AVANT, Vol. V, No. 3/2014
ISSN: 2082-6710 avant.edu.pl/en
DOI: 10.26913/50302014.0112.0004

Bayesian inference, predictive coding and delusions

Rick A Adams, **Harriet R Brown**, and **Karl J Friston**[*]
The Wellcome Trust Centre for Neuroimaging, University College London
[*] *k.friston[]ucl.ac.uk*

Received May 2014; accepted December 2014; published winter 2014/2015.

Abstract

This paper considers psychotic symptoms in terms of false inferences or beliefs. It is based on the notion that the brain is an organ of inference that actively constructs hypotheses to explain or predict its sensations. This perspective provides a normative (Bayes optimal) account of action and perception that emphasises probabilistic representations; in particular, the confidence or *precision* of beliefs about the world. We consider sensory attenuation deficits, catatonia and delusions as various expressions of the same core pathology: namely, an aberrant encoding of precision in a predictive coding hierarchy. In predictive coding, precision is thought to be encoded by the postsynaptic gain of neurons reporting prediction error. This suggests that both pervasive *trait* abnormalities and florid failures of inference in the psychotic *state* can be linked to factors controlling postsynaptic gain—such as NMDA receptor function and (dopaminergic) neuromodulation. We illustrate these points using a biologically plausible simulation of attribution of agency—showing how a reduction in the precision of prior beliefs, relative to sensory evidence, can lead to false inference.

Keywords: free energy; active inference; precision; sensory attenuation; illusions; psychosis; schizophrenia.

Introduction

This paper comprises five sections. We start with a brief review of Bayesian inference and predictive coding, and the importance of precision in this context. We then discuss the symptoms and signs of schizophrenia, with a special focus on how *trait* and *state* abnormalities can be cast in terms of false inference. The third section reviews the psychopharmacology of psychosis with an emphasis on the synaptic (neuromodulatory) mechanisms that we suppose underlie false inference. The fourth section establishes the normative theory

(active inference) and its biological instantiation in the brain (generalised Bayesian filtering or predictive coding). The fifth section provides an illustrative example of the approach by simulating abnormalities of active inference in the context of sensory attenuation and the attribution of agency.

Bayesian inference, predictive coding and precision

In what follows, we will refer to beliefs, inference, priors and precision in a Bayesian sense. In this setting, a *belief* is a probability distribution over some unknown state or attribute of the world. Beliefs, in this sense, may or may not be consciously accessible. A belief can be held with great *precision*, such that the probability distribution is concentrated over the most likely value—the mean or *expectation*. This means the precision (inverse variance) corresponds to the confidence or certainty associated with a belief. In Bayesian inference, beliefs prior to observing data are called *prior beliefs*, which are updated to form *posterior beliefs* after seeing the data. This updating rests upon combining a prior belief with sensory evidence or the *likelihood* of the data. In hierarchical Bayesian inference, the *sufficient statistics* of a belief (like the expectation and precision) are themselves treated as unknown quantities. This means that one can have beliefs about beliefs; for example, one can have an expectation about a precision (c.f., expected uncertainty). Beliefs about beliefs are inevitable in hierarchical inference and are sometimes referred to as *empirical priors*, because they provide constraints on beliefs at lower levels of the hierarchy. Behaviourally, precision and beliefs about precision (including subjective confidence in beliefs) are to some extent dissociable (Fleming, Dolan, and Frith 2012). Beliefs about precision are particularly important in hierarchical Bayesian inference, because they can have a profound effect on posterior expectations—and inappropriate beliefs about precision can easily lead to false inference:

The nature of this failure can be understood intuitively by considering classical statistical inference: Imagine that we are using a *t*-test to compare the mean of some data, against the null hypothesis that the mean is zero. The sample mean provides evidence against the null hypothesis in the form of a *prediction error*: namely, the sample mean minus the expectation under the null hypothesis. The sample mean provides evidence against the null—but how much evidence? This can only be quantified in relation to the precision of the prediction error. The *t*-statistic is simply the prediction error weighted by its precision (i.e., divided by its standard error). If this precision weighted prediction error is sufficiently large, one rejects the null hypothesis. Clearly, if we overestimate the precision of the data, the *t*-statistic will be too large and we expose ourselves to false positives. Analogous rules apply to Bayesian inference, in that the optimal combination of a prior belief with some evidence is a posterior belief whose mean is a mixture of the prior and data means,

weighted according to their precision. If the precision of the data is overestimated, or if the precision of the prior is underestimated, the posterior expectation will shift from the prior mean to the data mean (Figure 1).

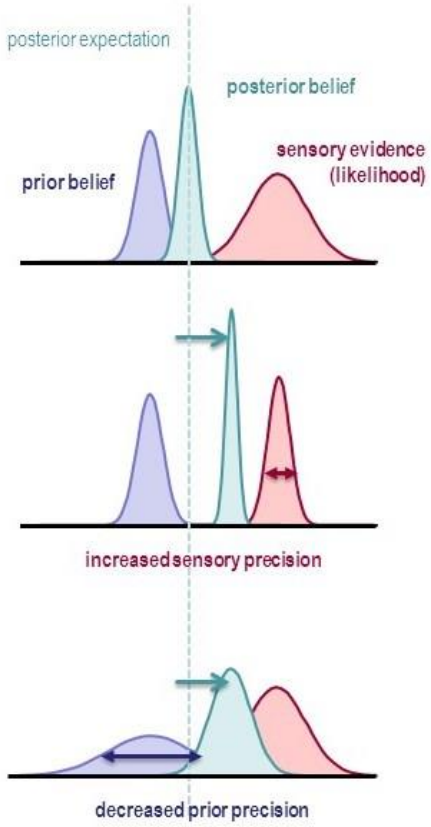

Figure 1: This schematic illustrates the importance of precision when forming posterior beliefs and expectations. The graphs show Gaussian probability distributions that represent prior beliefs, posterior beliefs and the likelihood of some data or sensory evidence as functions of some hidden (unknown) parameter. The dotted line corresponds to the posterior expectation, while the width of the distributions corresponds to their dispersion or variance. Precision is the inverse of this dispersion and can have a profound effect on posterior beliefs. Put simply, the posterior belief is biased towards the prior or sensory evidence in proportion to their relative precision. This means that the posterior expectation can be biased towards sensory evidence by either increasing sensory precision—or failing to attenuate it—or by decreasing prior precision. Reproduced from (Adams et al. 2013).

So how could this lead to false beliefs and delusions? The following scenario (Frith and Friston 2012) illustrates this: imagine the temperature warning light in your car is too sensitive (precise), reporting the slightest fluctuations (prediction errors) above some temperature. You naturally infer that there is something wrong with your car and take it to the garage. However, they find no fault—and yet the warning light continues to flash. Your first instinct may be to suspect the garage has failed to identify the fault—and even to start to question the Good Garage Guide that recommended it. From your point of view, these are all plausible hypotheses that accommodate the evidence available to you. However, from the perspective of somebody who has never seen your warning light, your suspicions would have an irrational and slightly paranoid flavour. This anecdote illustrates how delusional systems may be elaborated as a consequence of imbuing sensory evidence with too much precision. Crucially, there is no necessary impairment in forming predictions or predic-

tion errors—the problem lies in the way they are used to inform inference or hypotheses.

In what follows, we will consider the brain as performing inference using predictive coding, in which the evidence for hypotheses is reported by precision weighted prediction errors. In these schemes, certain neurons compare bottom-up inputs with top-down predictions to form a prediction error that is weighted in proportion to its expected precision. Crucially, this weighting corresponds to the gain or sensitivity of prediction error units. This means that abnormalities in the modulation of postsynaptic gain could, in principle, lead to false inferences of the sort described above. We illustrate this using a biologically plausible simulation of false inference, in a predictive coding scheme in which there is a decrease in the precision (postsynaptic gain of prediction error units) at higher levels of the cortical hierarchy, relative to the precision at sensory levels. These arguments and the simulations have been reported previously in different contexts (Adams et al. 2013; Brown et al. 2013).

This paper focuses on false inference. However, the normative principles we appeal to cover both *inference* and *learning*. The important thing here is that abnormal beliefs about precision also lead to *false learning*, which produces—and is produced by—false inference. In brief, a simple failure of neuromodulation (and implicit encoding of precision) can have far-reaching and knock-on effects that can be manifest at many different levels of perceptual inference, learning and consequent behaviour.

Psychosis and false inference

In this section, we briefly review the state and trait abnormalities of schizophrenia to emphasise a common theme; namely, a failure of inference about the world that arises from an imbalance in the precision or confidence attributed to beliefs. In this setting, state abnormalities include the florid (Schneiderian or first rank) symptoms of acute psychosis, while trait abnormalities are more pervasive and subtle. The diagnostic criteria for schizophrenia are based largely on state abnormalities, because they are easily and reliably detected. These include:

- Delusions and hallucinations: c.f., positive symptoms (Crow 1980) and the reality distortion of chronic schizophrenia (Liddle 1987).
- Thought disorder and catatonia (World Health Organization 1992; American Psychiatric Association 2000), where formal thought disorder is also characteristic of the disorganisation syndrome of chronic schizophrenia (Liddle 1987).

Other (as yet non-diagnostic) state abnormalities include:

- Abnormalities of perceptual organization: in particular, a decreased influence of context, leading to a loss of global (Gestalt) organization (Phillips and Silverstein 2003). These abnormalities have not been found in first degree relatives or before the first psychotic episode, and tend to covary with disorganisation symptoms (Silverstein and Keane 2011). A decreased influence of context can sometimes lead to perceptions that are more veridical than those of normal subjects. Important examples here include a resistance to the hollow mask illusion—which is also state-dependent (Keane et al. 2013)—and the size-weight illusion (Williams et al. 2010).

These symptoms can occur episodically and—with the possible exception of catatonia—respond well to anti-dopaminergic drugs in the majority of patients. We use the term 'trait' abnormalities to refer to more constant features of the disorder, which are less responsive to dopamine blockade (although these responses have not been explored as thoroughly as those of state symptoms). Some are found in first degree relatives and high-risk groups, and may qualify as endophenotypes of schizophrenia. Despite their prevalence, they are less diagnostic because they are found in other diagnostic categories (and to some extent in the normal population). They include (among others):

- Soft neurological signs: probably best exemplified by abnormalities of smooth pursuit eye movements (SPEM) as reviewed by (O'Driscoll and Callahan 2008). These abnormalities are present in first-degree relatives (Calkins et al. 2008) and in drug naive first episode schizophrenics (Hutton et al. 1998; Campion et al. 1992; Sweeney et al. 1994), and may even be exacerbated by dopamine blockade (Hutton et al. 2001).
- Abnormal event-related potentials: such as a larger P50 response to a repeated stimulus, and reduced P300 and mismatch negativity (MMN) responses to violations or oddball stimuli. Abnormal P50, P300 and MMN responses have also been demonstrated in first-degree relatives, and do not normalise with treatment(reviewed in (Winterer and McCarley 2011).
- Anhedonia, cognitive impairments, and negative symptoms: such as loss of normal affect, experience of pleasure, motivation and sociability are all found (subclinically) in first-degree relatives (Fanous et al. 2001; Jabben et al. 2010) to a greater or lesser degree (Mockler et al. 1997; Johnstone et al. 1987) and are notoriously resistant to anti-dopaminergic treatment.

We distinguish between state and trait abnormalities in part because the evidence suggests that trait abnormalities may be associated with a relative decrease in prior precision (or failure to attenuate sensory precision), while state abnormalities may have more complex and heterogeneous origins. For example, some state abnormalities may result from the straightforward exac-

erbation of trait abnormalities, e.g., progressive reduction in prior precision leading to perceptual disorganisation, or delusional mood (see below). Other state abnormalities can be explained by a (possibly compensatory) increase in prior precision (or reduction in sensory precision); e.g. somatic delusions resulting from compensations for loss of sensory attenuation (Adams et al. 2013).

Many trait abnormalities have been considered as the result of a failure to adequately predict sensory input, rendering all percepts surprising (e.g., the P50) and reducing differential responses to oddball stimuli (e.g., the MMN and P300). Predictive coding in particular has been used in recent formulations of these deficits in schizophrenia (Fletcher and Frith 2009). Specifically, it is suggested that the main problem in schizophrenia lies not with the prediction of sensory input *per se*, but in the delicate balance of precision ascribed to prior beliefs and sensory evidence (Friston 2005; Corlett et al. 2011).

In terms of cognitive paradigms, the 'beads task' has been used to characterise formal beliefs and probabilistic reasoning in schizophrenic subjects. In this paradigm, subjects are told that red and green beads are drawn at random from an urn that contains (for example) 85% of one colour and 15% of the other. The subject must decide which colour predominates. In reality, all subjects are shown the same sequence of beads. In the *draws to decision* version of the task, the subject has to answer as soon as they are certain. In the *probability estimates* version, the subject can continue to draw and change their answer. Interestingly, delusional patients 'jump to conclusions' in the first version, while they are more willing to revise their decision in light of contradictory evidence in the second (Garety and Freeman 1999). Bayesian modelling suggests that jumping to conclusions may reflect greater 'cognitive noise' in delusional patients (Moutoussis et al. 2011), which may speak to reduced precision of higher level representations and consequently a greater influence of new sensory evidence (Speechley et al. 2010).

Can state abnormalities also be explained by imbalances in the precisions of prior beliefs and sensations? The short answer is yes. For example, delusional mood describes a state in which patients feel the world is strange and has changed in some way—where their attention is drawn to apparently irrelevant stimuli and odd coincidences. A loss of precise prior beliefs is consistent with a sense of unpredictability and greater attention to sensory events. The notion of attentional deficits fits comfortably with recent formulations of attention in terms of the selective augmentation or attenuation of sensory precision or gains (see below). In other words, a failure of sensory attenuation in the perceptual domain would look very much like a disorder of attention.

State abnormalities include the cardinal psychotic symptoms, such as hallucinations and delusions. Hallucinations could be understood as the result of an *increase* in the relative precision of prior beliefs, such that the posterior be-

liefs are impervious to contradictory—but imprecise—sensory evidence. This has been discussed as an explanation for visual hallucinosis in organic psychosyndromes (Friston 2005). However, the hallucinations associated with psychosis may be better understood as a failure to attenuate the sensory consequences (corollary discharge) of self-made acts; for example, a failure to attenuate the auditory consequences of sub-vocal or inner speech (Frith et al. 1998; Allen et al. 2007). Delusions are probably more complex and their emergence may be better understood as secondary phenomena: several authors (Fletcher and Frith 2009) have proposed that they could arise as rational (Bayes optimal) posterior beliefs that explain away precise sensory prediction errors. These explanations relate to earlier 'empiricist' accounts (Maher 1974; Gray et al. 1991; Kapur 2003), that emphasised aberrant salience (c.f., sensory precision). Implicit in these secondary accounts is a compensatory increase in the precision of explanations for sensory cues that are imbued with too much precision or salience. This is consistent with their peculiar resistance to rational argument.

In summary, the symptoms and signs of schizophrenia are not inconsistent with a reduction of high-level precision or a failure of sensory attenuation (the top-down attenuation of sensory precision), with compensatory (secondary) changes in the precision of (empirical) prior beliefs. In particular, some psychotic states may reflect a compensatory response to trait abnormalities that bias inference towards sensory evidence, which is imbued with too much precision or salience. A further mechanistic dissociation between state and trait abnormalities is suggested by the fact that the former generally respond to antipsychotic (anti-dopaminergic) treatment, while trait abnormalities do not. Before considering the computational anatomy of hierarchical inference in the brain, we will briefly review the psychopharmacology and neuropathology of schizophrenia.

The psychopharmacology of precision

This section considers the neuromodulatory processes implicated in schizophrenia. Our premise here is that psychotic abnormalities are manifestations of false inference, caused by the aberrant encoding of precision. This precision is thought to be encoded by postsynaptic gain of neuronal populations reporting prediction errors—the principal or pyramidal cells of superficial cortical layers (Mumford 1992; Feldman and Friston 2010). Synaptic gain modulation is a change in the response amplitude of a neuron that is independent of its selectivity or receptive field characteristics (Salinas and Thier 2000). In other words, postsynaptic gain is a factor that quantifies the effect of a presynaptic input on postsynaptic output (e.g. depolarisation at the soma). Changes in synaptic gain are generally thought to be mediated

by nonlinear (e.g., multiplicative) synaptic mechanisms; for example, NMDA receptor activation.

Of all the receptors that determine synaptic gain, the most ubiquitous is the glutamatergic NMDA receptor (NMDA-R). NMDA-Rs have several important functions that are expressed over different timescales. First, they can drive (i.e. induce an excitatory postsynaptic potential) postsynaptic cells like other ionotropic glutamatergic (AMPA and Kainate) receptors. However, the driving effect of NMDA-Rs is only possible if the cell is already depolarised; otherwise, the NMDA-R is blocked by a magnesium ion. This nonlinear property makes them synaptic coincidence detectors or 'AND gates'. Second, NMDA-Rs have time constants that are much longer than that of AMPA-Rs and Kainate-Rs. This enables integration of synaptic inputs over tens to hundreds of milliseconds—increasing the gain of synaptic inputs to distal dendrites. Finally, NMDA-Rs are famous for their role in plasticity: at longer timescales, the influx of calcium ions through NMDA-R channels causes a cascade of intracellular events that result in long-term synaptic depression or potentiation (LTD or LTP). However, NMDA-Rs also have a major impact on the short-term plasticity of glutamatergic synapses. This is because they regulate the functional state and number of AMPA-Rs—by phosphorylation or by changing the trafficking of AMPA-R subunits to and from the cell membrane (Bagal et al. 2005; Montgomery and Madison 2004; Passafaro et al. 2001). Together, these properties make a significant contribution to the dynamics of neural networks, especially to oscillatory behaviour and sustained firing patterns (Durstewitz 2009).

Other key determinants of synaptic gain are the classical neuromodulator receptors; e.g., dopamine (DA-Rs), acetylcholine (in particular muscarinic AChRs) and serotonin, (5-HTRs). With the exception of nicotinic AChRs (which are ionotropic) these are all metabotropic receptors—they do not activate ion channels but are coupled to signal transduction mechanisms (via G proteins) that affect intracellular second messengers, such as cyclic adenosine monophosphate (cAMP) or cyclic guanosine monophosphate (cGMP). Fluctuations in cAMP/cGMP concentration affect the activity of protein kinases, which—through phosphorylation—alters neuronal excitability via changes in the production, surface expression or activity of voltage or ligand-gated ion channels, including the NMDA-R itself. It is important to note that DA-R subtypes have opposite effects on synaptic gain: D_1R activation stimulates cAMP production and increases the excitability of depolarised neurons, whereas D_2R activation inhibits cAMP production and reduces gain(reviewed in (Frank 2005).

Synaptic gain is not just determined by receptor activity but also by network dynamics, like the synchronization of fast oscillations, especially in the 40-100Hz or gamma frequencies (c.f., synchronous gain (Chawla et al. 1999)). The fast acting inhibitory γ-amino butyric acid receptor ($GABA_A$-R) is instrumental in this synchronization process. In the cortex, a GABAergic (parvalbumin-

positive basket cell or PVBC) interneuron contacts many pyramidal cells, which it transiently hyperpolarises. When this hyperpolarization wears off, all the cortical pyramidal cells can then fire together, leading to synchronous firing across the network and oscillations as the cycle recurs (Gonzalez-Burgos and Lewis 2008).

Abnormalities in at least three of these synaptic gain mechanisms have been proposed to be a primary pathology in schizophrenia—those of NMDA, GABA, and dopamine. NMDA-Rs play a central role in theories of schizophrenia (Olney and Farber 1995; Abi-Saab et al. 1998; Goff and Coyle 2001; Stephan et al. 2006; Corlett et al. 2011). Studies of genetic risk in schizophrenia have highlighted the role of genes related to glutamatergic transmission, with GABA and dopamine related genes implicated to a lesser extent (Stephan et al. 2006; Greenwood et al. 2012; Harrison and Weinberger 2005; Hall et al. 2009). Neuropathological evidence indicates abnormalities of the glutamate and GABA systems: both pre- and post-synaptic markers, morphometric and biochemical measures of glutamatergic transmission are reduced, as is the expression of the GABA synthesizing enzyme glutamic acid decarboxylase (GAD), parvalbumin-immunoreactive GABAergic interneurons and their synaptic markers (Harrison et al. 2011). These neuropathological changes are particularly apparent in hippocampus and frontal cortex, both at high levels in the cortical hierarchy (Felleman and Van Essen 1991).

Conversely, the evidence for dopaminergic abnormalities in schizophrenia is neither neuropathological nor structural, but functional. The most widely replicated abnormality is that of elevated striatal dopamine availability—in acute psychoses of both schizophrenia (Laruelle et al. 1996; Breier et al. 1997) and epilepsy (Reith et al. 1994). A recent review concluded that dopamine dysregulation is more closely linked to the state of psychosis than the trait of schizophrenia (Howes and Kapur 2009), although there are some important caveats: presynaptic dopamine is also raised to a lesser degree in those who are prone to psychosis but not floridly psychotic, and patients with symptoms resistant to dopamine blockade do not have elevated striatal dopamine synthesis (Demjaha et al. 2012).

Is aberrant glutamatergic and GABAergic transmission linked to the trait abnormalities of the previous section? The psychotomimetic effects of ketamine suggest a strong association. Ketamine blocks NMDA-Rs and also potentiates AMPA-R signalling, leading to decreased burst firing of pyramidal neurons, with subsequent impairment of activation of GABAergic interneurons (Shi and Zhang 2003). Ketamine administration can reproduce a whole spectrum of trait phenomena: such as smooth pursuit eye movement (SPEM) abnormalities (Radant et al. 1998; Weiler et al. 2000); impaired P50 suppression (Oranje et al. 2002); diminished P300 (Gunduz-Bruce et al. 2012); reduced MMN (Umbricht et al. 2000; Schmidt et al. 2012); cognitive impairments (Kantrowitz and

Javitt 2010) and negative symptoms (Krystal et al. 1994). In fact, the only trait phenomenon that ketamine does not reproduce is a reduced susceptibility to the hollow mask illusion (Passie et al. 2003). This is in contrast to dopaminergic agonists, which do not reproduce perceptual, SPEM (Reilly et al. 2008), P50 (B Oranje et al. 2004) or MMN (Leung et al. 2007) abnormalities—and have only small effects on the P300 (Luthringer et al. 1999). Indeed, prefrontal D_1R hypoactivity has been associated with cognitive deficits and negative symptoms in animal models (Goldman-Rakic et al. 2004).

Ketamine's reproduction of state symptoms is less consistent: its effects include loss of perceptual organization (Uhlhaas et al. 2007) and induction of a delusional mood (Corlett et al. 2011), but it does not cause a loss of attenuation of self-induced sensations (PC Fletcher, personal communication) or lead to auditory verbal hallucinations. It is interesting to note that while the negative symptoms induced by ketamine are correlated with its NMDA-R binding, the positive symptoms are not (Stone et al. 2008). Conversely, D_2R levels in cortical and striatal areas correlate with positive but not negative symptom scores (Kessler et al. 2009). Nevertheless, some trait-like phenomena can be reproduced by both ketamine and dopaminergic agonists, such as reduced latent inhibition (Razoux, Garcia, and Léna 2007; Young, Moran, and Joseph 2005), blocking (Freeman et al. 2013; O'Tuathaigh et al. 2003) and the body ownership illusion (Morgan et al. 2011; Albrecht et al. 2011). This is not surprising, as there are complex interactions between glutamatergic, GABAergic and dopaminergic neurotransmission, within and between the brainstem, striatum and prefrontal cortex (see Figure 2). For example, in the prefrontal cortex, NMDA-R impairments may lead to hypofunction of GABAergic PVBC's, disinhibition of pyramidal cells and reduced prefrontal gamma activity (Gonzalez-Burgos and Lewis 2012).

Crucially, the neuropathology of schizophrenia is associated with higher (in hierarchical terms) cortical systems; e.g. prefrontal cortex and the medial temporal lobe. For example, perceptual deficits in schizophrenics (and normal subjects) have been shown to correlate with frontal and temporal volume loss (Dazzan et al. 2006). This has important implications for the computational modelling of psychotic symptoms, because impairment in precision-encoding at higher cortical levels will reduce the influence of prior beliefs on inference, as we now illustrate.

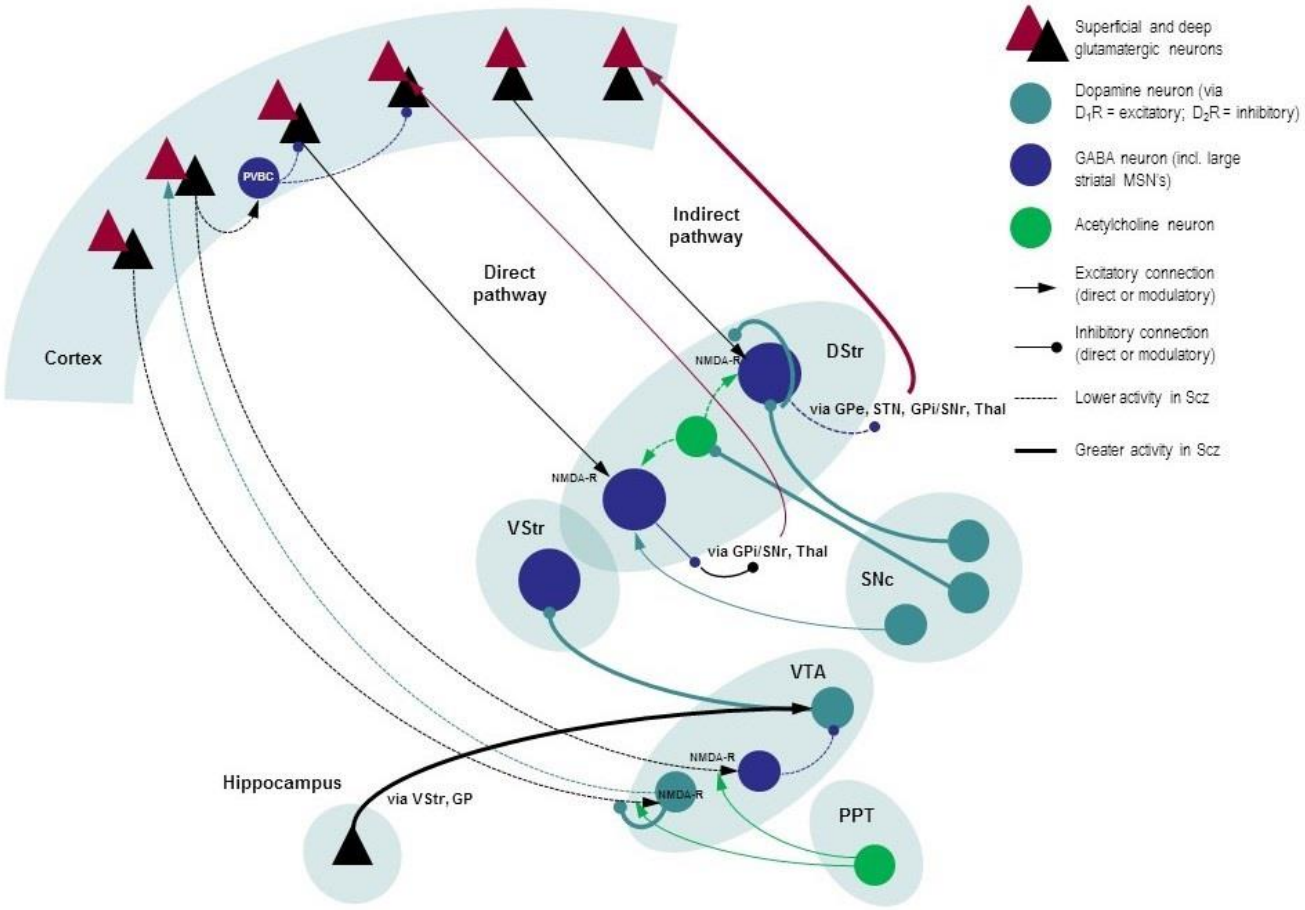

Figure 2: A schematic illustration of putative pathological processes in schizophrenia—emphasising the interactions among neuromodulatory mechanisms. These mechanisms include: (i) decreased prefrontal NMDA-R function that may reduce the stimulation of VTA-DA neurons that project back to prefrontal D_1Rs (decreasing cortical precision), and disinhibition of VTA-DA neurons that project to the striatum; (ii) increased dopamine release from SNc-DA neurons disinhibits the indirect pathway (by direct inhibition of striatal GABA neurons, inhibition of striatal cholinergic interneurons, and reduction of glutamate release in corticostriatal neurons); (iii) reduced NMDA-R stimulation of cortical PVBC's reduces activity of these GABAergic interneurons, impairing coordination of cortical oscillatory activity; and (iv) increased hippocampal drive to the VTA, leading to hyperdopaminergia in the VStr. Significant omissions (for clarity) include: the GP, SNr, STN and Thal, most connections of the VStr including its direct and indirect pathways and excitatory connections from the VTA (via D_1Rs), and circuitry within the VStr, two more inhibitory connections in the indirect pathway and both somatic and axonal dopamine neuron D_2 autoreceptors in SNc. As in other figures, descending projections are in black and ascending projections in red. Abbreviations: PPT—pedunculopontine tegmental nucleus, VTA—ventral tegmental area, VStr—ventral striatum, DStr—dorsal striatum, SNc/r—substantia nigra pars compacta/reticulata, GP—globus pallidus, Thal—thalamus, STN—subthalamic nucleus, PVBC—parvalbumin-positive basket cell. Sources: (Stephan, Friston, and Frith 2009; Morrison 2012; Carlsson, Waters, and Carlsson 1999; Lisman et al. 2008; Simpson, Kellendonk, and Kandel 2010), reproduced from (Adams et al. 2013).

Neurobiological implementation of active inference

This section introduces the theory behind inference in the brain. This normative account provides key constraints on the functional (computational) anatomy of action and perception. This allows one to understand (and simulate) inference in a principled way—that is also grounded in neuroanatomy and neurophysiology. We will use the formalism below to simulate some of the schizophrenic abnormalities reviewed above. These simulations rest on descriptions of the neuronal processes (differential equations) that underwrite inference in the brain. These equations are based on three assumptions:

- The brain minimises the free energy of sensory inputs defined by a generative model.
- The generative model used by the brain is hierarchical, nonlinear and dynamic.
- Neuronal firing rates encode the expected state of the world, under this model.

The first assumption is the free energy principle, which leads to active inference in the embodied setting of action (Friston et al. 2010). This provides a normative (Bayes-optimal) account of action and perception, in which both minimise a free energy bound on the (negative log) evidence for the brain's model of the world. Free energy is a quantity from statistics that measures the quality of a model in terms of the probability that it could have generated observed outcomes. This means that minimising free energy maximises the Bayesian evidence for the generative model (Ballard et al. 1983; Hinton and van Camp 1993; Dayan et al. 1995). The second assumption is motivated by noting that the world is both dynamic and nonlinear and that hierarchical causal structure emerges inevitably from a separation of temporal scales (Ginzburg 1955; Haken 1983). The final assumption is the Laplace assumption that, in terms of neural codes, leads to the *Laplace code* that is arguably the simplest and most flexible of all neural codes (Friston 2009).

Given these assumptions, one can simulate a whole variety of neuronal processes by specifying the particular equations that constitute the brain's generative model. Action and perception are then specified completely by the above assumptions and can be implemented in a biologically plausible fashion. In brief, these simulations use differential equations that minimise the free energy of sensory input using a generalised gradient descent (Friston et al. 2010).

(1)

$$\dot{\tilde{\mu}}(t) = D\tilde{\mu}(t) - \partial_{\tilde{\mu}} F(\tilde{s}, \tilde{\mu})$$

$$\dot{a}(t) = -\partial_a F(\tilde{s}, \tilde{\mu})$$

These coupled differential equations describe perception and action respectively. They say that neuronal activity encoding posterior expectations about (generalised) hidden states of the world $\tilde{\mu} = (\mu, \mu', \mu'', \dots)$ and action a reduce free energy—where free energy $F(\tilde{s}, \tilde{\mu})$ is a function of (generalised) sensory inputs $\tilde{s} = (s, s', s'', \dots)$ and neuronal activity. The first differential equation is known as generalised predictive coding or Bayesian filtering: see also Rao and Ballard (1999). The first term is a prediction based upon a differential matrix operator D that returns the generalised motion of expected hidden states. The second (correction) term is usually expressed as a mixture of prediction errors that ensures the changes in posterior expectations are Bayes-optimal predictions about hidden states of the world. The second differential equation says that action also minimises free energy. The differential equations above are coupled because sensory input depends upon action, which depends upon perception through the posterior expectations. This circular dependency leads to a sampling of sensory input that is both predicted and predictable, thereby minimising free energy and, implicitly, prediction errors.

To perform neuronal simulations under this scheme, it is only necessary to integrate or solve Equation (1) to simulate the neuronal dynamics that encode posterior expectations and associated action. Posterior expectations depend upon the brain's generative model of the world, which we assume has the following hierarchical form:

(2)

$$
\begin{aligned}
s &= g^{(1)}\left(x^{(1)}, v^{(1)}\right) + \omega_v^{(1)} \\
\dot{x}^{(1)} &= f^{(1)}\left(x^{(1)}, v^{(1)}\right) + \omega_x^{(1)} \\
&\vdots \\
v^{(i-1)} &= g^{(i)}\left(x^{(i)}, v^{(i)}\right) + \omega_v^{(i)} \\
\dot{x}^{(i)} &= f^{(i)}\left(x^{(i)}, v^{(i)}\right) + \omega_x^{(i)} \\
&\vdots \\
\omega_x^{(i)} &\sim N(0, \Pi_x^{(i)-1}) \\
\omega_v^{(i)} &\sim N(0, \Pi_x^{(i)-1}) \\
\Pi_x^{(i)} &= \exp(\pi_x^{(i)}\left(x^{(i)}, v^{(i)}\right)) \\
\Pi_v^{(i)} &= \exp(\pi_v^{(i)}\left(x^{(i)}, v^{(i)}\right))
\end{aligned}
$$

This equation describes a probability density over the sensory and hidden states that generate sensory input. Here, the hidden states have been divided into hidden states and causes $(x^{(i)}, v^{(i)})$, with (i) denoting their level within the hierarchical model. Hidden states and causes are abstract variables that the brain uses to explain or predict sensations—like the motion of an object in the field of view. In these models, hidden causes link hierarchical levels, whereas hidden states link dynamics over time. Here, $(g^{(i)}, f^{(i)})$ are nonlinear functions of hidden states and causes that generate hidden causes for the level below and—at the lowest level—sensory inputs. Random fluctuations in the motion of hidden states and causes $(\omega_x^{(i)}, \omega_v^{(i)})$ enter each level of the hierarchy. Gaussian assumptions about these random fluctuations make the model probabilistic. They play the role of sensory noise at the first level and induce uncertainty at higher levels. The amplitudes of these random fluctuations are quantified by their precisions $(\Pi_x^{(i)}, \Pi_v^{(i)})$ that may depend upon the hidden states or causes through their log-precisions $(\pi_x^{(i)}, \pi_v^{(i)})$.

Perception and predictive coding

Given the form of the generative model (Equation 2) we can now write down the differential equations (Equation 1) describing neuronal dynamics in terms of (precision-weighted) prediction errors on the hidden causes and states. These errors represent the difference between posterior expectations and predicted values, under the generative model (using $\boldsymbol{A} \cdot \boldsymbol{B} := \boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{B}$ and omitting higher-order terms):

(3)

$$\dot{\tilde{\mu}}_x^{(i)} = D\tilde{\mu}_x^{(i)} + \left(\frac{\partial \tilde{g}^{(i)}}{\partial \tilde{\mu}_x^{(i)}} - \frac{1}{2}\tilde{\varepsilon}_v^{(i)}\frac{\partial \tilde{\pi}_v^{(i)}}{\partial \tilde{\mu}_x^{(i)}}\right)\cdot \xi_v^{(i)} + \left(\frac{\partial \tilde{f}^{(i)}}{\partial \tilde{\mu}_x^{(i)}} - \frac{1}{2}\tilde{\varepsilon}_x^{(i)}\frac{\partial \tilde{\pi}_x^{(i)}}{\partial \tilde{\mu}_x^{(i)}}\right)\cdot \xi_x^{(i)} + \frac{\partial tr\left(\tilde{\pi}_v^{(i)}+\tilde{\pi}_x^{(i)}\right)}{\partial \tilde{\mu}_x^{(i)}} - D^T\xi_x^{(i)}$$

$$\dot{\tilde{\mu}}_v^{(i)} = D\tilde{\mu}_v^{(i)} + \left(\frac{\partial \tilde{g}^{(i)}}{\partial \tilde{\mu}_v^{(i)}} - \frac{1}{2}\tilde{\varepsilon}_v^{(i)}\frac{\partial \tilde{\pi}_v^{(i)}}{\partial \tilde{\mu}_v^{(i)}}\right)\cdot \xi_v^{(i)} + \left(\frac{\partial \tilde{f}^{(i)}}{\partial \tilde{\mu}_x^{(i)}} - \frac{1}{2}\tilde{\varepsilon}_x^{(i)}\frac{\partial \tilde{\pi}_x^{(i)}}{\partial \tilde{\mu}_v^{(i)}}\right)\cdot \xi_x^{(i)} + \frac{\partial tr\left(\tilde{\pi}_v^{(i)}+\tilde{\pi}_x^{(i)}\right)}{\partial \tilde{\mu}_v^{(i)}} - \xi_x^{(i+1)}$$

$$\xi_x^{(i)} = \tilde{\Pi}_x^{(i)}\tilde{\varepsilon}_x^{(i)} = \tilde{\Pi}_x^{(i)}(D\tilde{\mu}_x^{(i)} - \tilde{f}^{(i)}(\tilde{\mu}_x^{(i)}, \tilde{\mu}_v^{(i)}))$$

$$\xi_v^{(i)} = \tilde{\Pi}_v^{(i)}\tilde{\varepsilon}_v^{(i)} = \tilde{\Pi}_v^{(i)}(D\tilde{\mu}_v^{(i-1)} - \tilde{g}^{(i)}(\tilde{\mu}_x^{(i)}, \tilde{\mu}_v^{(i)}))$$

Equation (3) can be derived by computing the free energy for the hierarchical model in Equation (2) and inserting its gradients into Equation (1). This produces a relatively simple update scheme, in which posterior expectations are driven by a mixture of prediction errors, where prediction errors are defined by the equations of the generative model.

It is difficult to overstate the generality of Equation (3): its solutions grandfather nearly every known statistical estimation scheme, under parametric assumptions about additive or multiplicative noise (Friston, 2008). These range

from ordinary least squares to advanced variational deconvolution schemes. The scheme is called *generalised Bayesian filtering* or *predictive coding* (Friston et al. 2010). In neural network terms, Equation (3) says that error-units ($\xi_v^{(i)}$) compute the difference between expectations at one level ($\tilde{\mu}_v^{(i-1)}$) and predictions from the level above ($\tilde{g}^{(i)}(\tilde{\mu}_x^{(i)}, \tilde{\mu}_v^{(i)})$). Conversely, posterior expectations (encoded by the activity of state units) are driven by prediction errors from the same level and the level below. These constitute bottom-up and lateral messages that drive posterior expectations towards a better prediction to reduce the prediction error in the level below. This is the essence of recurrent message passing between hierarchical levels to optimise free energy or suppress prediction error: see Friston and Kiebel (2009b) and Feldman and Friston (2010) for a more detailed discussion. Crucially, in neurobiological implementations of this scheme, the sources of bottom-up prediction errors have to be superficial pyramidal cells, because it is these—and only these—cells that send forward (ascending) connections to higher cortical areas. Conversely, predictions are conveyed from deep pyramidal cells, by backward (descending) connections, to target the superficial pyramidal cells encoding prediction error (Mumford 1992; Bastos et al. 2012): see Figure 3.

Note that the precisions depend on the expected hidden causes and states. We have proposed that this dependency mediates attention and action selection in hierarchical processing (Feldman and Friston 2010; Friston et al. 2012). Equation (3) tells us that the (state-dependent) precisions ($\tilde{\Pi}_v^{(i)}, \tilde{\Pi}_x^{(i)}$) modulate the responses of prediction error units to their presynaptic inputs. This modulation depends on the posterior expectations about the states and suggests something intuitive—attention is mediated by activity-dependent modulation of the synaptic gain of principal cells that convey sensory information (prediction error) from one cortical level to the next. This translates into a top-down control of synaptic gain in principal (superficial pyramidal) cells elaborating prediction errors and fits comfortably with the modulatory effects of top-down connections in cortical hierarchies that have been associated with attention and action selection.

Action

In active inference, posterior expectations elicit behaviour by sending top-down predictions down the hierarchy that are unpacked into proprioceptive predictions at the level of the cranial nerve nuclei and spinal cord. These engage classical reflex arcs to suppress proprioceptive prediction errors and produce the predicted motor trajectory

$$\dot{a} = \frac{\partial}{\partial a} F = -\frac{\partial \tilde{s}}{\partial a} \cdot \xi_v^{(i)} \qquad (4)$$

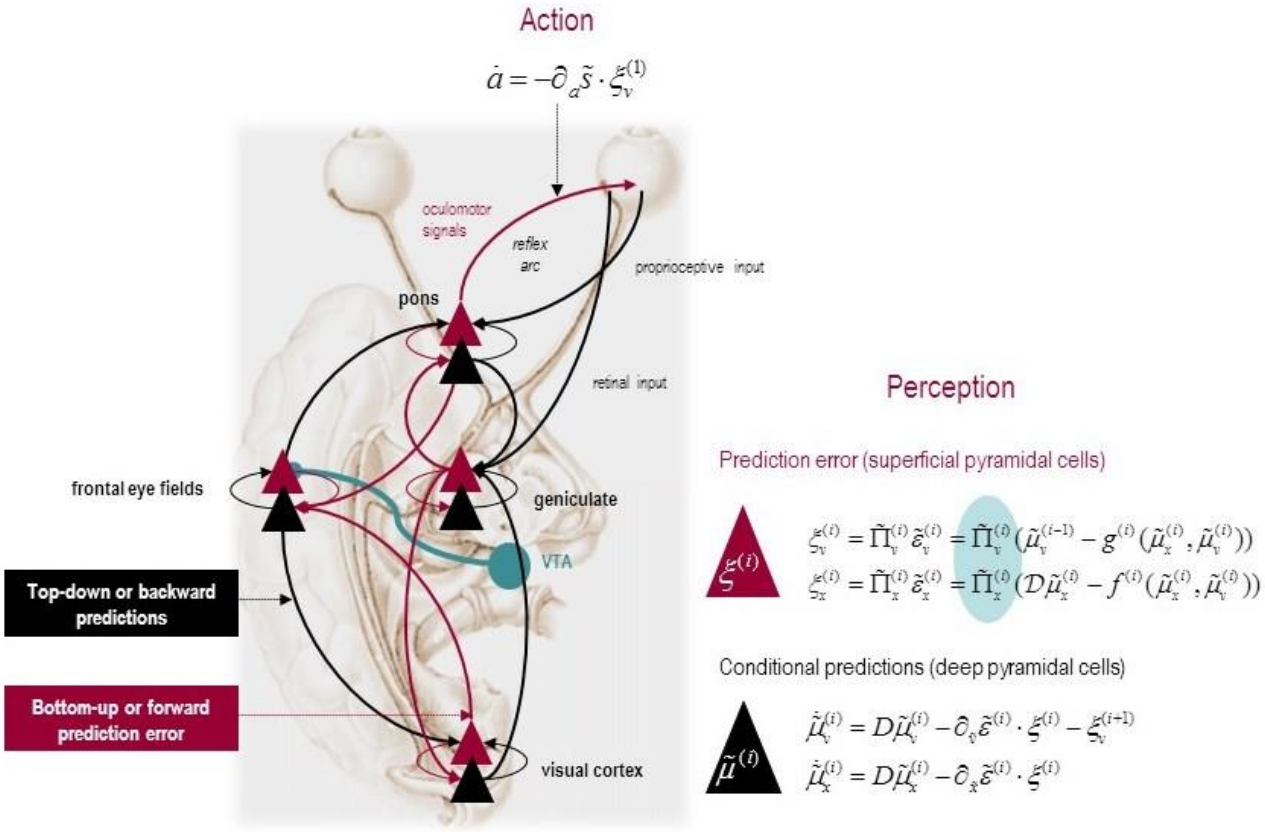

Figure 3: Hierarchical message passing in the visual-oculomotor system: the schematic illustrates a neuronal message passing scheme (generalised Bayesian filtering or predictive coding) that optimises posterior expectations about hidden states of the world, given sensory (visual) data and the active (oculomotor) sampling of those data. It shows the speculative cells of origin of forward driving connections (in red) that convey prediction errors from a lower area to a higher area and the backward connections (in black) that construct predictions. These predictions try to explain away prediction error in lower levels. The sources of forward and backward connections are superficial (red) and deep (black) pyramidal cells respectively. The cyan connection denotes a neuromodulatory connection from the ventral tegmental area (VTA) which mediates estimates of precision. The equations on the right represent a generalised descent on free energy under a hierarchical model—this can be regarded as a generalisation of predictive coding or Bayesian (e.g., Kalman-Bucy) filtering. In neural network terms, the equations say that error-units ($\xi_v^{(i)}$) compute the difference between expectations at one level ($\tilde{\mu}_v^{(i-1)}$) and a non-linear function of predictions from the level above ($\tilde{g}^{(i)}(\tilde{\mu}_x^{(i)}, \tilde{\mu}_v^{(i)})$). Conversely, posterior expectations ($\dot{\tilde{\mu}}_v^{(i)}$) in state units are driven by prediction errors from the same level ($\partial_{\tilde{v}} \tilde{\varepsilon}^{(i)} \cdot \xi^{(i)}$) and the level below ($\xi_v^{(i+1)}$).This is the essence of recurrent message passing between hierarchical levels to optimise free energy or suppress prediction error. State-units are in black and error-units are in red. The cyan circle highlights where precisions ($\tilde{\Pi}_v^{(i)}, \tilde{\Pi}_x^{(i)}$) enter these equations—to modulate prediction error units (superficial pyramidal cells) such that they report precision weighted prediction errors. In this schematic, we have placed different levels of a hierarchical model within the visual-oculomotor system. Visual input arrives in an intrinsic (retinal) frame of reference that depends on the direction of gaze. Exteroceptive input is then passed to the lateral geniculate nuclei (LGN) and to higher visual and prefrontal (e.g., frontal eye fields) areas in the form of prediction errors. Crucially, proprioceptive sensations are also predicted, creating prediction errors at the level of the cranial nerve nuclei (pons). The special aspect of these proprioceptive prediction errors is that

they can be resolved in one of two ways: top-down predictions can change or the errors can be resolved through classical reflex arcs—in other words, they can elicit action to change the direction of gaze and close the visual–oculomotor loop. Reproduced from (Adams et al. 2013).

The reduction of action to classical reflexes follows because the only way that action can minimise free energy is to change sensory (proprioceptive) prediction errors by changing sensory signals; cf., the equilibrium point formulation of motor control (Feldman and Levin 1995). In short, active inference can be regarded as equipping a generalised predictive coding scheme with classical reflex arcs: see Friston et al. (2010), Friston et al. (2009) and Adams et al. (2012) for details. The actual movements produced clearly depend upon top-down predictions that can have a deep and complex structure, as we will see later.

Summary

In summary, starting with the assumption that the brain is trying to maximise the evidence for its model of the world, one can derive plausible equations describing neuronal dynamics in terms of message passing among different levels of a (cortical) hierarchical model. These messages comprise precision weighted prediction errors that are passed forward from one level to the next and top-down predictions that are reciprocated to minimise prediction error. In this scheme, precision is encoded by the gain of superficial pyramidal cells reporting prediction error, which is implicated in the synaptic pathology of schizophrenia. This is a straightforward consequence of the mathematical form of predictive coding and the fact that superficial pyramidal cells are the source of ascending connections in the brain. At the proprioceptive level, prediction errors can be reduced either by changing predictions (perception) or by changing sensations (action). In the last three sections, we use Equations (3) and (4) to simulate active inference under a number of generative models, while manipulating the precision at different hierarchical levels. These models are described completely by the equations in (2), which are provided in figures that summarise the generative model used in each example.

Sensory attenuation, attribution of agency and delusions

This section uses a generative model of (somatosensory) sensations that could be generated internally or externally. This model is used to illustrate the perceptual consequences of sensory attenuation, in terms of estimating the magnitude of externally and internally generated events. In brief, we reproduce the force matching illusion (Shergill et al. 2003; Shergill et al. 2005) by yoking externally applied forces to the perceived level of self-generated forces. Final-

ly, we demonstrate the disappearance of the illusion and the emergence of false inferences about (antagonistic) external forces, when there is a failure to attenuate sensory precision and a compensatory increase in the precision of empirical prior beliefs.

Active inference and sensory attenuation

Sensory attenuation refers to a decrease in the intensity of a perceived stimulus when it is self-generated (Blakemore et al. 1998). We have suggested that sensory attenuation is necessary to allow reflex arcs to operate (Brown et al. 2013). The argument is simple: proprioceptive prediction errors can only be resolved by moving—via motor reflexes—or by changing predictions. This means the effects of ascending prediction errors on posterior expectations must be attenuated to allow movement: if proprioceptive sensations are conveyed by ascending primary (Ia and Ib) sensory afferents with too much precision, then they would subvert descending predictions that create prediction errors and therefore prevent movement. It is therefore necessary to temporarily suspend the precision of sensory reafference to permit movement. If we associate the perceived intensity or detectability of the sensory consequences of action with a lower bound on their posterior confidence interval, attenuation of sensory precision provides a simple explanation for the attenuation of the perceived intensity of self-generated sensations. In what follows, we present simulations of sensory attenuation by simulating the force-match illusion and then demonstrate how overly precise prior beliefs can compensate for a failure of sensory attenuation but expose the actor to somatic delusions.

The generative process and model

Figure 4 summarises the generative process and model (using the form of Equation 2). This model is as simple as we could make it, while retaining the key ingredients that are required to demonstrate inference about or attribution of agency. The equations on the left describe the real world, while the equations on the right constitute the subject's generative model. In the real world, there is one hidden state x_i modelling self-generated force that is registered by both proprioceptive s_p and somatosensory s_s inputs. Externally generated forces v_e are added to internally generated forces to provide somatosensory input. The key thing about this model is that somatosensory sensations are caused ambiguously, by either internally or externally generated forces: $s_s = x_i + v_e$. The only way that the underlying cause of the sensations can be inferred is by reference to proprioceptive input—that is only generated internally. This is a very simple model, where the somatosensory input is used metaphorically to represent the sensory consequences of events that could be caused by self or others, while proprioceptive input represents signals that

can only be caused by self-made acts. Active inference now compels the subject to infer the causes of its sensations:

The generative model used for this inference is shown on the right. In this model, internally and externally generated forces (x_i, x_e) are modelled symmetrically, where changes in both are attributed to internal and external hidden causes (v_i, v_e). The hidden causes trigger the dynamics associated with the hidden states, much like the push that sets a swing in motion. This means that proprioceptive and somatosensory inputs are explained in terms of hidden causes, where proprioceptive sensations are caused by internally generated forces and somatosensory consequences report a mixture of internal and external forces. Crucially, the precision afforded sensory prediction errors depends upon the internally generated force (and its hidden cause). This dependency is controlled by a parameter γ that mediates the attenuation of sensory precision: as internally generated forces rise, sensory precision falls, thereby attenuating the amplitude of (precision weighted) sensory prediction errors. These context or state-dependent changes in precision enable the agent to attend to sensory input, or not—depending upon the relative precision of prediction errors at the sensory and higher levels. This context sensitive sensory precision is shown in Figure 5 as π (cyan circles).

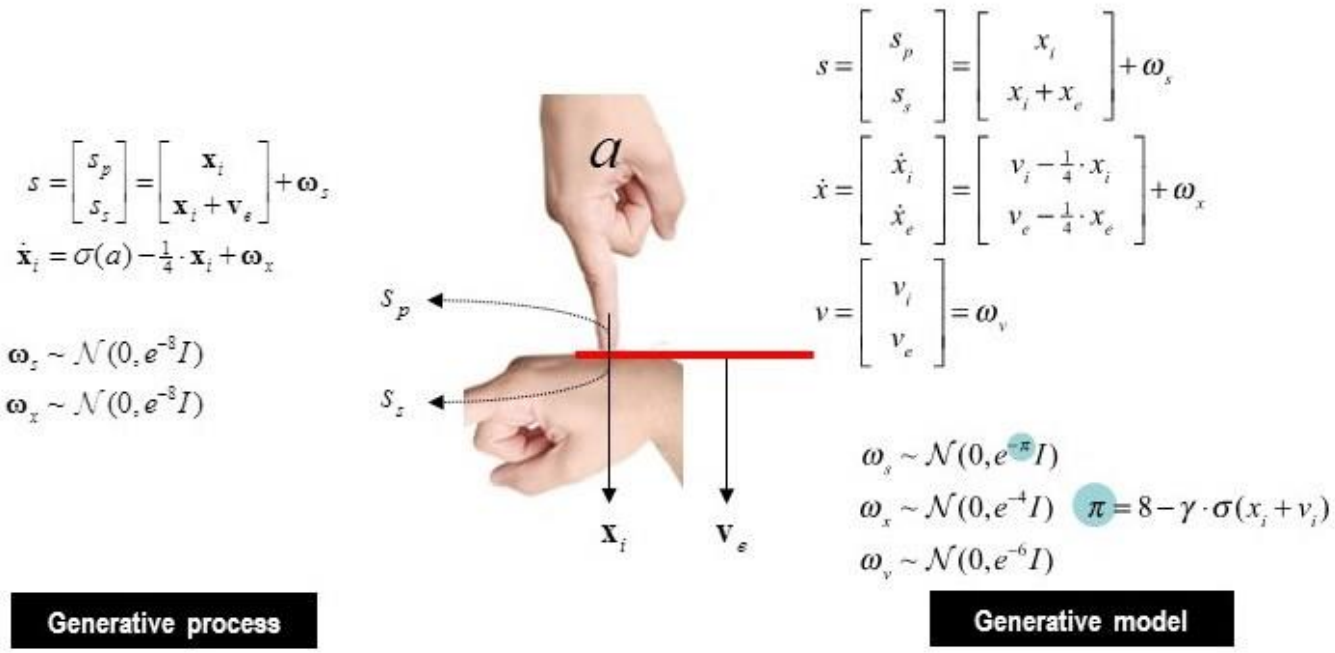

Figure 4: This figure shows the generative process and model used in the simulations of sensory attenuation. The generative process (on the left) models real-world states and causes, while the model on the right is the generative model used by the subject. In the real world, the hidden state x_i corresponds to self-generated pressures that are sensed by both somatosensory s_s and proprio-

ceptive s_p input channels. External forces are modelled with the hidden cause v_e and are sensed only by the somatosensory channel. Action causes the self-generated force x_i to increase and is modified by a sigmoid squashing function σ. The hidden state decays slowly over four time bins. In the generative model, causes of sensory data are divided into internal v_i and external causes v_e. The hidden cause excites dynamics in hidden states x_i and x_e, which decay slowly. Internal force is perceived by both proprioceptive and somatosensory receptors, as before, while external force is perceived only by somatosensory receptors. Crucially, the precision of the sensory input ω_s is influenced by the level of internal force, again modulated by a squashing function, and controlled by a parameter γ that governs the level of attenuation of precision. The generalised predictive coding scheme associated with this generative model is shown schematically in the next figure.

Functional anatomy

Figure 5 illustrates how this generative model could be transcribed into a plausible neuronal architecture. In this example, we have assigned sensory expectations and prediction errors to the thalamus, while corresponding expectations and prediction errors about hidden states (forces) are associated with the sensorimotor cortex. The expectations and prediction errors about the hidden causes of forces have been placed—somewhat agnostically—in the prefrontal cortex. Notice how proprioceptive predictions descend to the spinal-cord to elicit output from alpha motor neurons (playing the role of proprioceptive prediction error units) that cause movements through a classical reflex arc. Red connections denote ascending prediction errors, black connections descending predictions (posterior expectations), and the cyan connection denotes descending neuromodulatory effects that mediate sensory attenuation. The ensuing hierarchy conforms to the functional form of the predictive coding scheme in Equation (3). In this architecture, predictions based on expected states of the world can either be fulfilled by reflex arcs or they can be corrected by ascending sensory prediction errors. Which of these alternatives occurs depends on the relative precisions along each pathway—that are set by the descending modulatory connection to sensory prediction errors. We now use this model to demonstrate some key points:

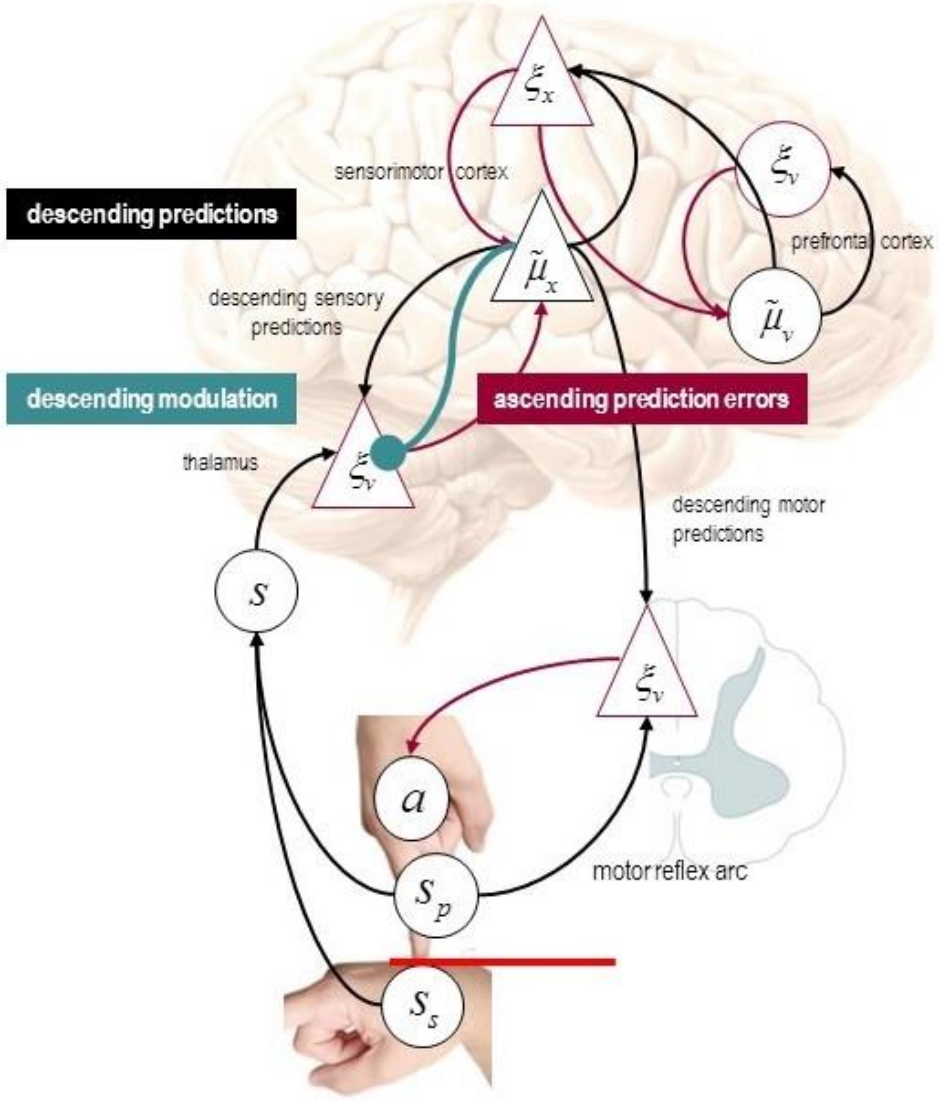

Figure 5: Speculative mapping of Equation (3)—for the generative model in the previous figure—onto neuroanatomy. Somatosensory and proprioceptive prediction errors are generated by the thalamus, while the expectations and prediction errors about hidden states (the forces) are placed in sensorimotor cortex. The expectations and prediction errors about the hidden causes of forces have been placed in the prefrontal cortex. Under active inference, proprioceptive predictions descend to the spinal cord and elicit output from alpha motor neurons (playing the role of proprioceptive prediction error units) via a classical reflex arc. Red connections originate from prediction error units— ξ cells—and can be regarded as intrinsic connections or ascending (forward) extrinsic connections (from superficial pyramidal cells). Conversely, the black connections represent intrinsic connections and descending (backward) efferents (from deep pyramidal cells) encoding posterior expectations—$\tilde{\mu}$ cells. The cyan connection denotes descending neuromodulatory effects that mediate sensory attenuation. The crucial point to take from this schematic is that conditional expectations of sensory states (encoded in the pyramidal cell $\tilde{\mu}_x$) can either be fulfilled by descending proprioceptive predictions (that recruit classical reflex arcs) or they can be corrected by ascending sensory prediction errors. In order for descending motor efferents to prevail, the precision of the sensory prediction errors must be attenuated.

Sensory attenuation and the force matching illusion

To produce internally generated movements, we simply supplied the subject with prior beliefs that the internal hidden cause increased transiently to a value of one, with high sensory attenuation $\gamma = 6$. We then followed this self-generated movement with an exogenously generated force that matched the self-generated force. The left-hand panels in Figure 6 show the results of this simulation. The lower left panel shows the internal hidden cause (blue line) with relatively tight 90% confidence intervals (grey areas). Prior beliefs about this hidden cause excite posterior beliefs about internally generated forces, while at the same time attenuating the precision of sensory prediction errors. This is reflected by the rise in the posterior expectation of the internal force (blue line in the upper right panel) and the transient increase in the confidence interval about this expectation. The resulting proprioceptive predictions are fulfilled by action (bottom right panel) to produce the predicted sensations (upper left panel). Note that proprioceptive prediction (blue line) corresponds to somatosensory prediction (green line) and that both are close to the real values (broken black line). This simulation shows normal self-generated movement under permissive sensory attenuation.

The right-hand panels of Figure 6 show exactly the same results as in the left hand panels; however here, we have yoked the exogenous force x_e to the self-generated force x_i perceived at 90% confidence (dotted line in the top right graph)—as opposed to the true force exerted by the subject. In other words, the external force corresponds to the force that would be reported by the subject to match the perceived force at 90% confidence. The 90% confidence interval was chosen as a proxy for the percept to reconcile the perceived intensity literature with results from signal detection paradigms (Cardoso-Leite et al. 2010). Experimental work in the auditory domain has demonstrated that perceived intensity can be attenuated by increasing sensory noise (decreasing precision) (Richards 1968; Lochner and Burger 1961). When coupled to the 90% confidence interval, the internally generated force is now much greater than the matched external force (shown on the upper left graph). This is the key finding in the force matching illusion and is entirely consistent with sensory attenuation. In this setting, the loss of confidence in posterior estimates of hidden states that are self-generated translates into an illusory increase in the force applied, relative to the equivalent force in the absence of sensory attenuation.

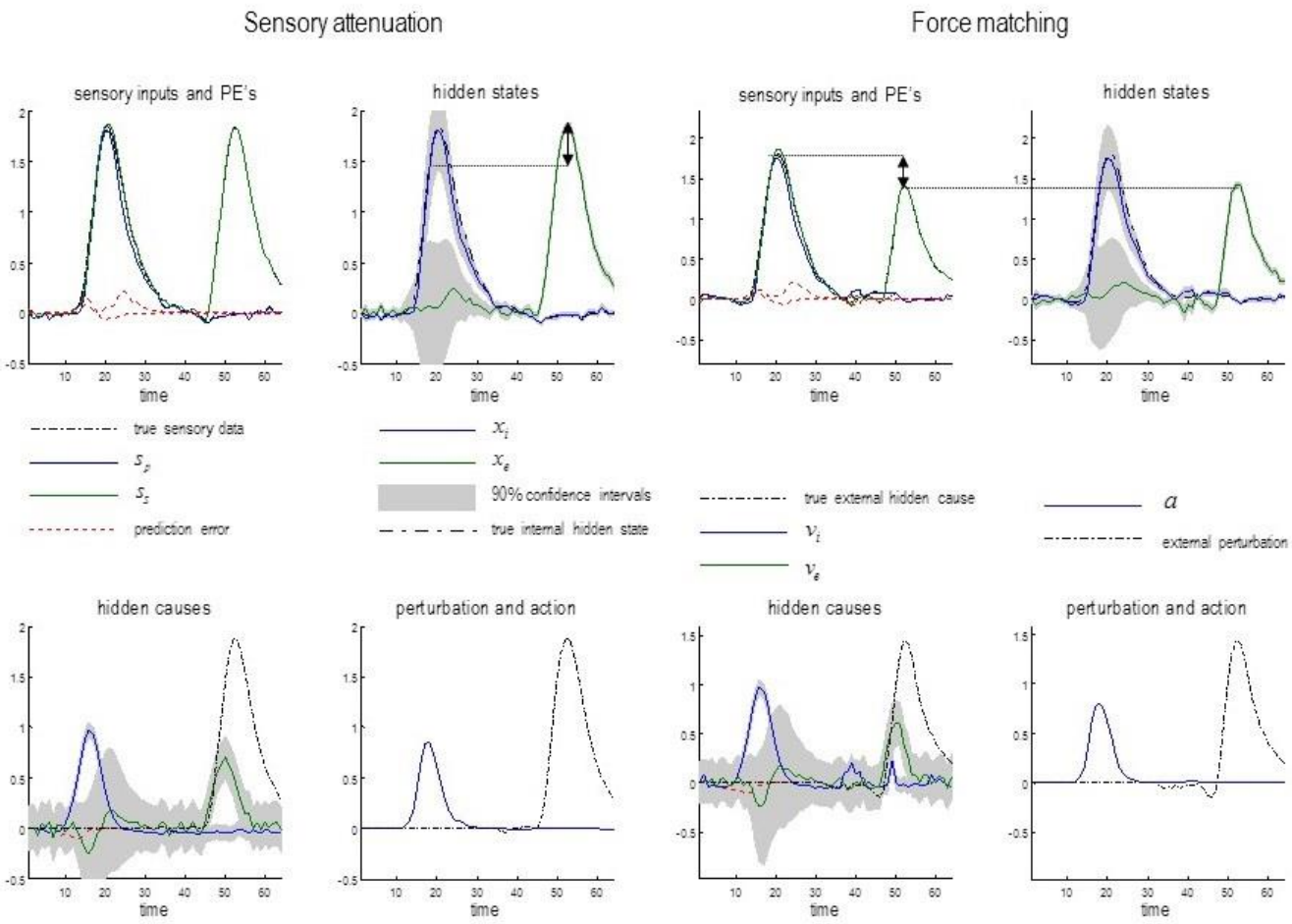

Figure 6: Simulation of the force matching task. The x axes denote time in 100ms time bins; the y axes force in Newtons. Left panels: in the first part of this simulation an internal force is generated from a prior belief about the cause v_i, followed by the presentation of an external force. Posterior beliefs about the hidden states (upper right panel) are similar, but the confidence interval around the force for the internally generated state is much broader. This is because sensory level precision must be attenuated in order to allow proprioceptive predictions to be fulfilled by reflex arcs instead of being corrected by sensory input: i.e. the confidence intervals around v_i must be narrower than those around x_i to allow movement to proceed. If perceived intensity of the sensation is associated with the lower 90% confidence bound of the estimate of hidden state (highlighted by the dotted line), it will be lower when the force is self-generated than when the force is exogenous (the difference is highlighted by the arrow). Right panels: the simulation was repeated but the external force was matched to the lower bound of the 90% confidence interval of the internal force. This means that internally generated force is now greater than the externally applied force (double-headed arrow, upper left panel). This reproduces the normal psychophysics of the force matching illusion that can be regarded as entirely Bayes optimal, under appropriate levels of precision.

We repeated these simulations under different levels of self-generated forces by modulating the prior beliefs about the internal hidden cause (from a half to twice the normal amplitude). The results are shown as the blue circles in the left panel of Figure 7, which plots the self-generated force against the yoked or matched external force with a corresponding 90% confidence interval. These results are remarkably similar to those obtained empirically (right panel—

reproduced from (Shergill et al. 2005) and reveal sensory attenuation through an illusory increase in the self-generated force, relative to matched forces over a wide range of forces. The red line in the left panel comes from the final simulations, in which we asked what would happen if subjects compensated for a failure in sensory attenuation by increasing the precision of their prior beliefs?

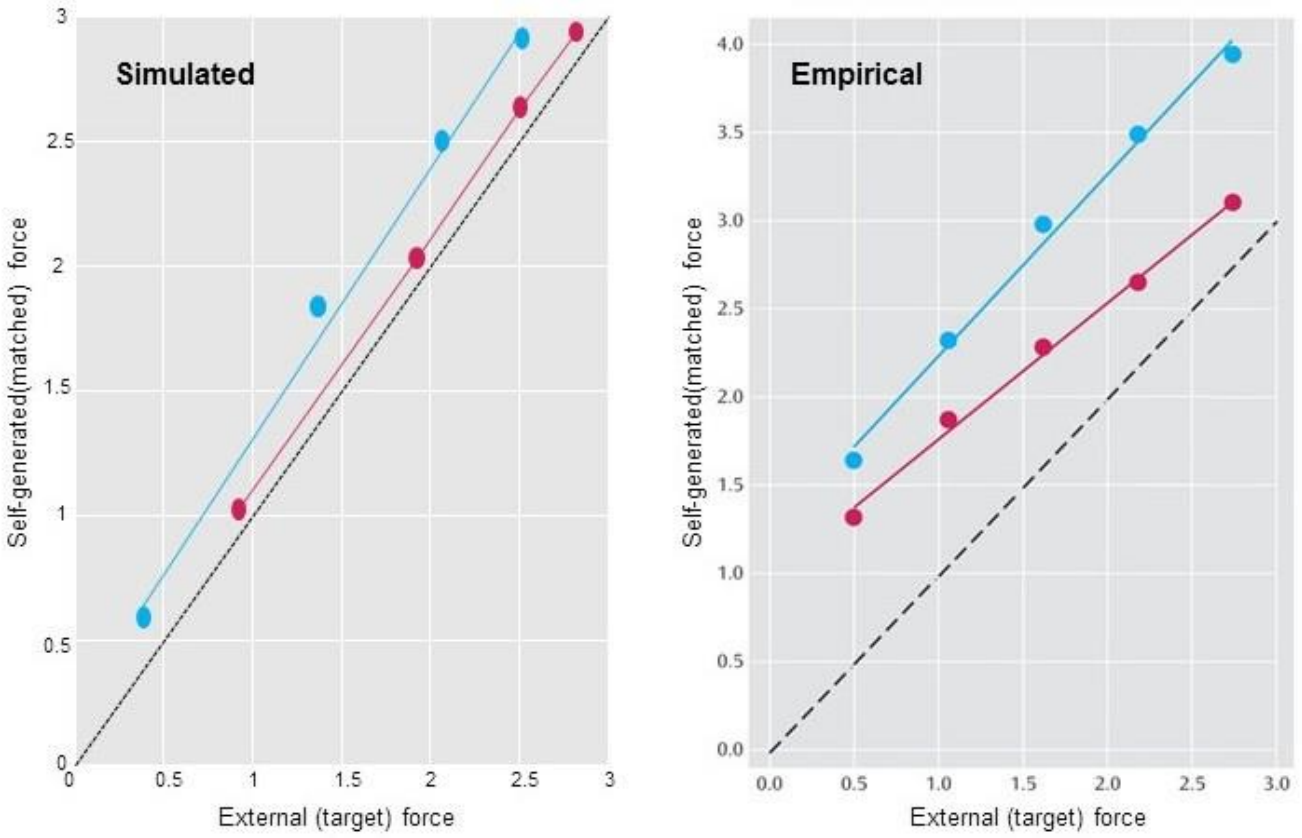

Figure 7: Left panel: the force matching simulation was repeated under different levels of self-generated force. For normal levels of sensory attenuation (blue circles), internally produced force is higher than externally generated force at all levels. Data from patients with schizophrenia was simulated by attenuating sensory precision and increasing the precision of prediction errors at higher levels of the hierarchy. This resulted in a more veridical perception of internally generated force (red circles). Right panel: the empirical data from the force matching task, with normal subjects' forces in blue, and schizophrenics' forces in red (reproduced from (Shergill et al. 2005).

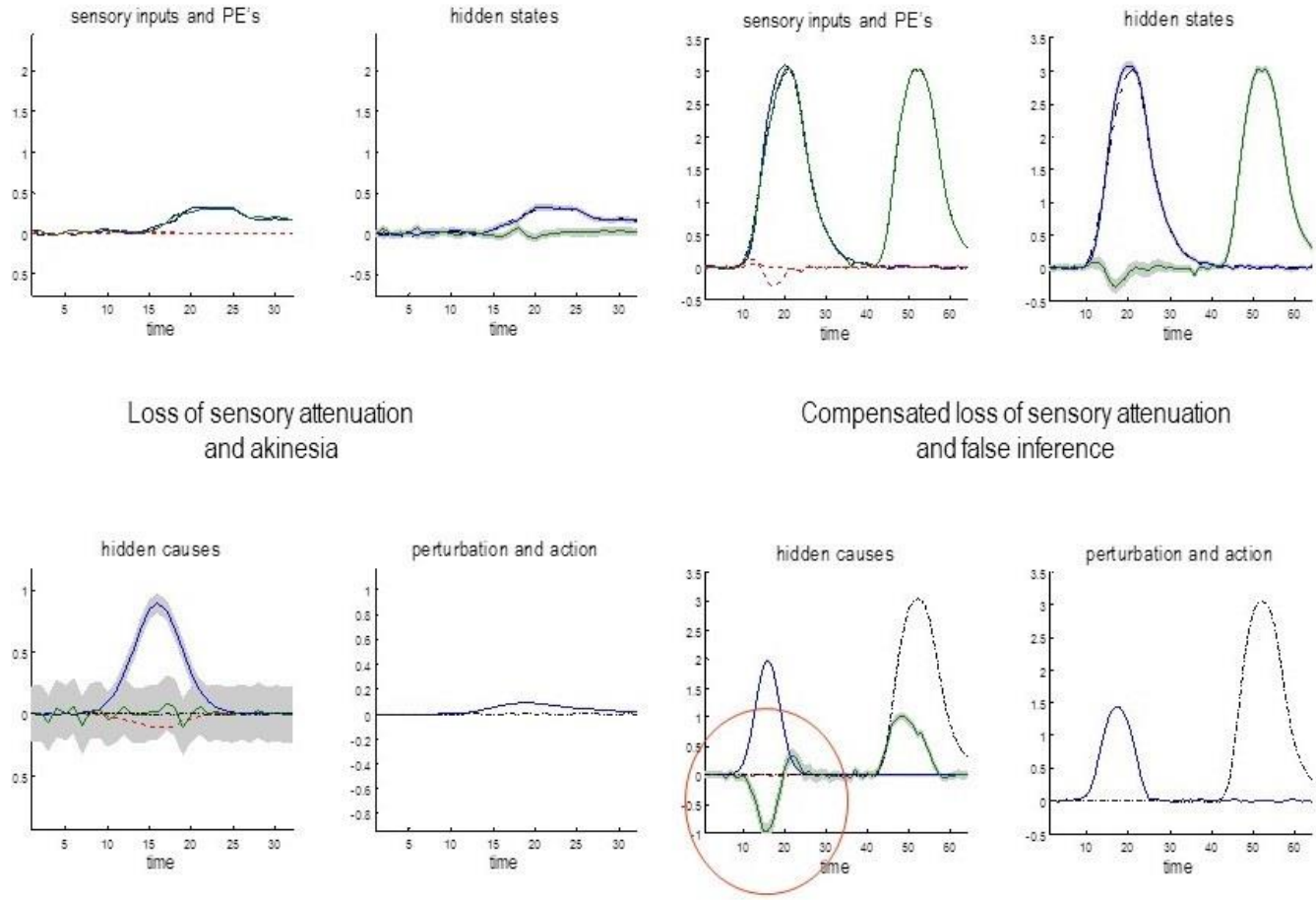

Figure 8: Pathology of sensory attenuation. Left panel: Here sensory attenuation is much lower ($\gamma = 2$). In this case, bottom-up prediction errors have a higher precision than top-down predictions: the confidence intervals around v_i (bottom left panel) are now broader than those around x_i (upper right panel). The expected hidden state is thus profoundly suppressed (upper right panel), meaning proprioceptive prediction errors are not produced (upper left panel) and action is suppressed (lower right panel) resulting in akinesia. Right panels: To simulate the force-matching results seen in schizophrenia, precision at the second level of the hierarchy was increased to allow movement. The underlying failure of sensory attenuation still enables a precise and accurate perception of internally and externally generated sensations (upper left panel). However, the causes of sensory data are not accurately inferred: a false (delusional) cause (lower left panel) is perceived during internally generated movement that is antagonistic to the movement. This is because the proprioceptive prediction errors driving action are rendered overly precise, meaning higher levels of the hierarchy must be harnessed to explain them, resulting in a delusion that exogenous forces are opposing the expected outcome (encircled in red).

False inference and failures of sensory attenuation

We now demonstrate two pathologies of sensory attenuation: first, a loss of sensory attenuation resulting in a catatonic state and second, how compensation for such a loss could allow movement but result in a somatic delusion. The consequences of reducing sensory attenuation (from six to two) are illustrated in the left panels of Figure 8. Here, the loss of sensory attenuation maintains the precision of the hidden states above the precision of prior be-

liefs about hidden causes (lower left panel). This means that bottom-up sensory prediction errors predominate over top-down predictions and expectations about internally generated forces are profoundly suppressed. Because there are no predictions about proprioceptive changes, there is a consequent akinesia. This state is reminiscent of the catatonic symptoms of schizophrenia such as immobility, mutism, catalepsy and waxy flexibility, in which the patient may maintain a fixed posture for a long time, even though (in the case of waxy flexibility) their limbs can be moved easily by someone else.

We shall now examine how a loss of sensory attenuation might be compensated for by increasing the precision of prediction errors at higher levels in the hierarchy (by increasing the log-precision of prediction errors on hidden states and causes by four log units). This compensatory increase is necessary for movement and ensures the precision of top-down predictions is greater than bottom-up sensory prediction errors. These manipulations permit movement but abolish the force matching illusion, as indicated by the line of red circles in the left panel of Figure 7. One might ask—why don't subjects adopt this strategy and use precise prior beliefs about hidden causes all the time?

The answer is evident in the right panels of Figure 8, which show the results of a simulation with low sensory attenuation and compensatory increases in precision at higher levels. Here, there is an almost perfect and precise inference about internally and externally generated sensations. However, there is a failure of inference about their hidden causes. This can be seen on the lower left, where the subject has falsely inferred an antagonistic external hidden cause that mirrors the internal hidden causes. Note that this false inference does not occur during normal sensory attenuation (see Figure 6), where the true external hidden cause always lies within the 90% confidence intervals. The reason for this false inference or delusion is simple: action is driven by proprioceptive prediction errors that always report less force than that predicted. However, when these prediction errors are very precise they need to be explained—and can only be explained by falsely inferring an opposing exogenous force. This only occurs when both the predictions and their consequences are deemed to be very precise. This false inference could be interpreted as a delusion in the same sense that the sensory attenuation is an illusion. Having said this, it should be noted that—from the point of view of the subject—its inferences are Bayes-optimal. It is only our attribution of the inference as false that gives it an illusory or delusionary aspect.

This simulation has some face validity in relation to empirical studies of the force matching illusion. The illusion is attenuated in normal subjects that score highly on ratings of delusional beliefs (Teufel et al. 2010). Furthermore, subjects with schizophrenia—who are prone to positive symptoms like delusions—do not show the force matching illusion (Shergill et al. 2005).

In other words, there may be a trade-off between illusions at a perceptual level and delusions at a conceptual level that is mediated by a (failure of) sensory attenuation.

Summary

The ideas reviewed in this section suggest that attribution of agency—in an ambiguous situation—can be resolved by attenuating the precision of sensory evidence during movement: in other words, attending away from the consequences of self-made acts. When implemented in the context of active inference, this provides a Bayes-optimal explanation for sensory attenuation and attending illusions. The simulations show how exacerbations of a trait loss of sensory attenuation could subvert movement and even cause catatonia. This can be ameliorated by compensatory increases in high-level precision, which in turn necessarily induce false (delusional) inferences about agency. This is important, given the negative correlation between sensory attenuation and predisposition to delusional beliefs in normal subjects and the reduced force matching illusion in schizophrenia. On a physiological level, increased dopaminergic transmission in the striatum could reflect a putative increase in high-level precision, compensating for hypofunction of cortical NMDA-Rs. In summary, we have shown how active inference can explain the fundamental role of sensory attenuation, and how its failure could lead to not only catatonic states but also compensatory changes that induce delusions. This is one illustration of how psychotic state abnormalities might be secondary compensations for trait abnormalities.

Other symptoms

In recent work (Adams et al. 2012), we have simulated various characteristic deficits in pursuit eye movements in schizophrenia by simply reducing the precision at a high hierarchical level of the generative model of target (and eye) movement. The deficits which this change reproduces include the increased slowing of eye movements during the occlusion of a target, a 'paradoxical' improvement in tracking a sudden unexpected change of target trajectory, and a problem in inferring high level causes (in this case the frequency) of target motion, despite apparently normal eye movements. We have also simulated abnormal ERP responses to predictable and unpredictable stimuli in an auditory paradigm (Adams et al. 2013; Kiebel et al. 2009). This and other examples illustrate the complicated interplay between expectations and precision in optimising our beliefs about, and exchange with the world—and how sensitive this interplay is to abnormalities in precision or neuromodulatory gain control.

Conclusion

Bayesian computations enable inference and learning under uncertainty. Furthermore, they prescribe the optimal integration of prior expectations (amassed over a lifetime or indeed evolution) with the sensory evidence of a moment. This integration is optimal because it embodies the relative uncertainty (precision) of each source of information. For this reason, the accurate representation of precision in a hierarchical Bayesian scheme is crucial for informed and veridical inference. The aberrant encoding of precision can therefore lead to false inference by overweighting prior expectations or sensory evidence.

This paper has described how various trait abnormalities in schizophrenia could result from a decrease in prior precision (or a failure to attenuate sensory precision); and how some psychotic states could result from compensatory increases in prior precision (or decreases in sensory precision). We have outlined several physiological mechanisms for encoding precision (such as neuromodulation and neuronal oscillations) that are abnormal in schizophrenia. Genetic and neuropathological evidence suggest that NMDA-R (and GABA to some extent) may play a role in trait abnormalities, whereas the physiological evidence points towards dopaminergic pathology in the psychotic state. Clearly, a strict dichotomy is unlikely, since these neurotransmitter systems have complex interactions. Using a biologically plausible predictive coding scheme, we have simulated the normal phenomenon of movement-related 'sensory attenuation' as a transient reduction in the precision of sensory evidence during movement, and the consequent suspension of attention to the sensory consequences of self-made acts. We have also shown how a failure to attenuate sensory precision might explain a resistance to (force matching) illusions and (in severe cases) catatonia. Using these model systems, we were able to explain the delusional and hallucinatory inference characteristic of the psychotic state by compensatory increases (resp. decreases) in prior (resp. sensory) precision.

Simulations of the sort used above clearly require empirical validation. This should be possible as the models make quantitative predictions about the dynamics of cortical populations that can be tested with careful modelling of neurophysiological and behavioural timeseries (Friston et al. 2003). Indeed, dynamic causal modelling studies of schizophrenic subjects have already demonstrated changes in effective connectivity consistent with decreased high level—and increased low level—precision in the hollow mask paradigm (Dima et al. 2009; Dima et al. 2010).

Acknowledgements: This work was funded by the Wellcome Trust.

References

Abi-Saab, W., D'Souza, D. C., Moghaddam, B., and Krystal, J. H. 1998. The NMDA Antagonist Model for Schizophrenia: Promise and Pitfalls. *Pharmacopsychiatry*, 31(2): 104-9. doi:10.1055/s-2007-979354.

Adams, R. A, Perrinet, L. U., and Friston, K. 2012. Smooth Pursuit and Visual Occlusion: Active Inference and Oculomotor Control in Schizophrenia. *PloS One*, 7(10). e47502. doi:10.1371/journal.pone.0047502.

Adams, R. A., Shipp, S., and Friston, K. 2012. Predictions Not Commands: Active Inference in the Motor System. *Brain Structure & Function*, 218(3). doi:10.1007/s00429-012-0475-5.

Adams, R. A, Stephan, K. E., Brown, H. R., Frith, C. D., and Friston, K. 2013. The Computational Anatomy of Psychosis. *Frontiers in Psychiatry*, 4: 47. doi:10.3389/fpsyt.2013.00047.

Albrecht, M. A, Martin-Iverson, M. T., Price, G., Lee, J., Iyyalol, R., and Waters, F. 2011. Dexamphetamine Effects on Separate Constructs in the Rubber Hand Illusion Test. *Psychopharmacology*, 217(1): 39-50. doi:10.1007/s00213-011-2255-y.

Allen, P., Aleman, A., and McGuire, P. K. 2007. Inner Speech Models of Auditory Verbal Hallucinations: Evidence from Behavioural and Neuroimaging Studies. *International Review of Psychiatry (Abingdon, England)*, 19(4): 407-15. doi:10.1080/09540260701486498.

American Psychiatric Association. 2000. *Diagnostic and Statistical Manual of Mental Disorders: DSM-IV-TR*. 4th ed. Washington, DC: Author.

Bagal, A., Joseph, A., Kao, P. Y., Tang, C.-M., and Thompson, S. M. 2005. Long-Term Potentiation of Exogenous Glutamate Responses at Single Dendritic Spines. *Proceedings of the National Academy of Sciences of the United States of America*, 102(40): 14434-39. doi:10.1073/pnas.0501956102.

Ballard, D. H., Hinton, G. E., and Sejnowski, T. J. 1983. Parallel Visual Computation. *Nature*, 306(5938): 21-26.

Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., and Friston, K. 2012. Canonical Microcircuits for Predictive Coding. *Neuron*, 76(4): 695-711. doi:10.1016/j.neuron.2012.10.038.

Blakemore, S. J., Wolpert, D. M., and Frith, C. D. 1998. Central Cancellation of Self-Produced Tickle Sensation. *Nature Neuroscience*, 1(7): 635-40. doi:10.1038/2870.

Breier, A., Su, T. P., Saunders, R., Carson, R. E., Kolachana, B. S., de Bartolomeis, A., Weinberger, D. R., et al. 1997. Schizophrenia Is Associated with Elevated Amphetamine-Induced Synaptic Dopamine Concentrations: Evidence from a Novel Positron Emission Tomography Method. *Proceedings of the National Academy of Sciences of the United States of America*, 94(6): 2569-2574.

Brown, H., Adams, R. A., Parees, I., Edwards, M., and Friston, K. 2013. Active Inference, Sensory Attenuation and Illusions. *Cognitive Processing*, 14(4): 411-27. doi:10.1007/s10339-013-0571-3.

Calkins, M. E., Iacono, W. G., Ones, D. S. 2008. Eye Movement Dysfunction in First-Degree Relatives of Patients with Schizophrenia: A Meta-Analytic Evaluation of Candidate Endophenotypes. *Brain and Cognition*, 68(3): 436-61. doi:10.1016/j.bandc.2008.09.001.

Campion, D., Thibaut, F., Denise, P., Courtin, P., Pottier, M., and Levillain, D. 1992. SPEM Impairment in Drug-Naive Schizophrenic Patients: Evidence for a Trait Marker. *Biological Psychiatry*, 32(10): 891-902.

Cardoso-Leite, P., Mamassian, P., Schütz-Bosbach, S., and Waszak, F. 2010. A New Look at Sensory Attenuation. Action-Effect Anticipation Affects Sensitivity, Not Response Bias. *Psychological Science*, 21(12): 1740-45. doi:10.1177/0956797610389187.

Carlsson, A., Waters, N., and Carlsson, M. L. 1999. Neurotransmitter Interactions in Schizophrenia-Therapeutic Implications. *European Archives of Psychiatry and Clinical Neuroscience* 249 Suppl. 4: 37-43.

Chawla, D., Lumer, E. D., and Friston, K. 1999. The Relationship between Synchronization among Neuronal Populations and Their Mean Activity Levels. *Neural Computation*, 11(6): 1389-1411.

Corlett, P. R., Honey, G. D., Krystal, J. H., and Fletcher, P. C. 2011. Glutamatergic Model Psychoses: Prediction Error, Learning, and Inference. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, 36(1): 294-315. doi:10.1038/npp.2010.163.

Crow, T. J. 1980. Molecular Pathology of Schizophrenia: More than One Disease Process? *British Medical Journal*, 280(6207): 66-68.

Dayan, P., Hinton, G. E., Neal, R. M., and Zemel, R. S. 1995. The Helmholtz Machine. *Neural Computation*, 7(5): 889-904.

Dazzan, P., Morgan, K. D., Chitnis, X., Suckling, J., Morgan, C., Fearon, P., McGuire, P.K., Jones, P. B., Leff, J., and Murray, R. M. 2006. The Structural Brain Correlates of Neurological Soft Signs in Healthy Individuals. *Cerebral Cortex (New York, N.Y.: 1991)*, 16(8): 1225-31. doi:10.1093/cercor/bhj063.

Demjaha, A., Murray, R. M., McGuire, P. K., Kapur, S., and Howes, O. D. 2012. Dopamine Synthesis Capacity in Patients With Treatment-Resistant Schizophrenia. *The American Journal of Psychiatry*, 169(11): 1203-10. doi:10.1176/appi.ajp.2012.12010144.

Dima, D., Dietrich, D. E., Dillo, W., and Emrich, H. M. 2010. Impaired Top-down Processes in Schizophrenia: A DCM Study of ERPs. *NeuroImage*, 52(3): 824-32. doi:10.1016/j.neuroimage.2009.12.086.

Dima, D., Roiser, J. P., Dietrich, D. E., Bonnemann, C., Lanfermann, H., Emrich, H. M., and Dillo, W. 2009. Understanding Why Patients with Schizophrenia Do Not Perceive the Hollow-Mask Illusion Using Dynamic Causal Modelling. *NeuroImage*, 46(4): 1180-86. doi:10.1016/j.neuroimage.2009.03.033.

Durstewitz, D. 2009. Implications of Synaptic Biophysics for Recurrent Network Dynamics and Active Memory. *Neural Networks: The Official Journal of the International Neural Network Society*, 22(8): 1189-1200. doi:10.1016/j.neunet.2009.07.016.

Fanous, A., Gardner, C., Walsh, D., and Kendler, K. S. 2001. Relationship between Positive and Negative Symptoms of Schizophrenia and Schizotypal Symptoms in Nonpsychotic Relatives. *Archives of General Psychiatry*, 58(7): 669-73.

Feldman, A. G., and Levin, M. F. 1995. The Origin and Use of Positional Frames of Reference in Motor Control. *Behavioral and Brain Sciences*, 18(04): 723-44. doi:10.1017/S0140525X0004070X.

Feldman, H., and Friston, K. 2010. Attention, Uncertainty, and Free-Energy. *Frontiers in Human Neuroscience*, 4: 215. doi:10.3389/fnhum.2010.00215.

Felleman, D. J., and Van Essen, D. C. 1991. Distributed Hierarchical Processing in the Primate Cerebral Cortex. *Cerebral Cortex (New York, N.Y.: 1991)*, 1(1): 1-47.

Fleming, S. M., Dolan, R. J., and Frith, C. D. 2012. Metacognition: Computation, Biology and Function. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1594): 1280-86. doi:10.1098/rstb.2012.0021.

Fletcher, P. C, and Frith, C. D. 2009. Perceiving Is Believing: A Bayesian Approach to Explaining the Positive Symptoms of Schizophrenia. *Nature Reviews. Neuroscience*, 10(1): 48-58. doi:10.1038/nrn2536.

Frank, M. J. 2005. Dynamic Dopamine Modulation in the Basal Ganglia: A Neurocomputational Account of Cognitive Deficits in Medicated and Nonmedicated Parkinsonism. *Journal of Cognitive Neuroscience*, 17(1): 51-72. doi:10.1162/089892905 2880093.

Freeman, T. P., Morgan, C. J. A., Pepper, F., Howes, O. D., Stone, J. M., and Curran, H. V. 2013. Associative Blocking to Reward-Predicting Cues Is Attenuated in Ketamine Users but Can Be Modulated by Images Associated with Drug Use. *Psychopharmacology*, 225(1): 41-50. doi:10.1007/s00213-012-2791-0.

Friston, K., Harrison, L., and Penny, W. 2003. Dynamic Causal Modelling. *NeuroImage*, 19(4): 1273-1302. doi:10.1016/S1053-8119(03)00202-7.

Friston, K. 2009. The Free-Energy Principle: A Rough Guide to the Brain? *Trends in Cognitive Sciences*, 13(7): 293-301. doi:10.1016/j.tics.2009.04.005.

Friston, K., Daunizeau, J., and Kiebel, S. J. 2009. Reinforcement Learning or Active Inference? *PloS One*, 4(7): e6421. doi:10.1371/journal.pone.0006421.

Friston, K., Daunizeau, J., Kilner, J., and Kiebel, S. J. 2010. