EXPERT REVIEW   OPEN

Check for updates

Computational psychiatry: from synapses to sentience

Karl Friston [ID][1][✉]

This review considers computational psychiatry from a particular viewpoint: namely, a commitment to explaining psychopathology in terms of pathophysiology. It rests on the notion of a *generative model* as underwriting (i) sentient processing in the brain, and (ii) the scientific process in psychiatry. The story starts with a view of the brain—from cognitive and computational neuroscience—as an organ of inference and prediction. This offers a formal description of neuronal message passing, distributed processing and belief propagation in neuronal networks; and how certain kinds of dysconnection lead to aberrant belief updating and false inference. The dysconnections in question can be read as a pernicious synaptopathy that fits comfortably with formal notions of how we—or our brains—encode uncertainty or its complement, *precision*. It then considers how the ensuing process theories are tested empirically, with an emphasis on the computational modelling of neuronal circuits and synaptic gain control that mediates attentional set, active inference, learning and planning. The opportunities afforded by this sort of modelling are considered in light of in silico experiments; namely, computational neuropsychology, computational phenotyping and the promises of a computational nosology for psychiatry. The resulting survey of computational approaches is not scholarly or exhaustive. Rather, its aim is to review a theoretical narrative that is emerging across subdisciplines within psychiatry and empirical scales of investigation. These range from epilepsy research to neurodegenerative disorders; from post-traumatic stress disorder to the management of chronic pain, from schizophrenia to functional medical symptoms.

INTRODUCTION

The inception of computational psychiatry—a decade ago— reflected a consensus that a formal, crisp and clear approach to psychiatry could be furnished by computational advances in data modelling, neuroscience and machine learning [1–5]. There are now annual workshops in computational psychiatry and an eponymous journal. However, introducing computational psychiatry was a risky move: most editors know that that pre-pending any journal title with "Computational" is likely to drop its impact factor by at least one point. So, has computational psychiatry delivered? And where is it going?

Thirty years ago, state-of-the-art computational techniques meant the application of factor analysis to discern patterns of symptoms and signs and schizophrenia—e.g. [6]—or, at least for me, the use of canonical correlation analysis to describe patterns of cerebral blood flow in brain imaging [7]. Nowadays, things are very different. I spend my time discussing the finer details of modelling synaptic efficacy with researchers, whose expertise spans magnetoencephalography, functional genomics, optogenetics and cell cultures. I sit there—pretending to know all the acronyms—and marvel at how a common, computational, narrative can bring molecular biologists and clinical neuroscientists to the same table. This review considers some of the foundational tenets of this narrative, illustrating the 'work in progress' with a few select examples.

Our focus will be on computational ideas and procedures that have endured, sometimes for centuries, and are currently [re] emerging across theoretical neurobiology. This limits us to a

theory or hypothesis-led approaches, where hypotheses are articulated as a *generative model*, for which empirical evidence can be sought. This demarcates computational approaches with 'strong' explanatory power from 'weak', descriptive or theory-free approaches. Many 'weak' approaches have come and gone, leaving an underwhelming legacy. Examples of weak approaches include behaviourism [8]; namely, a description of behaviour that is reinforced by an outcome that is defined (tautologically) by its ability to reinforce behavior [Clearly, behaviourism is not theory free. I am using 'behaviourism' here as a euphemism for simply collecting and summarising dependencies in observable behaviour (e.g., with an autoregression or Rescorla-Wagner model)]. Strong alternatives rest upon the explanatory power of active inference and learning [9], that emphasise beliefs, uncertainty, preferences and curiosity [10–14]. In imaging neuroscience a weak example would be correlation patterns in functional magnetic resonance imaging (fMRI) or electrophysiology, a.k.a., functional connectivity [15]; where its strong counterpart would be the directed connectivity in neuronal circuits that cause the correlations, a.k.a., effective connectivity [16–22]. In the data sciences, weak computational approaches rest on 'big data' and machine (e.g., deep) learning procedures that preclude generative or world models and, as a consequence, any mechanistic interpretability or explainability [20, 23, 24]. The strong complement, in machine learning, rests on generative modelling; e.g., generative adversarial networks and automatic variational inference [25–27]. [The bright line between strong (hypothesis-led) and weak (data-led) approaches is, of course, fragile. For example, weak approaches

are necessary to furnish evidence for strong approaches (e.g., genome-wide association studies GWAS that point to particular synaptopathies), while hypotheses are necessary to make the best sense of data (e.g., enrichment analysis in prioritizing variants from GWAS)].

The common theme that distinguishes weak from strong approaches is reference to a generative model: namely, a hypothesis about causal mechanisms, specified formally as a joint probability distribution over some (unobservable, latent or hidden) causes and their (observable, measurable) consequences. What follows is, at every level, undergirded by a generative model that places (strong) computational psychiatry—and perhaps our own sentience—firmly in the realm of mechanistic, evidenced-based science; in the sense, that one can compare the evidence for one generative model, relative to another.

OVERVIEW
The first section sets the scene for a mechanistic formulation of psychopathology under current ideas about the brain as an organ of inference, prediction and planning. These ideas can be articulated in terms of the computational processes that accompany belief updating in the Bayesian brain [28–32]. The second section turns to psychopathology—and the underlying pathophysiology—by asking how aberrant belief updating and predictive processing might manifest? In brief, the conclusion is that many psychiatric and neurological disorders can be framed in terms of a synaptopathy that confounds the encoding of uncertainty or precision in a (Bayesian) brain [33–35].

The third section considers the imperatives for computational modelling of empirical data that inherit from the previous sections. This leads to a discussion of how to assess the modulation of synaptic efficacy in terms of changes in effective connectivity, with an emphasis on dynamic causal modelling in psychiatry [20, 36, 37]. The final sections deal with *computational phenotyping*; namely, using computational or in silico models of belief updating—and psychophysical or behavioural concomitants—to explain the responses of a given subject. They briefly consider *computational nosology*, by asking if there is a generative model of pathophysiology and psychopathology that can be used to best explain the trajectories of individual patients using high-density longitudinal data. Finally, the opportunity afforded by a formal and mechanistic understanding of functional brain architectures is exemplified by *computational neuropsychology*; namely, performing in silico lesion or pharmacological experiments.

THE BAYESIAN BRAIN, PRECISION ENGINEERING AND ACTIVE INFERENCE
The narrative starts with the simple premise that if psychology—read as belief updating in the brain—can be cast as a computational process of inference, it follows that psychopathology just is false inference. False inference is meant in the usual sense of false positives (i.e., type I errors); namely, inferring something is there when it is not. Cardinal examples here include hallucinations, delusions and other features of reality distortion seen in psychosis. False negatives (i.e., type II errors) mean inferring something is not there when it is; for example, dissociative disorders, neglect syndromes, derealisation phenomena, *et cetera*. Indeed, when one thinks about psychiatric and neurological disorders, most can be framed as false inference: ranging from dysmorphophobia in eating disorders, through to delusional systems in paranoid schizophrenia; from phobias through to Parkinson's disease. It may seem odd to include Parkinson's disease; however, the pathognomonic bradykinesia—and failure to initiate movement—can, on one reading, be seen as a failure to realise motor planning as inference (see below). So, what licences the assumption that psychology is inference?

[Clearly, asserting that all of psychopathology is 'false inference' requires some qualification. The basic argument—that licences this assertion—is that any symptom (or sign) of psychopathology must, at some level, entail aberrant belief updating. The kind of beliefs here are Bayesian beliefs; namely, subpersonal or propositional representations that possess the attribute of uncertainty, and are conditional in a well-specified sense. Under this reading of beliefs, sentient behaviour—and beliefs about that behaviour—can then be cast as inference (e.g., perceptual inference, planning as inference, attribution of agency, emotional recognition, et cetera). Reducing everything to inference then provides a mechanics to understand neuronal dynamics and development, in terms of belief updating and the requisite message in the brain].

The foundations of this account of sentient behaviour can be traced back to the days of Plato and were most clearly articulated in the 19th-century by Helmholtz as unconscious inference [38, 39]—ideas that are reminiscent of Kantian philosophy. These ideas endured through the 20th-century in several flavours; for example, analysis by synthesis [40, 41], epistemological automata [42], perception as hypothesis testing [43, 44] and, in machine learning, the Helmholtz machine [45]. The inference narrative supervened at the turn of the century, with a resurgence of interest in enactivist approaches [46–51] that now predominate in the cognitive and systems neurosciences, in the form of things like predictive processing and active inference [52–60].

Active inference can be read as an enactive version of the Bayesian brain hypothesis [61, 62] that subsumes sentience (perceptual inference) and behaviour by treating control and planning as inference [28–31]. So, what is inference? In this setting, inference just refers to a process that maximises the evidence for some (generative) model or hypothesis about the causes of (sensory) data. Model evidence is also known as marginal likelihood; namely, the likelihood of some data, under a model of how those data were generated. Maximising the evidence for our own generative model is sometimes called self-evidencing [63]. In brief, active inference casts the brain as a fantastic organ: a generator of fantasies, hypotheses and predictions that are tested against sensory evidence. One might ask how this account of sentient behaviour speaks to neuroanatomy, and the functional architectures implied by neurophysiology.

The answer is relatively straightforward. Given a generative model, there are well described belief updating or propagation schemes that specify the requisite message passing that must, in some form, be implemented by neuronal networks. For generative models based upon continuous states of the world, these schemes are known as Bayesian filters or predictive coding [56, 64–67]. In generative models of discrete states (e.g., "I am in the kitchen", as opposed to "I am at these continuous GPS coordinates") the message passing schemes are variously known as belief propagation or variational message passing [68–71]. All of these schemes can be cast as a gradient ascent on model evidence or marginal likelihood [72]. In short, neuronal dynamics just are a process of inference. See Fig. 1 for a schematic description of predictive coding.

Crucially, the gradients that subtend neuronal dynamics—and consequent belief updating—can always be formulated as a prediction error [This may sound like a sweeping statement; however, any random dynamical system (including neuronal dynamics) can be cast as a gradient flow, where the gradients can always be expressed as a difference in log probabilities, which can be read as prediction errors of one kind or another] [70]. In other words, the divergence between predictions of sensory input and the observed sensations. In predictive coding schemes, it is thought that prediction errors are represented explicitly: e.g., by superficial pyramidal cells in the upper layers of the cortex [66, 73–76].

