

# Computational Psychosomatics and Computational Psychiatry: Toward a Joint Framework for Differential Diagnosis

Frederike H. Petzschner, Lilian A.E. Weber, Tim Gard, and Klaas E. Stephan

## ABSTRACT

This article outlines how a core concept from theories of homeostasis and cybernetics, the inference-control loop, may be used to guide differential diagnosis in computational psychiatry and computational psychosomatics. In particular, we discuss 1) how conceptualizing perception and action as inference-control loops yields a joint computational perspective on brain-world and brain-body interactions and 2) how the concrete formulation of this loop as a hierarchical Bayesian model points to key computational quantities that inform a taxonomy of potential disease mechanisms. We consider the utility of this perspective for differential diagnosis in concrete clinical applications.

**Keywords:** Allostasis, Cybernetics, Hierarchical Bayesian model, Homeostasis, Inference, Metacognition, Prediction error

<http://dx.doi.org/10.1016/j.biopsych.2017.05.012>

Psychiatry faces major challenges: its nosology is agnostic about mechanisms, lacks predictive validity, and leads to trial-and-error treatment (1,2). Strikingly, neuroscientific advances have hardly affected nosology or clinical practice (3). One response to this disconnect is computational psychiatry, with its emerging focus on clinical applications (4–11). One strategy for computational psychiatry is to learn from internal medicine, where mechanistic frameworks for differential diagnosis enable targeted treatment decisions for individual patients. Importantly, differential diagnosis does not necessarily require molecular mechanisms. Much coarser distinctions—inflammatory, infectious, vascular, neoplastic, autoimmunological, or hereditary causes of disease—can provide crucial guidance for treatment, as they disclose fundamentally distinct disease processes.

This article outlines a framework for differential diagnosis that is motivated by a general computational perspective on brain function. While not the first attempt of its kind (4,12–14), this article makes three contributions. First, we adopt a disease-independent motif—the inference-control loop as fundament of cybernetic theories (15–18)—and consider how this may help in systematizing computational perspectives on brain-world and brain-body interactions. Second, we consider a hierarchical Bayesian implementation that suggests three possible computational quantities (predictions, prediction errors [PEs], and their precisions) at five potential failure loci (sensation, perception, metacognition, forecasting, action). Third, we discuss the potential clinical utility of this taxonomy for differential diagnosis in computational psychiatry and psychosomatics [compare (14,19–21)].

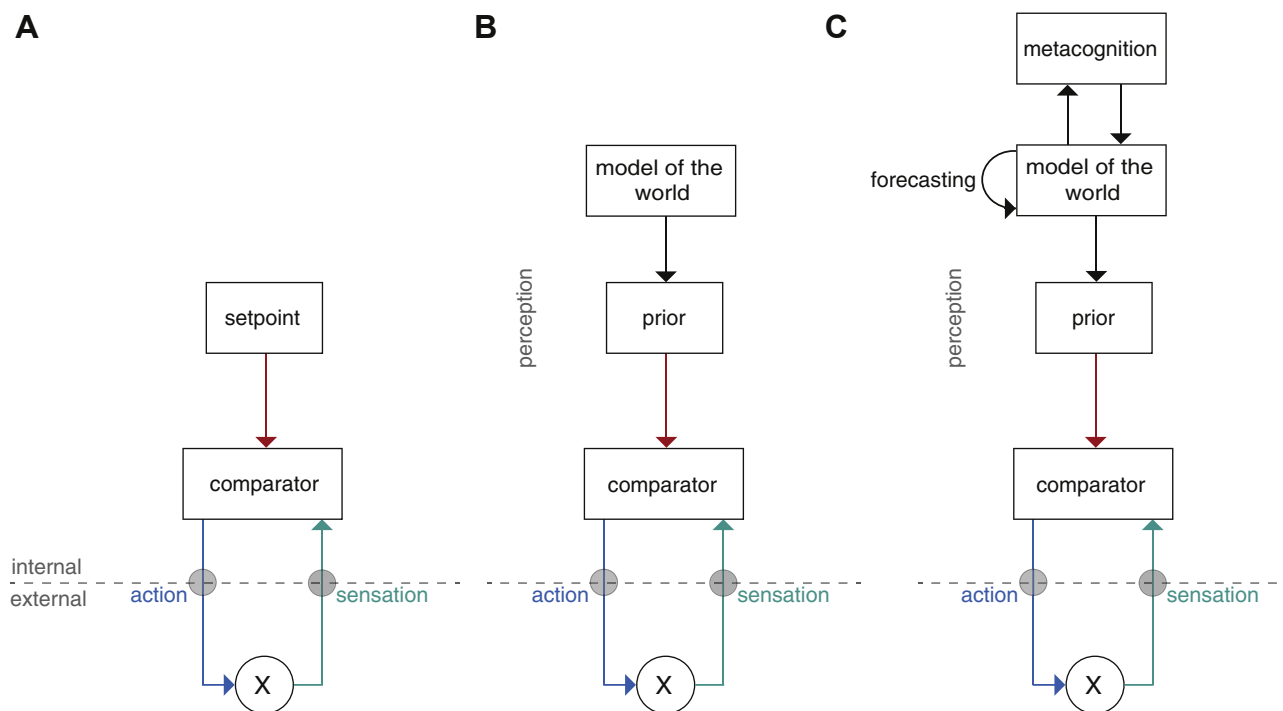
## INFERENCE-CONTROL LOOPS

Different theories of adaptive behavior exist, but they share common themes. We focus on the closed loop between sensations and actions that is at the core of classical cybernetic theories (15,16) and homeostatic principles (22,23). We first summarize extended cybernetic/homeostatic theories (Figure 1) before considering one particular implementation as a foundation for a joint taxonomy of disease mechanisms in psychiatry and psychosomatics.

A useful starting point to reflect on adaptive behavior is the observation that it must be constrained by requirements of bodily homeostasis. In the simplest case, actions can be purely reactive. For instance, to maintain constant body temperature, sensor information can be compared with a pre-defined set-point (e.g., 37°C). Actions, such as heating or cooling the body, are then selected to bring sensory inputs closer to that set-point. This reflex arc—which implements the same feedback control as a simple thermostat—is illustrated in Figure 1A.

If biological systems were like thermostats, with unambiguous sensory inputs and purely reactive in nature, simple feedback control would be sufficient. However, biological systems face three major challenges.

First, sensations (inputs from sensory channels) (see Glossary in the Supplement) are noisy and often highly ambiguous because the world's states (body or environment) that excite sensors can interact nonlinearly and/or hierarchically (24,25). It has long been recognized that the world's true state is not directly accessible for the brain and needs to be inferred (26,27). This notion of perception as inference



**Figure 1.** (A) Simple example of a homeostatic reflex arc as described by classical cybernetics. Sensory inputs (sensations) about an environmental quantity “X” (e.g., current body temperature) are compared with a predefined set-point (e.g., ideal body temperature). Corrective actions occur as a function of the mismatch between input and set-point, such that “X” is moving closer to the set-point (e.g., heating or cooling the body). (B) Extension to an inference-control loop, where perception (inference of environmental states) under an individual’s generative model of the world updates beliefs that change the reflex arc’s set-point (e.g., allostatic control of bodily states); in other cases, actions might be chosen based on the perception rather than the sensation (not shown here). (C) Further extension of the inference-control loop to include forecasting and metacognition. We wish to emphasize that this plot is highly schematic and provides a core summary of different types of inference-control loops; it should not be misunderstood as a detailed circuit proposal.

renders perception an interpretation of sensations, guided by prior beliefs and a model of the world (28,29). Among many findings supporting this notion, illusions prominently illustrate how learned physical regularities can shape perception profoundly (Figure 2A) (30,31). When sensations are ambiguous, perception can expand the capacity for control, particularly when action selection requires information about hierarchically deep states of the world that relate nonlinearly to sensations. For example, in social interactions, inferring the nature of others’ acts that generated visual input may not be sufficient; instead, inference on deeper states, such as the intentions of others that generated their acts, may be required (32,33).

