Honey CJ. 2015 Hierarchical process memory: memory as an integral component of information processing. *Trends Cogn. Sci.* **19**, 304–313. (doi:10.1016/j.tics.2015.04.006)
97. Lee JH, Whittington MA, Kopell NJ. 2013 Top-down beta rhythms support selective attention via interlaminar interaction: a model. *PLoS Comput. Biol.* **9**, e1003164. (doi:10.1371/journal.pcbi.1003164)
98. Arnal LH, Giraud A-L. 2012 Cortical oscillations and sensory predictions. *Trends Cogn. Sci.* **16**, 390–398. (doi:10.1016/j.tics.2012.05.003)
99. Bressler SL, Richter CG. 2015 Interareal oscillatory synchronization in top-down neocortical processing. *Curr. Opin. Neurobiol.* **31**, 62–66. (doi:10.1016/j.conb.2014.08.010)
100. Chanes L, Zhang J, Güell M, Touroutoglou A, Sepulcre J, Barrett LF. Submitted. Multilevel analysis of human limbic cortices reveals a high-level domain-general neural workspace.
101. García-Cabezas MÁ, Barbas H. 2014 Area 4 has layer IV in adult primates. *Eur. J. Neurosci.* **39**, 1824–1834. (doi:10.1111/ejn.12585)
102. Damasio A, Carvalho GB. 2013 The nature of feelings: evolutionary and neurobiological origins. *Nat. Rev. Neurosci.* **14**, 143–152. (doi:10.1038/nrn3403)
103. Beckmann M, Johansen-Berg H, Rushworth MF. 2009 Connectivity-based parcellation of human cingulate cortex and its relation to functional specialization. *J. Neurosci.* **29**, 1175–1190. (doi:10.1523/JNEUROSCI.3328-08.2009)
104. Buckner RL, Krienen FM, Castellanos A, Diaz JC, Yeo BT. 2011 The organization of the human cerebellum estimated by intrinsic functional connectivity. *J. Neurophysiol.* **106**, 2322–2345. (doi:10.1152/jn.00339.2011)
105. Haber SN, Behrens TE. 2014 The neural network underlying incentive-based learning: implications for interpreting circuit disruptions in psychiatric disorders. *Neuron* **83**, 1019–1039. (doi:10.1016/j.neuron.2014.08.031)
106. Whalen PJ. 1998 Fear, vigilance, and ambiguity: initial neuroimaging studies of the human amygdala. *Curr. Dir. Psychol. Sci.* **7**, 177–188. (doi:10.1111/1467-8721.ep10836912)
107. Counts SE, Mufson EJ. 2012 Locus coeruleus. In *The human nervous system* (eds JK Mai, G Paxinos), pp. 425–438. London, UK: Academic Press.
108. Wilson-Mendenhall C, Barrett LF, Barsalou LW. 2013 Neural evidence that human emotions share core affective properties. *Psychol. Sci.* **24**, 947–956. (doi:10.1177/0956797612464242)
109. Li SSY, McNally GP. 2014 The conditions that promote fear learning: prediction error and Pavlovian fear conditioning. *Neurobiol. Learn. Mem.* **108**, 14–21. (doi:10.1016/j.nlm.2013.05.002)
110. Schultz W. 2016 Dopamine reward prediction-error signalling: a two-component response. *Nat. Rev. Neurosci.* **17**, 183–195. (doi:10.1038/nrn.2015.26)
111. Guitart-Masip M, Duzel E, Dolan R, Dayan P. 2014 Action versus valence in decision making. *Trends Cogn. Sci.* **18**, 194–202. (doi:10.1016/j.tics.2014.01.003)
112. Salamone JD, Correa M. 2012 The mysterious motivational functions of mesolimbic dopamine. *Neuron* **76**, 470–485. (doi:10.1016/j.neuron.2012.10.021)
113. Fields HL, Margolis EB. 2015 Understanding opioid reward. *Trends Neurosci.* **38**, 217–225. (doi:10.1016/j.tins.2015.01.002)
114. Schmahmann JD, Pandya DN. 1997 The cerebrocerebellar system. *Int. Rev. Neurobiol.* **41**, 31–60. (doi:10.1016/S0074-7742(08)60346-3)
115. Strick PL, Dum RP, Fiez JA. 2009 Cerebellum and nonmotor function. *Annu. Rev. Neurosci.* **32**, 413–434. (doi:10.1146/annurev.neuro.31.060407.125606)
116. Schmahmann JD. 2010 The role of the cerebellum in cognition and emotion: personal reflections since 1982 on the dysmetria of thought hypothesis, and its historical evolution from theory to therapy. *Neuropsychol. Rev.* **20**, 236–260. (doi:10.1007/s11065-010-9142-x)
117. Power JD *et al.* 2011 Functional network organization of the human brain. *Neuron* **72**, 665–678. (doi:10.1016/j.neuron.2011.09.006)
118. Touroutoglou A, Hollenbeck M, Dickerson BC, Barrett LF. 2012 Dissociable large-scale networks anchored in the right anterior insula subserved affective experience and attention. *Neuroimage* **60**, 1947–1958. (doi:10.1016/j.neuroimage.2012.02.012)

119. Uddin LQ. 2015 Salience processing and insular cortical function and dysfunction. *Neuron* **16**, 55–61. (doi:10.1038/nrn3857)
120. Ullsperger M, Danielmeier C, Jocham G. 2014 Neurophysiology of performance monitoring and adaptive behavior. *Physiol. Rev.* **94**, 35–79. (doi:10.1152/physrev.00041.2012)
121. Lisman JE, Grace AA. 2005 The hippocampal-VTA loop: controlling the entry of information into long-term memory. *Neuron* **46**, 703–713. (doi:10.1016/j.neuron.2005.05.002)
122. Dunsmoor JE, Murty VP, Davachi L, Phelps EA. 2015 Emotional learning selectively and retroactively strengthens memories for related events. *Nature* **520**, 345–348. (doi:10.1038/nature14106)
123. Clark A. 2013 The many faces of precision (Replies to commentaries on 'Whatever next? Neural prediction, situated agents, and the future of cognitive science'). *Front. Psychol.* **4**, 270. (doi:10.3389/fpsyg.2013.00270)
124. Feldman H, Friston KJ. 2010 Attention, uncertainty, and free-energy. *Front. Hum. Neurosci.* **4**, 215. (doi:10.3389/fnhum.2010.00215)
125. 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)
126. 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)
127. John YJ, Zikopoulos B, Bullock D, Barbas H. 2016 The emotional gatekeeper: a computational model of attention selection and suppression through the pathway from the amygdala to the inhibitory thalamic reticular nucleus. *PLoS Comput. Biol.* **12**, e1004722. (doi:10.1371/journal.pcbi.1004722)
128. Zikopoulos B, Barbas H. 2006 Prefrontal projections to the thalamic reticular nucleus form a unique circuit for attentional mechanisms. *J. Neurosci.* **26**, 7348–7361. (doi:10.1523/JNEUROSCI.5511-05.2006)
129. Zikopoulos B, Barbas H. 2012 Pathways for emotions and attention converge on the thalamic reticular nucleus in primates. *J. Neurosci.* **32**, 5338–5350. (doi:10.1523/JNEUROSCI.4793-11.2012)
130. Wilson CR, Gaffan D, Browning PG, Baxter MG. 2010 Functional localization within the prefrontal cortex: missing the forest for the trees? *Trends Neurosci.* **33**, 533–540. (doi:10.1016/j.tins.2010.08.001)
131. Barrett LF. 2006 Solving the emotion paradox: categorization and the experience of emotion. *Pers. Soc. Psychol. Rev.* **10**, 20–46. (doi:10.1207/s15327957pspr1001_2)
132. Barrett LF. 2011 Bridging token identity theory and supervenience theory through psychological construction. *Psychol. Inq.* **22**, 115–127. (doi:10.1080/1047840X.2011.555216)
133. Barrett LF. 2012 Emotions are real. *Emotion* **12**, 413–429. (doi:10.1037/a0027555)
134. Barrett LF. 2013 Psychological construction: a Darwinian approach to the science of emotion. *Emot. Rev.* **5**, 379–389. (doi:10.1177/1754073913489753)
135. Barrett LF. 2014 The conceptual act theory: a précis. *Emot. Rev.* **6**, 292–297. (doi:10.1177/1754073914534479)
136. Murphy G. 2002 *The big book of concepts*. Cambridge, MA: MIT Press.
137. Gallivan JP, Logan L, Wolpert DM, Flanagan JR. 2016 Parallel specification of competing sensorimotor control policies for alternative action options. *Nat. Neurosci.* **19**, 320–326. (doi:10.1038/nn.4214)
138. Barsalou LW. 1983 *Ad hoc* categories. *Mem. Cogn.* **11**, 211–227. (doi:10.3758/BF03196968)
139. Barsalou LW. 2003 Situated simulation in the human conceptual system. *Lang. Cogn. Process.* **18**, 513–562. (doi:10.1080/01690960344000026)
140. Binder JR, Desai RH, Graves WW, Conant LL. 2009 Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb. Cortex* **19**, 2767–2796. (doi:10.1093/cercor/bhp055)
141. Binder JR, Desai RH. 2011 The neurobiology of semantic memory. *Trends Cogn. Sci.* **15**, 527–536. (doi:10.1016/j.tics.2011.10.001)
142. Barsalou LW. 2016 On staying grounded and avoiding Quixotic dead ends. *Psychon. Bull. Rev.* **23**, 1122–1142. (doi:10.3758/s13423-016-1028-3)
143. Fernandino L, Humphries CJ, Seidenberg MS, Gross WL, Conant LL, Binder JR. 2015 Predicting brain activation patterns associated with individual lexical concepts based on five sensory-motor attributes. *Neuropsychologia* **76**, 17–26. (doi:10.1016/j.neuropsychologia.2015.04.009)
144. Fernandino L, Binder JR, Desai RH, Pendl SL, Humphries CJ, Gross WL, Conant LL, Seidenberg MS. 2016 Concept representation reflects multimodal abstraction: a framework for embodied semantics. *Cereb. Cortex* **26**, 2018–2034. (doi:10.1093/cercor/bhv020)
145. Pulvermüller F. 2013 How neurons make meaning: brain mechanisms for embodied and abstract-symbolic semantics. *Trends Cogn. Sci.* **17**, 458–470. (doi:10.1016/j.tics.2013.06.004)
146. Hsu L-M *et al.* 2016 Constituents and functional implications of the rat default mode network. *Proc. Natl Acad. Sci. USA* **113**, E4541–E4547. (doi:10.1073/pnas.1601485113)
147. Bar M. 2009 A cognitive neuroscience hypothesis of mood and depression. *Trends Cogn. Sci.* **13**, 456–463. (doi:10.1016/j.tics.2009.08.009)
148. Bar M. 2009 The proactive brain: memory for predictions. *Phil. Trans. R. Soc. B* **364**, 1235–1243. (doi:10.1098/rstb.2008.0310)
149. Finlay BL, Uchiyama R. 2015 Developmental mechanisms channeling cortical evolution. *Trends Neurosci.* **38**, 69–76. (doi:10.1016/j.tins.2014.11.004)
150. Skerry AE, Saxe R. 2015 Neural representations of emotion are organized around abstract event features. *Current biology* **25**, 1945–1954. (doi:10.1016/j.cub.2015.06.009)
151. Etkin A, Buchel C, Gross JJ. 2015 The neural bases of emotion regulation. *Nat. Rev. Neurosci.* **16**, 693–700. (doi:10.1038/nrn4044)
152. Gross JJ. 2015 Emotion regulation: current status and future prospects. *Psychol. Inq.* **26**, 1–26. (doi:10.1080/1047840X.2014.940781)
153. Lecours S, Robert G, Desruisseaux F. 2009 Alexithymia and verbal elaboration of affect in adults suffering from a respiratory disorder. *Eur. Rev. Appl. Psychol.* **59**, 187–195. (doi:10.1016/j.erap.2009.03.001)
154. Lindquist KA, Barrett LF. 2008 Emotional complexity. In *The handbook of emotion* (eds M Lewis, JM Haviland-Jones, LF Barrett), pp. 513–530. New York, NY: Guilford.
155. Salminen JK, Saarijärvi S, Äärelä E, Toikka T, Kauhanen J. 1999 Prevalence of alexithymia and its association with sociodemographic variables in the general population of Finland. *J. Psychosom. Res.* **46**, 75–82. (doi:10.1016/S0022-3999(98)00053-1)
156. Farb N, Daubenmier J, Price CJ, Gard T, Kerr C, Dunn BD, Klein AC, Paulus MP, Mehling WE. 2015 Interoception, contemplative practice, and health. *Front. Psychol.* **6**, 886. (doi:10.3389/fpsyg.2015.00763)
157. Schacter DL, Addis DR, Buckner RL. 2007 Remembering the past to imagine the future: the prospective brain. *Nat. Rev. Neurosci.* **8**, 657–661. (doi:10.1038/nrn2213)
158. Wilson TD, Gilbert DT. 2003 Affective forecasting. *Adv. Exp. Soc. Psychol.* **35**, 345–411. (doi:10.1016/S0065-2601(03)01006-2)
159. Edelman GM. 1990 *The remembered present: a biological theory of consciousness*. New York, NY: Basic Books.
160. McGrath CL, Kelley ME, Dunlop BW, Holtzheimer III, PE, Craighead WE, Mayberg HS. 2014 Pretreatment brain states identify likely nonresponse to standard treatments for depression. *Biol. Psychiatry* **76**, 527–535. (doi:10.1016/j.biopsych.2013.12.005)
161. Zeng L-L, Shen H, Liu L, Wang L, Li B, Fang P, Zhou Z, Li Y, Hu D. 2012 Identifying major depression using whole-brain functional connectivity: a multivariate pattern analysis. *Brain* **135**, 1498–1507. (doi:10.1093/brain/aws059)
162. Lamers F, Vogelzangs N, Merikangas KR, de Jonge P, Beekman AT, Penninx BW. 2013 Evidence for a differential role of HPA-axis function, inflammation and metabolic syndrome in melancholic versus atypical depression. *Mol. Psychiatry* **18**, 692–699. (doi:10.1038/mp.2012.144)
163. Onyewuenyi IC, Muldoon MF, Christie IC, Erickson KI, Gianaros PJ. 2014 Basal ganglia morphology links the metabolic syndrome and depressive symptoms. *Physiol. Behav.* **123**, 214–222. (doi:10.1016/j.physbeh.2013.09.014)
164. Pan A, Keum N, Okereke OI, Sun DL, Kivimaki M, Rubin RR, Hu FB. 2012 Bidirectional association between depression and metabolic syndrome. *Diab. Care* **35**, 1171–1180. (doi:10.2337/dc11-2055)
165. Shelton RC, Miller AH. 2010 Eating ourselves to death (and despair): the contribution of adiposity and inflammation to depression. *Prog. Neurobiol.* **91**, 275–299. (doi:10.1016/j.pneurobio.2010.04.004)
166. Dowlati Y, Herrmann N, Swardfager W, Liu H, Sham L, Reim EK, Lanctôt KL. 2010 A meta-analysis of

- cytokines in major depression. *Biol. Psychiatry* **67**, 446–457. (doi:10.1016/j.biopsych.2009.09.033)
167. Grosse L *et al.* 2015 Clinical characteristics of inflammation-associated depression: monocyte gene expression is age-related in major depressive disorder. *Brain Behav. Immun.* **44**, 48–56. (doi:10.1016/j.bbi.2014.08.004)
168. Miller AH, Haroon E, Raison CL, Felger JC. 2013 Cytokine targets in the brain: impact on neurotransmitters and neurocircuits. *Depress. Anxiety* **30**, 297–306. (doi:10.1002/da.22084)
169. Swardfager W, Rosenblatt JD, Benlamri M, McIntyre RS. 2016 Mapping inflammation onto mood: inflammatory mediators of anhedonia. *Neurosci. Biobehav. Rev.* **64**, 148–166. (doi:10.1016/j.neubiorev.2016.02.017)
170. Walker A, Kavelaars A, Heijnen C, Dantzer R. 2014 Neuroinflammation and comorbidity of pain and depression. *Pharmacol. Rev.* **66**, 80–101. (doi:10.1124/pr.113.008144)
171. Binder EB, Nemeroff CB. 2010 The CRF system, stress, depression and anxiety—insights from human genetic studies. *Mol. Psychiatry* **15**, 574–588. (doi:10.1038/mp.2009.141)
172. Gold PW. 2015 The organization of the stress system and its dysregulation in depressive illness. *Mol. Psychiatry* **20**, 32–47. (doi:10.1038/mp.2014.163)
173. Juruena MF. 2014 Early-life stress and HPA axis trigger recurrent adulthood depression. *Epilepsy Behav.* **38**, 148–159. (doi:10.1016/j.yebeh.2013.10.020)
174. Mayberg HS *et al.* 1999 Reciprocal limbic-cortical function and negative mood: converging PET findings in depression and normal sadness. *Am. J. Psychiatry* **156**, 675–682. (doi:10.1176/ajp.156.5.675)
175. Alloy LB, Ahrens AH. 1987 Depression and pessimism for the future: biased use of statistically relevant information in predictions for self versus others. *J. Pers. Soc. Psychol.* **52**, 366–378. (doi:10.1037/0022-3514.52.2.366)
176. Harrison NA, Brydon L, Walker C, Gray MA, Steptoe A, Critchley HD. 2009 Inflammation causes mood changes through alterations in subgenual cingulate activity and mesolimbic connectivity. *Biol. Psychiatry* **66**, 407–414. (doi:10.1016/j.biopsych.2009.03.015)
177. Harrison NA, Brydon L, Walker C, Gray MA, Steptoe A, Dolan RJ, Critchley HD. 2009 Neural origins of human sickness in interoceptive responses to inflammation. *Biol. Psychiatry* **66**, 415–422. (doi:10.1016/j.biopsych.2009.03.007)
178. Weierich MR, Wright CI, Negreira A, Dickerson BC, Barrett LF. 2010 Novelty as a dimension of the affective brain. *Neuroimage* **49**, 2871–2878. (doi:10.1016/j.neuroimage.2009.09.047)
179. Edelman GM, Gally JA. 2001 Degeneracy and complexity in biological systems. *Proc. Natl Acad. Sci. USA* **98**, 13 763–13 768. (doi:10.1073/pnas.231499798)
180. Drevets WC, Price JL, Simpson JR, Todd RD, Reich T, Vannier M, Raichle ME. 1997 Subgenual prefrontal cortex abnormalities in mood disorders. *Nature* **386**, 824–827. (doi:10.1038/386824a0)
181. Skaf CR, Yamada A, Garrido GEJ, Buchpiguel CA, Akamine S, Castro CC, Busatto GF. 2002 Psychotic symptoms in major depressive disorder are associated with reduced regional cerebral blood flow in the subgenual anterior cingulate cortex: a voxel-based single photon emission computed tomography (SPECT) study. *J. Affect. Disord.* **68**, 295–305. (doi:10.1016/S0165-0327(00)00365-7)
182. Vardi N, Freedman N, Lester H, Gomori JM, Chisin R, Lerer B, Bonne O. 2011 Hyperintensities on T2-weighted images in the basal ganglia of patients with major depression: cerebral perfusion and clinical implications. *Psychiatry Res. Neuroimaging* **192**, 125–130. (doi:10.1016/j.psychres.2010.11.010)
183. Drevets WC, Öngür D, Price JL. 1998 Neuroimaging abnormalities in the subgenual prefrontal cortex: implications for the pathophysiology of familial mood disorders. *Mol. Psychiatry* **3**, 220–226. (doi:10.1038/sj.mp.4000370)
184. Critchley HD, Mathias CJ, Josephs O, O'Doherty J, Zanini S, Dewar BK, Cipolotti L, Shallice T, Dolan RJ. 2003 Human cingulate cortex and autonomic control: converging neuroimaging and clinical evidence. *Brain* **126**, 2139–2152. (doi:10.1093/brain/awg216)
185. Dunlop BW, Kelley ME, McGrath CL, Craighead WE, Mayberg HS. 2015 Preliminary findings supporting insula metabolic activity as a predictor of outcome to psychotherapy and medication treatments for depression. *J. Neuropsychiatry Clin. Neurosci.* **27**, 237–239. (doi:10.1176/appi.neuropsych.14030048)
186. Cummings DE. 2006 Ghrelin and the short- and long-term regulation of appetite and body weight. *Physiol. Behav.* **89**, 71–84. (doi:10.1016/j.physbeh.2006.05.022)
187. Laposky AD, Bass J, Kojima A, Turek FW. 2008 Sleep and circadian rhythms: key components in the regulation of energy metabolism. *FEBS Lett.* **582**, 142–151. (doi:10.1016/j.febslet.2007.06.079)
188. Havel PJ. 2001 Peripheral signals conveying metabolic information to the brain: short-term and long-term regulation of food intake and energy homeostasis. *Exp. Biol. Med.* **226**, 963–977.
189. Hagobian TA, Sharoff CG, Braun B. 2008 Effects of short-term exercise and energy surplus on hormones related to regulation of energy balance. *Metabolism* **57**, 393–398. (doi:10.1016/j.metabol.2007.10.016)
190. Hagobian TA, Sharoff CG, Stephens BR, Wade GN, Silva JE, Chipkin SR, Braun B. 2008 Effects of exercise on energy-regulating hormones and appetite in men and women. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **296**, R233–R242. (doi:10.1152/ajpregu.90671.2008)
191. Eisenberger NI. 2012 The pain of social disconnection: examining the shared neural underpinnings of physical and social pain. *Nat. Rev. Neurosci.* **13**, 421–434. (doi:10.1038/nrn3239)
192. Eisenberger NI, Cole SW. 2012 Social neuroscience and health: neurophysiological mechanisms linking social ties with physical health. *Nat. Neurosci.* **15**, 669–674. (doi:10.1038/nn.3086)
193. Palumbo RV, Marraccini ME, Weyandt LL, Wilder-Smith O, McGee HA, Liu S, Goodwin MS. In press. Interpersonal autonomic physiology: a systematic review of the literature. *Pers. Soc. Psychol. Rev.*
194. Sbarra DA, Hazan C. 2008 Coregulation, dysregulation, self-regulation: an integrative analysis and empirical agenda for understanding adult attachment, separation, loss, and recovery. *Pers. Soc. Psychol. Rev.* **12**, 141–167. (doi:10.1177/1088868308315702)
195. McLaughlin KA, Sheridan MA, Lambert HK. 2014 Childhood adversity and neural development: deprivation and threat as distinct dimensions of early experience. *Neurosci. Biobehav. Rev.* **47**, 578–591. (doi:10.1016/j.neubiorev.2014.10.012)
196. Sheridan MA, McLaughlin KA. 2014 Dimensions of early experience and neural development: deprivation and threat. *Trends Cogn. Sci.* **18**, 580–585. (doi:10.1016/j.tics.2014.09.001)
197. Ansell EB, Rando K, Tuit K, Guarnaccia J, Sinha R. 2012 Cumulative adversity and smaller gray matter volume in medial prefrontal, anterior cingulate, and insula regions. *Biol. Psychiatry* **72**, 57–64. (doi:10.1016/j.biopsych.2011.11.022)
198. Boes AD, McCormick LM, Coryell WH, Nopoulos P. 2008 Rostral anterior cingulate cortex volume correlates with depressed mood in normal healthy children. *Biol. Psychiatry* **63**, 391–397. (doi:10.1016/j.biopsych.2007.07.018)
199. Levesque ML *et al.* 2011 Altered patterns of brain activity during transient sadness in children at familial risk for major depression. *J. Affect. Disord.* **135**, 410–413. (doi:10.1016/j.jad.2011.08.010)
200. Nolen-Hoeksema S, Wisco BE, Lyubomirsky S. 2008 Rethinking rumination. *Pers. Psychol. Sci.* **3**, 400–424. (doi:10.1111/j.1745-6924.2008.00088.x)
201. Harel E, Tennyson R, Fava M, Bar M. 2016 Linking major depression and the neural substrates of associative processing. *Cogn. Affect. Behav. Neurosci.* (doi:10.3758/s13415-016-0449-9)
202. Gotlib IH, Krasnoperova E, Yue DN, Joermann J. 2004 Attentional biases for negative interpersonal stimuli in clinical depression. *J. Abnormal Psychol.* **113**, 127. (doi:10.1037/0021-843X.113.1.121)
203. Roberts SB, Kendler KS. 1999 Neuroticism and self-esteem as indices of the vulnerability to major depression in women. *Psychol. Med.* **29**, 1101–1109. (doi:10.1017/S0033291799008739)
204. Mazzoni P, Krakauer JW. 2006 An implicit plan overrides an explicit strategy during visuomotor adaptation. *J. Neurosci.* **26**, 3642–3645. (doi:10.1523/JNEUROSCI.5317-05.2006)
205. Craig AD. 2002 How do you feel? Interoception: the sense of the physiological condition of the body. *Nat. Rev. Neurosci.* **3**, 655–666. (doi:10.1038/nrn894)
206. Oppenheimer S, Cechetto D. 2011 The insular cortex and the regulation of cardiac function. *Compr. Physiol.* **6**, 1081–1133. (doi:10.1002/cphy.c140076)
207. Paulus MP, Stein MB. 2010 Interoception in anxiety and depression. *Brain Struct. Funct.* **214**, 451–463. (doi:10.1007/s00429-010-0258-9)

208. Fields H. 2004 State-dependent opioid control of pain. *Nat. Rev. Neurosci.* **5**, 565–575. (doi:10.1038/nrn1431)
209. Tseng YW, Diedrichsen J, Krakauer JW, Shadmehr R, Bastian AJ. 2007 Sensory prediction errors drive cerebellum-dependent adaptation of reaching. *J. Neurophysiol.* **98**, 54–62. (doi:10.1152/jn.00266.2007)
210. Raison CL, Miller AH. 2013 Malaise, melancholia and madness: the evolutionary legacy of an inflammatory bias. *Brain Behav. Immun.* **31**, 1–8. (doi:10.1016/j.bbi.2013.04.009)
211. Maes M, Berk M, Goehler L, Song C, Anderson G, Galecki P, Leonard B. 2012 Depression and sickness behavior are Janus-faced responses to shared inflammatory pathways. *BMC Med.* **10**, 66. (doi:10.1186/1741-7015-10-66)
212. Slavich GM, Irwin MR. 2014 From stress to inflammation and major depressive disorder: a social signal transduction theory of depression. *Psychol. Bull.* **140**, 774–815. (doi:10.1037/a0035302)
213. Marchetti I, Koster EH, Sonuga-Barke EJ, De Raedt R. 2012 The default mode network and recurrent depression: a neurobiological model of cognitive risk factors. *Neuropsychol. Rev.* **22**, 229–251. (doi:10.1007/s11065-012-9199-9)
214. Hamilton JP, Farmer M, Fogelman P, Gotlib IH. 2015 Depressive rumination, the default-mode network, and the dark matter of clinical neuroscience. *Biol. Psychiatry* **78**, 224–230. (doi:10.1016/j.biopsych.2015.02.020)
215. Pizzagalli DA, Iosifescu D, Hallett LA, Ratner KG, Fava M. 2008 Reduced hedonic capacity in major depressive disorder: evidence from a probabilistic reward task. *J. Psychiatr. Res.* **43**, 76–87. (doi:10.1016/j.jpsychires.2008.03.001)
216. Bromberg-Martin ES, Hikosaka O, Nakamura K. 2010 Coding of task reward value in the dorsal raphe nucleus. *J. Neurosci.* **30**, 6262–6272. (doi:10.1523/JNEUROSCI.0015-10.2010)
217. Cools R, Nakamura K, Daw ND. 2011 Serotonin and dopamine: unifying affective, activational, and decision functions. *Neuropsychopharmacology* **36**, 98–113. (doi:10.1038/npp.2010.121)
218. Miyazaki K, Miyazaki KW, Doya K. 2011 Activation of dorsal raphe serotonin neurons underlies waiting for delayed rewards. *J. Neurosci.* **31**, 469–479. (doi:10.1523/JNEUROSCI.3714-10.2011)
219. Doya K. 2008 Modulators of decision making. *Nat. Neurosci.* **11**, 410–416. (doi:10.1038/nn2077)
220. Dunlop BW, Nemeroff CB. 2007 The role of dopamine in the pathophysiology of depression. *Arch. Gen. Psychiatry* **64**, 327–337. (doi:10.1001/archpsyc.64.3.327)
221. Choi KS, Riva-Posse P, Gross RE, Mayberg HS. 2015 Mapping the ‘depression switch’ during intraoperative testing of subcallosal cingulate deep brain stimulation. *JAMA Neurol.* **72**, 1252–1260. (doi:10.1001/jamaneurol.2015.2564)
222. Kennedy SH, Giacobbe P, Rizvi SJ, Placenza FM, Nishikawa Y, Mayberg HS, Lozano AM. 2011 Deep brain stimulation for treatment-resistant depression: follow-up after 3 to 6 years. *Am. J. Psychiatry* **168**, 502–510. (doi:10.1176/appi.ajp.2010.10081187)
223. Mayberg HS, Lozano AM, Voon V, McNeely HE, Seminowicz D, Hamani C, Schwab JM, Kennedy SH. 2005 Deep brain stimulation for treatment-resistant depression. *Neuron* **45**, 651–660. (doi:10.1016/j.neuron.2005.02.014)
224. Hamani C, Mayberg H, Snyder B, Giacobbe P, Kennedy S, Lozano AM. 2009 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. (doi:10.3171/2008.10.JNS08763)
225. Riva-Posse P *et al.* 2014 Defining critical white matter pathways mediating successful subcallosal cingulate deep brain stimulation for treatment-resistant depression. *Biol. Psychiatry* **76**, 963–969. (doi:10.1016/j.biopsych.2014.03.029)
226. Sigalas PD, Garg H, Watson S, McAllister-Williams RH, Ferrier IN. 2012 Metyrapone in treatment-resistant depression. *Ther. Adv. Psychopharmacol.* **2**, 139–149. (doi:10.1177/2045125312436597)
227. Ferrier IN *et al.* 2015 Randomised controlled trial of antigluocorticoid augmentation (metyrapone) of antidepressants in depression (ADD Study). *Efficacy and Mechanism Evaluation*. 2.4. Southampton, UK: NIHR Journals Library.
228. Miller AH, Raison CL. 2015 Are anti-inflammatory therapies viable treatments for psychiatric disorders? *JAMA Psychiatry* **72**, 527–528. (doi:10.1001/jamapsychiatry.2015.22)
229. Eyre HA, Air T, Proctor S, Rositano S, Baune BT. 2015 A critical review of the efficacy of non-steroidal anti-inflammatory drugs in depression. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **57**, 11–16. (doi:10.1016/j.pnpbp.2014.10.003)
230. Zunsain PA. 2015 Improving the treatment for depressive symptoms and major depression with anti-inflammatory drugs. *Evid. Based Ment. Health* **18**, 116. (doi:10.1136/eb-2015-102063)
231. Mammen G, Faulkner G. 2013 Physical activity and the prevention of depression: a systematic review of prospective studies. *Am. J. Prev. Med.* **45**, 649–657. (doi:10.1016/j.amepre.2013.08.001)
232. Cooney G, Dwan K, Mead G. 2014 Exercise for depression. *JAMA* **311**, 2432–2433. (doi:10.1001/jama.2014.4930)
233. Nemeroff CB, Mayberg HS, Krahl SE, McNamara J, Frazer A, Henry TR, George MS, Charney DS, Brannan SK. 2006 VNS therapy in treatment-resistant depression: clinical evidence and putative neurobiological mechanisms. *Neuropsychopharmacology* **31**, 1345–1355. (doi:10.1038/sj.npp.1301082)
234. Khalsa SS, Feinstein JS, Li W, Feusner JD, Adolphs R, Hurlmann R. 2016 Panic anxiety in humans with bilateral amygdala lesions: pharmacological induction via cardiorespiratory interoceptive pathways. *J. Neurosci.* **36**, 3559–3566. (doi:10.1523/JNEUROSCI.4109-15.2016)
235. Khalsa SS, Craske MG, Vangala S, Strober M, Feusner JD. 2016. Altered interoceptive awareness in anorexia nervosa: effects of meal anticipation, consumption and bodily arousal. *Int. J. Eating Disord.* **48**, 889–897. (doi:10.1002/eat.22387)
236. Kjellgren A, Westman J. 2014 Beneficial effects of treatment with sensory isolation in flotation-tank as a preventive health-care intervention—a randomized controlled pilot trial. *BMC Complement. Altern. Med.* **14**, 383. (doi:10.1186/1472-6882-14-417)
237. Aupperle RL, Allard CB, Simmons AN, Flagan T, Thorp SR, Norman SB, Paulus MP, Stein MB. 2013 Neural responses during emotional processing before and after cognitive trauma therapy for battered women. *Psychiatry Res.* **214**, 48–55. (doi:10.1016/j.psychres.2013.05.001)
238. Goldapple K, Segal Z, Garson C, Lau M, Bieling P, Kennedy S, Mayberg H. 2004 Modulation of cortical-limbic pathways in major depression: treatment-specific effects of cognitive behavior therapy. *Arch. Gen. Psychiatry* **61**, 34–41. (doi:10.1001/archpsyc.61.1.34)
239. McGrath CL, Kelley ME, Holtzheimer PE, Dunlop BW, Craighead WE, Franco AR, Craddock RC, Mayberg HS. 2013 Toward a neuroimaging treatment selection biomarker for major depressive disorder. *JAMA Psychiatry* **70**, 821–829. (doi:10.1001/jamapsychiatry.2013.143)
240. McCormick LM, Ponto LLB, Pierson RK, Johnson HJ, Magnotta V, Brumm MC. 2007 Metabolic correlates of antidepressant and antipsychotic response in patients with psychotic depression undergoing electroconvulsive therapy. *J. ECT* **23**, 265–273. (doi:10.1097/lyct.0b013e318150d56d)