An integrative framework for perceptual disturbances in psychosis

Guillermo Horga[1] * *and Anissa Abi-Dargham*[2] *

Abstract | Perceptual disturbances in psychosis, such as auditory verbal hallucinations, are associated with increased baseline activity in the associative auditory cortex and increased dopamine transmission in the associative striatum. Perceptual disturbances are also associated with perceptual biases that suggest increased reliance on prior expectations. We review theoretical models of perceptual inference and key supporting physiological evidence, as well as the anatomy of associative cortico–striatal loops that may be relevant to auditory perceptual inference. Integrating recent findings, we outline a working framework that bridges neurobiology and the phenomenology of perceptual disturbances via theoretical models of perceptual inference.

Schizophrenia
A psychiatric illness characterized by a variety of symptoms, including positive symptoms, negative symptoms (for example, apathy and amotivation) and cognitive impairments (for example, memory deficits).

Psychotic disorders
A group of disorders, including schizophrenia and other disorders, such as bipolar disorder, with psychotic features, that present with psychotic or positive symptoms.

Auditory verbal hallucinations
(AVH). Percepts of speech or voices without corresponding speech stimuli.

[1]*Deparment of Psychiatry, Columbia University, and New York State Psychiatric Institute, New York, NY, USA.*

[2]*Department of Psychiatry, Stony Brook University, Stony Brook, NY, USA.*

**e-mail: HorgaG@nyspi.columbia.edu; anissa.abi-dargham@stonybrookmedicine.edu*

https://doi.org/10.1038/s41583-019-0234-1

Explanations such as 'hallucinations are caused by overactive dopamine receptors' are unsatisfactory because they leave an explanatory gap between the mental and the physical. How can dopamine cause a voice or a belief? — Paul C. Fletcher and Chris D. Frith[1]

Perceptual disturbances, including hallucinations or false percepts, are a defining and prevalent feature of schizophrenia and other psychotic disorders.

Perception is the subjective process by which neural systems represent, disambiguate and interpret sensory inputs from the environment. Because sensory inputs are always conveyed with some level of noise or uncertainty, a key function of sensory systems is to reduce sensory uncertainty in order to facilitate accurate judgments that permit successful adaptation to the environment. To do so in an optimal way, sensory systems leverage the prior knowledge that an agent has acquired through experience. Optimal sensory processing, however, does not imply error-free perceptual judgments. As we later discuss in more detail, incorporating prior knowledge into percepts, although generally advantageous, biases perception towards more likely sensory events: that is, it implies perceptual-judgment errors on the side of likely scenarios. This predisposes healthy individuals to a plethora of perceptual distortions, such as seeing faces in clouds or experiencing a cell phone vibration in the absence of true vibration. Likewise, such incorporation of prior knowledge may provide a point of vulnerability that, under the strain of pathological states, may result in qualitatively similar but more extreme biases. These extreme biases, in turn, may constitute the basic phenomenon behind perceptual disturbances such as auditory verbal hallucinations (AVH), or percepts of voices in the absence of speech stimuli. A critical aspect of our proposed framework is thus centred on the notion that the perceptual disturbances in psychosis represent an extreme variant of the perceptual biases that arise naturally from the incorporation of prior knowledge.

Here we present a selective review of the literature on the pathophysiology of psychotic perceptual disturbances, and outline an integrative framework grounded in models of normal perception that has the potential to advance current conceptualizations of these puzzling phenomena. In addition to our focus on perceptual models, we attempt to bridge a number of relevant fields that historically have developed in parallel. Most relevantly, an extensive body of pharmacological and molecular-imaging studies have definitively established a central role of excessive dopamine transmission in the striatum in psychosis[2]. However, previous views of sensory processing in schizophrenia have mostly focused on local cortical circuits and have de-emphasized the contributions to perceptual disturbances in schizophrenia of neuromodulators such as dopamine and of their effects on basal ganglia circuits. These views may be partly influenced by traditional conceptions of sensory systems as isolated brain modules, specialized for passive stimulus detection through local processing, that are immune to learning and cognitive influences[3]. In turn, dopamine transmission in the striatum has mostly been cast in terms of reward-based learning, despite the mounting evidence supporting a more general role for striatal dopamine in perception and cognition beyond reward-related processes. Similarly, striatal function has been classically portrayed in terms of reward learning and action selection, and only recently has its role in perceptual learning and perceptual decision making come into focus. The view of perceptual disturbances we present here instead aims to integrate recent insights into the roles of dopamine and the basal ganglia within the context of current theoretical models of perception.

Positive symptoms
Also known as psychotic symptoms or psychosis; symptoms that are added to the repertoire of usual experiences and that represent a loss of contact with reality (that is, subjective experiences that substantially deviate from what most perceive as objective evidence), including hallucinations and delusions.

Prodromal
Related to the psychosis prodrome or prodromal phase, terms that refer to the phase preceding the development of full-blown symptoms of a psychotic disorder; typically defined by the expression of attenuated forms of positive symptoms.

Given the clinical predominance of auditory disturbances, relative to disturbances in other sensory modalities, in psychotic disorders, we focus mainly on relevant pathways at the intersection between dopamine, basal ganglia circuits and the auditory system. We build upon prior theoretical accounts of psychosis that have capitalized on computational modelling to bridge phenomenology and neurobiology[1,4,5]. We place a special emphasis on the physiological basis for auditory perception and the candidate pathophysiological mechanisms that may lead to altered perceptual experiences in psychosis.

Specifically, some prominent models of psychosis emphasize aberrant reward signalling as the central factor linking dopamine dysfunction and psychotic symptoms[6,7], given the well-established role of dopamine in reward prediction-error signalling[8–10]. However, it is generally accepted that dopamine has additional functions[11]. In this vein, we focus on recent evidence for a role of dopamine in perceptual decision making and learning, and present an alternative view of psychotic perceptual disturbances that emphasizes alterations in perceptual inference.

Phenomenology

Patients with schizophrenia and other psychotic disorders report a wide range of perceptual disturbances. These phenomena, in particular hallucinations, have long been documented and have become emblematic of mental illness from a sociocultural standpoint. Like other positive symptoms of psychosis, perceptual distortions tend to develop progressively during adolescence or early adulthood[12]. They typically begin in attenuated forms, during the prodromal pre-psychosis phase, and gradually evolve into overt distortions and florid hallucinations, with the transition to clinical psychosis.

Illusions and related phenomena. Illusions refer to percepts that represent a distorted version of objective stimuli. Individuals with psychosis tend to experience illusory percepts such as alterations in stimulus intensity (bright colours or loud noises) or distortions in shapes or patterns (for example, faces). In many cases, particularly in the early stages of psychosis, patients experience stimuli as more salient than usual, as if the stimuli feel unfamiliar or strange, and tend to seek interpretations with personal meaning to explain the perceived changes. This is known as 'delusional mood'.

Hallucinations. Hallucinations are false percepts that do not correspond with objective stimuli. Patients with schizophrenia spectrum disorders experience hallucinations in various sensory modalities, including auditory, visual, olfactory, somatosensory and gustatory, but auditory hallucinations are the most prevalent and severe[13,14]. Unlike other symptoms of schizophrenia, the discrete on–off (that is, intermittent) nature of hallucinations allows patients to report the occurrence of hallucinations and makes their clinical assessment highly reliable[15]. Although some auditory hallucinations manifest as low-complexity sounds (for example, stepping sounds), most frequently they consist of voices with higher levels of complexity (from single words to complete sentences or dialogues), such as a running commentary on the patient's behaviour or a conversation between different familiar or unfamiliar voices. The content of hallucinated voices (that is, AVH) is variable, but commonly it exhibits a negative emotional valence (for example, a derogatory tone) and is relevant to the voice-hearer's personal experience[16]. In acute states, AVH are typically interpreted as real voices originating in the external environment (that is, patients have no insight into their internal and dysfunctional origin) and consequently provoke hallucination-related behaviours, such as out-loud verbal responses to the voices (soliloquy) or other congruent behaviours (for example, checking whether the perceived voices are the next-door neighbours' voices heard through the wall). Other forms of psychosis, such as those induced by pro-dopaminergic drugs in Parkinson disease, contrast with schizophrenia in that they feature visual hallucinations more frequently than auditory hallucinations[17], suggesting that auditory pathways are particularly relevant in the pathophysiology of psychotic disorders in the schizophrenia spectrum[18].

Delusions and psychotic syndrome. In psychosis, perceptual disturbances such as hallucinations almost always accompany other psychotic symptoms. They tend to co-occur and evolve in parallel with delusions, which are defined as tenacious false beliefs maintained in the face of contradictory evidence. This clustering of hallucinations with delusions is stronger than the correlation of either of these individual symptoms with other symptoms of schizophrenia (for example, amotivation or memory impairments), and thus defines a distinct clinical construct called the 'psychotic syndrome'[19]. Yet, seldom do patients present only or predominantly with hallucinations in the absence of delusions, or vice versa. These clinical observations thus suggest the existence of a common pathophysiological mechanism underlying the psychotic syndrome, as well as symptom-specific pathways.

Alleviating and aggravating factors. Ample evidence from randomized controlled clinical trials has demonstrated that the psychotic syndrome of schizophrenia, including hallucinations, responds to various antipsychotic medications that act by blocking D2 receptors[20], a type of dopamine receptor that is most abundant in the striatum[21,22]. Even though psychotic symptoms can improve rapidly with antipsychotic medications[23], gradual benefits typically build up over time, with a substantial proportion of responders achieving remission only after several weeks of treatment[24]. Importantly, these clinical improvements are specific to positive or psychotic symptoms rather than to other symptoms of schizophrenia, hinting at a specific involvement of striatal stimulation of D2 receptors by dopamine in psychosis, a notion solidified by the molecular-imaging studies reviewed below. Other than cognitive behavioural therapy[25] and transcranial magnetic stimulation applied over left temporo–parietal cortex[26], no other treatments have consistently shown improvements for perceptual disturbances in psychosis. In turn, environmental stressors, several drugs of abuse (for example, ketamine,

cannabis and cocaine) and pro-dopaminergic drugs such as amphetamine can trigger or worsen perceptual disturbances and other psychotic symptoms. Given that stress and most of the drugs of abuse that worsen psychosis increase dopamine transmission[27,28], as well as the causal evidence that D2-receptor-blocking drugs specifically improve psychosis[20], these observations provide strong support for a central role of dopamine in the pathophysiology of psychosis and its perceptual disturbances, as well as for the psychotic syndrome as a distinct neurobiological construct. This does not imply that other factors (for example, glutamatergic or GABAergic signalling) are irrelevant in psychosis (for recent reviews of the role of glutamatergic and excitation/inhibition imbalance in schizophrenia, which is beyond the scope of the present review, see[29,30]). Indeed, about one third of hallucinating patients do not experience substantial improvements with anti-dopaminergic medication[31,32], particularly those whose psychotic symptoms have remained untreated for longer periods of time[33]. However, a variety of non-dopaminergic drugs tested in schizophrenia so far have failed to consistently ameliorate perceptual disturbances in psychosis.

Cognitive mechanisms

From a computational perspective, perceptual judgments or decisions can be seen as a process aimed at maximizing the correct identification and classification of sensory inputs along multiple dimensions, or at minimizing perceptual-judgment errors. In this section, we focus on candidate algorithmic processes that neural systems are thought to deploy in solving this computational problem, drawing from a large body of literature in perceptual decision making, and discuss algorithmic alterations proposed to explain perceptual disturbances. In the next section, and in line with Marr's three levels of analysis, we discuss the neurobiological implementation of these processes in brain circuits, as well as abnormalities in these circuits in psychosis.

Two main concepts are central to current models of perception[34]. The first is that to minimize errors, optimal perceptual decisions need to overcome uncertainty in both the external stimuli (for example, low-intensity sounds in the context of background noise) and the internal representations of those stimuli (for example, intrinsic variability in the firing of neurons encoding sensory information). The second is that, in naturalistic environments, stimuli rarely occur in isolation; they typically occur in the context of other stimuli, which concurrently or consecutively predict each other. For this reason, leveraging prior knowledge about the contextual relationships between stimuli can be advantageous for resolving sensory uncertainty and minimizing errors in perceptual decisions.

Signal detection theory. Consider a situation in which a commuter is waiting for her train to arrive in a noisy subway station after a long day of work. She hears an ambiguous sound stimulus that resembles her name being called out, and quickly needs to decide whether someone has actually called her name or not — that is, whether the stimulus comprised signal or only noise, respectively. In this scenario (FIG. 1), signal detection theory posits that decisions respect the relative likelihood of the evidence provided by the internal stimulus representations for two alternative propositions (signal present or absent), a quantity typically expressed as a log-likelihood ratio:

$$\mathrm{LLR} = \log \frac{P(\text{evidence}|\text{signal present})}{P(\text{evidence}|\text{signal absent})}.$$

Under this framework, the listener will judge signal as being present or absent in the stimulus with equal probabilities if the strength of the evidence for each of the alternatives is equivalent (LLR = 0), or with different probabilities otherwise (judging the signal as more likely to be present if LLR > 0, or to be absent if LLR < 0). In other words, the listener will exhibit an unbiased response if the stimulus itself is completely ambiguous or if the internal representation of the stimulus is very equivocal, rendering sensory uncertainty maximal. Shifts of the response criterion along the evidence scale determine a trade-off between hits and misses — correct versus incorrect detection of a present signal, respectively — and between correct rejections and false alarms — correct versus incorrect detection of an absent signal, respectively (FIG. 1).

Bayesian inference. As we mentioned above, contextual information can be used to resolve uncertainty. Prior expectations derived from the context are incorporated in the decision process following Bayesian principles, in which prior knowledge provides an estimate for the prior probability of, or prior beliefs about, the signal. The listener in our example may have received a friend's text message indicating that he was on his way to the subway station, a scenario that would increase the probability that her friend did call out her name. In this way, prior beliefs should bias responses towards the more likely alternative, as psychophysics research has established[34]. Critically, this response bias favouring the expected alternative is both optimal, in that it minimizes errors overall, and induces a pattern of asymmetric errors towards the expected alternative. In the case of a bias towards the signal being present, the optimal response will translate into increases in both hits and false alarms (FIG. 1). (It is worth noting that the expectations related to prior knowledge that we discuss here may originate from recent information samples collected across different sensory modalities that are maintained in memory, or from longer-term information obtained from implicit memories.)

Bayesian inference can also be seen as a prescription for optimally combining two or more sources of information under uncertainty, a concept referred to as 'reliability weighting'. In the context of perception, the key information sources are the prior beliefs or expectations (source 1) and the stimulus likelihood (source 2), and they are combined to produce a posterior — or updated — belief, which corresponds to the resulting percept. This prescription for optimal 'belief updating' is rather intuitive, as it simply consists of weighting more heavily the source of information that

Marr's three levels of analysis
A framework whereby information-processing systems can be understood at three distinct, complementary levels: computational (the problem that is solved), algorithmic (what representations and processes are used to solve this problem) and implementational (the physical and biological substrates through which the solution is realized).

Bayesian inference
A statistical algorithm for probabilistic estimation that relies on the optimal combination of prior knowledge and new data.

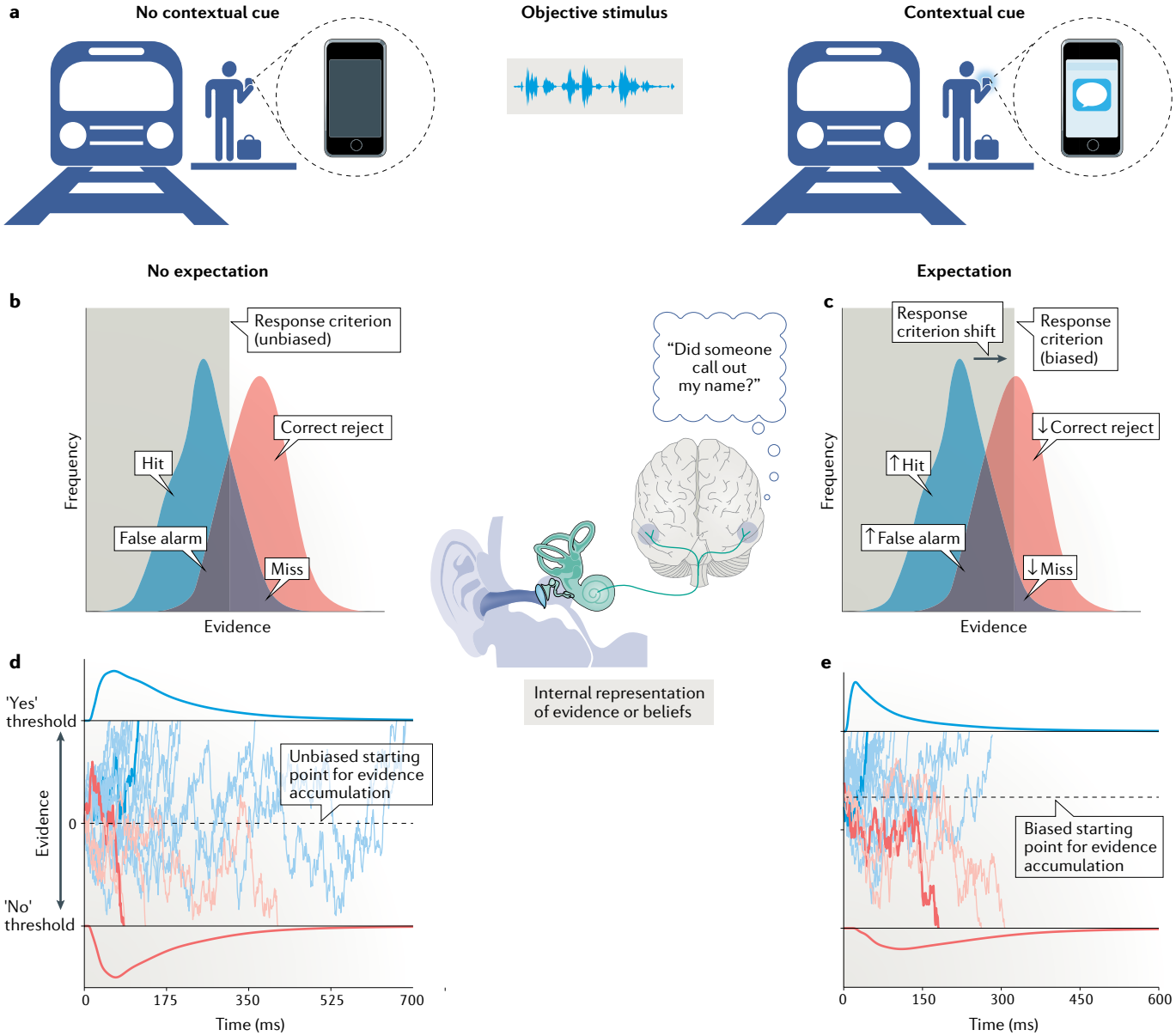


Fig. 1 | **Examples of perceptual decision making without and with an expectation bias that is contextually appropriate and adaptive. a** | The effect of expectation bias can be understood in the context of a commuter waiting for a train and hearing an ambiguous sound (an 'external objective stimulus', top centre) that resembles her name being called out. In this scenario, the perceptual decision (the commuter deciding she has heard her name called out — a 'yes' decision — or not — a 'no' decision) varies depending on expectations that reflect the internal representations of the evidence for each of two alternative propositions (signal [her name actually being called out] present or absent). **b,c** | The external objective stimulus is processed through the cochlea (in the inner ear), which responds to sound vibrations and translates them into neuronal impulses in the auditory pathways (brainstem nuclei, auditory thalamus and auditory cortex; part **b**, right panel), which represent the internal evidence associated with the external objective stimulus, here represented on the *x*-axis of the graphs. The blue histograms indicate the likelihood of the internal evidence given the presence of a signal (the commuter's name actually being called out), and red histograms indicate the likelihood of the internal evidence given the absence of a signal (the commuter's name not being called out and the stimulus comprising only noise). If the commuter had received a text message predictive of signal (part **a**, right), this would constitute a contextual cue that would shift her response criterion towards more liberally accepting evidence for a 'yes' decision (that is, the edge of the grey shaded area shifts along the *x*-axis). This would bias her response towards correctly identifying the presence of signal (a 'hit', or the area of the blue histogram to the left of the edge of the grey shaded area), but also increases the likelihood of her incorrectly identifying an absent signal (part **c**; that is, increased rate of a 'false alarm', or the area of the red histogram to the left of the edge of the grey shaded area). Conversely, in the absence of such a cue, the response criterion would be unbiased, and the likelihood of both hits and false alarms would be comparatively reduced (part **b**, left). **d,e** | Dynamics of how an internal evidence accumulation process transpires over time, according to the drift diffusion model. The internal evidence is accumulated via a noise-corrupted process that terminates when the evidence for one alternative (signal present or absent) hits the threshold associated with that alternative (yielding a 'yes' or 'no' decision, respectively). Example trajectories for evidence accumulation leading to 'yes' (blue) and 'no' (red) decisions are highlighted against a background of other possible trajectories (faded colours), given the same model parameters. The marginal distributions above and below the graphs represent the distribution of termination times associated with each of the responses (which vary due to the noisy accumulation process, as illustrated by the different possible trajectories of evidence accumulation). The evidence accumulation process has an unbiased (part **d**) or a biased (part**e**) starting point, associated with the absence or presence of prior expectations, respectively.

is more reliable or less uncertain. It follows that prior beliefs should be weighted more strongly in situations in which the relative likelihood of the evidence provided by the stimulus is equivocal — that is, under high sensory uncertainty — or in situations in which prior beliefs are more certain. In our example, the commuter may be more likely to judge that someone is definitely calling out her name if the sound stimulus was highly ambiguous and she was certain she would meet her friend at the station.

In real-life settings, information is rarely accrued all at once. More often, it is acquired bit by bit, through sequential samples of information. For instance, the commuter in our example would likely search her surroundings for additional information, such as evidence of someone approaching her, to aid in her judgement. Under sequential sampling, Bayesian inference iteratively produces updated beliefs that incorporate the new stimulus likelihood from each sample, thereby capturing a process of progressive integration of evidence over time. This process is approximated by 'bounded evidence accumulation' models. One such model, the drift diffusion model (DDM), describes a noisy process of sensory-evidence accumulation that evolves from a starting point until one of two symmetric decision bounds, or thresholds, is reached. Evidence accumulation is controlled by a drift rate parameter that scales the sensory evidence via multiplicative gain. Within this sequential framework, expectations derived from prior knowledge (in the example, the friend's text message; expectations could also derive from a variety of other sensory or mnemonic sources) can be incorporated into this process by shifting the starting point (and, partly, by increasing the gain of the evidence)[34].

Predictive coding. Although they are mathematically related to evidence accumulation models[35], predictive-coding models emphasize a neural architecture of reciprocal message-passing up and down different levels of a sensory-processing hierarchy in which higher levels are capable of more complex representations[36,37]. Evidence accumulation in these models is cast as a belief-updating process aimed at forming an explanatory model of the causes of sensory inputs. To this end, the probability distributions representing beliefs about different possible causes are iteratively refined via message-passing, to settle on the most plausible explanations for sensory inputs. The basic computational motif, which repeats at each level of the hierarchy, consists of the integration of top-down messages, encoding predictions based on prior knowledge, with bottom-up messages, which here encode errors in those predictions (prediction errors). These models can be thought of as a sensory analogue of hierarchical reinforcement-learning models, in which the key teaching signal prompting learning, or belief updates, is a sensory rather than a reward prediction error. To implement a form of reliability weighting that modulates the impact of errors on beliefs, sensory prediction errors are weighted by the respective reliabilities of the top-down and bottom-up messages, such that weighted prediction errors ultimately determine the degree of belief updating.

To sum up, although they are distinct, the different frameworks reviewed above share the common assumption that, during perceptual decisions, prior expectations are integrated with new sensory information through a process of evidence accumulation or belief updating.

Algorithmic models of hallucinations. Friston proposed an algorithmic model for hallucinations that consists of an excessive bias towards prior expectations, itself resulting from overly certain prior beliefs about sensory states[38] (FIG. 2). In other words, he proposed that hallucinations are percepts driven mostly by strong expectations. The basic premise of this model is analogous to longstanding Bayesian models of context-driven laboratory illusions, but Friston's model suggests that expectation biases could also explain de novo percepts arising in the absence of sensory stimuli. Although it may be somewhat counterintuitive to argue that perceptual inference may explain false, and not just distorted, percepts, it is important to bear in mind that the likelihood in signal detection and Bayesian models refers to the internal representation of the sensory stimulus, rather than to the external stimulus itself. Because sensory neurons always exhibit some level of spontaneous activity (that is, they have non-zero firing rates in the absence of their preferred stimuli), this means that internally there is never a complete absence of sensory evidence, which blurs the line between illusions and hallucinations. This is an appealing aspect of this model, in that it portrays hallucinations as extreme illusions, which fits with the notion of a perceptual-disturbance continuum as well as with the usual progression from perceptual distortions to overt hallucinations during the psychosis prodrome. Critically, this model suggests that the same neural machinery that facilitates adaptive biases in perceptual inference can provide a point of vulnerability that, under pathological states, results in more extreme and maladaptive biases. More recent extensions have framed this basic model in the context of action affordance[39] and hierarchical predictive coding[4]. Hierarchical models further suggest complex interactions between certainty about higher-level prior beliefs and certainty about sensory evidence at lower processing levels. A particularly appealing notion in this work[4] is that relatively increased certainty about higher-level prior beliefs compared with lower-level sensory evidence, and the ensuing false percepts driven by expectation biases, may itself stem from compensation deriving from diametrically opposite alterations: patients with schizophrenia may exhibit a more stable, trait-like alteration, consisting of a relatively decreased certainty about higher-level prior beliefs that may be compensated for by decreased certainty about the sensory evidence at lower levels, a compensation that may ultimately drive a state of hallucinatory psychosis. See BOX 1 for other theories and models related to perceptual disturbances in psychosis.

Behavioural studies on hallucinations. Although alterations in a wide array of perceptual processes have been reported in individuals with schizophrenia spectrum disorders, only a handful of these processes seem to correlate with the severity of hallucinations or other

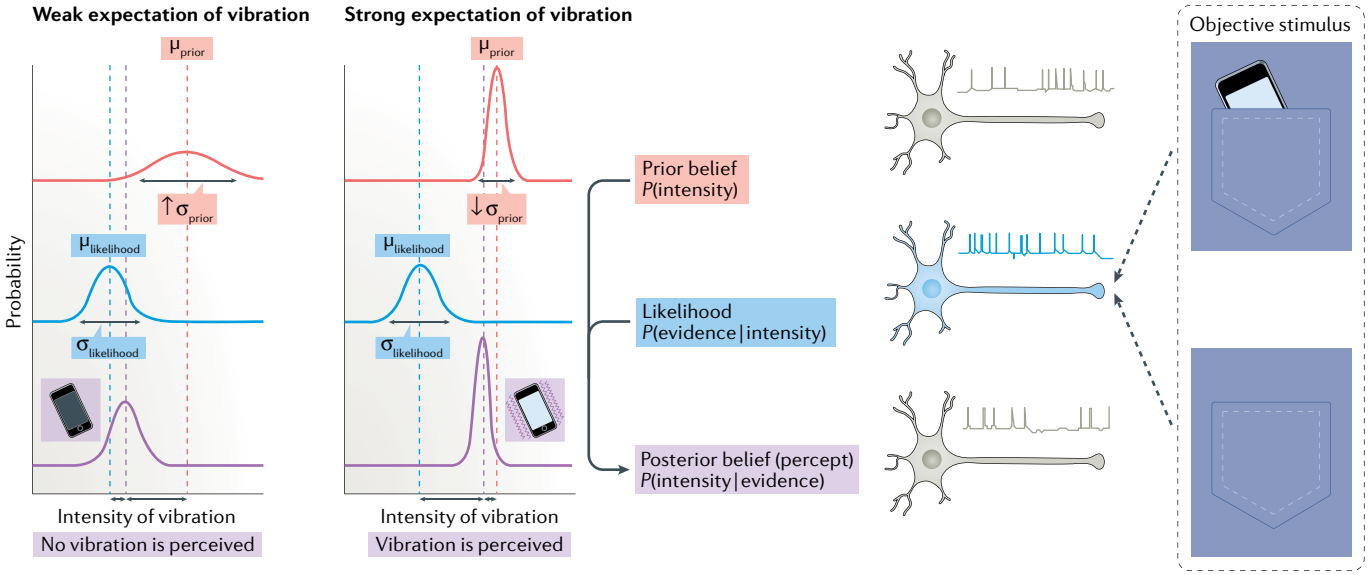


Fig. 2 | **Bayesian model of perceptual inference, illustrating the expectation biases related to common perceptual distortions and pathological hallucinations.** Illustration of the 'phantom pocket vibration' phenomenon, a commonly experienced phenomenon that can be thought of as a hallucination. An external objective stimulus (either the mild friction produced by a cell phone slightly shifting in one's pocket (top right) or no vibrotactile stimulus at all (bottom right)) is associated with (arrows) an internal sensory representation (evidence) in the brain — for instance, a neuron encoding such evidence by modulations of its firing rate. Critically, this firing rate, and thus the internal evidence, is non-zero even in the complete absence of an external objective stimulus. For a given level of internal evidence, the resulting percept depends on the strength (certainty or precision) of prior beliefs — here, how strong the expectation of a high-intensity cell phone vibration (for example, caused by a phone call) is. The belief-updating process that gives rise to the resulting, final percept is illustrated as a curved arrow between boxes of changing colours, as the prior belief (red, top) is combined with a likelihood distribution (blue, middle), giving rise to the posterior belief (purple, bottom). A weak expectation (left graph) is characterized by a prior belief with high uncertainty (high σ_{prior}); a strong expectation (right graph), by one with low variance (low σ_{prior}). Otherwise, here the expected amount of vibration (μ_{prior}) and the likelihood distribution ($\mu_{likelihood}$ and $\sigma_{likelihood}$) are, on average, the same for the left and right scenarios. The resulting percept is represented by a posterior belief. This posterior belief differs substantially between the two scenarios: under a strong expectation, vibration is perceived subjectively (right) — the posterior belief is close to the prior belief and far from the likelihood — and under a weak expectation, it is not (left) — the posterior belief is far from the prior belief and close to the likelihood. Thus, a false cell phone vibration can be perceived in the presence of a slight friction (illusion) or in the complete absence of any vibrotactile stimulus (hallucination), given strong expectations of the phone ringing (that is, a more certain or precise expectation about an upcoming high-intensity vibration).

perceptual disturbances. Within the signal detection framework, hallucinations are by definition analogous to false alarms. In signal detection tasks using auditory stimuli embedded in noise, patients who experience auditory hallucinations tend to exhibit biased responses towards reporting speech and other auditory signals (that is, increased hits and false alarms), either at baseline[40–43] or in response to conditioned cues[44].

The few studies that have used the DDM framework in psychosis provide preliminary evidence for algorithmic alterations in evidence accumulation. In Parkinson disease, impaired gain of evidence accumulation during a visual task has been linked to visual hallucinations[45]. In schizophrenia, preliminary evidence suggests alterations in the starting point of the accumulation process in some tasks[46], which points to differences in the integration of prior expectations, although the relationship of this parameter to hallucinations is unclear.

Excessive influence of prior expectations on the perceptual sensitivity of ambiguous visual stimuli, but not on response bias, was shown in individuals at risk for psychosis[47]. Importantly, this effect was specifically related to perceptual disturbances in a subclinical population. In a tone duration reproduction task, unmedicated psychotic patients with severe hallucinations also exhibited excessive influence of prior expectations, derived from context tones, on their subjective perception of target tones with invariant duration[48]. This was specific to hallucinations, even compared with other psychotic symptoms, and was accompanied by an inability to adjust the weight of expectations in more variable (uncertain) contexts, suggesting alterations in reliability weighting due to the excessive weight of prior expectations. In line with this finding, computational modelling of behaviour during a Pavlovian conditioning paradigm showed that hallucinators, with or without a diagnosis of schizophrenia, exhibited increased expectations for hearing tones predicted by a light cue, as well as an excessive weight of prior expectations[44].

In summary, the reviewed research is compatible with a model of hallucinations wherein this symptom results from undue influences of prior expectations on perception. Some of this work is consistent with the clinically inspired notion of a symptom-specific mechanism

Box 1 | **Models of perceptual disturbances in psychosis**

Bottom-up theories. It has been proposed that disrupted sensory signals arising at early levels of sensory processing may be responsible for perceptual disturbances[127]. This notion partially rests on the assumption that 'pure' sensory inputs, unaltered by top-down modulation, are anatomically dissociable at early stages of sensory processing, in contrast with the abundance of feedback connections and top-down modulatory effects at these stages[128,129]. A potential way of reconciling the bottom-up and top-down views is that excessive sensory uncertainty may secondarily result in overreliance on prior expectations[47].

Sensory-gating theories. Sensory gating implies an active (for example, attentional) process through which sensory systems filter out irrelevant sensory inputs in favour of relevant ones[130]. A disruption in this filter leads to flooding of sensory systems with irrelevant sensory information.

Salience misattribution theories. Abnormalities in dopamine signals have been proposed to alter the motivational or incentive salience of stimuli[6,7]. Objectively neutral stimuli (including internal signals such as memories or inner speech) become imbued with abnormal significance, driving abnormal interpretations. Although these theories do not clearly explain false percepts per se and lack a formal operationalization[131], recent views have incorporated elements of salience misattribution into top-down models[5].

Predictive-coding, Bayesian inference and related top-down models. These models generally cast sensory disturbances as an extreme version of the perceptual biases that arise naturally in optimal models of perception[4,5,126], accounting for expectation biases. Corollary-discharge models are a special case, in which hallucinations represent inner speech misattributed to an external agent due to failures in the attenuation of (bottom-up) sensory inputs through efference copies of self-generated motor commands[132–134] (which can be thought of as a top-down predictive signal). Other models implement hierarchical inference through different belief propagation algorithms. Specifically, the circular-inference model posits alterations in biophysically realistic interneurons that lead to information reverberation in schizophrenia[135].

for perceptual disturbances, as the relevant findings supporting the said mechanism seem more closely related to hallucinations than to delusions or other symptoms. Also notably, the observed correlations between the severity of hallucinations outside the laboratory and laboratory measures of illusory percepts in these studies support the continuum implied by Friston's model of hallucinations. A possible explanation for excessive expectation biases in psychotic perceptual disturbances is the faulty incorporation of contextual uncertainty into prior expectations[48], or abnormal beliefs about contextual volatility[44], with overweighting of prior expectations even in situations in which high contextual uncertainty renders expectations uninformative. So how might this mechanism be implemented in the brain?

Neurobiological implementation

In this section we focus mainly on the perceptual processes relevant to auditory disturbances. However, we first provide a brief description of relevant findings during visual decision making (BOX 2, FIG. 3). This literature suggests that a network of associative cortico–subcortical regions implements a form of evidence accumulation consistent with the DDM[49]. Importantly, the integration of prior expectations that is central to perceptual inference may be implemented in this network through biases in the baseline (pre-stimulus) activity of neurons tuned to more likely stimulus features, which provide sensory inputs to downstream associative regions that instantiate evidence accumulation[34]. In what follows, we review evidence suggesting that a network of associative regions involved in auditory processes, including the associative auditory cortex and a downstream region in the associative striatum, is relevant to perceptual inference, a circuit we suggest may be central to perceptual disturbances in psychosis. In line with this general notion, it should be noted that current models of consciousness generally posit that associative regions downstream from sensory cortex — rather than sensory cortex alone — are involved in conscious perception. These models invoke recurrent network activity involving associative regions during conscious perceptual reports[50,51], consistent with primate data showing that associative downstream regions are involved in perceptual reports (including involvement during false alarms) and exhibit baseline biases in neural activity associated with response bias[52–54].

Auditory system. Cochlear neurons projecting to brainstem nuclei represent the first stage of processing in auditory pathways (FIG. 1). These are followed by neurons in the inferior colliculus, medial geniculate nucleus of the thalamus, and primary auditory cortex, which in humans is located in the transverse temporal, or Heschl's, gyrus. This primary or core area of the auditory cortex, which, like the processing levels preceding it, exhibits sound-frequency tuning or tonotopy, projects to surrounding areas in belt and parabelt areas of non-primary auditory cortex[55]. The latter areas comprise secondary and associative regions of auditory cortex that reside along the superior temporal gyrus and exhibit more complex response patterns, including representations of human speech and other species-specific vocalizations, as well as emotional (for example, prosody) and non-emotional features of speech[56]. Classic studies in patients with epilepsy have shown that, upon electrical stimulation of associative auditory cortex, patients without a history of psychosis reported hearing voices (for example, familiar voices) and other complex sounds[57]. Other patients with a rare type of partial epilepsy affecting associative auditory cortex[58] often present with ictal auditory hallucinations, including AVH[59]. In primates, associative auditory cortex has distinct efferent projections that distinguish it from primary auditory cortex, including its direct projections to prefrontal cortex[55] and associative striatum[60], which are absent or very limited for primary auditory cortex (FIG. 4). The afferent projections for primary versus non-primary auditory cortex also differ, with the latter receiving more projections from parts of the medial geniculate complex other than its primary relay nucleus, as well as from other thalamic nuclei, including the medial pulvinar[61]. In addition to these feedforward connections, each level of processing receives dense feedback projections[55] that are thought to be critical for the top-down regulation of perceptual processes, consistent with the hierarchical architecture of predictive-coding models. Some feedback loops involve direct connections from lateral prefrontal cortex back to associative auditory cortex[55], and others may follow an indirect route via basal ganglia–thalamo–cortical circuits. Although the latter are not fully characterized for associative auditory cortex, similar basal ganglia–thalamo–cortical circuits have been described for primate inferior temporal cortex that are likely relevant to

Inner speech
A person's inner dialogue, expressed as a silent conscious stream of thoughts in a coherent linguistic form.

Box 2 | **Neurobiological implementation of visual perceptual decisions**

A well-studied neural circuit for visual–perceptual decision making is the one that comprises the middle temporal (MT) area (also known as visual area V5) and the lateral intraparietal cortex (LIP) (see FIG. 3). The physiology of this circuit has been extensively and meticulously studied in monkeys performing visual-motion tasks, in particular the random-dot-motion task[49,136]. Transient changes in the activity of MT neurons encode the momentary evidence for motion in a specific direction. LIP, a downstream region that receives direct projections from MT, instead exhibits persistent activity that supports an evidence accumulation process analogous to the DDM. Specifically, LIP neurons exhibit ramping activity that scales with the strength of sensory evidence (typically manipulated through changes in motion coherence). Critically, manipulations of the prior probability of motion direction are associated with increased baseline (pre-stimulus) activity in MT neurons tuned to the more likely direction[137], as well as increased baseline activity in LIP[138]. Consistent with this finding, human fMRI studies have shown biased patterns of baseline activity in human MT in this scenario[139]. Notably, activity profiles similar to that in LIP have been demonstrated in other associative regions downstream from LIP, including frontal regions and associative striatum (see the main text), which suggests that evidence accumulation may depend on a widespread network of associative cortico–subcortical regions. Thus, the integration of prior expectations may be implemented in a circuit supporting evidence accumulation through biases in the baseline activity of neurons tuned to the stimulus features that are more likely to occur in a given context[34]. While the drivers for biased baseline activity are not fully understood, some evidence is consistent with a prefrontal top-down enhancement of activity in lower-level regions encoding the expected sensory stimuli[140].

higher-order visual processes, and perhaps contribute to visual hallucinations[62]. Similarly, outputs from rodent basal ganglia to the posterior thalamic nuclei that project back to the associative auditory cortical regions[63], including the same cortical regions that send forward projections to the associative striatum, suggest the existence of a closed subcortical loop that is likely relevant for the modulation of higher-order auditory processes and is an important candidate circuit for auditory hallucinations (FIG. 4).

Auditory alterations in hallucinations. In patients with psychosis, symptom-capture studies using fMRI have captured a relatively consistent pattern of activity associated with auditory hallucinations. During scanning periods in which patients endorse hallucinatory auditory–verbal percepts (in the absence of true speech stimuli; AVH), they generally exhibit increased neural activation in associative auditory cortex, along the superior temporal gyrus[64] and in the proximity of Wernicke's area[65]. Some of the studies have also reported changes in activity in these regions at baseline[64,66,67], in the absence of concurrent hallucinations. This evidence from clinical neuroimaging studies suggests an involvement of associative auditory cortex — a region involved in the perception of speech and other higher-order auditory processes — in AVH in psychosis. Furthermore, it raises the intriguing possibility that baseline increases in activity in speech-selective neuronal ensembles may represent a baseline bias similar to that implementing biases towards expected stimuli (BOX 2).

Dopamine and basal ganglia circuits

The striatum is centrally involved in decision making and action selection. Key to these striatal functions is a gating mechanism that operates within basal ganglia circuits to facilitate or impede thalamo–cortical output of the striatal inputs to these circuits[68,69]. Two distinct pathways are critical for gating: a direct (Go) pathway controlled by striatal medium-spiny neurons (MSNs) expressing dopamine D1 receptors (D1-MSNs), the activation of which induces behavioural activation by facilitating thalamo–cortical outputs, and an indirect (NoGo) pathway controlled by striatal MSNs expressing dopamine D2 receptors (D2-MSNs), the activation of which induces behavioural inhibition by impeding thalamo–cortical outputs. Dopamine neurons located in the midbrain — mostly in the ventral tegmental area (VTA) and substantia nigra, pars compacta (SNc) — send dense projections to the striatum following a topographic gradient[2]: the VTA projects to the ventral striatum via the mesolimbic pathway, and different tiers within SNc project to different aspects of the dorsal striatum, including associative and sensorimotor striatum, via nigrostriatal pathways. The striatum also receives cortical projection neurons that synapse onto MSNs, following a topographic organization into partially segregated cortico–striatal–thalamo–cortical loops[70,71]. In the associative striatum, cortico–striatal inputs from diverse associative cortical regions additionally overlap within small convergence zones that abound in the anterior caudate[72].

Phasic dopamine release into the striatum plays a central role in reinforcement learning by progressively shaping this circuit on the basis of experience. At the cellular level, dopamine transients modulate plasticity in cortico–striatal synapses[73,74], inducing long-term potentiation of D1-MSNs and long-term depression of D2-MSNs. In addition to these progressive changes in plasticity, dopamine transients modulate the excitability of this circuit instantaneously during decision making[69]. Dopamine increases excitability in D1-MSNs and decreases excitability in D2-MSNs, effectively modulating the gain of striatal inputs.

More generally, dopamine signalling has long been thought to be relevant to sensory processing[75], but only recent work has produced unambiguous evidence for its role in sensory learning, beyond its involvement in reward-based learning. Uncued visual and auditory stimuli are known to elicit burst activity in dopamine neurons[76]. Phasic responses in these neurons have also been shown to reflect perceptual uncertainty in a way that cannot be explained simply by reward expectations[77,78]. Furthermore, subjective percepts in the absence of sensory stimuli are associated with phasic dopamine signals in non-human primates[77,79]. Most critically, a recent study in rodents showed that dopamine transients are sufficient and necessary for learning stimulus–stimulus associations, which depend on sensory prediction errors, even when those associations occur between neutral sensory stimuli unrelated to reward[80]. Measuring dopamine transients directly, another rodent study revealed that whereas dopamine signals in the ventral striatum conformed with reward prediction errors, dopamine signals in parts of the dorsal striatum, in contrast, conformed with a type of sensory prediction error encoding stimulus unpredictability, independent of reward value[81]. Altogether, these studies therefore suggest that in addition to the established role of dopamine in reward prediction-error signalling, dopamine transients — at least in parts of the nigrostriatal system

Gating
A process by which the passage of information is actively controlled, thereby facilitating or impeding information flow.

— may encode reward-unrelated information relevant to perceptual inference.

In humans, some work with pharmacological manipulations using pro-dopaminergic drugs also suggests a role for dopamine in perceptual decision making. A study of Parkinson disease showed that patients on L-DOPA, compared to when they were off L-DOPA, exhibited decreased gain of evidence accumulation (that is, a reduced drift rate in the DDM) while they made perceptual decisions on a random-dot-motion

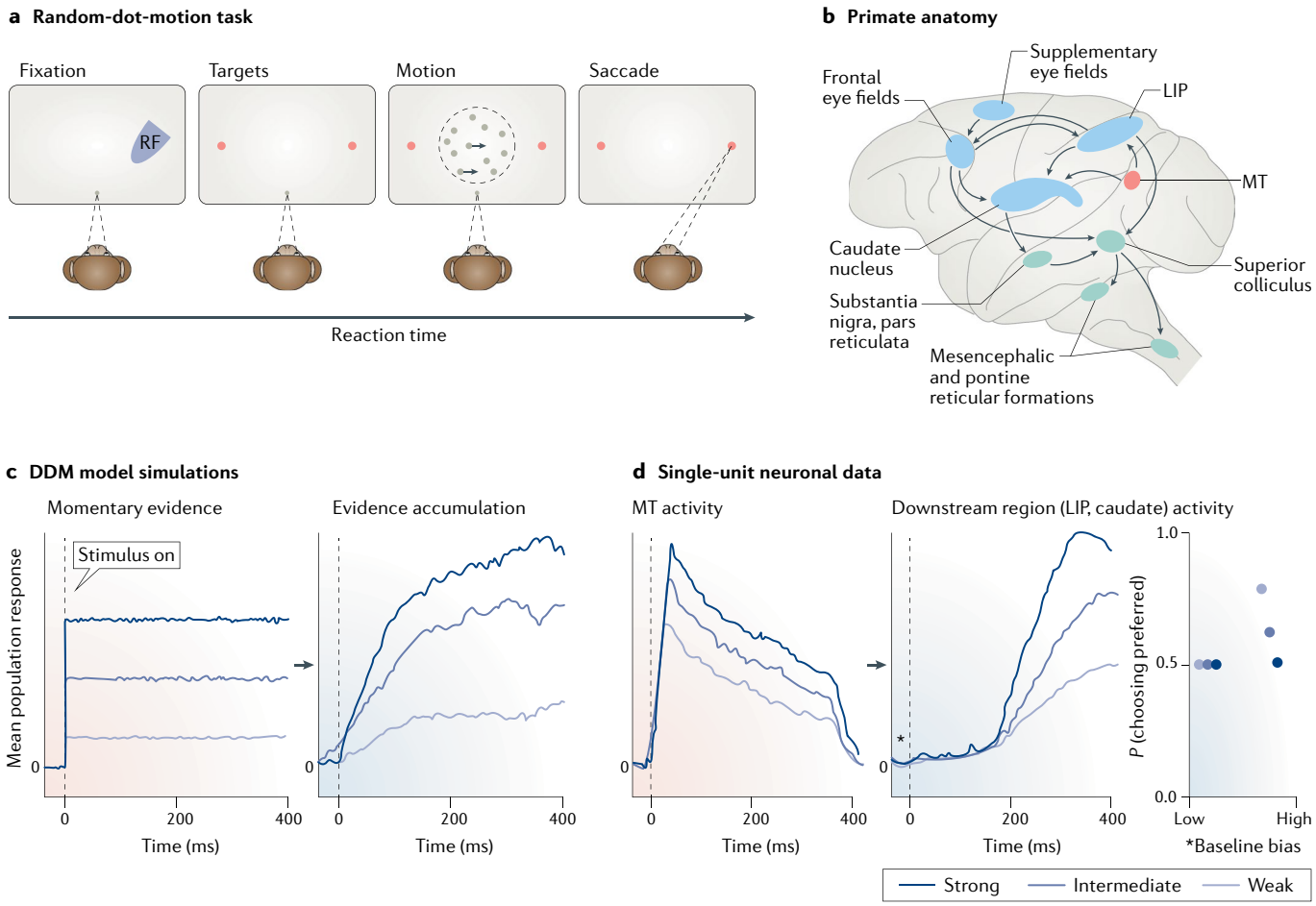
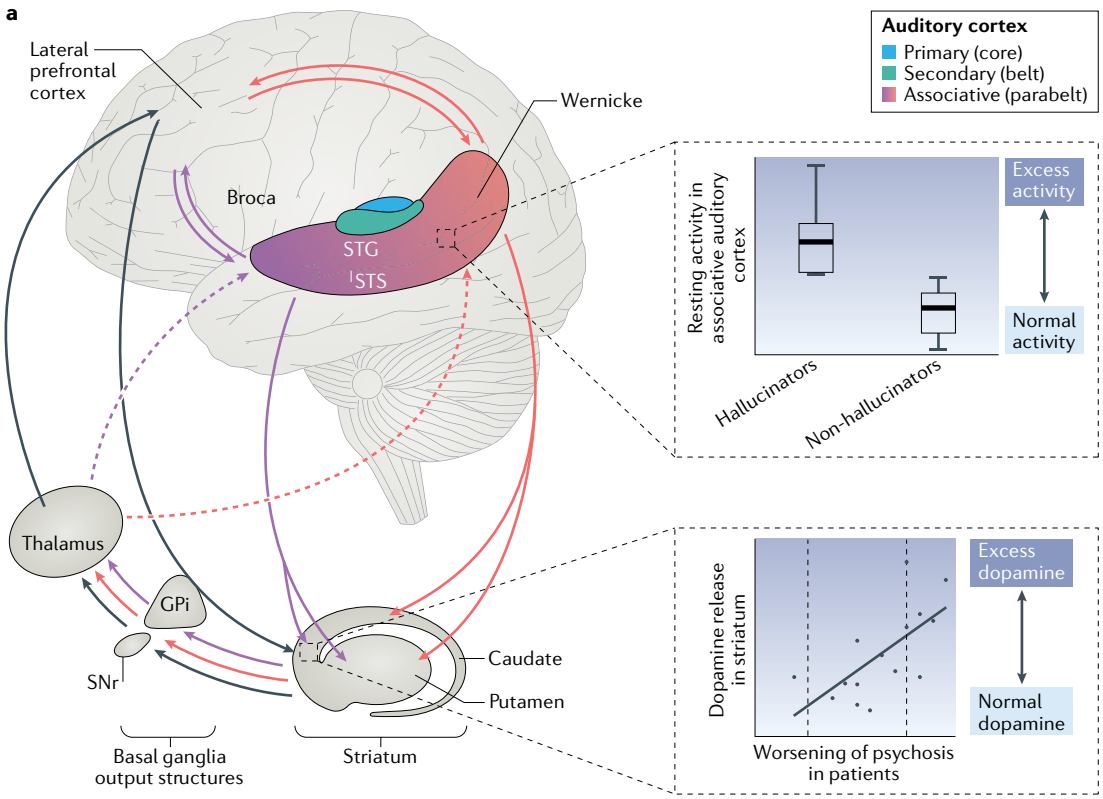


Fig. 3 | **Neural implementation of perceptual decision making in the saccade generation system.** Illustration of the drift diffusion model (DDM) of evidence accumulation and its neural correlates during visual-motion discrimination. **a** | The stimulus in the task consists of a cloud of moving dots, and the goal is to decide accurately what is the predominant motion direction. The monkey fixates, two targets are presented (one in the receptive field, RF, of a neuron being recorded, and one opposite it), the motion stimulus is presented, and the monkey decides by making a saccade to one of the targets. Motion strength is varied as the fraction of dots moving in the same direction (coherence). **b** | Primate anatomy, illustrating the brain regions involved in the saccade generation system. The middle temporal area (MT, pink) is a higher-order visual region with neurons tuned to motion direction and that projects to downstream, associative regions (blue) involved in integrating evidence (including lateral intraparietal cortex (LIP) and the caudate, which is part of the associative striatum), which in turn project to regions (green) directly responsible for saccades through their action on the eye muscles. Arrowheads indicate the direction of projections between upstream and downstream regions. **c** | DDM model simulations of an input functional unit — the model equivalent of a neuronal ensemble — representing momentary evidence through changes in the mean population response at different strengths of sensory evidence associated with the stimulus (weak, intermediate or strong motion coherence; left panel) and of a downstream functional unit accumulating this sensory evidence over time (right panel). The grey dashed line indicates the time of stimulus presentation (stimulus on). Evidence accumulation manifests as a ramping mean population response, with a steeper slope for stronger evidence. **d** | Representative single-unit neuronal data — mean population responses (that is, firing rates) for actual neuronal ensembles — from monkey physiology experiments indicate that visual area MT (left) represents momentary evidence, whereas a network of regions including LIP and caudate (associative striatum) exhibits a profile consistent with evidence accumulation (middle) during stimulus presentation. Only activity for stimuli consistent with the neurons' preferred direction — that is, its RF — is depicted, for simplicity. The rightmost panel shows that greater biases in baseline, pre-stimulus activity in both LIP and associative striatum, indicated by an asterisk in the period immediately preceding stimulus presentation, have stronger effects on decisions (a higher probability of choosing the direction that motion-direction-tuned neurons are tuned to — that is, 'choosing preferred'). This is particularly evident for weaker stimuli, consistent with the Bayesian notion that expectation biases should have a stronger impact in conditions of higher sensory uncertainty. Panel **a** is adapted with permission of Annual Reviews from 'The neural basis of decision making', Gold & Shadlen, *Annu. Rev. Neurosci.* **30**, 535–574 (2007)[49], permission conveyed through Copyright Clearance Center. Panel **c** is adapted from REF.[49]. Panel **d** is adapted with permission of *Journal of Neuroscience* from 'Caudate encodes multiple computations for perceptual decisions', Ding & Gold, *J. Neurosci.* **30**, 15747–15759 (2010)[109], permission conveyed through Copyright Clearance Center, Inc.; and from REF.[155], Springer Nature Limited.



task (BOX 3), suggesting that relatively increased dopamine levels could lead to decreased reliance on sensory evidence. Another study using a force-matching paradigm showed that higher L-DOPA dosage was associated with increased reliance on sensorimotor predictions at the expense of decreased reliance on sensory evidence[82], consistent with the previous study. However, in another study, L-DOPA treatment in patients with Parkinson disease shifted perception towards increasing the weight of sensory evidence relative to prior beliefs during a visual–spatial inference task[83]. Other work in healthy individuals showed that increased striatal dopamine correlated with decreased belief-updating signals related to task-relevant information in the striatum[84]. These results are thus consistent with a role for dopamine in perceptual inference and suggest that dopamine could have opposing effects on different cortico–striatal systems[82].

Dopaminergic alterations in psychosis. In schizophrenia, dopamine dysfunction plays a central role in the manifestation and treatment of perceptual disturbances and other psychotic symptoms[2,85]. In vivo molecular-imaging studies have been used to probe dopaminergic receptors, in particular D1 and D2 receptors, dopamine transporters and, indirectly, dopamine levels. This is achieved by pharmacologically inducing an acute change in synaptic and perisynaptic dopamine levels and examining the impact of these changes on the

◀ Fig. 4 | **Circuitry of associative auditory cortex and basal ganglia relevant to perceptual disturbances in psychosis. a** | The anatomy of auditory cortex and the downstream projections of associative auditory cortex to prefrontal cortex and striatum, as well as the cortico–cortical loops and cortico–basal ganglia–thalamo–cortical loops proposed to be relevant for perceptual disturbances. In the centre, a lateral surface of the brain is shown, indicating the different areas of auditory cortex in different colours. The associative auditory cortex features a gradient from anterior–ventral to posterior–dorsal aspects. Connections from these aspects of the associative auditory cortex to the dorsal striatum and lateral prefrontal cortex are indicated by arrows reflecting the colours of the areas of origin. Note that the striatum, basal ganglia and thalamus output structures are in the centre of the brain, but here they are shown separately for illustrative purposes. Prefrontal inputs are also shown as grey arrows. Outputs from the striatum to basal ganglia output structures and the thalamus are also illustrated, in addition to thalamo–cortical projections back to the areas where the inputs originate. Dashed arrows indicate connections that have not been fully described in primates. Findings of increased activity in associative auditory cortex in relation to hallucinations and increased striatal dopamine in psychosis are depicted as inserts. **b** | Schematic illustrating hypothesized changes in the cortico–basal ganglia–thalamo–cortical circuits due to dopamine excess in the Maia and Frank model of schizophrenia. The reference scenario (normal dopamine levels (small dark-blue box) and a normal balance of the striatal medium-spiny neuron pathways expressing dopamine D1 receptor (D1-MSN; Go, blue) versus D2-MSN (NoGo, red)) is depicted on the left, compared with the hypothesized imbalance in the Go versus NoGo pathways due to an increase in striatal dopamine transients (large dark-blue boxes) in psychosis (middle and right panels). The middle panel shows short-term, instantaneous effects of increased dopamine transients, which can enhance Go activation (large light-blue arrow) and decrease NoGo activation (small red arrow) by acting directly on striatal MSNs (note the intact cortico–striatal inputs, denoted by arrows pointing to the Go and NoGo boxes whose size is equal to that in the reference scenario). The right panel illustrates long-term effects of increased dopamine on cortico–striatal plasticity, which results in progressive potentiation of the cortico–striatal inputs onto Go and depression of the cortico–striatal inputs onto NoGo striatal neurons (note the increased size of the arrow depicting the cortico–striatal input to the Go pathway and the decreased size of the arrow depicting the cortico–striatal input to the NoGo pathway). Both abnormal scenarios (middle and right) result in increased output — that is, increased Go gating — from basal ganglia, via the thalamus, to associative (auditory) cortex. GPi, internal globus pallidus; SNr, substantia nigra, pars reticulata; STG, superior temporal gyrus; STS, superior temporal sulcus. Part **a** is modified with permission from REF.[60], Wiley-VCH; REF.[106] (Horga et al., 'Differential brain glucose metabolic patterns in antipsychotic-naive first-episode schizophrenia with and without auditory verbal hallucinations', *J. Psychiatry Neurosci.* **36**, 312–321. © Canadian Medical Association (2011)); REF.[156], reprinted with permission from the *American Journal of Psychiatry* (© 1998). American Psychiatric Association. All rights reserved.

uptake of a D2 radiotracer, which changes as a direct result of competition for binding between dopamine and the radiotracer, although some of this effect may also be due to D2 trafficking after agonist exposure[86]. These molecular-imaging studies have shown that the density of D1 receptors and dopamine transporters is normal in the striatum[87–90]. D2 receptors show upregulation in the striatum in previously treated patients but not in those who have never received antipsychotic treatment[2]. Most notably, abnormalities in indices of presynaptic dopamine have been consistently shown, including increases in striatal dopamine release[91,92], intrasynaptic levels[93,94] and synthesis rates[95,96]. These increases are especially evident in the associative striatum, part of the nigrostriatal system, compared to other parts of the striatum, including its mesolimbic aspects[2]. Furthermore, excess dopamine in the associative striatum relates to psychosis severity, including hallucination severity, both in established disease and in subjects at high risk for the disease[97,98]. Even in schizophrenia patients with comorbid addiction, who exhibit overall decreases in striatal dopamine release, the correlation between higher dopamine and more severe psychosis is apparent[99]. Excess dopamine in the associative striatum also manifests as an early warning sign that forecasts the development of psychotic disorders. Furthermore, excess striatal dopamine is most apparent in acutely psychotic patients, compared to more stable patients[98], and is also manifest in affective psychotic individuals without schizophrenia[100], suggesting a phenotype more related to psychotic states than to schizophrenia as a whole. Thus, alterations in dopamine transmission in the striatum represent the most established neurobiological correlate of psychosis, including its perceptual disturbances. In addition, earlier molecular-imaging studies showed a dose-dependent relationship between D2-receptor occupancy and the clinical efficacy of antipsychotics[101,102], which in combination with numerous controlled clinical trials with antipsychotics has established a role for D2 in the treatment of psychotic symptoms. This evidence thus provides robust support for a contribution of excess presynaptic dopamine to psychosis and suggests an additional role for postsynaptic striatal D2 pathways. Therefore, despite the clear involvement of striatal dopamine in psychosis, the exact nature of the dopamine dysfunction requires further investigation.

Striatal dysfunction in psychosis. In addition to evidence for dopaminergic dysregulation in the striatum of psychotic patients, neurological insults to the striatum can cause psychotic perceptual disturbances. Several case reports have described hallucinations in patients suffering strokes localized to the dorsal striatum[103,104]. Also, some functional-imaging studies have shown increased metabolic activity associated with hallucinations in the dorsal striatum of hallucinating compared to non-hallucinating patients[105–107], although this has not generally been observed in fMRI symptom-capture studies that have evaluated changes in haemodynamic signals on a shorter timescale.

Dopamine model of psychosis. A recent model of dopamine dysfunction in schizophrenia[69] posits that a disruption in phasic dopamine signals or transients, possibly in conjunction with an increase in tonic dopamine levels, can explain a wide variety of imaging and behavioural findings in schizophrenia through the lens of reinforcement-learning theory. Central to this model are the dual effects of dopamine dysregulation, in terms of longer-term changes in reinforcement learning via cortico–striatal plasticity and shorter-term changes during decision making, via instantaneous modulations of MSN excitability, as we discussed above. Striatal dopamine dysregulation was proposed to induce an imbalance in the D1/D2 pathways, a regime with relative D1-MSN pathway activation and/or D2-MSN pathway deactivation, resulting in a preponderance of D1-MSN long-term potentiation and/or D2-MSN long-term depression. Critically, the net effect of this D1/D2 imbalance would be to facilitate increased basal ganglia outputs ('Go gating') to cortical regions. Based on the notion that basal ganglia circuits gate motor, sensory and cognitive information through different cortico–basal ganglia–thalamo–cortical loops, this model proposes

Box 3 | **Other candidate neuromodulatory systems**

Extra-striatal dopamine systems

Cortical dopamine. Cortical dopamine release is decreased in schizophrenia[141], but a relationship to perceptual disturbances has not been reported. Dopaminergic modulation of recurrent activity in cortical networks may be relevant to auditory working memory processes and perceptual disturbances. Stimulation of auditory-cortex-projecting dopamine neurons reshapes cortical tuning in auditory cortex[142], so a similar process may contribute to perceptual disturbances through an overrepresentation of speech features.

Thalamic dopamine. Some evidence suggests that abnormal thalamo–cortical inputs may be relevant to altered sensory processing in psychotic disorders[143]. In particular, enhancement of thalamo–cortical inputs to primary auditory cortex due to elevated D2 receptors in auditory thalamus may be relevant.

Cholinergic systems

The cholinergic system is involved in multiple aspects of sensory processing. In rodents, cholinergic transients in prefrontal cortex increase hits and induce false alarms[144]. Stimulation of the cholinergic basal forebrain in rodents also induces reorganization of plasticity in auditory cortex[145]. Computational models suggest a role for cortical acetylcholine in signalling of (expected) prior uncertainty in perceptual inference[146], as supported by some human fMRI data[147]. But the relevance of acetylcholine to psychosis remains unclear and insufficiently examined. Acetylcholine has been implied in psychosis in both schizophrenia[148] and dementia[149], although clinical benefits of the cholinergic agents tested so far have been dubious at best[150,151]. Altogether, cholinergic systems remain important candidate systems in psychosis, and much work, including molecular-imaging studies with recently developed cholinergic radiotracers, is needed. Specifically, cholinergic interneurons in the striatum may be especially relevant, given their ability to regulate local release of dopamine in this region.

Norepinephrine system

The locus coeruleus–norepinephrine system supports arousal and attention[152,153] and is thought to modulate perceptual inference by signalling unexpected uncertainty[154]. So far, only indirect and scarce evidence suggests the potential involvement of this system in psychosis.

facilitated gating of striatal inputs to auditory cortex as a potential driver of auditory hallucinations.

Auditory–basal ganglia interactions

In primates, the associative auditory cortex sends monosynaptic projections to the striatum[60], including associative regions of the dorsal striatum. This anatomical topography suggests an involvement of associative striatum in higher-order perceptual inference, as prior theoretical models have suggested[108] and recent physiology work in non-human primates and rodents has shown. Single-unit recordings in associative striatum have demonstrated that persistent activity in this region instantiates evidence accumulation during perceptual decisions in visual and auditory tasks[109,110]. In addition, baseline (pre-stimulus) activity in response-tuned neurons within this region produces biases towards the preferred response in subsequent perceptual decisions, particularly in the presence of weak sensory evidence[111], consistent with the implementation of optimal incorporation of prior expectations in Bayesian models of perceptual inference (FIG. 3d).

In rodents, the associative striatum has been less studied in the context of auditory decision making. However, a posterior region of dorsal sensorimotor striatum that in rodents receives projections from primary auditory cortex and dopaminergic midbrain — the so-called 'auditory striatum' — has drawn increasing attention in recent years. Neurons in this region contribute to auditory decisions[112] and encode fine-grained information about auditory stimulus features (for example, sound frequency) in a stable manner, even in the absence of signals encoding reward-predictive actions[113]. Furthermore, learning in the context of auditory decision making has been shown to depend on plasticity in auditory cortico–striatal projections to this area, specifically in neurons tuned to task-relevant auditory features, such as high-frequency sounds predictive of reward[74].

The human associative striatum has also been shown to be involved in perceptual decision making in fMRI studies. Striatal activation during a random-dot motion task correlated with the presentation of valid predictive cues, suggesting that this region mediates the incorporation of prior information underlying response biases[114]. Striatal signals were also shown to scale with the degree of uncertainty in prior beliefs during a visual–spatial inference task[115].

Auditory–striatal changes in psychosis. Preliminary support for dysfunction of the associative striatum in perceptual disturbances in psychosis comes from clinical fMRI work. Some studies have shown the involvement of networks including the striatum in perceptual inference in hallucinating subjects, who exhibited deficient sensory prediction signals in putamen that correlated with deficient auditory prediction-error signals during a speech-discrimination task[64]. In a study using a conditioned-hallucination paradigm that demonstrated a behavioural response bias, with increased false alarms in hallucinators compared to non-hallucinators, striatal activations were observed during the false alarms[44]. Finally, increased dopamine in the associative striatum correlated with hallucination severity and with the inability to downweigh prior expectations in a tone-duration reproduction task in unmedicated patients with schizophrenia[48].

A potential integrative framework

Given the evidence reviewed above, and drawing from previous models[1,4,69], we speculate that the associative-striatum-dependent and dopamine-dependent processes involved in perceptual decisions represent a candidate neurobiological mechanism for perceptual disturbances in psychosis. We propose that this may be a sufficient but not necessary mechanism underlying at least some cases of hallucinations in schizophrenia, a mechanism that is unlikely to explain all hallucination-related phenomena across clinical and subclinical populations. Importantly, this dopamine-dependent candidate mechanism is biologically plausible, understandable in algorithmic terms and parsimonious in that it accommodates the findings of striatal dopamine dysfunction in psychosis. As we reviewed above, perceptual-inference models of hallucinations posit an excessive influence of prior expectations on perception, whereby prior beliefs produce an excessive bias of percepts towards expected stimulus features and away from objective ones. In the case of AVH, such an expectation bias could be implemented via increases in the baseline activity in speech-selective neurons in associative auditory cortex and in its downstream projections to associative striatum and prefrontal cortex (similar to the findings reviewed in BOX 2, for a

different system). This could bias a belief-updating process instantiated in the associative striatum[109] towards the threshold for speech detection, resulting in speech false alarms (that is, AVH; FIG. 4).

Although this framework could provide a basis for expectation biases underlying AVH, it does not yet address the presumed origin of increases in baseline activity in speech-selective neurons in auditory cortex. It may, however, help with understanding how this auditory hyperactivity could originate from excess dopamine in the associative striatum, as we postulate. In the early stages of psychosis, enhanced dopamine transients[69] encoding perceptual uncertainty[77,78] or sensory prediction errors[81] could result in a state of perpetual sensory surprise consistent with the phenomenology of the 'delusional mood' stage, in line with the notion of hyperdopaminergic states fomenting salience misattribution[6,7] or increased readiness to associate incidental events[1]. In individuals who go on to develop AVH, such increased dopamine transients could randomly reinforce cortico–striatal plasticity and produce an imbalance in the D1/D2 pathways that would relatively favour long-term potentiation in speech-selective inputs to striatal D1-MSN pathways and/or long-term depression in inputs to D2-MSN pathways, through the mechanisms reviewed above[69]. Instead of a random event, this could also occur if increased dopamine transmission predominantly affects a part of the associative striatum that is a main target of speech-selective inputs from associative auditory cortex, or if speech inputs are overrepresented or enhanced, given the ubiquity and prominence of speech in social communication. Alternatively, a stressful state associated with 'delusional mood' may itself predispose individuals to seek dreadful and personally meaningful interpretations. This could enhance speech-selective auditory neurons that encode negative emotional valence — such as derogatory social commentary — via prefrontal top-down selection of inputs conforming with these negative social expectations or drive enhanced attention to this type of stimuli. Regardless, excess dopamine transients could reinforce speech inputs to the associative striatum and facilitate basal ganglia outputs to the associative auditory cortex via this imbalance in D1/D2 pathways[69], which could drive increased baseline activity in these speech-selective regions of associative auditory cortex. This is consistent with preliminary work in rodents[69] and may receive some support from functional-imaging studies showing dopamine-dependent striatal dysconnectivity in psychosis[116]. By inducing biases in baseline activity in speech-selective neurons as well as biasing of striatal belief-updating computations downstream, selective long-term facilitation of speech inputs would result in chronic biases in perception towards these speech inputs and chronic AVH.

Furthermore, similar to prior suggestions[69], during a given perceptual event, ongoing dopamine transients may report on perceptual uncertainty by modulating the gain of speech-related cortico–striatal inputs relative to other inputs, via instantaneous changes in the relative excitability of D1-MSN versus D2-MSN pathways, which in algorithmic terms could correspond to weighting of prior expectations. In psychosis, excessive and noisy dopamine transients could thus imply a reduction in the dynamic range for encoding perceptual uncertainty in a context-appropriate manner, such that patients would unduly rely on prior expectations even in variable environments in which prior knowledge may be effectively uninformative.

This working model, although speculative, has several appealing aspects. It could simultaneously explain an expectation bias and deficient adjustments of this bias in variable contexts with high uncertainty, both of which agree with recent findings[44,47,48]. Furthermore, although under this model both of these processes may be driven by striatal dopamine dysfunction, the expectation bias itself could arise gradually via changes in cortico–striatal connectivity, perhaps consistent with the progression from illusory to hallucinatory percepts in the prodrome, whereas deficient context-related adjustments of this bias may be more closely related to concurrent dopamine levels. This could explain our initial observation that only deficient adjustments in the expectation bias, but not the strength of the bias itself, correlated with increased dopamine release into the associative striatum in hallucinating patients[48]. It could also explain the bimodal time course observed in the clinical response to antipsychotic medication, which is thought to rebalance the D1/D2 pathways, by blocking D2 receptors and consequently increasing short-term activity and long-term plasticity in D2-MSN pathways[117–119]. This could induce a short-term normalization of the adjustment in the expectation bias that might mitigate hallucinations to some degree, as well as a longer-term reshaping of cortico–striatal plasticity that could correct the expectation bias — and hallucinations — to a fuller extent. Furthermore, this working model could potentially provide a basis for the lack of treatment effectiveness in some patients with longer durations of untreated psychosis, in whom deep-rooted changes in cortico–striatal plasticity could be less malleable to treatment, especially when the pathology begins at earlier developmental stages[120]. While this pathophysiological framework does not deal with the aetiology of dopamine dysregulation, such dysregulation has been proposed to arise through a combination of genetic and environmental factors, partly related to stress, that affect upstream regions that modulate midbrain or striatal function[121,122].

Extensions of this working model in the context of hierarchical inference[1,5] may explain delusions as a consequence of alterations that are computationally similar to and intertwined with those driving hallucinations, where alterations may induce overly certain and rigid high-level beliefs about abstract hidden states[123]. These delusion-related beliefs are likely represented in distinct high-level circuits that are hierarchically interconnected with but partially dissociable from hallucination-related circuits, explaining the typical clustering and partial dissociation of the two sets of symptoms. Some recent work provides preliminary support for these notions[123–125], although how this model may be reconciled with a long body of work suggesting jumping-to-conclusion biases in schizophrenia and the model's exact neurobiological implementation remain open questions. Future work

will also be required in order to directly pit our proposed framework — focused on a role for nigrostriatal dopamine dysfunction in perceptual inference — against other models that emphasize a role of dopamine-related disruptions to either reward learning or the predictive motor signals used for agency attribution (see REF.[126] for a possible way of reconciling the latter models with a perceptual-inference framework). Finally, as stated above, we believe that striatal dopaminergic dysfunction may be a sufficient but not a necessary mechanism for perceptual disturbances in schizophrenia, with similar computational phenotypes potentially arising from non-dopaminergic mechanisms (for example, imbalances in glutamatergic and GABAergic signalling[29,30] or alterations in other neuromodulatory systems; see BOX 3).

Concluding remarks

In this article, we have reviewed perceptual disturbances in psychosis from the standpoint of perceptual decision making and have proposed a dopamine-dependent implementation of perceptual biases in the associative striatum that may contribute to such phenomena. Recent work suggests that dopamine plays a key role in perceptual learning and perceptual decision making, separate from its more established role in reward learning and reward-based decision making. Other work in non-human primates and rodents has highlighted the role of associative striatum in perceptual decision making, including its involvement in response biases towards prior expectations. Given the importance of expectation biases in current models of perception, and given the well-established alterations in associative-striatum dopamine in individuals with psychosis, a parsimonious model of psychotic perceptual disturbances positions expectation biases as a central algorithmic process and striatal dopamine excess as a key neurobiological candidate for their implementation. Although many critical questions regarding the pathophysiology of perceptual psychotic symptoms remain to be answered, we contend that theory-driven, integrative approaches such as the one illustrated here hold promise in elucidating such enigmatic phenomena as hallucinations, which have perplexed humanity for generations. We thus believe that the development and iterative refinement of theory-driven models bridging biology and subjective experience are key to advancing our knowledge about psychosis, and ultimately may help destigmatize this condition and contribute to the development of next-generation treatments.

Published online 11 November 2019

1. Fletcher, P. C. & Frith, C. D. Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia. *Nat. Rev. Neurosci.* **10**, 48–58 (2009).
This in-depth review proposes an influential framework based on Bayesian inference to bridge the known dopamine dysregulation and individual experiential aspects associated with psychosis.
2. Weinstein, J. J. et al. Pathway-specific dopamine abnormalities in schizophrenia. *Biol. Psychiatry* **81**, 31–42 (2017).
This study reviews the literature on dopamine alterations in schizophrenia with special emphasis on the distinct anatomical pathways that comprise the dopamine system.
3. Powers, A. R. III, Kelley, M. & Corlett, P. R. Hallucinations as top-down effects on perception. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* **1**, 393–400 (2016).
4. Adams, R. A., Stephan, K. E., Brown, H. R., Frith, C. D. & Friston, K. J. The computational anatomy of psychosis. *Front. Psychiatry* **4**, 47 (2013).
This report provides a comprehensive computational account of various phenomenological and neurophysiological aspects of schizophrenia.
5. Sterzer, P. et al. The predictive coding account of psychosis. *Biol. Psychiatry* **84**, 634–643 (2018).
6. Heinz, A. Dopaminergic dysfunction in alcoholism and schizophrenia—psychopathological and behavioral correlates. *Eur. Psychiatry* **17**, 9–16 (2002).
7. Kapur, S. Psychosis as a state of aberrant salience: a framework linking biology, phenomenology, and pharmacology in schizophrenia. *Am. J. Psychiatry* **160**, 13–23 (2003).
8. Glimcher, P. W. Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis. *Proc. Natl Acad. Sci. USA* **108** (Suppl. 3), 15647–15654 (2011).
9. Schultz, W. Predictive reward signal of dopamine neurons. *J. Neurophysiol.* **80**, 1–27 (1998).
10. Schultz, W. Getting formal with dopamine and reward. *Neuron* **36**, 241–263 (2002).
11. Schultz, W. Multiple dopamine functions at different time courses. *Annu. Rev. Neurosci.* **30**, 259–288 (2007).
12. Yung, A. R. & McGorry, P. D. The prodromal phase of first-episode psychosis: past and current conceptualizations. *Schizophr. Bull.* **22**, 353–370 (1996).
13. Lim, A., Hoek, H. W., Deen, M. L., Blom, J. D. & GROUP Investigators. Prevalence and classification of hallucinations in multiple sensory modalities in schizophrenia spectrum disorders. *Schizophr. Res.* **176**, 493–499 (2016).
14. Waters, F. & Fernyhough, C. Hallucinations: a systematic review of points of similarity and difference across diagnostic classes. *Schizophr. Bull.* **43**, 32–43 (2017).
15. Andreasen, N. C. & Flaum, M. Schizophrenia: the characteristic symptoms. *Schizophr. Bull.* **17**, 27–49 (1991).
16. Nayani, T. H. & David, A. S. The auditory hallucination: a phenomenological survey. *Psychol. Med.* **26**, 177–189 (1996).
17. Llorca, P. M. et al. Hallucinations in schizophrenia and Parkinson's disease: an analysis of sensory modalities involved and the repercussion on patients. *Sci. Rep.* **6**, 38152 (2016).
18. Lehembre-Shiah, E. et al. Distinct relationships between visual and auditory perceptual abnormalities and conversion to psychosis in a clinical high-risk population. *JAMA Psychiatry* **74**, 104–106 (2017).
19. Emsley, R., Rabinowitz, J., Torreman, M. & RIS-INT-35 Early Psychosis Global Working Group. The factor structure for the Positive and Negative Syndrome Scale (PANSS) in recent-onset psychosis. *Schizophr. Res.* **61**, 47–57 (2003).
20. Freedman, R. Schizophrenia. *N. Engl. J. Med.* **349**, 1738–1749 (2003).
21. De Keyser, J., De Backer, J. P., Ebinger, G. & Vauquelin, G. Regional distribution of the dopamine D2 receptors in the mesotelencephalic dopamine neuron system of human brain. *J. Neurol. Sci.* **71**, 119–127 (1985).
22. Palacios, J. M., Camps, M., Cortes, R. & Probst, A. Mapping dopamine receptors in the human brain. *J. Neural. Transm. Suppl.* **27**, 227–235 (1988).
23. Agid, O., Seeman, P. & Kapur, S. The 'delayed onset' of antipsychotic action—an idea whose time has come and gone. *J. Psychiatry Neurosci.* **31**, 93–100 (2006).
24. Emsley, R., Rabinowitz, J. & Medori, R. Time course for antipsychotic treatment response in first-episode schizophrenia. *Am. J. Psychiatry* **163**, 743–745 (2006).
25. Rector, N. A. & Beck, A. T. Cognitive behavioral therapy for schizophrenia: an empirical review. *J. Nerv. Ment. Dis.* **189**, 278–287 (2001).
26. Slotema, C. W., Aleman, A., Daskalakis, Z. J. & Sommer, I. E. Meta-analysis of repetitive transcranial magnetic stimulation in the treatment of auditory verbal hallucinations: update and effects after one month. *Schizophr. Res.* **142**, 40–45 (2012).
27. Pruessner, J. C., Champagne, F., Meaney, M. J. & Dagher, A. Dopamine release in response to a psychological stress in humans and its relationship to early life maternal care: a positron emission tomography study using [^{11}C]raclopride. *J. Neurosci.* **24**, 2825–2831 (2004).
28. Volkow, N. D., Fowler, J. S., Wang, G. J. & Swanson, J. M. Dopamine in drug abuse and addiction: results from imaging studies and treatment implications. *Mol. Psychiatry* **9**, 557–569 (2004).
29. Jardri, R. et al. Are hallucinations due to an imbalance between excitatory and inhibitory influences on the brain? *Schizophr. Bull.* **42**, 1124–1134 (2016).
30. Krystal, J. H. et al. impaired tuning of neural ensembles and the pathophysiology of schizophrenia: a translational and computational neuroscience perspective. *Biol. Psychiatry* **81**, 874–885 (2017).
31. Bohlken, M. M., Hugdahl, K. & Sommer, I. E. Auditory verbal hallucinations: neuroimaging and treatment. *Psychol. Med.* **47**, 199–208 (2017).
32. Goghari, V. M., Harrow, M., Grossman, L. S. & Rosen, C. A 20-year multi-follow-up of hallucinations in schizophrenia, other psychotic, and mood disorders. *Psychol. Med.* **43**, 1151–1160 (2013).
33. Marshall, M. et al. Association between duration of untreated psychosis and outcome in cohorts of first-episode patients: a systematic review. *Arch. Gen. Psychiatry* **62**, 975–983 (2005).
34. Summerfield, C. & de Lange, F. P. Expectation in perceptual decision making: neural and computational mechanisms. *Nat. Rev. Neurosci.* **15**, 745–756 (2014).
This review integrates non-primate electrophysiology and human neuroimaging findings that represent possible neural implementations of expectation effects in perceptual decision-making in the brain.
35. Bitzer, S., Park, H., Blankenburg, F. & Kiebel, S. J. Perceptual decision making: drift-diffusion model is equivalent to a Bayesian model. *Front. Hum. Neurosci.* **8**, 102 (2014).
36. Friston, K. & Kiebel, S. Predictive coding under the free-energy principle. *Philos Trans. R. Soc. Lond. B. Biol. Sci.* **364**, 1211–1221 (2009).
37. Rao, R. P. & Ballard, D. H. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* **2**, 79–87 (1999).
38. Friston, K. Hallucinations and perceptual inference. *Behavioral and Brain Sciences* **28**, 764–766 (2005).
In this short commentary, Karl Friston delineates for the first time a perceptual-inference model of hallucinations.
39. Benrimoh, D., Parr, T., Vincent, P., Adams, R. A. & Friston, K. Active inference and auditory hallucinations. *Comput. Psychiatr.* **2**, 183–204 (2018).

40. Bentall, R. P. & Slade, P. D. Reality testing and auditory hallucinations: a signal detection analysis. *Br. J. Clin. Psychol.* **24** (Pt 3), 159–169 (1985).
41. Brookwell, M. L., Bentall, R. P. & Varese, F. Externalizing biases and hallucinations in source-monitoring, self-monitoring and signal detection studies: a meta-analytic review. *Psychol. Med.* **43**, 2465–2475 (2013).
42. Varese, F., Barkus, E. & Bentall, R. P. Dissociation mediates the relationship between childhood trauma and hallucination-proneness. *Psychol. Med.* **42**, 1025–1036 (2012).
43. Vercammen, A., de Haan, E. H. & Aleman, A. Hearing a voice in the noise: auditory hallucinations and speech perception. *Psychol. Med.* **38**, 1177–1184 (2008).
44. Powers, A. R., Mathys, C. & Corlett, P. R. Pavlovian conditioning-induced hallucinations result from overweighting of perceptual priors. *Science* **357**, 596–600 (2017).
45. O'Callaghan, C. et al. Visual hallucinations are characterized by impaired sensory evidence accumulation: insights from hierarchical drift diffusion modeling in Parkinson's Disease. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* **2**, 680–688 (2017).
46. Limongi, R., Bohaterewicz, B., Nowicka, M., Plewka, A. & Friston, K. J. Knowing when to stop: aberrant precision and evidence accumulation in schizophrenia. *Schizophr. Res.* **197**, 386–391 (2018).
47. Teufel, C. et al. Shift toward prior knowledge confers a perceptual advantage in early psychosis and psychosis-prone healthy individuals. *Proc. Natl Acad. Sci. USA* **112**, 13401–13406 (2015).
This behavioural human study provided a first demonstration of exaggerated expectation biases specifically relevant to perceptual disturbances in populations at risk for psychosis.
48. Cassidy, C. M. et al. A perceptual inference mechanism for hallucinations linked to striatal dopamine. *Curr. Biol.* **28**, 503–514 (2018).
This study is the first demonstration of a correlation between striatal dopamine excess and altered expectation biases in relation to hallucinations in patients with schizophrenia.
49. Gold, J. I. & Shadlen, M. N. The neural basis of decision making. *Annu. Rev. Neurosci.* **30**, 535–574 (2007).
This general review delves into the electrophysiology literature on perceptual decision making and its links to computational models of evidence accumulation.
50. Dehaene, S. & Changeux, J. P. Experimental and theoretical approaches to conscious processing. *Neuron* **70**, 200–227 (2011).
51. Maia, T. V. & Cleeremans, A. Consciousness: converging insights from connectionist modeling and neuroscience. *Trends Cogn. Sci.* **9**, 397–404 (2005).
52. de Lafuente, V. & Romo, R. Neuronal correlates of subjective sensory experience. *Nat. Neurosci.* **8**, 1698–1703 (2005).
53. de Lafuente, V. & Romo, R. Neural correlate of subjective sensory experience gradually builds up across cortical areas. *Proc. Natl Acad. Sci. USA* **103**, 14266–14271 (2006).
54. Van Vugt, B. et al. The threshold for conscious report: signal loss and response bias in visual and frontal cortex. *Science* **360**, 537–542 (2018).
55. Kaas, J. H. & Hackett, T. A. Subdivisions of auditory cortex and processing streams in primates. *Proc. Natl Acad. Sci. USA* **97**, 11793–11799 (2000).
56. Tang, C., Hamilton, L. S. & Chang, E. F. Intonational speech prosody encoding in the human auditory cortex. *Science* **357**, 797–801 (2017).
57. Penfield, W. & Perot, P. The brain's record of auditory and visual experience: a final summary and discussion. *Brain* **86**, 595–696 (1963).
58. Kobayashi, E. et al. Magnetic resonance imaging abnormalities in familial temporal lobe epilepsy with auditory auras. *Arch. Neurol.* **60**, 1546–1551 (2003).
59. Winawer, M. R., Ottman, R., Hauser, W. A. & Pedley, T. A. Autosomal dominant partial epilepsy with auditory features: defining the phenotype. *Neurology* **54**, 2173–2176 (2000).
60. Yeterian, E. H. & Pandya, D. N. Corticostriatal connections of the superior temporal region in rhesus monkeys. *J. Comp. Neurol.* **399**, 384–402 (1998).
This tracing study in non-human primates describes the anatomical downstream connections of the superior temporal cortex.
61. Hackett, T. A., Stepniewska, I. & Kaas, J. H. Thalamocortical connections of the parabelt auditory cortex in macaque monkeys. *J. Comp. Neurol.* **400**, 271–286 (1998).
62. Middleton, F. A. & Strick, P. L. The temporal lobe is a target of output from the basal ganglia. *Proc. Natl Acad. Sci. USA* **93**, 8683–8687 (1996).
63. Shammah-Lagnado, S. J., Alheid, G. F. & Heimer, L. Efferent connections of the caudal part of the globus pallidus in the rat. *J. Comp. Neurol.* **376**, 489–507 (1996).
64. Horga, G., Schatz, K. C., Abi-Dargham, A. & Peterson, B. S. Deficits in predictive coding underlie hallucinations in schizophrenia. *J. Neurosci.* **34**, 8072–8082 (2014).
65. Jardri, R., Pouchet, A., Pins, D. & Thomas, P. Cortical activations during auditory verbal hallucinations in schizophrenia: a coordinate-based meta-analysis. *Am. J. Psychiatry* **168**, 73–81 (2011).
66. Diederen, K. M. et al. Deactivation of the parahippocampal gyrus preceding auditory hallucinations in schizophrenia. *Am. J. Psychiatry* **167**, 427–435 (2010).
67. Hoffman, R. E., Pittman, B., Constable, R. T., Bhagwagar, Z. & Hampson, M. Time course of regional brain activity accompanying auditory verbal hallucinations in schizophrenia. *Br. J. Psychiatry* **198**, 277–283 (2011).
68. Gerfen, C. R. & Surmeier, D. J. Modulation of striatal projection systems by dopamine. *Annu. Rev. Neurosci.* **34**, 441–466 (2011).
69. Maia, T. V. & Frank, M. J. An integrative perspective on the role of dopamine in schizophrenia. *Biol. Psychiatry* **81**, 52–66 (2017).
This article provides a comprehensive model of how dopamine dysregulation in schizophrenia may explain different symptom domains via reinforcement-learning theory.
70. Alexander, G. E., DeLong, M. R. & Strick, P. L. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu. Rev. Neurosci.* **9**, 357–381 (1986).
71. Haber, S. N. The primate basal ganglia: parallel and integrative networks. *J. Chem. Neuroanat.* **26**, 317–330 (2003).
72. Choi, E. Y., Tanimura, Y., Vage, P. R., Yates, E. H. & Haber, S. N. Convergence of prefrontal and parietal anatomical projections in a connectional hub in the striatum. *Neuroimage* **146**, 821–832 (2017).
73. Wickens, J. R., Reynolds, J. N. & Hyland, B. I. Neural mechanisms of reward-related motor learning. *Curr. Opin. Neurobiol.* **13**, 685–690 (2003).
74. Xiong, Q., Znamenskiy, P. & Zador, A. M. Selective corticostriatal plasticity during acquisition of an auditory discrimination task. *Nature* **521**, 348–351 (2015).
75. Redgrave, P., Prescott, T. J. & Gurney, K. Is the short-latency dopamine response too short to signal reward error? *Trends Neurosci.* **22**, 146–151 (1999).
76. Horvitz, J. C., Stewart, T. & Jacobs, B. L. Burst activity of ventral tegmental dopamine neurons is elicited by sensory stimuli in the awake cat. *Brain Res.* **759**, 251–258 (1997).
77. de Lafuente, V. & Romo, R. Dopamine neurons code subjective sensory experience and uncertainty of perceptual decisions. *Proc. Natl Acad. Sci. USA* **108**, 19767–19771 (2011).
This non-human primate study shows that dopamine neurons encode perceptual uncertainty during a decision-making task.
78. Lak, A., Nomoto, K., Keramati, M., Sakagami, M. & Kepecs, A. Midbrain dopamine neurons signal belief in choice accuracy during a perceptual decision. *Curr. Biol.* **27**, 821–832 (2017).
79. Sarno, S., de Lafuente, V., Romo, R. & Parga, N. Dopamine reward prediction error signal codes the temporal evaluation of a perceptual decision report. *Proc. Natl Acad. Sci. USA* **114**, E10494–E10503 (2017).
80. Sharpe, M. J. et al. Dopamine transients are sufficient and necessary for acquisition of model-based associations. *Nat. Neurosci.* **20**, 735–742 (2017).
This rodent study uses a sensory preconditioning paradigm and optogenetics to show that dopamine is necessary for stimulus–stimulus learning, separate from reward learning.
81. Menegas, W., Babayan, B. M., Uchida, N. & Watabe-Uchida, M. Opposite initialization to novel cues in dopamine signaling in ventral and posterior striatum in mice. *eLife* **6**, e21886 (2017).
This rodent study shows that, in contrast with dopamine signals in ventral striatum, those in dorsal striatum encode a type of sensory prediction error.
82. Wolpe, N. et al. Sensory attenuation in Parkinson's disease is related to disease severity and dopamine dose. *Sci. Rep.* **8**, 15643 (2018).
83. Vilares, I. & Kording, K. P. Dopaminergic medication increases reliance on current information in Parkinson's Disease. *Nat. Hum. Behav.* **1**, 0129 (2017).
84. Nour, M. M. et al. Dopaminergic basis for signaling belief updates, but not surprise, and the link to paranoia. *Proc. Natl Acad. Sci. USA* **115**, E10167–E10176 (2018).
85. Howes, O. D. et al. The nature of dopamine dysfunction in schizophrenia and what this means for treatment. *Arch. Gen. Psychiatry* **69**, 776–786 (2012).
86. Laruelle, M. Imaging synaptic neurotransmission with in vivo binding competition techniques: a critical review. *J. Cereb. Blood Flow Metab.* **20**, 423–451 (2000).
87. Abi-Dargham, A. et al. Prefrontal dopamine D1 receptors and working memory in schizophrenia. *J. Neurosci.* **22**, 3708–3719 (2002).
88. Abi-Dargham, A. et al. Increased prefrontal cortical D_1 receptors in drug naive patients with schizophrenia: a PET study with [^{11}C]NNC112. *J. Psychopharmacol.* **26**, 794–805 (2012).
89. Okubo, Y. et al. Decreased prefrontal dopamine D1 receptors in schizophrenia revealed by PET. *Nature* **385**, 634–636 (1997).
90. Okubo, Y., Suhara, T., Sudo, Y. & Toru, M. Possible role of dopamine D1 receptors in schizophrenia. *Mol. Psychiatr.* **2**, 291–292 (1997).
91. Laruelle, M. et al. Single photon emission computerized tomography imaging of amphetamine-induced dopamine release in drug free schizophrenic subjects. *Proc. Natl Acad. Sci. USA* **93**, 9235–9240 (1996).
92. Abi-Dargham, A. et al. Increased striatal dopamine transmission in schizophrenia: confirmation in a second cohort. *Am. J. Psychiatry* **155**, 761–767 (1998).
93. Abi-Dargham, A. et al. Increased baseline occupancy of D2 receptors by dopamine in schizophrenia. *Proc. Natl Acad. Sci. USA* **97**, 8104–8109 (2000).
94. Kegeles, L. et al. Increased synaptic dopamine in associative regions of the striatum in schizophrenia. *Archives of General Psychiatry* **67**, 231–239 (2010).
95. Howes, O. D. et al. Dopamine synthesis capacity before onset of psychosis: a prospective [18f]-dopa pet imaging study. *Am. J. Psychiatry* **168**, 1311–1317 (2011).
96. Howes, O. et al. Progressive increase in striatal dopamine synthesis capacity as patients develop psychosis: a PET study. *Mol Psychiatry* **16**, 885–886 (2011).
97. Howes, O. D. et al. Elevated striatal dopamine function linked to prodromal signs of schizophrenia. *Arch. Gen. Psychiatry* **66**, 13–20 (2009).
98. Laruelle, M., Abi-Dargham, A., Gil, R., Kegeles, L. & Innis, R. Increased dopamine transmission in schizophrenia: relationship to illness phases. *Biol. Psychiatry* **46**, 56–72 (1999).
99. Thompson, J. L. et al. Striatal dopamine release in schizophrenia comorbid with substance dependence. *Mol. Psychiatry* **18**, 909–915 (2013).
100. Jauhar, S. et al. A test of the transdiagnostic dopamine hypothesis of psychosis using positron emission tomographic imaging in bipolar affective disorder and schizophrenia. *JAMA Psychiatry* **74**, 1206–1213 (2017).
101. Farde, L., Hall, H., Ehrin, E. & Sedvall, G. Quantitative analysis of D2 dopamine receptor binding in the living human brain by PET. *Science* **231**, 258–261 (1986).
102. Farde, L. et al. Positron emission tomographic analysis of central D1 and D2 dopamine receptor occupancy in patients treated with classical neuroleptics and clozapine: relation to extrapyramidal side effects. *Arch. Gen. Psychiatry* **49**, 538–544 (1992).
103. Farid, F. & Mahadun, P. Schizophrenia-like psychosis following left putamen infarct: a case report. *J. Med. Case Rep.* **3**, 7337 (2009).
104. Kitabayashi, Y. et al. Schizophrenia-like psychosis following right putaminal infarction. *J. Neuropsychiatry Clin. Neurosci.* **18**, 561–562 (2006).
105. Cleghorn, J. M. et al. Toward a brain map of auditory hallucinations. *Am. J. Psychiatry* **149**, 1062–1069 (1992).
106. Horga, G. et al. Differential brain glucose metabolic patterns in antipsychotic-naive first-episode schizophrenia with and without auditory verbal hallucinations. *J. Psychiatry Neurosci.* **36**, 312–321 (2011).
107. Zhuo, C. et al. Cerebral blood flow alterations specific to auditory verbal hallucinations in schizophrenia. *Br. J. Psychiatry* **210**, 209–215 (2017).

108. Samejima, K. & Doya, K. Multiple representations of belief states and action values in corticobasal ganglia loops. *Ann. N. Y. Acad. Sci.* **1104**, 213–228 (2007).
109. Ding, L. & Gold, J. I. Caudate encodes multiple computations for perceptual decisions. *J. Neurosci.* **30**, 15747–15759 (2010).
110. Yartsev, M. M., Hanks, T. D., Yoon, A. M. & Brody, C. D. Causal contribution and dynamical encoding in the striatum during evidence accumulation. *eLife* **7**, e34929 (2018).
111. Ding, L. & Gold, J. I. Separate, causal roles of the caudate in saccadic choice and execution in a perceptual decision task. *Neuron* **75**, 865–874 (2012).
112. Znamenskiy, P. & Zador, A. M. Corticostriatal neurons in auditory cortex drive decisions during auditory discrimination. *Nature* **497**, 482–485 (2013).
113. Guo, L., Walker, W. I., Ponvert, N. D., Penix, P. L. & Jaramillo, S. Stable representation of sounds in the posterior striatum during flexible auditory decisions. *Nat. Commun.* **9**, 1534 (2018).
114. Forstmann, B. U., Brown, S., Dutilh, G., Neumann, J. & Wagenmakers, E. J. The neural substrate of prior information in perceptual decision making: a model-based analysis. *Front. Hum. Neurosci.* **4**, 40 (2010).
115. Vilares, I., Howard, J. D., Fernandes, H. L., Gottfried, J. A. & Kording, K. P. Differential representations of prior and likelihood uncertainty in the human brain. *Curr. Biol.* **22**, 1641–1648 (2012).
116. Horga, G. et al. Dopamine-related disruption of functional topography of striatal connections in unmedicated patients with schizophrenia. *JAMA Psychiatry* **73**, 862–870 (2016).
117. Bateup, H. S. et al. Distinct subclasses of medium spiny neurons differentially regulate striatal motor behaviors. *Proc. Natl Acad. Sci. USA* **107**, 14845–14850 (2010).
118. Beeler, J. A. et al. A role for dopamine-mediated learning in the pathophysiology and treatment of Parkinson's disease. *Cell Rep.* **2**, 1747–1761 (2012).
119. Durieux, P. F., Schiffmann, S. N. & De Kerchove d'Exaerde, A. Differential regulation of motor control and response to dopaminergic drugs by D1R and D2R neurons in distinct dorsal striatum subregions. *EMBO J.* **31**, 640–653 (2012).
120. Mowery, T. M. et al. The sensory striatum is permanently impaired by transient developmental deprivation. *Cell Rep.* **19**, 2462–2468 (2017).
121. Grace, A. A. Dysregulation of the dopamine system in the pathophysiology of schizophrenia and depression. *Nat. Rev. Neurosci.* **17**, 524–532 (2016).
122. Howes, O. D. & Murray, R. M. Schizophrenia: an integrated sociodevelopmental–cognitive model. *Lancet* **383**, 1677–1687 (2014).
123. Baker, S. C., Konova, A. B., Daw, N. D. & Horga, G. A distinct inferential mechanism for delusions in schizophrenia. *Brain* **142**, 1797–1812 (2019).
124. Corlett, P. R. et al. Disrupted prediction-error signal in psychosis: evidence for an associative account of delusions. *Brain* **130**, 2387–2400 (2007).
125. Davies, D. J., Teufel, C. & Fletcher, P. C. Anomalous perceptions and beliefs are associated with shifts toward different types of prior knowledge in perceptual inference. *Schizophr. Bull.* **44**, 1245–1253 (2018).
126. Corlett, P. R. et al. Hallucinations and strong priors. *Trends Cogn. Sci.* **23**, 114–127 (2019).
127. Javitt, D. C. & Sweet, R. A. Auditory dysfunction in schizophrenia: integrating clinical and basic features. *Nat. Rev. Neurosci.* **16**, 535–550 (2015).
128. Hosoya, T., Baccus, S. A. & Meister, M. Dynamic predictive coding by the retina. *Nature* **436**, 71–77 (2005).
129. Singla, S., Dempsey, C., Warren, R., Enikolopov, A. G. & Sawtell, N. B. A cerebellum-like circuit in the auditory system cancels responses to self-generated sounds. *Nat. Neurosci.* **20**, 943–950 (2017).
130. Freedman, R. et al. Neurobiological studies of sensory gating in schizophrenia. *Schizophr. Bull.* **13**, 669–678 (1987).
131. Jensen, J. & Kapur, S. Salience and psychosis: moving from theory to practise. *Psychol. Med.* **39**, 197–198 (2009).
132. Feinberg, I. Efference copy and corollary discharge: implications for thinking and its disorders. *Schizophr. Bull.* **4**, 636–640 (1978).
133. Ford, J. M. & Mathalon, D. H. Corollary discharge dysfunction in schizophrenia: can it explain auditory hallucinations? *Int. J. Psychophysiol.* **58**, 179–189 (2005).
134. Frith, C. D. The positive and negative symptoms of schizophrenia reflect impairments in the perception and initiation of action. *Psychol. Med.* **17**, 631–648 (1987).
135. Leptourgos, P., Deneve, S. & Jardri, R. Can circular inference relate the neuropathological and behavioral aspects of schizophrenia? *Curr. Opin. Neurobiol.* **46**, 154–161 (2017).
136. Gold, J. I. & Shadlen, M. N. Banburismus and the brain: decoding the relationship between sensory stimuli, decisions, and reward. *Neuron* **36**, 299–308 (2002).
137. Schlack, A. & Albright, T. D. Remembering visual motion: neural correlates of associative plasticity and motion recall in cortical area MT. *Neuron* **53**, 881–890 (2007).
138. Hanks, T. D., Mazurek, M. E., Kiani, R., Hopp, E. & Shadlen, M. N. Elapsed decision time affects the weighting of prior probability in a perceptual decision task. *J. Neurosci.* **31**, 6339–6352 (2011).
139. Kok, P., Brouwer, G. J., van Gerven, M. A. & de Lange, F. P. Prior expectations bias sensory representations in visual cortex. *J. Neurosci.* **33**, 16275–16284 (2013).
140. Summerfield, C. & Egner, T. Expectation (and attention) in visual cognition. *Trends Cogn. Sci.* **13**, 403–409 (2009).
141. Slifstein, M. et al. Deficits in prefrontal cortical and extrastriatal dopamine release in schizophrenia: a positron emission tomographic functional magnetic resonance imaging study. *JAMA Psychiatry* **72**, 316–324 (2015).
142. Bao, S., Chan, V. T. & Merzenich, M. M. Cortical remodelling induced by activity of ventral tegmental dopamine neurons. *Nature* **412**, 79–83 (2001).
143. Chun, S. et al. Specific disruption of thalamic inputs to the auditory cortex in schizophrenia models. *Science* **344**, 1178–1182 (2014).
144. Gritton, H. J. et al. Cortical cholinergic signaling controls the detection of cues. *Proc. Natl Acad. Sci. USA* **113**, E1089–E1097 (2016).
145. Kilgard, M. P. & Merzenich, M. M. Cortical map reorganization enabled by nucleus basalis activity. *Science* **279**, 1714–1718 (1998).
146. Yu, A. J. & Dayan, P. Acetylcholine in cortical inference. *Neural Netw.* **15**, 719–730 (2002).
147. Iglesias, S. et al. Hierarchical prediction errors in midbrain and basal forebrain during sensory learning. *Neuron* **80**, 519–530 (2013).
148. Rowntree, D. W., Nevin, S. & Wilson, A. The effects of diisopropylfluorophosphonate in schizophrenia and manic depressive psychosis. *J. Neurol. Neurosurg. Psychiatry* **13**, 47–62 (1950).
149. Perry, E. K. & Perry, R. H. Acetylcholine and hallucinations: disease-related compared to drug-induced alterations in human consciousness. *Brain Cognit.* **28**, 240–258 (1995).
150. Buchanan, R. W. et al. Galantamine for the treatment of cognitive impairments in people with schizophrenia. *Am. J. Psychiatry* **165**, 82–89 (2008).
151. Keefe, R. S. et al. Efficacy and safety of donepezil in patients with schizophrenia or schizoaffective disorder: significant placebo/practice effects in a 12-week, randomized, double-blind, placebo-controlled trial. *Neuropsychopharmacology* **33**, 1217–1228 (2008).
152. Aston-Jones, G. & Cohen, J. D. An integrative theory of locus coeruleus–norepinephrine function: adaptive gain and optimal performance. *Annu. Rev. Neurosci.* **28**, 403–450 (2005).
153. Yu, A. J. & Dayan, P. Uncertainty, neuromodulation, and attention. *Neuron* **46**, 681–692 (2005).
154. Krishnamurthy, K., Nassar, M. R., Sarode, S. & Gold, J. I. Arousal-related adjustments of perceptual biases optimize perception in dynamic environments. *Nat. Hum. Behav.* **1**, 0107 (2017).
155. Yates, J. L., Park, I. M., Katz, L. N., Pillow, J. W. & Huk, A. C. Functional dissection of signal and noise in MT and LIP during decision-making. *Nat. Neurosci.* **20**, 1285–1292 (2017).
157. Abi-Dargham, A. et al. Increased striatal dopamine transmission in schizophrenia: confirmation in a second cohort. *Am. J. Psychiatry* **155**, 761–767 (1998).

Acknowledgements
The authors thank T. Maia, S. Haber and K. Schmack for insightful discussions on the ideas presented in this article. The authors acknowledge funding sources: G.H.: R01MH117323, R01MH114965;A.A-D.: R01MH109635.

Author contributions
G.H. researched data for the article, and both authors made substantial contributions to discussion of the content and writing, reviewing and editing the manuscript before submission.

Competing interests
The authors declare no competing interests.

Peer review information
Nature Reviews Neuroscience thanks P. Corlett, P. Fletcher and the other, anonymous, reviewer(s), for their contribution to the peer review of this work.