Second, inference on current states of the world can only finesse reactive control. By contrast, prospective control requires predicting the world’s future states (forecasting), taking into account both the influence of possible actions (34,35) and the world’s endogenous dynamics (Figure 1C) (36).

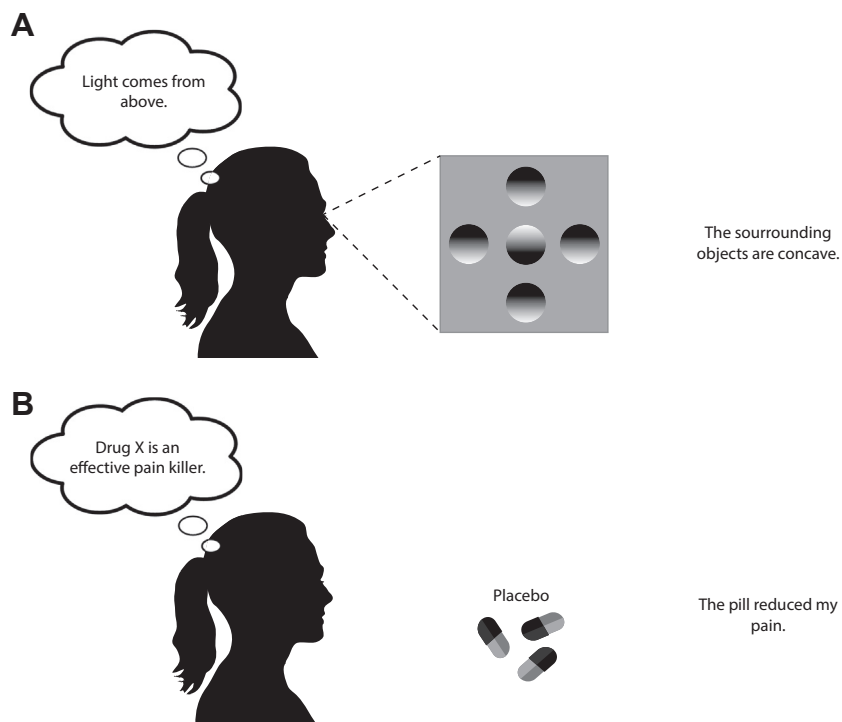
Third, action selection and execution are influenced by beliefs about one’s abilities (37). This self-monitoring of one’s level of mastery in acting on the world is part of metacognition and can be seen as a high-level form of inference about one’s capacity for control (Figure 1C) (19,38).

Finally, given an inferred (or forecast) state of the world, actions can be selected to achieve a particular goal (optimize

some objective function). This objective function can be defined differently—in terms of utility (39), reward (34), cost (40), loss (41), or surprise (42). Figure 1C depicts a schematic illustration of the extended inference-control loop. Importantly, any given action alters the world, thus shaping future sensory input. In other words, sensation, perception, forecasting, and actions form a closed loop between the brain and its external world. For brevity, we refer to this entire cycle as inference-control loop. Its closed-loop nature is fundamentally important, as it creates problems of circular causality that are at the core of diagnostic challenges we examine below.

### COMPUTATIONAL MODELING OF INFERENCE-CONTROL LOOPS

We now consider how inference-control loops can be formalized as concrete computational models. We adopt hierarchical Bayesian models (HBMs) here but emphasize that this is not the only possible perspective; for forecasting and control in particular, alternative (and arguably more established) modeling approaches exist [e.g., (34,43–45)]. We prefer a hierarchical Bayesian view for two main reasons. First, it uses the same formalism and quantities—precision-weighted predictions and precision-weighted PEs (pwPEs)—for implementing perception, forecasting, reactive/prospective control,



**Figure 2.** Two examples of perceptual inference. From left to right: prior belief, sensory data, resulting perception (posterior). **(A)** A classical example of a visual illusion. We perceive the surrounding objects in the image as concave and the center object as convex, even though the sensory data stem from a two-dimensional gray-scale image. The reason is that humans (likely resulting from experience) hold an implicit belief that light comes from above. If light comes from above, the shadow of a concave object should be located at the top, while the shadow of a convex object should be located at the bottom. The resulting percept is thus a reinterpretation of current sensory input based on an implicit a priori belief about lights and shadows. **(B)** Example of the placebo effect. Treatment with drugs that contain no therapeutic ingredient can alter the perception of a physical condition (e.g., reduce physical pain) and elicit autonomic reactions (e.g., an immune response). Again, the change in perception depends on a prior (implicit) belief—here, that the treatment will be effective. Notably, the placebo effect scales with the predicted efficacy of the intervention (for example, syringes are typically considered more potent than pills).

and metacognition. This suggests a compact taxonomy of computational dysfunctions and their differential diagnosis [compare (2,14)]. Second, the formulation of control in HBMs is intimately connected to concepts of homeostatic (reactive) and allostatic (prospective) control, which are of central importance for psychosomatics.

### Bayesian Inference

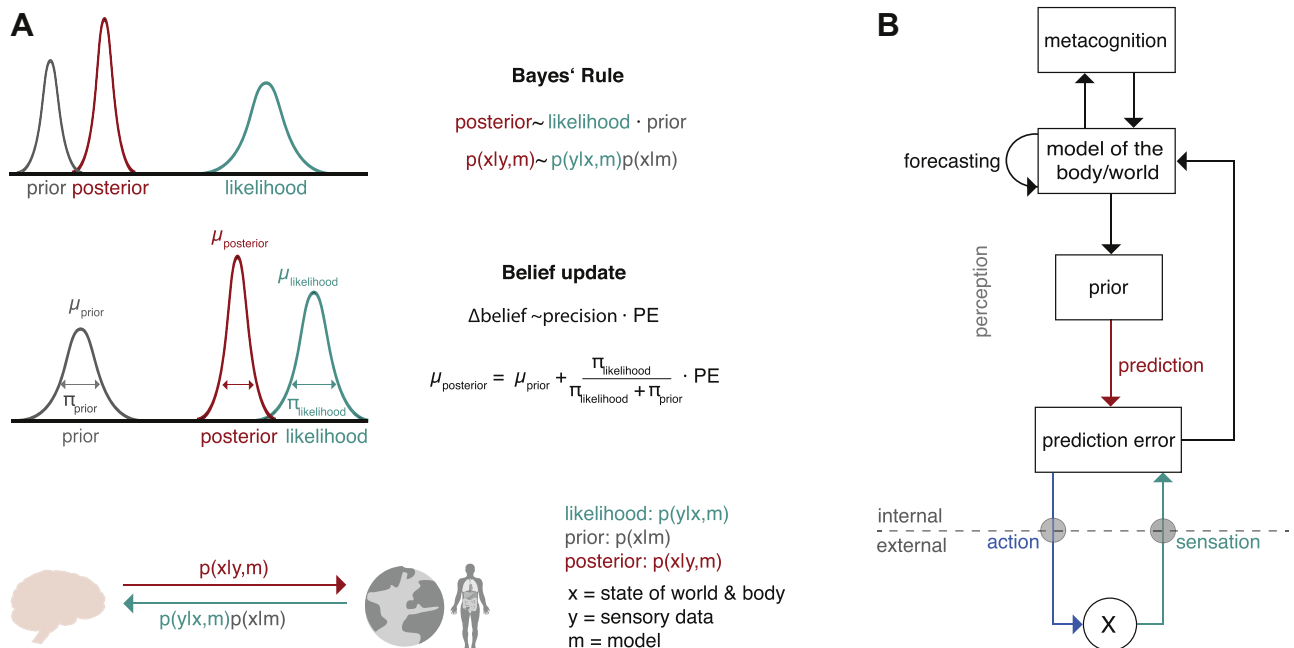
A widely adopted concept of perception is the Bayesian framework (27,28,46,47). This casts perception as inference, where prior beliefs about hidden states of the world are updated in the light of sensory data to yield a posterior belief (Figure 3A) (24,27). A popular notion is that this computation rests on a (hierarchical) “generative model” of how sensory data are caused by hidden states of the world (Figure 3A) (28,29,48,49). Inverting this model under beliefs about the states’ a priori probability allows for inferring the causes of sensations.

Bayesian models explain phenomena across the spectrum of perception; for example, how humans combine multisensory information (50,51) and how biases and illusions result from prior beliefs and experience (24,30,52,53). A key point for this article is that Bayesian belief updates have, for most probability distributions, a generic form: the change in belief is proportional to PE—the difference between actual (sensory) data and predicted data (under the prior)—weighted by a precision ratio (54). The latter is critical, as it determines the relative influence of prior and sensory data: precise predictions (priors reduce, while precise sensory inputs increase, belief updates) (Figure 3A). Generally, abnormal computations and/or signaling of any of these three quantities—PEs, predictions, and precisions—could disrupt inference.

### Hierarchical Bayesian Models

The hierarchical structure of the external world suggests an equivalent (mirrored) structure of the brain’s generative model (28,48). Anatomically, this “hierarchical Bayesian” idea is supported by structural hierarchies in cortex (55–57). Popular HBMs include hierarchical filtering (54,58) and predictive coding (28,49). In these models, each level holds a belief (prediction) about the state of the level below (in predictive coding) or its rate of change (in hierarchical filtering). This prediction is signaled to the lower level, where it is compared against the actual state, resulting in a PE. This PE is sent back up the hierarchy to update the prediction—and thus reduce future PEs. Critically, again, this update is weighted by a precision ratio (Figure 3A): higher precision of bottom-up signals (sensory inputs or PEs) or lower precision of predictions leads to more pronounced belief updates. Neurobiologically, in cortex, predictions are likely signaled via *N*-methyl-D-aspartate receptors at descending connections, and PEs are likely signaled via alpha-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid receptors (and possibly *N*-methyl-D-aspartate receptors) at ascending connections, while precision weighting depends on postsynaptic gain; this is determined by neuromodulators [e.g., dopamine, acetylcholine (59)] and gamma-aminobutyric acidergic inhibition [for reviews, see (60–64)].

In an HBM context, the brain’s objective function can be seen as minimizing PEs (as a proxy to surprise) under its generative model (65). Notably, PEs can be reduced not only by updating the generative model (as above) but also by changing the precision of sensory channels (attention) or by actions that fulfill predictions. The latter is “active inference” (35,42,66), a concept in line with the cybernetic notion



**Figure 3.** Schematic of inference-control in a Bayesian framework. **(A)** (Top panel) Illustration of Bayes' rule using Gaussian distributions as an example. Bayes' rule describes how different information sources—prior beliefs (predictions based on a model of the environment and the body within) and new sensory data (likelihood)—are combined to update the belief (posterior). The amount of belief update is proportional to the prediction error (PE)—the difference between predicted (prior) and actual (sensory) data—weighted by a precision ratio ( $\pi$ , inverse variance) of prior beliefs and sensory inputs (likelihood), respectively. Simply speaking, precise prior beliefs diminish and precise sensory data increase the impact of PEs on belief updates. (Bottom panel) Illustration of the concept of a generative model. A generative model infers hidden states of the world (environment or body) by inverting a probabilistic forward model from those states to possible sensory data (likelihood), under prior beliefs about the values of the hidden states. Inverting a generative model thus corresponds to the application of Bayes' rule. Notably, the mapping from states to data can be mechanistically interpretable (e.g., biophysical models of neuronal responses) or descriptive, such as noisy fluctuations around a constant value or a periodic function (compare circadian rhythms of bodily states). **(B)** Example of an inference-control loop that is cast as a hierarchical Bayesian model. This figure is not meant to provide a detailed description, nor does it claim to represent the only possible layout. Briefly, the key premise here is that the brain represents and updates generative models (“model of the body/world”), with hierarchically structured beliefs. A low-level belief about a bodily/environmental state “ $x$ ” (“prior”) is displayed separately from the rest of the model. The expected sensory inputs (under this prior) can be compared against actual sensations to yield a PE; this PE can be sent up the inference hierarchy and update the model. Switching from perception to action requires (temporarily) abolishing sensory precision [see (19,42) for details]. Actions can then be implemented in two main ways. Homeostatic (reactive) control unfolds as a direct function of PE and serves to fulfill beliefs about sensory input [as encoded by the prior; this can be seen as a probabilistic set-point (19)]. Allostatic (predictive) control prospectively shifts this probabilistic set-point to elicit actions; this requires predicting future states as a function of actions and bodily/environmental dynamics (“forecasting”). Finally, metacognition could be implemented as an additional layer in the model that holds (and updates) expectations with regard to the amount of PE at the top of the inference hierarchy (19).

that “... control systems control what they sense, not what they do” (17).

### Forecasting, Action, and Metacognition in HBMs

While HBMs are popular models of perceptual inference, they can also implement forecasting, action, and metacognition; again, this rests on pwPEs. Switching from inference to forecasting and actions requires “switching off” sensory precision (sensory attenuation) (66–68); this abolishes belief updates, while PEs are now used as simulation or action signals (19,36,42).

While different formalisms of forecasting exist (34,43,44,69), their common theme is a “forward simulation” under a given model. Bayesian implementations of forecasting include “planning by inference” (36,70) and inference on trajectories of states (generalized coordinates) (42,71). One challenge for psychiatric and psychosomatic applications is that the model often needs to predict not only the effects of chosen actions

but also the intrinsic dynamics of environment and body (36,45).

Turning to action, HBMs can implement both reactive and prospective control. The former occurs through a reflex arc at the bottom of the hierarchy (Figure 3B). Specifically, by replacing classical cybernetic set-points with beliefs about hidden states that cause sensory inputs, reactive control can be cast as a reflex where PEs elicit corrective actions that minimize surprise about sensory inputs (19,42). Importantly, the precision of the belief determines the vigor of these actions (19)—a property that allows for new explanations of psychosomatic phenomena and placebo effects (see below). Prospective control can be implemented by dynamically adjusting this belief (e.g., its mean or precision) as a function of predicted future states (19,72). These predictions could be signaled from higher levels in the HBM that implement forecasting.

Action selection in HBMs could, in principle, proceed with respect to optimizing any chosen objective function, e.g., a subject-specific utility function (73). We focus on active

inference (35,42,66) as a specific proposal. Simply speaking, this postulates that actions serve to minimize PEs by changing the world (environment or body) to fulfill the brain's expectation of sensory inputs. We focus on this idea because it is closely related to cybernetics [e.g., perceptual control theory (17)] and represents a probabilistic formulation of the core principle of homeostasis—that regulatory actions minimize discrepancies between expected and actual inputs. It thus provides a basis for formal models of brain-body interactions (19) and a bridge to psychosomatics.

Finally, metacognition could be incorporated into HBMs through an additional layer that holds expectations about the level of PEs throughout an inference hierarchy (Figure 1) (19). This layer infers the performance of the inference-control loop as a whole, enabling a representation (and updating) of mastery or self-efficacy beliefs.

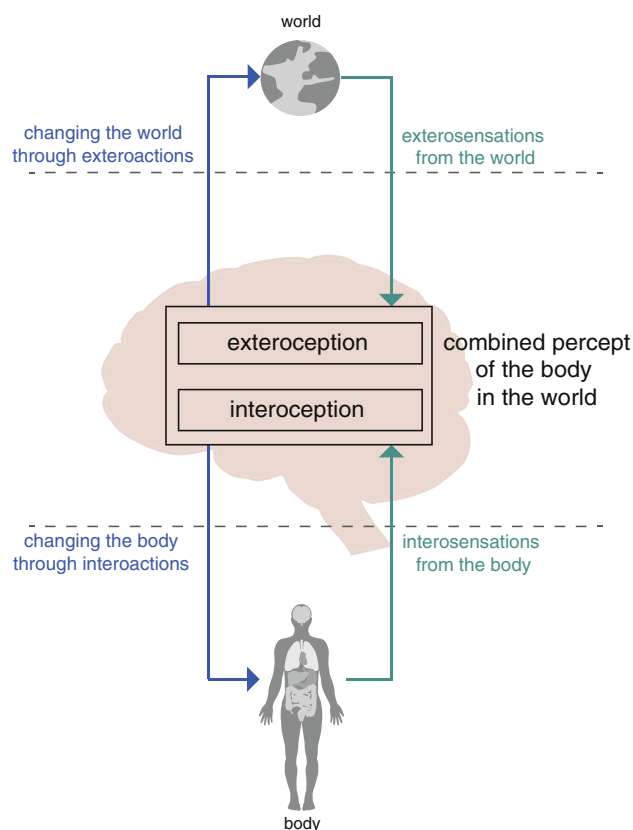
### INTEROCEPTION AND HOMEOSTATIC/ALLOSTATIC CONTROL

HBMs have been used for more than 2 decades to investigate perceptual inference on environmental states (exteroception) (48,49,54,74,75). However, the same inference challenge exists with regard to bodily states (interoception) (76–78). Signals from bodily sensors (interosensations)—blood oxygenation and osmolality, temperature, pain, heart rate, or plasma concentrations of metabolites and hormones—reach the brain through various afferent pathways that converge on posterior and/or mid insula cortex (79–81), a region regarded as interoceptive cortex (82,83). Several lines of evidence—in particular from pain and placebo research [for reviews, see (79,84–86)]—indicate that interosensations are not processed “raw” but are shaped by prior beliefs (Figure 2B).

Supported by anatomical and physiological findings [for reviews, see (77,87)], it has been proposed that perception and control of bodily states follow the same hierarchical Bayesian principles as for environmental states (76–78). This implies a joint computational approach to characterizing disease mechanisms in exteroceptive (psychiatry) and interoceptive (psychosomatics) domains (Figure 4).

Notably, regulation of bodily states comes in two forms. Homeostatic control (22) is a form of reactive control (23) that is classically formalized as cybernetic feedback control (15). The more recent concept of allostasis (“stability through change”) (88) refers to prospective control, where actions are taken before homeostasis is violated. Put differently, allostasis is a self-initiated temporary change in homeostatic set-points to prepare for a predicted external perturbation (89). When replacing classical set-points with homeostatic beliefs (expectations about bodily states), both can be cast formally as active inference (19). Homeostatic control can then be understood as reflex-like emission of corrective actions that fulfill beliefs about bodily states, and allostatic control can be understood as changing homeostatic beliefs under guidance by higher beliefs or forecasts about future perturbations of bodily states (19).

Neuroanatomically concrete circuits for interoception and homeostatic/allostatic control have been suggested (19,78,80,81,87). Anterior insula (AI) and anterior cingulate cortex (ACC) play a central role in these proposals, as they are thought to represent current and predicted states of the body



**Figure 4.** Highly schematic illustration of the inference-control loop for interoception and exteroception. For exteroception, exterosensations (sensory inputs caused by states of the external environment) originate from receptors (e.g., mechanoreceptors, proprioceptors, photoreceptors) and are transmitted via the classical sensory channels (vision, audition, touch, taste, smell) to reach the brain's primary sensory areas. From the perspective of perception as inference, exterosensations are combined with a priori beliefs, based on a model of the environment, resulting in a perception of the environment that is referred to as exteroception. For interoception, interosensations (sensory inputs caused by bodily states) originate from various bodily receptors (baroreceptors, chemoreceptors, thermoreceptors, etc.). Interosensations carry information about bodily states, such as temperature, pain, itch, blood oxygenation, intestinal tension, heart rate, hormonal concentration, etc., and reach the brain via two major afferent pathways: small-diameter, modality-specific afferent fibers in lamina 1 of the spinal cord that project to specific thalamocortical nuclei and the vagus and glossopharyngeal nerves projecting to the nucleus of the solitary tract. Both pathways converge on the posterior insula cortex. From the perspective of perception as inference, interosensations are combined with a priori beliefs, based on a model of the body, resulting in a perception of the body that is referred to as interoception. Interoception and exteroception combined yield the percept of the body within its environment that informs action selection with regard to both internally directed (autonomic) and externally directed (motor) actions.

within the external world (80,90,91). Equipped with projections to regions with homeostatic reflex arcs (e.g., hypothalamus, brainstem), AI and ACC may signal the forecasts that guide allostatic control (19,80,87). Furthermore, they likely interface interoceptive and exteroceptive systems and mediate their interactions, such as the influence of interoceptive signals on exteroceptive judgments (Figure 4) (92–94).

## TAXONOMY OF FAILURE LOCI AND COMPUTATIONAL DYSFUNCTIONS

Our general thesis is that conceptualizing adaptive behavior in terms of inference-control loops and their concrete implementation as HBMs systematizes potential failure loci and associated computational dysfunctions. The ensuing taxonomy of disease mechanisms could guide differential diagnosis, in analogous ways for computational psychiatry and computational psychosomatics. That is, in the general inference-control loop outlined above, maladaptive behavior could arise from primary disruptions at five major loci (Figure 3B): 1) sensory inputs (sensations), 2) inference (perception), 3) forecasting, 4) control (action), and 5) metacognition.

Clearly, each of these processes could be conceptualized under different computational frameworks. In the specific case of HBMs, failures at any of these levels can arise from disturbances in a small set of computational quantities (Figure 3A): 1) bottom-up signals (sensory input or PEs), 2) top-down signals (expectations or predictions), and 3) their precision (inverse uncertainty).

These two axes may lend useful overarching structure to pathogenetic considerations and provide a conceptual grid for classifying disease mechanisms in computational psychiatry and computational psychosomatics. However, this requires that the above levels and quantities can be inferred non-invasively in individual patients, using computational assays that can be applied to behavioral, (neuro)physiological, and neuroimaging data (6,95). Suitable techniques for model-based inference on pwPE signaling in cortical hierarchies exist (96); owing to space limitations, we discuss them in the Supplement.

## Computational Psychosomatics

Psychosomatic medicine is concerned with somatic diseases that are caused or influenced by mental processes (97), for example, bodily symptoms caused by beliefs. Classic examples for the influence of beliefs on bodily states are placebo and nocebo effects (Figure 2B) (84–86). In placebo and nocebo effects, expectations about the effects of an intervention trigger reactions that fulfill the expectation. Importantly, the strength of placebo is known to depend not only on beliefs about effect amplitude but also on the precision of this belief (85). Our framework offers a formal explanation for this empirical phenomenon because in HBM implementations of homeostatic control, the vigor of belief-fulfilling actions depends on the precision of the beliefs (19).

Computational treatments of psychosomatic disorders are rare [but see (98,99)]. This may be due to the (perceived) lack of a comprehensive framework that formalizes interoception and homeostatic/allostatic control and makes them measurable in individual patients. In the following section, we consider one concrete problem of differential diagnosis and describe how the conceptual grid described above may guide the search for the locus of the primary (initial) abnormality.

## Example: Depression and Somatic Symptoms

Many patients with depression have somatic abnormalities, including cardiac (100), immunological (101), and metabolic disturbances (102). One long-standing explanation of this

association highlights maladaptive beliefs. For example, false high-level beliefs about volatility of the world could cause prolonged allostatic responses, with persistent sympathetic activation and ensuing damage to cardiovascular, immunological, and metabolic health (“allostatic load”) (89,103). In our framework, the influence of high-level beliefs could be mediated via projections from allostatic control regions (e.g., AI, ACC) on sympathetic effector regions (e.g., hypothalamus, amygdala, or periaqueductal gray) where they elicit autonomic actions by altering homeostatic set-points. Notably, HBMs can infer fluctuations in beliefs about environmental volatility from behavioral and peripheral physiological measurements (32,58,75,98). These belief trajectories could be integrated into physiological models [e.g., dynamic causal models (96,104)] of the above connection strengths and, by comparing models with and without modulatory effects of these beliefs, identify patients in whom bodily symptoms are possible consequences of beliefs. One might also hypothesize that these connection strengths correlate with peripheral indices of sympathetic activation (105).

An opposite interpretation views depression as “reactive” to initial somatic disease. In our framework, this can be formalized as a metacognitive response to (real or perceived) chronic dyshomeostasis. One implementation of metacognition in HBMs is through a top-level layer that holds beliefs about the performance of the inference-control loop. In this “allostatic self-efficacy” (19) concept, persistently elevated PEs decrease one’s beliefs of mastery over bodily states; this metacognitive “diagnosis” of lack of control may lead to depression as a form of learned helplessness. This proposal could be tested by correlating model-based indices of interoceptive PE signaling with questionnaire measures of self-efficacy and helplessness.

Critically, our framework emphasizes that dyshomeostasis could be real or perceived and could exist independently from the brain or be caused by it:

1. A real bodily source of dyshomeostasis (that evades cerebral attempts of regulation).
2. Sensations—altered bodily receptors (“broken sensor”) [e.g., visceral hypersensitivity (106)].
3. Inference—illusionary dyshomeostasis, owing to impairments of the afferent branch of the inference-control loop, for example, atrophic (107) or inflammatory (108) processes within the insula or functional pathologies of *N*-methyl-D-aspartate receptors and/or neuromodulators that alter the signaling of pwPEs [for reviews, see (62,109)]. For example, abnormally high precision of beliefs about bodily states could render unremarkable events, such as normal sensory noise, meaningful; this is an interoceptive analog to “aberrant salience” (110) in schizophrenia.
4. Control—inadequate deployment of autonomic, endocrine, and immunological actions; for example, owing to inflammatory changes in allostatic control regions [AI, ACC (108)], regions implementing homeostatic reflex arcs [e.g., hypothalamus (111)], or their projections (112) or owing to inadequately shifted set-points as a result of false beliefs/forecasts (see above).

Distinguishing these options is hard: the closed-loop nature of the inference-control cycle means that any primary

disturbance will cause compensatory changes downstream. Inflammation-sensitive imaging (108,113) could help but covers only a few possible causes. Instead, we propose that model-based inference [from behavior and functional magnetic resonance imaging data (96)] (Supplement) on pwPE signaling in brainstem-hypothalamic-insular-cingulate circuitry could help identify a primary dysfunction. For example, under experimentally controlled perturbations of a (yet undisturbed) bodily state, pwPE signals in posterior and/or mid insula to predictable and unpredictable interoceptions should differ, depending on whether the pathology is located at inference or control levels.

### Computational Psychiatry

Bayesian perspectives of inference-control impairments feature frequently in computational concepts of depression (19,87), autism (20,114–117), schizophrenia (12,21,60,61,118,119), and anxiety (58,120,121). We briefly discuss one application of this framework to distinguish disease mechanisms in autism spectrum disorder (ASD).

#### Example: ASD

Hierarchical Bayesian theories of ASD revisit long-standing observations of perceptual anomalies in patients, including the excessive processing of irrelevant details and concomitant difficulties of abstraction. They suggest two competing explanations (114–117): sensory inputs of overwhelming precision or higher-order beliefs that are too imprecise for providing generalizable predictions. In either case, a child with ASD would incessantly experience large PEs during perception (see equation in Figure 3A). Typical symptoms, such as repetitive behaviors and avoidance of complex and volatile situations (e.g., social interactions), can then be interpreted as coping mechanisms to reduce PEs [see (20) for discussion].

Additionally, individuals with ASD show various interoceptive disturbances (122,123) that may equally result from an increase in sensory precision from the viscera or a failure to attenuate it (7). Viscerosensory precision weighting has been linked to oxytocin; associated disturbances during development might compromise the construction of generative models that attribute self versus other agency to interoceptive experiences (7,124).

The competing explanations of high sensory versus low belief precision (116) could be disambiguated by psychophysical experiments in combination with Bayesian models of perception. These have previously been used to assess individual sensory processing (50,125,126) in healthy volunteers, as have been electroencephalography-based circuit models of precision weighting in auditory cortex (127). These models could be used in ASD to detect (sub)groups with exaggerated precision estimates of sensory inputs and insufficiently precise predictions, respectively (20).

### CHALLENGES AND OPPORTUNITIES

Assessing the computational anatomy of circuit dysfunctions follows principles of homeostatic thinking, as is commonplace in medicine, and holds great diagnostic potential. However, clinical translation faces nontrivial challenges, particularly in application to psychosomatics.

### Chicken and Egg Problems

The inference-control loop represents the conceptual heart of theories of homeostasis, allostasis, and cybernetics (Figure 1). Its closed-loop nature means that a dysfunction in one domain typically invokes a cascade of changes throughout the circuit, making it difficult to differentiate cause from consequence. However, different primary disturbances induce distinct patterns of change that might be discriminable statistically—as commonly done in fields familiar with compensatory changes throughout dyshomeostatic systems, such as internal medicine (compare differential diagnosis of hypothalamic, pituitary, and glandular disturbances in endocrinology). Computational psychiatry and psychosomatics could finesse this by statistical comparison of models embodying alternative disease processes (95). Additionally, in medicine, challenge (perturbation) approaches are often crucial for diagnosis. Combining designed perturbations with model selection and prospective assessments of disease trajectories (10,96) represents a promising approach to resolve ambiguity created by circular causality.

One central challenge for computational psychosomatics concerns availability of somatic perturbation techniques. In contrast to computational psychiatry, where we can adopt methods for manipulating beliefs and precisions from psychology and psychophysics, manipulating the somatocerebral branch of psychosomatics has access to only a few techniques. These include cardiac challenges with short-acting sympathomimetics (128), manipulating inspiratory breathing load or air composition (129,130), transcutaneous vagus nerve stimulation (131), baroreceptor stimulation (132), acute induction of inflammation by vaccination (133), or C-fiber stimulation under capsaicin (134). Developing further challenges that are noninvasive and provide temporal control should become a priority topic for computational psychosomatics.

### Universality versus Specificity

The HBM framework suggests pwPEs as a central computational quantity for inference, forecasting, action, and meta-cognition. This generalizing view has pros and cons. On one hand, it suggests a conceptual grid for differential diagnosis and implies that computational differentiation of pwPE abnormalities could find broad diagnostic application. On the other hand, one may be concerned that we portray cortex as a “nonspecific hierarchical Bayesian machine” (as put by one of our reviewers) without neuroanatomical specificity. We do not wish to convey this impression. The inference problems the brain faces vary, for example, depending on the sensory channels involved and the depth of hierarchical coupling among environmental states. Different tasks require different types of (cortically represented) generative models and thus distinct circuits; compare proposed circuits for interoception/allostasis (19,87) and vision/oculomotor control (42,72). Empirically, in tasks using the same sensory modality but requiring inference on concrete versus abstract social quantities, pwPEs were reflected by activity in partially overlapping and partially distinct circuits (75,135).

Furthermore, we do not claim that the framework presented covers all existing psychiatric and psychosomatic phenomena. Not all symptoms relate to perception, forecasting, action, or

metacognition as the core components of our framework. However, where this relation exists, our framework may provide useful guidance in establishing analogous schemes for differential diagnostics in computational psychiatry and computational psychosomatics. Combined with models that can infer pwPE signaling in cortical hierarchies from neuroimaging or electrophysiological data (Supplement), this could allow for noninvasive readouts of circuit function that may support differentiation of potential failure loci. The promise and limitations of this approach require prospective patient studies that evaluate its predictive validity.

### ACKNOWLEDGMENTS AND DISCLOSURES

This work was supported by the René and Susanne Braginsky Foundation (KES), University of Zurich (FHP, KES), Deutsche Forschungsgemeinschaft Grant No. TR-SFB 134 (KES), and University of Zurich Clinical Research Priority Programs Multiple Sclerosis and Molecular Imaging Network Zurich (KES).

The authors report no biomedical financial interests or potential conflicts of interest.

### ARTICLE INFORMATION

From the Translational Neuromodeling Unit (FHP, LAEW, TG, KES), Institute for Biomedical Engineering, University of Zurich and Swiss Federal Institute of Technology Zurich; Center for Complementary and Integrative Medicine (TG), University Hospital Zurich, Zurich, Switzerland; Max Planck Institute for Metabolism Research (KES), Cologne, Germany; and Wellcome Trust Centre for Neuroimaging (KES), University College London, London, United Kingdom.

Address correspondence to Klaas E. Stephan, M.D., Dr.med., Ph.D., Translational Neuromodeling Unit, Institute for Biomedical Engineering, University of Zurich and Swiss Federal Institute of Technology Zurich, Wilfriedstrasse 6, CH-8032 Zurich, Switzerland; E-mail: [stephan@biomed.ethz.ch](mailto:stephan@biomed.ethz.ch).

Received Jan 1, 2017; revised Apr 14, 2017; accepted May 15, 2017.

Supplementary material cited in this article is available online at <http://dx.doi.org/10.1016/j.biopsych.2017.05.012>.

### REFERENCES

- Casey BJ, Craddock N, Cuthbert BN, Hyman SE, Lee FS, Ressler KJ (2013): DSM-5 and RDoC: Progress in psychiatry research? *Nat Rev Neurosci* 14:810–814.
- Stephan KE, Bach DR, Fletcher PC, Flint J, Frank MJ, Friston KJ, *et al.* (2016): Charting the landscape of priority problems in psychiatry, part 1: Classification and diagnosis. *Lancet Psychiatry* 3:77–83.
- Kapur S, Phillips AG, Insel TR (2012): Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it? *Mol Psychiatry* 17:1174–1179.
- Maia TV, Frank MJ (2011): From reinforcement learning models to psychiatric and neurological disorders. *Nat Neurosci* 14:154–162.
- Montague PR, Dolan RJ, Friston KJ, Dayan P (2012): Computational psychiatry. *Trends Cogn Sci* 16:72–80.
- Stephan KE, Mathys C (2014): Computational approaches to psychiatry. *Curr Opin Neurobiol* 25:85–92.
- Friston KJ, Stephan KE, Montague R, Dolan RJ (2014): Computational psychiatry: The brain as a phantastic organ. *Lancet Psychiatry* 1:148–158.
- Wang XJ, Krystal JH (2014): Computational psychiatry. *Neuron* 84:638–654.
- Huys QJ, Maia TV, Frank MJ (2016): Computational psychiatry as a bridge from neuroscience to clinical applications. *Nat Neurosci* 19:404–413.
- Paulus MP, Huys QJ, Maia TV (2016): A roadmap for the development of applied computational psychiatry. *Biol Psychiatry Cogn Neurosci Neuroimaging* 1:386–392.
- Corlett PR, Fletcher PC (2014): Computational psychiatry: A Rosetta Stone linking the brain to mental illness. *Lancet Psychiatry* 1:399–402.
- Stephan KE, Friston KJ, Frith CD (2009): Dysconnection in schizophrenia: From abnormal synaptic plasticity to failures of self-monitoring. *Schizophr Bull* 35:509–527.
- Huys QJ, Guitart-Masip M, Dolan RJ, Dayan P (2015): Decision-theoretic psychiatry. *Clin Psychol Sci* 3:400–421.
- Mathys C (2016): How could we get nosology from computation? In: Redish AD, Gordon JA, editors. *Computational Psychiatry: What Can Theoretical Neuroscience and Psychiatry Teach Each Other?* Cambridge, MA: MIT Press.
- Wiener N (1948): *Cybernetics*. New York: Wiley.
- Ashby WR (1956): *An Introduction to Cybernetics*. London: Chapman & Hall.
- Powers WT (1973): Feedback: Beyond behaviorism. *Science* 179:351–356.
- 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.
- Stephan KE, Manjaly ZM, Mathys C, Weber LAE, Paliwal S, Gard T, *et al.* (2016): Allostatic self-efficacy: A metacognitive theory of dyshomeostasis-induced fatigue and depression. *Front Hum Neurosci* 10:550.
- Haker H, Schneebeli M, Stephan KE (2016): Can Bayesian theories of autism spectrum disorder help improve clinical practice? *Front Psychiatry* 7:107.
- Stephan KE, Baldeweg T, Friston KJ (2006): Synaptic plasticity and dysconnection in schizophrenia. *Biol Psychiatry* 59:929–939.
- Cannon WB (1929): Organization for physiological homeostasis. *Physiol Rev* 9:399–431.
- Modell H, Cliff W, Michael J, McFarland J, Wenderoth MP, Wright A (2015): A physiologist's view of homeostasis. *Adv Physiol Educ* 39:259–266.
- Kersten D, Mamassian P, Yuille A (2004): Object perception as Bayesian inference. *Annu Rev Psychol* 55:271–304.
- Friston K (2003): Learning and inference in the brain. *Neural Netw* 16:1325–1352.
- Helmholtz H (1860/1962): *Handbuch der Physiologischen Optik* (English trans., Southall JPC, Ed.). New York: Dover.
- Knill D, Richards W (1996): *Perception as Bayesian Inference*. Cambridge: Cambridge University Press.
- Friston K (2005): A theory of cortical responses. *Philos Trans R Soc Lond B Biol Sci* 360:815–836.
- Dayan P, Hinton GE, Neal RM, Zemel RS (1995): The Helmholtz machine. *Neural Comput* 7:889–904.
- Adams WJ, Graf EW, Ernst MO (2004): Experience can change the "light-from-above" prior. *Nat Neurosci* 7:1057–1058.
- Geisler WS, Kersten D (2002): Illusions, perception and Bayes. *Nat Neurosci* 5:508–510.
- Diaconescu AO, Mathys C, Weber LA, Daunizeau J, Kasper L, Lomakina EI, *et al.* (2014): Inferring on the intentions of others by hierarchical Bayesian learning. *PLoS Comput Biol* 10:e1003810.
- Devaine M, Hollard G, Daunizeau J (2014): The social Bayesian brain: Does mentalizing make a difference when we learn? *PLoS Comput Biol* 10:e1003992.
- Daw ND, Dayan P (2014): The algorithmic anatomy of model-based evaluation. *Philos Trans R Soc Lond B Biol Sci* 369(1655).
- Friston K, Schwartenbeck P, Fitzgerald T, Moutoussis M, Behrens T, Dolan RJ (2013): The anatomy of choice: Active inference and agency. *Front Hum Neurosci* 7:598.
- Penny W, Stephan KE (2014): A dynamic Bayesian model of homeostatic control. *Lecture Notes in Computer Science* 8779:60–69.
- Chambon V, Sidarus N, Haggard P (2014): From action intentions to action effects: How does the sense of agency come about? *Front Hum Neurosci* 8:320.
- Fleming SM, Daw ND (2017): Self-evaluation of decision-making: A general Bayesian framework for metacognitive computation. *Psychol Rev* 124:91–114.
- Ortega PA, Braun DA (2013): Thermodynamics as a theory of decision-making with information-processing costs. *Proc Roy Soc A* 469:20120683.



40. Todorov E (2009): Efficient computation of optimal actions. *Proc Natl Acad Sci U S A* 106:11478–11483.
41. Kording KP, Wolpert DM (2004): The loss function of sensorimotor learning. *Proc Natl Acad Sci U S A* 101:9839–9842.
42. Friston KJ, Daunizeau J, Kilner J, Kiebel SJ (2010): Action and behavior: A free-energy formulation. *Biol Cybern* 102:227–260.
43. Huys QJ, Lally N, Faulkner P, Eshel N, Seifritz E, Gershman SJ, *et al.* (2015): Interplay of approximate planning strategies. *Proc Natl Acad Sci U S A* 112:3098–3103.
44. Kaelbling LP, Littman ML, Cassandra AR (1998): Planning and acting in partially observable stochastic domains. *Artif Intell* 101: 99–134.
45. Keramati M, Gutkin B (2014): Homeostatic reinforcement learning for integrating reward collection and physiological stability. *Elife* 3.
46. Doya K, Ishii S, Pouget A, Rao RP (2011): *Bayesian Brain: Probabilistic Approaches to Neural Coding*. Cambridge, MA: MIT Press.
47. Friston K (2010): The free-energy principle: A unified brain theory? *Nat Rev Neurosci* 11:127–138.
48. Lee TS, Mumford D (2003): Hierarchical Bayesian inference in the visual cortex. *J Opt Soc Am A Opt Image Sci Vis* 20:1434–1448.
49. Rao RP, Ballard DH (1999): Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nat Neurosci* 2:79–87.
50. Ernst MO, Banks MS (2002): Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415:429–433.
51. Angelaki DE, Gu Y, DeAngelis GC (2009): Multisensory integration: Psychophysics, neurophysiology, and computation. *Curr Opin Neurobiol* 19:452–458.
52. Petzschner FH, Glasauer S, Stephan KE (2015): A Bayesian perspective on magnitude estimation. *Trends Cogn Sci* 19:285–293.
53. Jazayeri M, Shadlen MN (2010): Temporal context calibrates interval timing. *Nat Neurosci* 13:1020–1026.
54. Mathys CD, Lomakina EI, Daunizeau J, Iglesias S, Brodersen KH, Friston KJ, *et al.* (2014): Uncertainty in perception and the Hierarchical Gaussian Filter. *Front Hum Neurosci* 8:825.
55. Felleman DJ, Van Essen DC (1991): Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex* 1:1–47.
56. Mumford D (1992): On the computational architecture of the neocortex. II. The role of cortico-cortical loops. *Biol Cybern* 66: 241–251.
57. Hilgetag CC, Medalla M, Beul SF, Barbas H (2016): The primate connectome in context: Principles of connections of the cortical visual system. *Neuroimage* 134:685–702.
58. Browning M, Behrens TE, Jocham G, O'Reilly JX, Bishop SJ (2015): Anxious individuals have difficulty learning the causal statistics of aversive environments. *Nat Neurosci* 18:590–596.
59. McCormick DA, Wang Z, Huguenard J (1993): Neurotransmitter control of neocortical neuronal activity and excitability. *Cereb Cortex* 3:387–398.
60. Corlett PR, Taylor JR, Wang XJ, Fletcher PC, Krystal JH (2010): Toward a neurobiology of delusions. *Prog Neurobiol* 92:345–369.
61. Adams RA, Stephan KE, Brown HR, Frith CD, Friston KJ (2013): The computational anatomy of psychosis. *Front Psychiatry* 4:47.
62. Corlett PR, Honey GD, Krystal JH, Fletcher PC (2011): Glutamatergic model psychoses: Prediction error, learning, and inference. *Neuropsychopharmacology* 36:294–315.
63. Bastos AM, Usrey WM, Adams RA, Mangun GR, Fries P, Friston KJ (2012): Canonical microcircuits for predictive coding. *Neuron* 76: 695–711.
64. Friston K, Kiebel S (2009): Predictive coding under the free-energy principle. *Philos Trans R Soc Lond B Biol Sci* 364:1211–1221.
65. Friston K (2009): The free-energy principle: A rough guide to the brain? *Trends Cogn Sci* 13:293–301.
66. Adams RA, Shipp S, Friston KJ (2013): Predictions not commands: Active inference in the motor system. *Brain Struct Funct* 218: 611–643.
67. Brown H, Adams RA, Parees I, Edwards M, Friston K (2013): Active inference, sensory attenuation and illusions. *Cogn Process* 14: 411–427.
68. Voss M, Ingram JN, Wolpert DM, Haggard P (2008): Mere expectation to move causes attenuation of sensory signals. *PLoS One* 3:e2866.
69. Friston K, Samothrakis S, Montague R (2012): Active inference and agency: Optimal control without cost functions. *Biol Cybern* 106:523–541.
70. Donnarumma F, Maisto D, Pezzulo G (2016): Problem solving as probabilistic inference with subgoal: Explaining human successes and pitfalls in the Tower of Hanoi. *PLoS Comput Biol* 12: e1004864.
71. Kiebel SJ, Daunizeau J, Friston KJ (2008): A hierarchy of time-scales and the brain. *PLoS Comput Biol* 4:e1000209.
72. Perrinet LU, Adams RA, Friston KJ (2014): Active inference, eye movements and oculomotor delays. *Biol Cybern* 108:777–801.
73. Daunizeau J, den Ouden HE, Pessiglione M, Kiebel SJ, Friston KJ, Stephan KE (2010): Observing the observer (II): Deciding when to decide. *PLoS One* 5:e15555.
74. Behrens TE, Woolrich MW, Walton ME, Rushworth MF (2007): Learning the value of information in an uncertain world. *Nat Neurosci* 10:1214–1221.
75. Iglesias S, Mathys C, Brodersen KH, Kasper L, Piccirelli M, den Ouden HE, *et al.* (2013): Hierarchical prediction errors in midbrain and basal forebrain during sensory learning. *Neuron* 80:519–530.
76. Seth AK, Suzuki K, Critchley HD (2011): An interoceptive predictive coding model of conscious presence. *Front Psychol* 2:395.
77. Seth AK (2013): Interoceptive inference, emotion, and the embodied self. *Trends Cogn Sci* 17:565–573.
78. Gu X, Hof PR, Friston KJ, Fan J (2013): Anterior insular cortex and emotional awareness. *J Comp Neurol* 521:3371–3388.
79. Craig AD (2002): How do you feel? Interoception: The sense of the physiological condition of the body. *Nat Rev Neurosci* 3:655–666.
80. Craig AD (2003): Interoception: The sense of the physiological condition of the body. *Curr Opin Neurobiol* 13:500–505.
81. Critchley HD, Harrison NA (2013): Visceral influences on brain and behavior. *Neuron* 77:624–638.
82. Cechetti DF, Saper CB (1987): Evidence for a viscerotopic sensory representation in the cortex and thalamus in the rat. *J Comp Neurol* 262:27–45.
83. Evrard HC, Logothetis NK, Craig AD (2014): Modular architectonic organization of the insula in the macaque monkey. *J Comp Neurol* 522:64–97.
84. Wiech K (2016): Deconstructing the sensation of pain: The influence of cognitive processes on pain perception. *Science* 354:584–587.
85. Buchel C, Geuter S, Sprenger C, Eippert F (2014): Placebo analgesia: A predictive coding perspective. *Neuron* 81:1223–1239.
86. Wager TD, Atlas LY (2015): The neuroscience of placebo effects: Connecting context, learning and health. *Nat Rev Neurosci* 16: 403–418.
87. Feldman-Barrett LF, Simmons WK (2015): Interoceptive predictions in the brain. *Nat Rev Neurosci* 16:419–429.
88. Sterling P (2014): Homeostasis vs allostasis: Implications for brain function and mental disorders. *JAMA Psychiatry* 71:1192–1193.
89. Sterling P (2012): Allostasis: A model of predictive regulation. *Physiol Behav* 106:5–15.
90. Simmons WK, Avery JA, Barcalow JC, Bodurka J, Drevets WC, Bellgowan P (2013): Keeping the body in mind: Insula functional organization and functional connectivity integrate interoceptive, exteroceptive, and emotional awareness. *Hum Brain Mapp* 34: 2944–2958.
91. Nguyen VT, Breakspear M, Hu X, Guo CC (2016): The integration of the internal and external milieu in the insula during dynamic emotional experiences. *Neuroimage* 124:455–463.
92. Allen M, Frank D, Schwarzkopf DS, Fardo F, Winston JS, Hauser TU, *et al.* (2016): Unexpected arousal modulates the influence of sensory noise on confidence. *Elife* 5.
93. Garfinkel SN, Minati L, Gray MA, Seth AK, Dolan RJ, Critchley HD (2014): Fear from the heart: Sensitivity to fear stimuli depends on individual heartbeats. *J Neurosci* 34:6573–6582.
94. Salomon R, Ronchi R, Donz J, Bello-Ruiz J, Herbelin B, Martet R, *et al.* (2016): The insula mediates access to awareness of visual

- stimuli presented synchronously to the heartbeat. *J Neurosci* 36:5115–5127.
95. Stephan KE, Schlagenhauf F, Huys QJ, Raman S, Aponte EA, Brodersen KH, *et al.* (2017): Computational neuroimaging strategies for single patient predictions. *Neuroimage* 145:180–199.
  96. Stephan KE, Iglesias S, Heinzle J, Diaconescu AO (2015): Translational perspectives for computational neuroimaging. *Neuron* 87:716–732.
  97. Fava GA, Sonino N (2010): Psychosomatic medicine. *Int J Clin Pract* 64:1155–1161.
  98. de Berker AO, Rutledge RB, Mathys C, Marshall L, Cross GF, Dolan RJ, *et al.* (2016): Computations of uncertainty mediate acute stress responses in humans. *Nat Commun* 7:10996.
  99. Fineberg SK, Steinfeld M, Brewer JA, Corlett PR (2014): A computational account of borderline personality disorder: Impaired predictive learning about self and others through bodily simulation. *Front Psychiatry* 5:111.
  100. Joynt KE, Whellan DJ, O'Connor CM (2003): Depression and cardiovascular disease: Mechanisms of interaction. *Biol Psychiatry* 54:248–261.
  101. Dowlati Y, Herrmann N, Swardfager W, Liu H, Sham L, Reim EK, *et al.* (2010): A meta-analysis of cytokines in major depression. *Biol Psychiatry* 67:446–457.
  102. Renn BN, Feliciano L, Segal DL (2011): The bidirectional relationship of depression and diabetes: A systematic review. *Clin Psychol Rev* 31:1239–1246.
  103. McEwen BS (1998): Stress, adaptation, and disease. Allostasis and allostatic load. *Ann N Y Acad Sci* 840:33–44.
  104. Friston KJ, Harrison L, Penny W (2003): Dynamic causal modelling. *Neuroimage* 19:1273–1302.
  105. Freeman R, Chapleau MW (2013): Testing the autonomic nervous system. *Handb Clin Neurol* 115:115–136.
  106. Holzer P (2015): Acid-sensing ion channels in gastrointestinal function. *Neuropharmacology* 94:72–79.
  107. Chatterjee SS, Mitra S (2015): “I do not exist” — Cotard syndrome in insular cortex atrophy. *Biol Psychiatry* 77:e52–e53.
  108. Setiawan E, Wilson AA, Mizrahi R, Rusjan PM, Miler L, Rajkowska G, *et al.* (2015): Role of translocator protein density, a marker of neuroinflammation, in the brain during major depressive episodes. *JAMA Psychiatry* 72:268–275.
  109. Stephan KE, Diaconescu AO, Iglesias S (2016): Bayesian inference, dysconnectivity and neuromodulation in schizophrenia. *Brain* 139:1874–1876.
  110. Kapur S (2003): Psychosis as a state of aberrant salience: A framework linking biology, phenomenology, and pharmacology in schizophrenia. *Am J Psychiatry* 160:13–23.
  111. Huitinga I, Erkut ZA, van Beurden D, Swaab DF (2004): Impaired hypothalamus-pituitary-adrenal axis activity and more severe multiple sclerosis with hypothalamic lesions. *Ann Neurol* 55:37–45.
  112. Hanken K, Eling P, Kastrup A, Klein J, Hildebrandt H (2015): Integrity of hypothalamic fibers and cognitive fatigue in multiple sclerosis. *Mult Scler Relat Disord* 4:39–46.
  113. Harrison NA, Cooper E, Dowell NG, Keramida G, Voon V, Critchley HD, *et al.* (2015): Quantitative magnetization transfer imaging as a biomarker for effects of systemic inflammation on the brain. *Biol Psychiatry* 78:49–57.
  114. Lawson RP, Rees G, Friston KJ (2014): An aberrant precision account of autism. *Front Hum Neurosci* 8:302.
  115. Pellicano E, Burr D (2012): When the world becomes “too real”: A Bayesian explanation of autistic perception. *Trends Cogn Sci* 16:504–510.
  116. Van de Cruys S, de-Wit L, Evers K, Boets B, Wagemans J (2013): Weak priors versus overfitting of predictions in autism: Reply to Pellicano and Burr (TICS, 2012). *Iperception* 4:95–97.
  117. Austerweil JL (2015): Contradictory “heuristic” theories of autism spectrum disorders: The case for theoretical precision using computational models. *Autism* 19:367–368.
  118. Fletcher PC, Frith CD (2009): Perceiving is believing: A Bayesian approach to explaining the positive symptoms of schizophrenia. *Nat Rev Neurosci* 10:48–58.
  119. Jardri R, Duvernois S, Litvinova AS, Deneve S (2017): Experimental evidence for circular inference in schizophrenia. *Nat Commun* 8:14218.
  120. Paulus MP, Stein MB (2006): An insular view of anxiety. *Biol Psychiatry* 60:383–387.
  121. Paulus MP, Stein MB (2010): Interoception in anxiety and depression. *Brain Struct Funct* 214:451–463.
  122. Garfinkel SN, Tiley C, O’Keefe S, Harrison NA, Seth AK, Critchley HD (2016): Discrepancies between dimensions of interoception in autism: Implications for emotion and anxiety. *Biol Psychol* 114:117–126.
  123. Gu X, Eilam-Stock T, Zhou T, Anagnostou E, Kolevzon A, Soorya L, *et al.* (2015): Autonomic and brain responses associated with empathy deficits in autism spectrum disorder. *Hum Brain Mapp* 36:3323–3338.
  124. Quattrocki E, Friston K (2014): Autism, oxytocin and interoception. *Neurosci Biobehav Rev* 47:410–430.
  125. Berniker M, Voss M, Kording K (2010): Learning priors for Bayesian computations in the nervous system. *PLoS One* 5.
  126. Vilares I, Howard JD, Fernandes HL, Gottfried JA, Kording KP (2012): Differential representations of prior and likelihood uncertainty in the human brain. *Curr Biol* 22:1641–1648.
  127. 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.
  128. Khalsa SS, Craske MG, Li W, Vangala S, Strober M, Feusner JD (2015): Altered interoceptive awareness in anorexia nervosa: Effects of meal anticipation, consumption and bodily arousal. *Int J Eat Disord* 48:889–897.
  129. Pappens M, Van den Bergh O, Vansteenwegen D, Ceunen E, De Peuter S, Van Diest I (2013): Learning to fear obstructed breathing: Comparing interoceptive and exteroceptive cues. *Biol Psychol* 92:36–42.
  130. Paulus MP, Flagan T, Simmons AN, Gillis K, Kotturi S, Thom N, *et al.* (2012): Subjecting elite athletes to inspiratory breathing load reveals behavioral and neural signatures of optimal performers in extreme environments. *PLoS One* 7:e29394.
  131. Brock C, Brock B, Aziz Q, Moller HJ, Pfeiffer Jensen M, Drewes AM, *et al.* (2017): Transcutaneous cervical vagal nerve stimulation modulates cardiac vagal tone and tumor necrosis factor-alpha. *Neurogastroenterol Motil* 29(5).
  132. Makovac E, Garfinkel SN, Bassi A, Basile B, Macaluso E, Cercignani M, *et al.* (2015): Effect of parasympathetic stimulation on brain activity during appraisal of fearful expressions. *Neuropsychopharmacology* 40:1649–1658.
  133. Harrison NA, Brydon L, Walker C, Gray MA, Steptoe A, Dolan RJ, *et al.* (2009): Neural origins of human sickness in interoceptive responses to inflammation. *Biol Psychiatry* 66:415–422.
  134. Baumann TK, Simone DA, Shain CN, LaMotte RH (1991): Neurogenic hyperalgesia: The search for the primary cutaneous afferent fibers that contribute to capsaicin-induced pain and hyperalgesia. *J Neurophysiol* 66:212–227.
  135. Diaconescu AO, Mathys C, Weber LA, Kasper L, Mauer J, Stephan KE (2017): Hierarchical prediction errors in midbrain and septum during social learning. *Soc Cogn Affect Neurosci* 12:618–634.