This leads to a picture of hierarchical inference in the brain as the reciprocal message passing between the levels of a cortical
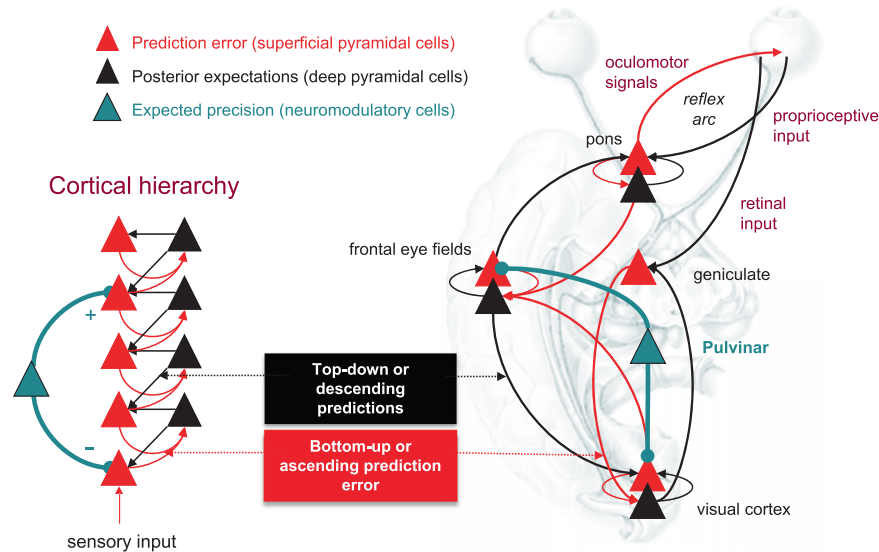
Fig. 1 This figure summarizes hierarchical message passing in predictive coding. Predictive coding is a scheme for generative models based upon continuous states. In these schemes, neuronal activity encodes expectations about the causes of sensory input that are continually updated to minimize prediction error. Prediction error is the difference between (ascending) sensory input and (descending) predictions of that input. This minimization rests upon recurrent neuronal interactions among different levels of a cortical hierarchy. It is thought that superficial pyramidal cells (red triangles) compare the expectations (at each hierarchical level) with top-down predictions from deep pyramidal cells (black triangles) of higher levels: see [66] for a review. Left panel: this schematic shows a simple cortical hierarchy with ascending prediction errors and descending predictions. It includes neuromodulatory gating or gain control (teal) of superficial pyramidal cells that determines their relative influence on deep pyramidal cells encoding expectations. Right panel: this provides a schematic example in the visual system: it shows the putative cells of origin of ascending or forward connections that convey prediction errors (red arrows) and descending or backward connections (black arrows) that furnish predictions. The prediction errors are weighted by their expected precision, which have been associated with projections from the pulvinar. In this example, the frontal eye fields send predictions to primary visual cortex, which projects to the lateral geniculate body. However, the frontal eye fields also send proprioceptive predictions to pontine nuclei, which are passed to the oculomotor system to induce movement, through classical oculomotor reflexes. This means that saccadic eye movements, for example, realise top-down proprioceptive predictions that are deeply informed by hierarchical processing of data from all sensory modalities. Similar schemes can be elaborated for discrete state space generative models, responsible for planning and motor control —or allostasis and autonomic function: see [54] for equivalent schematics and [70] for mixtures of continuous and discrete models.

hierarchy; in which prediction errors ascend from lower to higher levels to drive changes in neuronal populations encoding states of affairs in the world. These populations (e.g., deep pyramidal cells in lower layers of the cortex) then supply a counter stream of descending predictions that resolve or cancel prediction errors at lower levels [66, 75, 77]: e.g., by targeting inhibitory interneurons that are coupled to the superficial pyramidal cells broadcasting prediction errors [78, 79]. This architecture also plays out under discrete generative models and has become something of a workhorse for understanding recurrent message passing in cortical and subcortical hierarchies.

THE IMPORTANCE OF BEING PRECISE
Prediction errors (i.e., the gradients that drive belief updating) can be regarded as carrying the newsworthy information at any given hierarchical level to the level above. However, this is not the complete story. Higher levels have to select which prediction errors to listen to; much in the same way that we select our most trustworthy news channels or sources of information. This selection rests upon predictions of predictability or *precision*. Predicting or estimating precision is a universal requisite for making sense of data: from estimating the signal-to-noise ratio in some sensory apparatus, through to estimating standard error when making inferences via some Neyman-Pearson statistic. Affording certain prediction errors greater precision increases their influence on belief updating and has all the hallmarks of attentional selection [80–85]. Physiologically, this simply entails an increase in the excitability or postsynaptic gain of neuronal populations broadcasting prediction errors. On this view, there is

an intimate relationship between attention and the modulation of synaptic efficacy by classical neuromodulators and nonlinear postsynaptic responses responsible for mediating the exchange between fast-spiking inhibitory interneurons and (superficial) pyramidal cells [78, 79, 81, 86–91]. Please see Table 1 and [92] for a fuller discussion of neuromodulators and the representation of precision. Figures 2, 3 illustrate the neuronal circuitry implicated in precision or gain control, with a focus on canonical microcircuits and inhibitory interneurons, respectively.

One crucial aspect of this precision engineering is that it underwrites our ability to filter out—or ignore—certain prediction errors when they are deemed imprecise. A key example is the attenuation of sensory prediction errors that report the consequences of movement [93–98]. If we could not ignore the proprioceptive and somatosensory afferents—supplying evidence that we are not moving—then any beliefs about intended or predicted movement would be revised immediately, and we would not be able to initiate movement. This mandates a transient suspension of sensory precision during active sensing: c.f., saccadic suppression of optic flow signals during saccadic eye movements [99]. It also provides a glimpse of how one might explain Parkinson's disease in terms of dopaminergic failures of sensory attenuation [100]. More generally, it foregrounds the role of precision in orchestrating the perception action cycles [101] that underwrite active sensing [102]. For example, sensory attenuation suggests a particular scheduling of action and perception: although both work hand-in-hand to resolve prediction errors, this resolution may involve a fast (~4 Hz) alternation between sampling the world (e.g., visual palpation with saccades) and belief updating after each sample.

Table 1. The pharmacology of precision. Please see [92] for details.

| Neurotransmitter system	Functional role	Neuroanatomy
Cholinergic	Encoding the precision of outcomes given hidden states: (c.f., Attention and expected uncertainty [218, 219])	Nucleus basalis of Meynert
Noradrenergic	Encoding the precision of state transitions (c.f., Volatility and unexpected uncertainty [220])	Locus coeruleus
Dopaminergic	Encoding the precision of beliefs about policies (c.f. Action selection [221])	Substantia nigra pars compacta, Ventral tegmental area
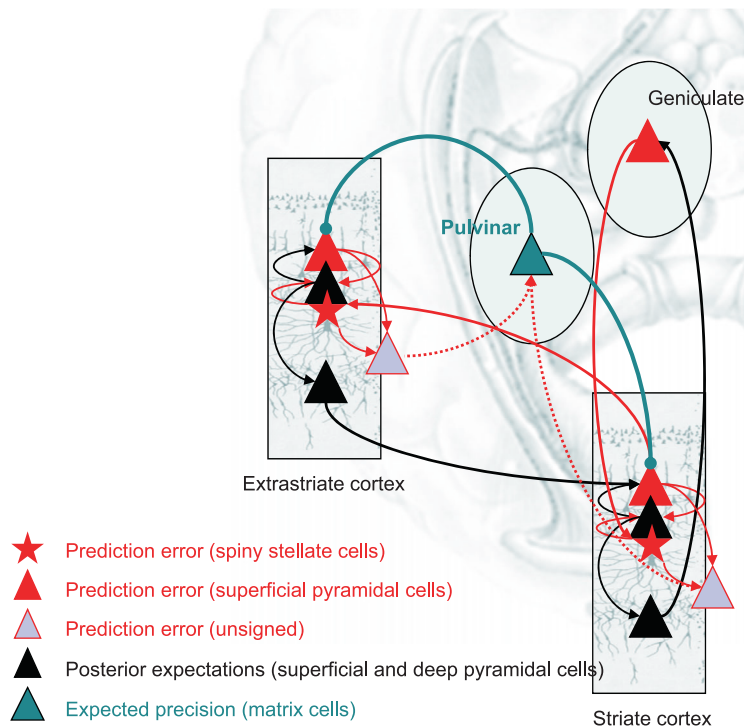
Fig. 2 Predicting precision. This schematic is a more detailed version of Fig. 1 that includes putative laminar-specific connections that are consistent with the precision-based predictive coding described in [176]. This architecture is based upon [66] and conforms roughly to the known neuroanatomy and physiology of canonical microcircuits in the visual system and the laminar specificity of extrinsic connections. The key aspect of this figure is the inclusion of deep pyramidal cells encoding the amplitude of (squared or unsigned) prediction error that inform posterior expectations about precision in the (matrix cells) of the pulvinar. These cells reciprocate descending projections to modulate the gain of superficial pyramidal cells in cortex. Forward connections are in red, and descending (backward) connections are in black. First-order prediction errors are shown as full red lines and second-order (unsigned) prediction errors are shown as broken red lines. This schematic ignores inhibitory interneurons that mediate some inferences vicariously. See subsequent figure.

In summary, the narrative so far is that psychopathology represents false inference or aberrant belief updating, under a view of the brain as a statistical organ, generating predictions and revising its (subpersonal Bayesian) beliefs on the basis of prediction errors. Crucially, these predictions are contextualised with *predictions of precision* or predictability that instantiate attentional or intentional set; allowing the selection of attenuation of prediction errors via a process of precision weighting. This precision weighting is nothing more than modulating the gain, postsynaptic sensitivity or excitability of appropriate neuronal populations. So, what does this computational architecture offer in terms of potential targets for pathology?

SYNAPTOPATHY, DYSCONNECTIONS AND FALSE INFERENCE
The narrative returns here to the 19th-century and early formulations of psychiatric disorders such as dementia praecox and schizophrenia. A common theme of these formulations is a disintegration or disruption of neuronal processing. From the perspective of the current narrative, these formulations came in two flavours. First, Wernicke's sejunction hypothesis [103] emphasised disruption to the organs of connection (i.e., white matter tracts in the brain), positing a form of *disconnection* syndrome [104]. A complementary perspective was provided by Bleuler's notion of disintegration of the psyche [105], cast in more functional terms that we might now read as a failure of synaptic integration. These ideas re-emerged with the advent of functional brain imaging in the 1990s, in the form of *disconnection* hypotheses, drawing analogies with things like metachromatic leukodystrophy [106] and the *dysconnection* hypothesis [107] that emphasised dysfunctional synaptic integration; specifically, neuromodulatory failures of synaptic function [34]. The latter formulation of dysfunctional connectivity is now arguably the prevalent view for most psychiatric and neurodegenerative conditions, while variants of the sejunction hypothesis would be apt for cerebrovascular accidents, space occupying lesions and other interruptions to white-matter fasciculi: for example, increases in conduction delays due to cerebral small vessel disease in white matter fasciculi.

In short, one could neatly summarise the pathophysiology of many psychiatric and neurological conditions in terms of one or more forms of synaptopathy [108]. Synaptopathy is taken to mean any failure of synaptic function due to a variety pathological

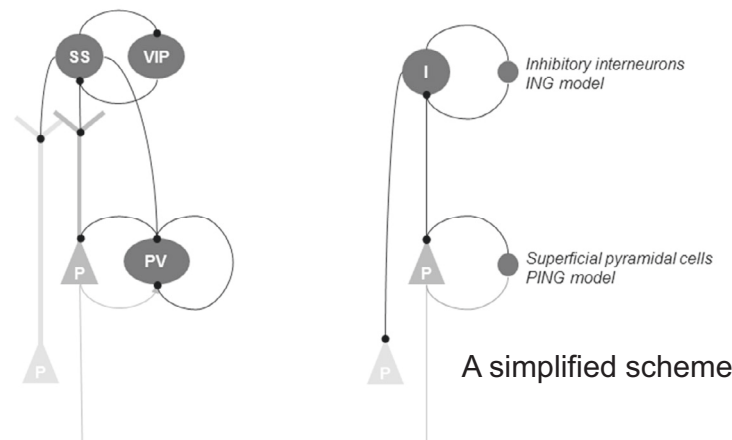## Interactions between pyramidal cells and inhibitory interneurons

Fig. 3 The left panel depicts interactions between (superficial and deep) pyramidal cells with inhibitory interneurons. Inhibitory interneurons are partitioned into three subtypes (Parvalbumin positive PV, somatostatin SST, and vasoactive intestinal peptide expressing interneurons, VIP), based upon optogenetic studies [217]. This schematic illustrates the intimate relationship between pyramidal cells and inhibitory interneurons responsible for excitation-inhibition balance and ensuing excitability in canonical microcircuits. Here, it is assumed that PV interneurons are reciprocally connected to the pyramidal cells through perisomatic compartments, while SST cells form synapses on their dendrites. The right panel shows a simplified architecture used in many modelling studies. Here, the recurrent inhibitory (PV/pyramidal cell) dynamics have been absorbed into an inhibitory recurrent connection, while the SST/VIP interneurons provide (dendritic) inhibitory drive. This allows one to map models of Interneuron Network Gamma (ING) oscillations onto canonical microcircuits used in dynamic causal modeling (see next figure). In this setting, the Pyramidal ING (PING) model emphasizes recurrent interactions among PV cells, as modeled by the inhibitory recurrent connections on superficial pyramidal cells. In contrast, the ING model corresponds to the influence of (SST/VIP) inhibitory interneurons on pyramidal cells. Please see [81] for further details of how these simplified architectures are used to model attentional effects on evoked responses.

mechanisms (i.e., formation, structure, metabolism, etc.). For example, the severe loss of synapse density in human Alzheimer's disease, frontotemporal dementia, Parkinson's disease, Progressive Supranuclear Palsy, etc. [109].

But what kind of synaptopathy? A coarse-grained overview of the synaptic theories of schizophrenia suggests that the synaptopathy in question is of a neuromodulatory sort; implicating classical neuromodulatory (ascending) neurotransmitter systems [110–113], GABAergic neurotransmission and NMDA receptors [113–116], where synchronous interactions between fast spiking inhibitory interneurons and pyramidal cells may set the excitability of canonical microcircuits [88, 91, 116–119]. The same story emerges from the functional genomics of schizophrenia, which point to the mechanisms that underwrite neuromodulation, synaptic gain control and ensuing excitation-inhibition balance in neuronal circuits [120, 121]. For example, a recent genome-wide association study [122] of 76,755 individuals with schizophrenia highlights the importance of GRIN2A (which encodes a subunit of the NMDA receptor), while a recent meta-analysis of whole exomes [123] emphasises the role of GRIN2A, as well as GRIA3 (which encodes a subunit of the AMPA receptor). Finally, it is worth noting that nearly all psychoactive drugs, including psychedelics, target classical neuromodulatory or NMDA receptors: e.g., [124, 125].

In short, many psychiatric, neurological and neurodegenerative disorders could be described as arising from pernicious synaptopathies that lead to a functional disintegration of neuronal message passing in cortical and subcortical hierarchies. The particular synaptopathies in question implicate neuromodulatory gain control, of the sort required to deploy precision (i.e., selective attention and sensory attenuation) during inference and planning in the Bayesian brain.

TOO MUCH OR TOO LITTLE?
This completes the next step in the narrative; namely, the psychopathology (i.e., false inference) characteristic of psychiatric

disorders may be attributable to aberrant precision control, which inherits from synaptopathies that confound neuromodulation. This narrative first emerged in theoretical treatments of hallucinosis in synucleinopathies [126–128] and, from a computational perspective, was developed in schizophrenia [129–133] and autism research [134–136]. Since that time, it has become difficult to find a psychiatric condition that has not been considered through the lens of aberrant precision, in one form or another [21, 135, 137–161].

A crosscutting theme in many of these accounts is a putative failure of sensory attenuation [53, 93–95, 98, 100, 159, 162–165]. In other words, a neuromodulatory failure to attenuate the postsynaptic gain of neuronal populations reporting the consequence of action (in the motor domain or mediated by autonomic reflexes). A failure of sensory attenuation implies an imbalance between the precision afforded sensory prediction errors and the precision of prediction errors deeper in neuronal hierarchies that mediate or maintain prior beliefs of a subpersonal or propositional nature. The ensuing imbalance is often read as a loss of precision or confidence in prior beliefs, relative to the sensory evidence at hand. In autism, this imbalance is consistent with a failure of sensory attenuation and an inability to ignore the sensorium, which may or may not be associated with the neuromodulatory role of oxytocin [135, 166–170].

In schizophrenia, the story is a little more involved; in the sense that a failure of sensory attenuation provides an apt explanation for resistance to illusions (that normally depend upon precise prior beliefs) and failures to elicit mismatch oddball responses (because everything is surprising). Some people then interpret delusional ideation as the brain's attempt to make sense of unattenuated prediction errors: see also [152, 155, 171, 172]. However, hallucinatory phenomena require a slightly more delicate argument; usually along the lines of a compensatory increase in prior precision that could manifest as a form of paradoxical lesion. In other words, in attempt to override precise sensory prediction errors, higher levels learn to ignore sensory evidence and—

sequestered from the sensorium—elaborate false percepts. See [22, 133, 141, 158, 173] for related discussion.

There are many interesting instances of this formulation; ranging from failures of sensory attenuation in autonomic and interoceptive inference, through depression and emotional processing [82, 140, 149, 174], to functional medical symptoms and dissociative phenomena [137, 164]. These computational formulations are nice because they lend themselves to various lines of enquiry; ranging from simulating false inference; e.g., [149], through the careful analysis of psychophysics and choice behaviour e.g., [154, 155, 175]; and, crucially, neurophysiological measurements of synaptic function—to which we now turn.

SYNAPTOPATHY, DYSCONNECTIONS AND DYNAMIC CAUSAL MODELLING

The narrative so far is that psychopathology is a pernicious form of false inference that can be attributed to a synaptopathy that confounds the deployment of precision during perceptual inference and ensuing action. This narrative is based upon first principle (i.e., self evidencing and Bayesian) accounts of brain function. Crucially, these accounts have attendant process theories that specify various forms of neuronal dynamics and message passing that can be evaluated empirically [54, 66, 70, 71, 73, 74, 176, 177]. Note that the strong kind of computational psychiatry advocated here demands a process theory, which can be tested empirically, in terms of the evidence for one process theory or model, relative to another.

Although synaptopathy of a certain kind may be a common theme—in terms of computational and pathophysiological mechanisms—where and how these pathologies manifest is an open question: a question that may require a different answer for every condition (and possibly every patient). For example, at the level of (canonical) microcircuits, neuromodulatory abnormalities may manifest in different ways at different levels of cortical hierarchies and, indeed, in the (inter and intralaminar) message passing within microcircuits. This means that one needs a way of identifying functional brain architectures—and, particularly, context sensitive changes in synaptic efficacy—that report the presence or absence of synaptopathy. In short, this calls for in vivo and ex vivo assays of synaptic efficacy [178, 179] that can be deployed across scale; i.e., from molecular biology through to patients in the clinic (or bedside).

There have been amazing advances in the measurement of neuronal processes; ranging from laminar fMRI [180, 181] through to optogenetics [33, 91, 182–184]. From a computational perspective, there is one universal imperative for quantifying changes in synaptic efficacy: one has to be able to assess the evidence—in any empirical measurement—for changes in effective connectivity. In turn, one has to commit to generative models of how the data were generated, in order to assess the evidence for models with and without a change [185]. Again, we come back to the foundational role of generative models but now from the perspective of a scientist, as opposed to the brain (as a scientist). At its simplest, synaptic efficacy is modelled as a connection strength (or rate) with accompanying synaptic rate or time constants. These have to be estimated under some model of hidden or latent physiological fluctuations that generate observed data. This can be at a molecular scale or the scale of whole brain imaging.

In imaging neuroscience, there are two approaches to connectivity; namely the weak approach called *functional* connectivity and the strong approach called *effective* connectivity [186]. Functional connectivity is essentially looking for patterns or correlations between measured neuronal responses. In contrast, effective connectivity is the estimate of directed connection strengths, under some forward or generative model. The power of effective connectivity analyses rests on being able to compare the evidence for one model (sometimes, of functional connectivity as a data feature) relative to another. A prescient example of this is the estimation of intrinsic excitability or gain, usually associated with the neuromodulatory precision control above. To assess the intrinsic (within node or source) excitability of a node—in a distributed network of nodes coupled by extrinsic (between node or source) connections—one has to parameterise network models in terms of intrinsic or self-connections [187]. This foregrounds one of the weaknesses of descriptive or functional connectivity: in the sense that the correlation of a source with itself tells you nothing (it is always one).

In imaging neuroscience, the predominant analysis of effective connectivity is dynamic causal modelling, usually based upon neural mass models with a greater or lesser degree of biological plausibility [16, 17, 19, 20, 22, 36, 37, 188–190]. In brief, modelling neuronal dynamics in terms of neural mass models means that one can parameterise intrinsic and extrinsic connectivity—and context sensitive changes—in terms of model parameters and, crucially, assess the evidence for models with and without certain parameters, by fitting the models to empirical timeseries.

MEASURING SYNAPTIC EFFICACY IN VIVO

Pursuing the narrative of testing hypotheses about neuromodulatory synaptopathies, the three-way link between precision control, attention (and sensory attenuation) and neuromodulation means there has been considerable progress in characterising the synaptic basis of sentience using attentional and oddball paradigms. Indeed, in schizophrenia research, one of the best studied paradigms is the mismatch negativity that allows one to estimate excitability (i.e., intrinsic connectivity) induced by a changing context, signalled by an oddball stimulus; e.g., [22, 191]. Clearly, changes in synaptic efficacy over a hundred milliseconds to seconds requires electrophysiological measurements of the kind afforded by electrophysiology (e.g., EEG and MEG). This allows one to pinpoint the expression of aberrant gain or precision control to various neuronal populations in cortical hierarchies. See Fig. 4 for an early example. The dénouement of this kind of computational study of schizophrenia (at the time of writing) concludes: (i) EEG responses in three classic paradigms are all attributable to the same underlying synaptic change: greater self-inhibition in pyramidal cells; (ii) psychotic symptoms relate to disinhibition in neural circuits and "these findings are more commensurate with the hypothesis that a primary loss of synaptic gain on pyramidal cells is then compensated by interneuron downregulation (rather than the converse)" [192].

The construct validation of these synaptic assays speaks to integration across scales. Perhaps the most developed use cases here are in epilepsy research and the characterisation of neurodegenerative disorders [179, 193–196]. A compelling example of applying generative (i.e., dynamic causal) modelling across scales can be found when drilling down on the molecular basis of childhood epilepsy, using data from small animal models and patients. For example, local field potential recordings in a mouse model were used to validate a dynamic causal model of NMDAR-Ab effects on cortical microcircuitry. The ensuing DCM was then used to identify the synaptic parameters that best explain EEG paroxysms in paediatric patients with NMDAR-Ab encephalitis. The authors then returned to the mouse model to show that NMDAR-Ab-related changes render microcircuitry critically susceptible to overt EEG paroxysms [197]. See Fig. 5.

Another field that is pursuing a similar approach is the characterisation of synaptopathies in neurodegeneration; again, with convergence on the interactions between pyramidal cells and inhibitory interneurons. For example, [198] used the mismatch negativity paradigm to validate a dynamic causal model of laminar-specific microcircuitry and GABAergic neurotransmission. Bayesian model comparison identified an effect of tiagabine on
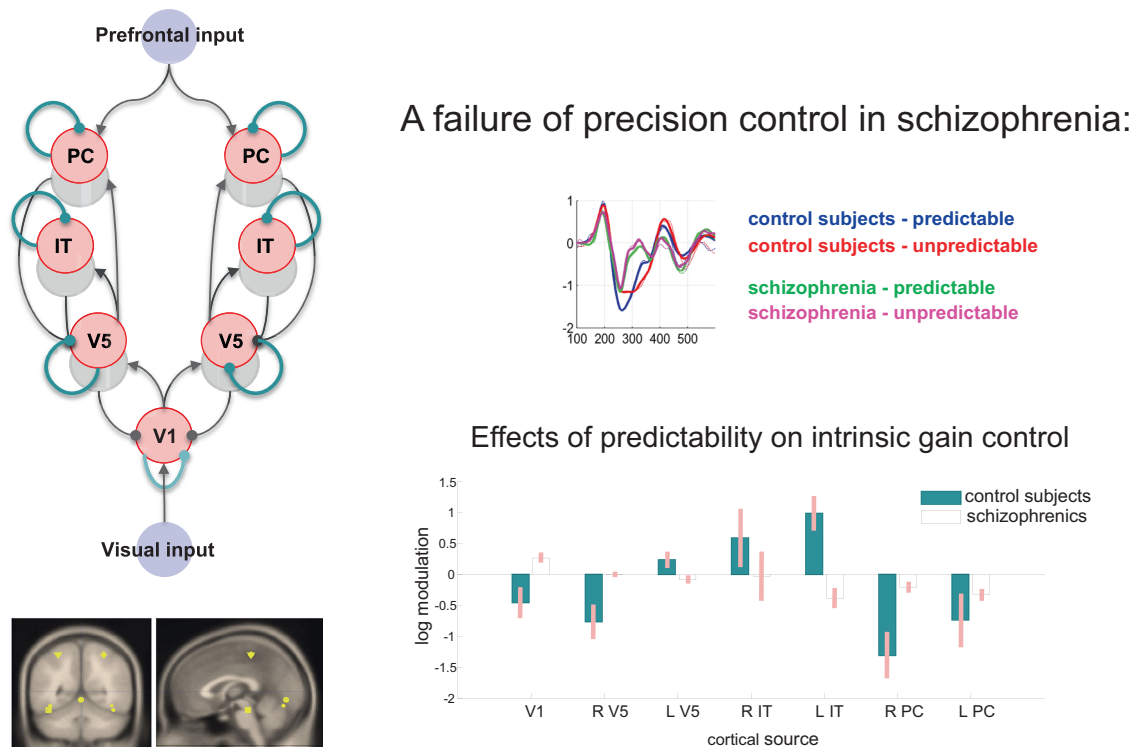
Fig. 4 Synaptic gain control in schizophrenia. This figure summarizes an early example of using DCM of event related responses (ERPs), to assess the modulation of synaptic efficacy in schizophrenia using EEG [191]. ERPs were elicited in a visual target detection paradigm, under predictable and unpredictable conditions, in neurotypical and schizophrenic subjects, respectively. Sources—modeled as small cortical patches (lower left panel)—include: a midline visual source (V1), right and left sources near the temporoparietal junction (V5), right and left inferotemporal sources (IT) and bilateral superior parietal sources (PC). The distributed network connecting these sources (upper right panel) entails top-down connections from PC and IT to V5 that send backward connections to V1 (black lines); reciprocal forward connections (red lines); and intrinsic connections for each source (black loops). The principal components of predicted and observed ERPs (upper right panel) show a pronounced difference in the evoked responses of normal subjects to predictable and unpredictable targets around 300 to 400 ms after stimulus onset (compare the red and blue traces). This difference is attenuated in the schizophrenia patients (green and magenta). The lower right panel reports (log scaling of) intrinsic connections and their 90% posterior confidence intervals, for predictable versus unpredictable targets, for patients (white) and controls (teal). The notable thing here is a failure of predictability to modulate the intrinsic gain (i.e., self-connectivity) throughout the hierarchy in the schizophrenic subjects.

GABAergic modulation of deep pyramidal populations and inhibitory interneurons. The authors then applied this model to patients with frontotemporal lobar degeneration, showing that "the phasic inhibition of deep cortico-cortical pyramidal neurons following tiagabine, but not placebo, was a function of GABA concentration" [199].

Similar approaches have been considered in Parkinson's disease and schizophrenia, spanning cell cultures to human subjects. The basic idea behind these proposals is to link scales through Bayesian model comparison. In other words, optimise the model of neuronal microcircuits in ex vivo cell cultures, in terms of synaptic time constants of intrinsic connections, using Bayesian model inversion and selection. One can then use the posterior estimates as priors for the next scale (e.g., small animal preparations, through to humans). Crucially, the superordinate scale introduces additional parameters (e.g., extrinsic connections among cortical areas) that are estimated more efficiently, having resolved uncertainty about the parameters that are shared with lower scales. In short, by recursively using yesterday's (one scale) posteriors as tomorrow's (next scale) priors, one has a seamless and principled approach to integrating across scales.

COMPUTATIONAL PHENOTYPING AND NOSOLOGY
The narrative so far has moved from the theoretical frameworks, within which to place synaptopathy, to the requisite modelling of empirical data to identify the synaptopathy that characterises

different disorders. This could be regarded as one kind of precision medicine [200]; however, there is another kind that speaks to characterising each individual in the sense of personalised medicine. Computational approaches in this setting are potentially important in stratifying and classifying various clinical phenotypes for the mechanistic studies outlined above.

This usually involves summarising a subject's behaviour in terms of a computational model of that behaviour. Traditionally, this has been a descriptive (e.g., reinforcement learning) model. This represents a weak kind of phenotyping because there are no well-developed process theories for reinforcement learning. A strong complement to computational phenotyping can however be formalised using something called the complete class theorem [201, 202]. This basically says that there for any pair of behaviours and reward functions, there exists some priors that render the behaviour Bayes optimal. This could be read as licensing any Bayesian 'just so' stories about choice behaviour [203]. However, it has a deeper and more pragmatic implication.

It means that any given patient can be fully characterised by her prior beliefs under ideal (active Bayesian inference) observer assumptions. This motivates the difficult game of inferring what generative model this subject is using, based purely upon their behavioural or neuronal responses [204–206]. Although at an early stage, this sort of computational phenotyping shows promise, in the sense that it can be more efficient than simply trying to classify or stratify patients on the basis of their responses per se [20]. This approach, sometimes called generative embedding,
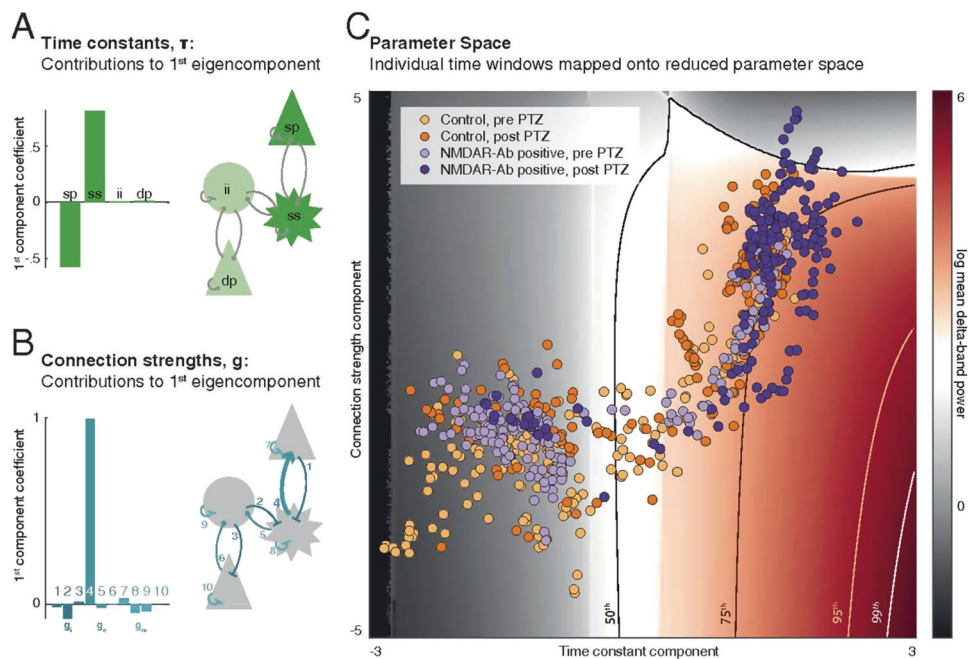
Fig. 5 Measuring synaptopathy in vivo. This figure summarises the results of a DCM study of seizure activity in mice—induced with PTZ—in NMDA receptor antibody (Ab) negative and positive cohorts, respectively [197]. Its agenda was to examine the effects of PTZ and receptor antibodies on synaptic efficacy; summarised here in terms of the principal [eigen] component of changes in intrinsic connectivity. The principal component of time constants implicated superficial pyramidal and spiny stellate cell changes (**A**), while the principal component of connectivity-strength reflects changes in the coupling of spiny stellate to superficial pyramidal cells (**B**). The two principal components (of variations in time constants and intrinsic connectivity) constitute a parameter space summarising the modulation of synaptic efficacy. For each point in this parameter space, one can simulate the spectral responses one would observe data space. **C** in this panel, estimates of synaptic efficacy for experimental subjects in the four conditions (pre-and post-PTZ, with and without antibodies) are shown as coloured dots. These are superimposed on the predicted log mean delta power (with selected centile isoclines). The basic message here is that seizure induction with PTZ shifts synaptic efficacy into regimes of high delta (c.f., slow-wave activity), and that this effect is more marked in the presence of NMDA receptor antibodies. dp, deep pyramidal cells; ii, inhibitory interneurons; sp, superficial pyramidal cells; ss, spiny stellate cells. Adapted with permission from the authors from.

has been applied both to the generative models a subject might use to specify her behaviour and the generative models the researcher uses to explain her physiological or neuronal responses [4, 207].

A related application of generative models to phenotyping involves constructing dynamic causal models of hidden variables or states that underwrite the signs and symptoms expressed by patients over weeks and years. This was a notable outcome of an Ernst Strüngmann forum on computational psychiatry [208], which concluded that conventional diagnoses—based upon traditional nosology—could play an important role in this kind of computational nosology. However, the role was not to stratify patients but rather as data features that supply evidence for or against models of the mapping between pathophysiology and psychopathology that are hidden from direct observation.

A key tenet of this approach is to cast pathophysiology and psychopathology as slowly fluctuating dynamical processes. This speaks to a separation of temporal scales between processes like active inference [206], active learning [32] and structure learning [209]. The mandatory coupling among these (fast and slow) processes may underwrite cyclothymic disorders, relapsing-remitting presentations and neurodevelopmental determinants. An illustrative example of modelling the dynamics of schizoaffective disorders can be found in [208].

An important aspect—of this application of computational psychiatry—is the quantification of uncertainty about how a patient will behave, and what is likely to happen to her in the future. Indeed, uncertainty quantification is a primary focus of related approaches in quantitative epidemiology and meteorology; especially in the context of forecasting and scenario modelling [210]. In principle, finding the right generative model for a patient guarantees the best predictions. This is because a model with the greatest evidence precludes overfitting and ensures generalisability [211]; namely, the generalisation from old (i.e., what has happened to the patient in the past) to new data (what will happen to her in the future).

COMPUTATIONAL NEUROPSYCHOLOGY

A final part of the narrative is that a generative model of pathophysiology, and ensuing psychopathology, allows one to perform in silico or synthetic experiments [212–214]. These entail optimising the parameters of a model of a particular subject or cohort, and seeing what would happen if one increased or decreased certain synaptic efficacies, e.g., [197]. Alternatively, one can emulate extremes of aberrant precision by simply deleting connections in active inference models of diagnostic paradigms, e.g., [215]. In these computational studies, the extrinsic (between node) connections are usually assigned to likelihood mappings relating sensory observations to representations or expectations about hidden states of affairs. Conversely, intrinsic (within node) connections are usually treated as embodying prior beliefs about state transitions and narratives that characterise the paradigm in question.

One obvious advantage of being able to create a 'digital twin' of a patient in silico, is the ability to perform synthetic lesion studies and psychopharmacology, to test various hypotheses about the effect of therapeutic interventions. This is probably best illustrated in the setting of epilepsy, particularly, in predicting the effects of pharmacological and surgical interventions; e.g., [216].

CONCLUSION

This review has taken a somewhat colloquial tour through the fundaments of one (strong) kind of computational psychiatry. Perhaps the take-home message is the foundational importance of a hypothesis or generative model—and being able to articulate this model in computational or formal terms. This is a difficult problem; both in terms of identifying or selecting the right sort of generative models that explain sentient behaviour and its pathologies, but also the biophysical (e.g., dynamic causal) models we, as researchers, use to make sense of our data. This difficulty should not be understated. For example, although dynamic causal modelling has enjoyed wide uptake in fMRI, the requisite models —necessary to explain fine-grained neuronal dynamics in electrophysiology—are notoriously difficult to build and explore using model comparison (a.k.a., structure learning). This usually calls for researchers that are well versed in the system and scale of enquiry, who are also fluent in Bayesian and variational modelling procedures.

The ultimate goal of this endeavour is a generative model or explanation for various psychiatric disorders that furnishes a parsimonious yet accurate account of all the data at hand. This truism reflects the fact that model evidence can be decomposed into accuracy minus complexity; speaking to the pressure to find minimally complex explanations that are as simple as possible but not too simple. In short, the mechanistic explanations we seek will be alluringly simple, but the journey may be difficult. One could argue that psychiatry has chosen the most difficult journey of all; namely, to explain ourselves.

REFERENCES

1. Huys QJ, Moutoussis M, Williams J. Are computational models of any use to psychiatry? Neural Netw. 2011;24:544–51.
2. Montague PR, Dolan RJ, Friston KJ, Dayan P. Computational psychiatry. Trends Cogn Sci. 2012;16:72–80.
3. Corlett PR, Fletcher PC. Computational psychiatry: a Rosetta Stone linking the brain to mental illness. Lancet Psychiatry. 2014;1:399–402.
4. Stephan KE, Mathys C. Computational approaches to psychiatry. Curr Opin Neurobiol. 2014;25:85–92.
5. Wang XJ, Krystal JH. Computational psychiatry. Neuron. 2014;84:638–54.
6. Liddle P, Carpenter WT, Crow T. Syndromes of Schizophrenia: Classic Literature. Br J Psychiatry. 1994;165:721–7.
7. Friston KJ, Liddle PF, Frith CD, Hirsch SR, Frackowiak RS. The left medial temporal region and schizophrenia. A PET study. Brain. 1992;115:367–82.
8. Sutton RS, Barto AG. *Reinforcement Learning: An Introduction*. MIT Press: Cambridge, MA, 1998.
9. MacKay DJC. Information-based objective functions for active data selection. Neural Comput. 1992;4:590–604.
10. Active Motor Babbling for Sensorimotor Learning. Proceedings of the 2008 Ieee International Conference on Robotics and Biomimetics, Vols 1–4. 2009.
11. Schmidhuber J. Formal Theory of Creativity, Fun, and Intrinsic Motivation (1990–2010). Ieee Trans Autonomous Ment Dev. 2010;2:230–47.
12. Barto A, Mirolli M, Baldassarre G. Novelty or surprise? Front Psychol. 2013;4:907.
13. Oudeyer P-Y, Kaplan F. What is intrinsic motivation? a typology of computational approaches. Front Neurorobotics. 2007;1:6.
14. Schwartenbeck P, Passecker J, Hauser TU, FitzGerald TH, Kronbichler M, Friston KJ. Computational mechanisms of curiosity and goal-directed exploration. eLife. 2019;8:e41703.
15. Friston KJ, Frith CD, Liddle PF, Frackowiak RS. Functional connectivity: the principal-component analysis of large (PET) data sets. J Cereb Blood Flow Metab. 1993;13:5–14.
16. Benetti S, Mechelli A, Picchioni M, Broome M, Williams S, McGuire P. Functional integration between the posterior hippocampus and prefrontal cortex is impaired in both first episode schizophrenia and the at risk mental state. Brain. 2009;132:2426–36.
17. Allen P, Stephan KE, Mechelli A, Day F, Ward N, Dalton J, et al. Cingulate activity and fronto-temporal connectivity in people with prodromal signs of psychosis. Neuroimage. 2010;49:947–55.
18. Tan HY, Chen AG, Kolachana B, Apud JA, Mattay VS, Callicott JH, et al. Effective connectivity of AKT1-mediated dopaminergic working memory networks and pharmacogenetics of anti-dopaminergic treatment. Brain. 2012;135:1436–45.
19. Dauvermann MR, Whalley HC, Romaniuk L, Valton V, Owens DG, Johnstone EC, et al. The application of nonlinear Dynamic Causal Modelling for fMRI in subjects at high genetic risk of schizophrenia. Neuroimage. 2013;73:16–29.
20. Brodersen KH, Deserno L, Schlagenhauf F, Lin Z, Penny WD, Buhmann JM, et al. Dissecting psychiatric spectrum disorders by generative embedding. NeuroImage Clin. 2014;4:98–111.
21. Kaplan CM, Saha D, Molina JL, Hockeimer WD, Postell EM, Apud JA, et al. Estimating changing contexts in schizophrenia. Brain. 2016;139:2082–95.
22. Ranlund S, Adams RA, Diez A, Constante M, Dutt A, Hall MH, et al. Impaired prefrontal synaptic gain in people with psychosis and their relatives during the mismatch negativity. Hum Brain Mapp. 2016;37:351–65.
23. Cortes-Briones JA, Tapia-Rivas NI, D'Souza DC, Estevez PA. Going deep into schizophrenia with artificial intelligence. Schizophrenia Res. 2022;245:122–40
24. Marcinkevičs R, Vogt JE. Interpretability and explainability: a machine learning Zoo Mini-tour. 2020, p arXiv:2012.01805.
25. Bai J, Song Q, Cheng G. Efficient variational inference for sparse deep learning with theoretical guarantee. 2020, p arXiv:2011.07439.
26. Millidge B. Deep active inference as Variational Policy Gradients. arXiv e-prints 2019: arXiv:1907.03876.
27. Adversarial Variational Bayes: Unifying Variational Autoencoders and Generative Adversarial Networks. Proceedings of the 34th International Conference on Machine Learning; Aug 06-11 2017; Sydney, AUSTRALIA2017.
28. Planning by Probabilistic Inference. Proceedings of the Proc. of the 9th Int. Workshop on Artificial Intelligence and Statistics 2003.
29. Botvinick M, Toussaint M. Planning as inference. Trends Cogn Sci. 2012;16:485–8.
30. Da Costa L, Parr T, Sajid N, Veselic S, Neacsu V, Friston K. Active inference on discrete state-spaces: a synthesis. J Math Psychol. 2020;99:102447.
31. Lanillos P, Meo C, Pezzato C, Meera AA, Baioumy M, Ohata W, et al. Active Inference in Robotics and Artificial Agents: Survey and Challenges. 2021, p arXiv:2112.01871.
32. MacKay DJC Information Theory, Inference and Learning Algorithms. Cambridge University Press: Cambridge, 2003.
33. Obi-Nagata K, Temma Y, Hayashi-Takagi A. Synaptic functions and their disruption in schizophrenia: From clinical evidence to synaptic optogenetics in an animal model. Proc Jpn Acad Ser B Phys Biol Sci. 2019;95:179–97.
34. Stephan KE, Friston KJ, Frith CD. Dysconnection in schizophrenia: from abnormal synaptic plasticity to failures of self-monitoring. Schizophr Bull. 2009;35:509–27.
35. Friston K, Brown HR, Siemerkus J, Stephan KE. The dysconnection hypothesis (2016). Schizophr Res. 2016;176:83–94.
36. Dima D, Frangou S, Burge L, Braeutigam S, James AC. Abnormal intrinsic and extrinsic connectivity within the magnetic mismatch negativity brain network in schizophrenia: a preliminary study. Schizophr Res. 2012;135:23–7.
37. Schmidt A, Smieskova R, Simon A, Allen P, Fusar-Poli P, McGuire PK, et al. Abnormal effective connectivity and psychopathological symptoms in the psychosis high-risk state. J Psychiatry Neurosci: JPN. 2014;39:239–48.
38. Helmholtz H. Concerning the perceptions in general. Treatise on physiological optics, vol. III. Dover: New York, 1866/1962.
39. Helmholtz H. The Facts of Perception. In: Middletown RK (ed). The Selected Writings of Hermann von Helmholtz. Wesleyan University Press: Connecticut, 1878, p 384.
40. Yuille A, Kersten D. Vision as Bayesian inference: analysis by synthesis? Trends Cogn Sci. 2006;10:301–8.
41. Neisser U. Cognitive psychology. Appleton-Century-Crofts: New York, 1967.
42. MacKay DMC. The epistemological problem for automata. In: Shannon C, McCarthy J (eds). Automata Studies. Princeton University Press: Princeton, NJ, 1956, pp 235–51.
43. Gregory RL. Perceptual illusions and brain models. Proc R Soc Lond B. 1968;171:179–96.
44. Gregory RL. Perceptions as hypotheses. Philos Trans R Soc Lond B. 1980;290:181–97.
45. Dayan P, Hinton GE, Neal RM, Zemel RS. The Helmholtz machine. Neural Comput. 1995;7:889–904.
46. Hutto D, Myin E. Radicalizing enactivism: Basic minds without content. MIT Press: Cambridge, MA, 2013.
47. Clark A, Chalmers DJ. The extended mind. Analysis. 1998;58:7–19.
48. Clark A. Being there putting brain, body, and world together again. MIT Press: Cambridge, Mass. 2001.
49. Goodwin C. Action and embodiment within situated human interaction. J Pragmat. 2000;32:1489–522.
50. Barsalou LW. Grounded cognition. Annu Rev Psychol. 2008;59:617–45.
51. Chemero A. Radical Embodied Cognitive Science, 2009, pp 1–252.
52. Seth AK. A predictive processing theory of sensorimotor contingencies: Explaining the puzzle of perceptual presence and its absence in synesthesia. Cogn Neurosci. 2014;5:97–118.

53. Wiese W. Action Is Enabled by Systematic Misrepresentations. Erkenntnis. 2017;82:1233–52.

54. Parr T, Friston KJ. The Anatomy of Inference: Generative Models and Brain Structure. Front Comput Neurosci. 2018;12:90

55. Hohwy J. New directions in predictive processing. Mind Language. 2020;35:209–33

56. Rao RP, Ballard DH. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nat Neurosci. 1999;2:79–87.

57. Ballard DH, Kit D, Rothkopf CA, Sullivan B. A Hierarchical Modular Architecture for Embodied Cognition. Multisens Res. 2013;26:177.

58. Friston KJ, Daunizeau J, Kilner J, Kiebel SJ. Action and behavior: a free-energy formulation. Biol Cybern. 2010;102:227–60.

59. Clark A. Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behav Brain Sci. 2013;36:181–204.

60. Hohwy J. The Predictive Mind. Oxford University Press: Oxford, 2013.

61. Knill DC, Pouget A. The Bayesian brain: the role of uncertainty in neural coding and computation. Trends Neurosci. 2004;27:712–9.

62. Doya K. Bayesian brain: Probabilistic approaches to neural coding. MIT press 2007.

63. Hohwy J. The Self-Evidencing. Brain Noûs. 2016;50:259–85.

64. Elias P. Predictive coding–I. IRE Trans Inf Theory. 1955;1:16–24.

65. Srinivasan MV, Laughlin SB, Dubs A. Predictive coding: a fresh view of inhibition in the retina. Proc R Soc Lond B Biol Sci. 1982;216:427–59.

66. Bastos AM, Usrey WM, Adams RA, Mangun GR, Fries P, Friston KJ. Canonical microcircuits for predictive coding. Neuron. 2012;76:695–711.

67. Egner T, Summerfield C. Grounding predictive coding models in empirical neuroscience research. Behav Brain Sci. 2013;36:210–1.

68. Winn J, Bishop CM. Variational message passing. J Mach Learn Res. 2005;6:661–94.

69. On Variational Message Passing on Factor Graphs. Proceedings of the 2007 IEEE International Symposium on Information Theory; 24–29 June 2007, 2007.

70. Friston K, Parr T, de Vries B. The graphical brain: Belief propagation and active inference. Netw Neurosci (Camb, Mass). 2017;1:381–414.

71. Parr T, Markovic D, Kiebel SJ, Friston KJ. Neuronal message passing using Mean-field, Bethe, and Marginal approximations. Sci Rep. 2019;9:1889.

72. Da Costa L, Parr T, Sengupta B, Friston K. Neural Dynamics under Active Inference: Plausibility and Efficiency of Information Processing. Entropy (Basel). 2021;23:454

73. Adams RA, Shipp S, Friston KJ. Predictions not commands: active inference in the motor system. Brain Struct Funct. 2013;218:611–43.

74. Shipp S. Neural Elements for Predictive Coding. Front Psychol. 2016;7:1792.

75. Mumford D. On the computational architecture of the neocortex. II. Biol Cyber. 1992;66:241–51.

76. Lee TS, Mumford D. Hierarchical Bayesian inference in the visual cortex. J Optical Soc Am A, Opt, image Sci, Vis. 2003;20:1434–48.

77. Markov NT, Ercsey-Ravasz M, Van Essen D, Knoblauch K, Toroczkai Z, Kennedy H. Cortical high-density counterstream architectures. Science. 2013;342:1238406.

78. Pinotsis DA, Brunet N, Bastos A, Bosman CA, Litvak V, Fries P, et al. Contrast gain control and horizontal interactions in V1: a DCM study. Neuroimage. 2014;92:143–55.

79. Shaw AD, Moran RJ, Muthukumaraswamy SD, Brealy J, Linden DE, Friston KJ, et al. Neurophysiologically-informed markers of individual variability and pharmacological manipulation of human cortical gamma. Neuroimage. 2017;161:19–31.

80. Kok P, Rahnev D, Jehee JFM, Lau HC, de Lange FP. Attention Reverses the Effect of Prediction in Silencing Sensory Signals. Cereb Cortex. 2012;22:2197–206.

81. Auksztulewicz R, Friston K. Attentional Enhancement of Auditory Mismatch Responses: a DCM/MEG Study. Cereb Cortex. 2015;25:4273–83.

82. Ainley V, Apps MA, Fotopoulou A, Tsakiris M. 'Bodily precision': a predictive coding account of individual differences in interoceptive accuracy. Philosophical transactions of the Royal Society of London Series B, Biological sciences. 2016;371:20160003

83. Limanowski J. Precision control for a flexible body representation. Neurosci Biobehavioral Rev. 2021;134:104401.

84. Clark A. The many faces of precision (Replies to commentaries on "Whatever next? Neural prediction, situated agents, and the future of cognitive science"). Front Psychol. 2013;4:270.

85. Feldman H, Friston KJ. Attention, uncertainty, and free-energy. Front Hum Neurosci. 2010;4:215.

86. Graboi D, Lisman J. Recognition by top-down and bottom-up processing in cortex: the control of selective attention. J Neurophysiol. 2003;90:798–810.

87. Lisman J, Buzsaki G. A neural coding scheme formed by the combined function of gamma and theta oscillations. Schizophr Bull. 2008;34:974–80.

88. Lisman J. Excitation, inhibition, local oscillations, or large-scale loops: what causes the symptoms of schizophrenia? Curr Opin Neurobiol. 2012;22:537–44.

89. Bauer M, Stenner MP, Friston KJ, Dolan RJ. Attentional modulation of alpha/beta and gamma oscillations reflect functionally distinct processes. J Neurosci. 2014;34:16117–25.

90. Spencer KM, Nestor PG, Niznikiewicz MA, Salisbury DF, Shenton ME, McCarley RW. Abnormal neural synchrony in schizophrenia. J Neurosci. 2003;23:7407–11.

91. Sohal VS, Zhang F, Yizhar O, Deisseroth K. Parvalbumin neurons and gamma rhythms enhance cortical circuit performance. Nature. 2009;459:698–702.

92. Parr T, Friston KJ. Uncertainty, epistemics and active inference. J R Soc Interface. 2017;14:20170376.

93. Brown H, Adams RA, Parees I, Edwards M, Friston K. Active inference, sensory attenuation and illusions. Cogn Process. 2013;14:411–27.

94. Oestreich LK, Mifsud NG, Ford JM, Roach BJ, Mathalon DH, Whitford TJ. Subnormal sensory attenuation to self-generated speech in schizotypy: Electrophysiological evidence for a 'continuum of psychosis'. Int J Psychophysiol: Off J Int Organ Psychophysiol. 2015;97:131–8.

95. Limanowski J (Dis-)Attending to the Body. In: Metzinger TK, Wiese W (eds). *Philosophy and Predictive Processing*. MIND Group: Frankfurt am Main, 2017.

96. Blakemore SJ, Frith CD, Wolpert DM. Spatio-temporal prediction modulates the perception of self-produced stimuli. J Cogn Neurosci. 1999;11:551–9.

97. Shergill SS, Samson G, Bays PM, Frith CD, Wolpert DM. Evidence for sensory prediction deficits in schizophrenia. Am J Psychiatry. 2005;162:2384–6.

98. Hughes G, Desantis A, Waszak F. Mechanisms of intentional binding and sensory attenuation: the role of temporal prediction, temporal control, identity prediction, and motor prediction. Psychol Bull. 2013;139:133–51.

99. Wurtz RH. Neuronal mechanisms of visual stability. Vis Res. 2008;48:2070–89.

100. Adams RA, Stephan KE, Brown HR, Frith CD, Friston KJ. The computational anatomy of psychosis. Front Psychiatry. 2013;4:47.

101. Fuster JM. Upper processing stages of the perception-action cycle. Trends Cogn Sci. 2004;8:143–5.

102. Wurtz RH, McAlonan K, Cavanaugh J, Berman RA. Thalamic pathways for active vision. Trends Cogn Sci. 2011;5:177–84.

103. Pillmann F. Carl Wernicke and the neurobiological paradigm in psychiatry. Acta Neuropsychologica. 2007;5:246–60.

104. Catani M, Ffytche DH. The rises and falls of disconnection syndromes. Brain. 2005;128:2224–39.

105. Bob P, Mashour GA. Schizophrenia, dissociation, and consciousness. Conscious Cogn. 2011;20:1042–9.

106. Weinberger DR. A connectionist approach to the prefrontal cortex. J Neuropsychiatry Clin Neurosci. 1993;5:241–53.

107. Friston KJ, Frith CD. Schizophrenia: a disconnection syndrome? Clin Neurosci. 1995;3:89–97.

108. Grant SG. Synaptopathies: diseases of the synaptome. Curr Opin Neurobiol. 2012;22:522–9.

109. Nicastro N, Holland N, Savulich G, Carter SF, Mak E, Hong YT, et al. (11)C-UCB-J synaptic PET and multimodal imaging in dementia with Lewy bodies. Eur J Hybrid Imaging. 2020;4:25–5.

110. Cohen JD, Servan-Schreiber D. Context, cortex, and dopamine: a connectionist approach to behavior and biology in schizophrenia. Psychol Rev. 1992;99:45–77.

111. Winterer G, Weinberger DR. Genes, dopamine and cortical signal-to-noise ratio in schizophrenia. Trends Neurosci. 2004;27:683–90.

112. Laruelle M. Schizophrenia: from dopaminergic to glutamatergic interventions. Curr Opin Pharm. 2014;14C:97–102.

113. Geyer MA, Krebs-Thomson K, Braff DL, Swerdlow NR. Pharmacological studies of prepulse inhibition models of sensorimotor gating deficits in schizophrenia: a decade in review. Psychopharmacology. 2001;156:117–54.

114. Krystal JH, Karper LP, Seibyl JP, Freeman GK, Delaney R, Bremner JD, et al. Subanesthetic effects of the noncompetitive NMDA antagonist, ketamine, in humans. Psychotomimetic, perceptual, cognitive, and neuroendocrine responses. Arch Gen Psychiatry. 1994;51:199–214.

115. Olney JW, Newcomer JW, Farber NB. NMDA receptor hypofunction model of schizophrenia. J Psychiatr Res. 1999;33:523–33.

116. Gonzalez-Burgos G, Lewis DA. NMDA receptor hypofunction, parvalbumin-positive neurons, and cortical gamma oscillations in schizophrenia. Schizophr Bull. 2012;38:950–7.

117. Lisman JE, Coyle JT, Green RW, Javitt DC, Benes FM, Heckers S, et al. Circuit-based framework for understanding neurotransmitter and risk gene interactions in schizophrenia. Trends Neurosci. 2008;31:234–42.

118. Uhlhaas PJ, Singer W. Abnormal neural oscillations and synchrony in schizophrenia. Nat Rev Neurosci. 2010;11:100–13.

119. Jardri R, Hugdahl K, Hughes M, Brunelin J, Waters F, Alderson-Day B, et al. Are Hallucinations Due to an Imbalance Between Excitatory and Inhibitory Influences on the Brain? Schizophrenia Bull. 2016;42:1124–34.

120. Schizophrenia Working Group of the Psychiatric Genomics C. Biological insights from 108 schizophrenia-associated genetic loci. Nature. 2014;511:421–7.

121. Hall J, Bray NJ. Schizophrenia Genomics: Convergence on Synaptic Development, Adult Synaptic Plasticity, or Both? Biol Psychiatry. 2022;91:709–17

122. Trubetskoy V, Pardiñas AF, Qi T, Panagiotaropoulou G, Awasthi S, Bigdeli TB, et al. Mapping genomic loci implicates genes and synaptic biology in schizophrenia. Nature. 2022;604:502–8.

123. Singh T, Poterba T, Curtis D, Akil H, Al Eissa M, Barchas JD, et al. Rare coding variants in ten genes confer substantial risk for schizophrenia. Nature. 2022;604:509–16.

124. Sanacora G, Treccani G, Popoli M. Towards a glutamate hypothesis of depression An emerging frontier of neuropsychopharmacology for mood disorders. Neuropharmacology. 2012;62:63–77.

125. Hamon M, Blier P. Monoamine neurocircuitry in depression and strategies for new treatments. Prog Neuropsychopharmacol Biol Psychiatry. 2013;45:54–63.

126. Russo M, Carrarini C, Dono F, Rispoli MG, Di Pietro M, Di Stefano V, et al. The pharmacology of visual hallucinations in synucleinopathies. Front Pharm. 2019;10:1379.

127. Collerton D, Perry E, McKeith I. Why people see things that are not there: a novel Perception and Attention Deficit model for recurrent complex visual hallucinations. Behav Brain Sci. 2005;28:737–57.

128. Friston KJ. Hallucinations and perceptual inference. Behav Brain Sci. 2005;28:764–6.

129. Fletcher PC, Frith CD. Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia. Nat Rev Neurosci. 2009;10:48–58.

130. Teufel C, Fletcher P, Davis G. Seeing other minds: attributed mental states influence perception. Trends Cogn Sci. 2010;14:376–82.

131. Corlett PR, Honey GD, Krystal JH, Fletcher PC. Glutamatergic model psychoses: prediction error, learning, and inference. Neuropsychopharmacology. 2011;36:294–315.

132. Haarsma J, Fletcher PC, Ziauddeen H, Spencer TJ, Diederen KMJ, Murray GK. Precision weighting of cortical unsigned prediction errors is mediated by dopamine and benefits learning. bioRxiv. 2018.

133. Sterzer P, Adams RA, Fletcher P, Frith C, Lawrie SM, Muckli L, et al. The predictive coding account of psychosis. Biol Psychiatry. 2018;84:634–43.

134. Pellicano E, Burr D. When the world becomes 'too real': a Bayesian explanation of autistic perception. Trends Cogn Sci. 2012;16:504–10.

135. Lawson RP, Rees G, Friston KJ. An aberrant precision account of autism. Front Hum Neurosci. 2014;8:302.

136. Van de Cruys S, Evers K, Van der Hallen R, Van Eylen L, Boets B, de-Wit L, et al. Precise minds in uncertain worlds: predictive coding in autism. Psychol Rev. 2014;121:649–75.

137. Edwards MJ, Adams RA, Brown H, Parees I, Friston KJ. A Bayesian account of 'hysteria'. Brain. 2012;135:3495–512.

138. Skewes JC, Jegindo EM, Gebauer L. Perceptual inference and autistic traits. Autism. 2014: [Epub ahead of print].

139. Gu X, Eilam-Stock T, Zhou T, Anagnostou E, Kolevzon A, Soorya L, et al. Autonomic and brain responses associated with empathy deficits in autism spectrum disorder. Hum Brain Mapp. 2015;36:3323–38.

140. Clark JE, Watson S, Friston KJ. What is mood? A computational perspective. Psychol Med. 2018;48:2277–84.

141. Benrimoh D, Parr T, Adams RA, Friston K. Hallucinations both in and out of context: an active inference account. PLoS One. 2019;14:e0212379.

142. Krahe C, Springer A, Weinman JA, Fotopoulou A. The social modulation of pain: others as predictive signals of salience - a systematic review. Front Hum Neurosci. 2013;7:386.

143. Hoskin R, Berzuini C, Acosta-Kane D, El-Deredy W, Guo H, Talmi D. Sensitivity to pain expectations: A Bayesian model of individual differences. Cognition. 2018;182:127–39.

144. Gerrans P. Pain Asymbolia as depersonalization for pain experience. An Interoceptive Active Inference Account. Front Psychol. 2020;11:523710

145. Seymour B, Mancini F. Hierarchical models of pain: Inference, information-seeking, and adaptive control. NeuroImage. 2020;222:117212.

146. Paulus MP, Stein MB. An insular view of anxiety. Biol Psychiatry. 2006;60:383–7.

147. Cornwell BR, Garrido MI, Overstreet C, Pine DS, Grillon C. The Unpredictive Brain Under Threat: A Neurocomputational Account of Anxious Hypervigilance. Biol Psychiatry. 2017;82:447–54.

148. Duquette P. Increasing Our Insular World View: Interoception and Psychopathology for Psychotherapists. Front Neurosci. 2017;11:135.

149. Smith R, Kuplicki R, Feinstein J, Forthman KL, Stewart JL, Paulus MP, et al. A Bayesian computational model reveals a failure to adapt interoceptive precision estimates across depression, anxiety, eating, and substance use disorders. PLOS Computational Biol. 2020;16:e1008484.

150. Notredame CE, Pins D, Deneve S, Jardri R. What visual illusions teach us about schizophrenia. Front Integr Neurosci. 2014;8:63.

151. Powers AR 3rd, Gancsos MG, Finn ES, Morgan PT, Corlett PR. Ketamine-Induced Hallucinations. Psychopathology. 2015;48:376–85.

152. Teufel C, Subramaniam N, Dobler V, Perez J, Finnemann J, Mehta PR, et al. Shift toward prior knowledge confers a perceptual advantage in early psychosis and psychosis-prone healthy individuals. Proc Natl Acad Sci USA. 2015;112:13401–6.

153. Powers AR, Mathys C, Corlett PR. Pavlovian conditioning–induced hallucinations result from overweighting of perceptual priors. Science. 2017;357:596–600.

154. Moutoussis M, Bentall RP, El-Deredy W, Dayan P. Bayesian modelling of Jumping-to-Conclusions bias in delusional patients. Cogn Neuropsychiatry. 2011;16:422–47.

155. Joyce DW, Averbeck BB, Frith CD, Shergill SS. Examining belief and confidence in schizophrenia. Psychol Med. 2013;43:2327–38.

156. Paulus MP, Feinstein JS, Khalsa SS. An Active Inference Approach to Interoceptive Psychopathology. In: Widiger T, Cannon TD (eds). Annual Review of Clinical Psychology, Vol 15, vol. 152019, pp 97–122.

157. Rae CL, Critchley HD, Seth AK. A Bayesian Account of the Sensory-Motor Interactions Underlying Symptoms of Tourette Syndrome. Front Psychiatry. 2019;10:29

158. Stuke H, Weilnhammer VA, Sterzer P, Schmack K. Delusion Proneness is Linked to a Reduced Usage of Prior Beliefs in Perceptual Decisions. Schizophrenia Bull. 2019;45:80–6.

159. Kube T, Rozenkrantz L, Rief W, Barsky A. Understanding persistent physical symptoms: Conceptual integration of psychological expectation models and predictive processing accounts. Clin Psychol Rev. 2020;76:101829.

160. Lyndon S, Corlett PR. Hallucinations in Posttraumatic Stress Disorder: Insights From Predictive Coding. J Abnorm Psychol. 2020;129:534–43.

161. Richards KL, Karvelis P, Lawrie SM, Series P. Visual statistical learning and integration of perceptual priors are intact in attention deficit hyperactivity disorder. Plos One. 2020;15:e0243100

162. Frith CD, Blakemore SJ, Wolpert DM. Abnormalities in the awareness and control of action. Philos Trans R Soc Lond Ser B, Biol Sci. 2000;355:1771–88.

163. Shergill SS, White TP, Joyce DW, Bays PM, Wolpert DM, Frith CD. Functional magnetic resonance imaging of impaired sensory prediction in schizophrenia. JAMA Psychiatry. 2014;71:28–35.

164. Parees I, Brown H, Nuruki A, Adams RA, Davare M, Bhatia KP, et al. Loss of sensory attenuation in patients with functional (psychogenic) movement disorders. Brain. 2014;137:2916–21.

165. Bhat A, Parr T, Ramstead M, Friston K. Immunoceptive inference: why are psychiatric disorders and immune responses intertwined? Biol Philos. 2021;36:27.

166. Quattrocki E, Friston K. Autism, oxytocin and interoception. Neurosci Biobehav Rev. 2014;47c:410–30.

167. Palmer CJ, Paton B, Kirkovski M, Enticott PG, Hohwy J. Context sensitivity in action decreases along the autism spectrum: a predictive processing perspective. Proc R Soc B: Biol Sci. 2015;282:20141557.

168. Palmer CJ, Seth AK, Hohwy J. The felt presence of other minds: Predictive processing, counterfactual predictions, and mentalising in autism. Conscious Cogn. 2015.

169. Lawson RP, Mathys C, Rees G. Adults with autism overestimate the volatility of the sensory environment. Nat Neurosci. 2017;20:1293–9.

170. Palmer CJ, Lawson RP, Hohwy J. Bayesian approaches to autism: Towards volatility, action, and behavior. Psychol Bull. 2017;143:521–42.

171. Coltheart M, Menzies P, Sutton J. Abductive inference and delusional belief. Cogn Neuropsychiatry. 2010;15:261–87.

172. Adams RA, Vincent P, Benrimoh D, Friston KJ, Parr T. Everything is connected: Inference and attractors in delusions. Schizophr Res. 2022;245:5–22.

173. Butler PD, Silverstein SM, Dakin SC. Visual perception and its impairment in schizophrenia. Biol Psychiatry. 2008;64:40–7.

174. Seth AK, Friston KJ. Active interoceptive inference and the emotional brain. Philos Trans R Soc Lon B Biol Sci. 2016;371:20160007.

175. Averbeck BB, Evans S, Chouhan V, Bristow E, Shergill SS. Probabilistic learning and inference in schizophrenia. Schizophr Res. 2011;127:115–22.

176. Kanai R, Komura Y, Shipp S, Friston K. Cerebral hierarchies: predictive processing, precision and the pulvinar. Philosophical Trans R Soc Lond B Biol Sci. 2015;370:20140169.

177. Friston K, FitzGerald T, Rigoli F, Schwartenbeck P, Pezzulo G. Active Inference: A Process Theory. Neural Comput. 2017;29:1–49.

178. Moran RJ, Symmonds M, Stephan KE, Friston KJ, Dolan RJ. An in vivo assay of synaptic function mediating human cognition. Curr Biol. 2011;21:1320–5.

179. Shaw AD, Hughes LE, Moran R, Coyle-Gilchrist I, Rittman T, Rowe JB. In Vivo Assay of Cortical Microcircuitry in Frontotemporal Dementia: A Platform for Experimental Medicine Studies. Cereb Cortex. 2021;31:1837–47.

180. Self MW, van Kerkoerle T, Goebel R, Roelfsema PR. Benchmarking laminar fMRI: Neuronal spiking and synaptic activity during top-down and bottom-up processing in the different layers of cortex. Neuroimage. 2019;197:806–17.

181. Haarsma J, Kok P, Browning M. The promise of layer-specific neuroimaging for testing predictive coding theories of psychosis. Schizophrenia Res. 2020.

182. Kravitz AV, Freeze BS, Parker PR, Kay K, Thwin MT, Deisseroth K, et al. Regulation of parkinsonian motor behaviours by optogenetic control of basal ganglia circuitry. Nature. 2010;466:622–6.

183. Anenberg E, Chan AW, Xie Y, LeDue JM, Murphy TH. Optogenetic stimulation of GABA neurons can decrease local neuronal activity while increasing cortical blood flow. J Cereb Blood Flow Metab: Off J Int Soc Cereb Blood Flow Metab. 2015;35:1579–86.

184. Okun M, Steinmetz NA, Cossell L, Iacaruso MF, Ko H, Bartho P, et al. Diverse coupling of neurons to populations in sensory cortex. Nature. 2015;521:511–5.

185. Penny WD. Comparing dynamic causal models using AIC, BIC and free energy. Neuroimage. 2012;59:319–30.

186. Friston KJ. Functional and effective connectivity: a review. Brain Connectivity. 2011;1:13–36.

187. Snyder AD, Ma L, Steinberg JL, Woisard K, Moeller FG. Dynamic Causal Modeling Self-Connectivity Findings in the Functional Magnetic Resonance Imaging Neuropsychiatric Literature. Front Neurosci 2021;15:636273.

188. Curcic-Blake B, van der Meer L, Pijnenborg GH, David AS, Aleman A. Insight and psychosis: Functional and anatomical brain connectivity and self-reflection in Schizophrenia. Hum Brain Mapp. 2015;36:4859–68.

189. Bastos-Leite AJ, Ridgway GR, Silveira C, Norton A, Reis S, Friston KJ. Dysconnectivity within the default mode in first-episode schizophrenia: a stochastic dynamic causal modeling study with functional magnetic resonance imaging. Schizophr Bull. 2015;41:144–53.

190. Ranlund S, Nottage J, Shaikh M, Dutt A, Constante M, Walshe M, et al. Resting EEG in psychosis and at-risk populations-a possible endophenotype? Schizophr Res. 2014;153:96–102.

191. Fogelson N, Litvak V, Peled A, Fernandez-del-Olmo M, Friston K. The functional anatomy of schizophrenia: A dynamic causal modeling study of predictive coding. Schizophr Res. 2014;158:204–12.

192. Adams RA, Pinotsis D, Tsirlis K, Unruh L, Mahajan A, Horas AM, et al. Computational Modeling of Electroencephalography and Functional Magnetic Resonance Imaging Paradigms Indicates a Consistent Loss of Pyramidal Cell Synaptic Gain in Schizophrenia. Biol Psychiatry. 2022;91:202–15.

193. Rosch R, Burrows DRW, Jones LB, Peters CH, Ruben P, Samarut E. Functional Genomics of Epilepsy and Associated Neurodevelopmental Disorders Using Simple Animal Models: From Genes, Molecules to Brain Networks. Front Cell Neurosci. 2019;13:556.

194. Rosch RE, Hunter PR, Baldeweg T, Friston KJ, Meyer MP. Calcium imaging and dynamic causal modelling reveal brain-wide changes in effective connectivity and synaptic dynamics during epileptic seizures. PLoS Comput Biol. 2018;14:e1006375.

195. Peters C, Rosch RE, Hughes E, Ruben PC. Temperature-dependent changes in neuronal dynamics in a patient with an SCN1A mutation and hyperthermia induced seizures. Sci Rep. 2016;6:31879.

196. Rowe JB, Hughes LE, Barker RA, Owen AM. Dynamic causal modelling of effective connectivity from fMRI: are results reproducible and sensitive to Parkinson's disease and its treatment? Neuroimage. 2010;52:1015–26.

197. Rosch RE, Wright S, Cooray G, Papadopoulou M, Goyal S, Lim M, et al. NMDA-receptor antibodies alter cortical microcircuit dynamics. Proc Natl Acad Sci USA. 2018;115:E9916–25.

198. Adams NE, Hughes LE, Phillips HN, Shaw AD, Murley AG, Nesbitt D, et al. GABA-ergic Dynamics in Human Frontotemporal Networks Confirmed by Pharmaco-Magnetoencephalography. J Neurosci. 2020;40:1640–9.

199. Adams NE, Hughes LE, Rouse MA, Phillips HN, Shaw AD, Murley AG, et al. GABAergic cortical network physiology in frontotemporal lobar degeneration. Brain. 2021;144:2135–45.

200. Kaufman J, Gelernter J, Hudziak JJ, Tyrka AR, Coplan JD. The Research Domain Criteria (RDoC) Project and Studies of Risk and Resilience in Maltreated Children. J Am Acad Child Adolesc Psychiatry. 2015;54:617–25.

201. Wald A. An Essentially Complete Class of Admissible Decision Functions. Ann Math Stat. 1947;18:549–55.

202. Brown LD. A Complete Class Theorem for Statistical Problems with Finite-Sample Spaces. Ann Stat. 1981;9:1289–300.

203. Bowers JS, Davis CJ. Bayesian just-so stories in psychology and neuroscience. Psychol Bull. 2012;138:389–414.

204. Schwartenbeck P, Friston K. Computational Phenotyping in Psychiatry: A Worked Example. eNeuro. 2016;3:0049-16.2016.

205. Smith R, Kirlic N, Stewart JL, Touthang J, Kuplicki R, McDermott TJ, et al. Long-term stability of computational parameters during approach-avoidance conflict in a transdiagnostic psychiatric patient sample. Sci Rep. 2021;11:11783.

206. Smith R, Kirlic N, Stewart JL, Touthang J, Kuplicki R, Khalsa SS, et al. Greater decision uncertainty characterizes a transdiagnostic patient sample during approach-avoidance conflict: a computational modelling approach. J Psychiatry Neurosci. 2021;46:E74–87.

207. Iglesias S, Mathys C, Brodersen KH, Kasper L, Piccirelli M, den Ouden HEM, et al. Hierarchical prediction errors in midbrain and basal forebrain during sensory learning. Neuron. 2013;80:519–30.

208. Friston KJ, Redish AD, Gordon JA. Computational Nosology and Precision Psychiatry. Comput Psychiatr. 2017;1:2–23.

209. Tervo DG, Tenenbaum JB, Gershman SJ. Toward the neural implementation of structure learning. Curr Opin Neurobiol. 2016;37:99–105.

210. Wu JT, Leung K, Leung GM. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. Lancet (Lond, Engl). 2020;395:689–97.

211. Friston K, Costello A, Pillay D. 'Dark matter', second waves and epidemiological modelling. BMJ Glob Health. 2020;5:e003978.

212. Testolin A, Zorzi M. Probabilistic Models and Generative Neural Networks: Towards an Unified Framework for Modeling Normal and Impaired Neurocognitive Functions. Front Computational Neurosci. 2016;10:73.

213. Parr T, Rees G, Friston KJ. Computational Neuropsychology and Bayesian Inference. Front Hum Neurosci. 2018;12:61.

214. Smith R, Khalsa SS, Paulus MP. An Active Inference Approach to Dissecting Reasons for Nonadherence to Antidepressants. Biol Psychiatry-Cogn Neurosci Neuroimaging. 2021;6:919–34.

215. Sajid N, Parr T, Gajardo-Vidal A, Price CJ, Friston KJ. Paradoxical lesions, plasticity and active inference. Brain Commun. 2020;2:fcaa164.

216. Proix T, Jirsa VK, Bartolomei F, Guye M, Truccolo W. Predicting the spatio-temporal diversity of seizure propagation and termination in human focal epilepsy. Nat Commun. 2018;9:1088.

217. Pfeffer CK, Xue M, He M, Huang ZJ, Scanziani M. Inhibition of inhibition in visual cortex: the logic of connections between molecularly distinct interneurons. Nat Neurosci. 2013;16:1068–76.

218. Yu AJ, Dayan P. Acetylcholine in cortical inference. Neural Netw. 2002;15:719–30.

219. Dayan P, Yu AJ. ACh, uncertainty, and cortical inference. In: Dietterich T, Becker S, Ghahramani Z (eds). Advances in Neural Information Processing Systems. Vol 14, MIT Press; 2001.

220. Dayan P, Yu AJ. Phasic norepinephrine: a neural interrupt signal for unexpected events. Netw: Comput Neural Syst. 2006;17:335–50.

221. Friston K, Schwartenbeck P, FitzGerald T, Moutoussis M, Behrens T, Dolan RJ. The anatomy of choice: dopamine and decision-making. Philos Trans R Soc B: Biol Sci. 2014;369:20130481.

GLOSSARY

Active inference	securing evidence for a generative model of the world through Bayesian inference and active sampling of (sensory) data. This is equivalent to resolving uncertainty and maximising **model evidence**.
Bayesian belief	a posterior probability distribution over a random variable, such as a latent cause or hidden state of the world causing (sensory) data.
Belief updating	the process of statistical inference, in which Bayes' theorem is used to update a **prior belief** to a **posterior belief** on the basis of new evidence.
Disconnection (syndrome)	a term for neuropsychiatric syndromes caused by damage to the white matter axons of associational or commissural pathways: i.e., a disrupted connection.
Dysconnection (syndrome)	a term for neuropsychiatric syndromes caused by a synaptopathy involving associational or commissural pathways: i.e., a dysfunctional connection.

Inference	updating beliefs to maximise **model evidence**. Variational Bayesian inference corresponds to minimising variational free energy, which provides an upper bound on log evidence.
Interoceptive	pertaining to internal (e.g., autonomic) states
Likelihood	the probability of observing some (sensory) data, given the causes of those data
Model evidence	the probability of some (sensory) data under a generative model. Also known as the marginal likelihood, i.e., having marginalised over the causes of the data.
Precision	The inverse variance of a random variable. More generally, the precision of a probability distribution decreases with its entropy.
Prediction	the prediction of some (sensory) data, based upon **posterior beliefs** about how those data were generated.
Prediction error	A quantity used in predictive coding to denote the difference between an observation or point estimate and its predicted value. Predictive coding uses precision weighted prediction errors to update **posterior beliefs** that generate **predictions**.
Prior belief	a belief prior to sampling (sensory) data.
Sejunction hypothesis	(from the Latin noun *seiunctio*: divorce, separation) was introduced by Carl Wernicke (1848-1904) to suggest that psychopathology results from interruption ("sejunction") of associative connections in the brain.
Variational inference	a tractable form of Bayesian belief updating based upon minimising variational free energy (a.k.a., an evidence bound). The gradients of variational free energy correspond to precision weighted prediction errors.

ADDITIONAL INFORMATION

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.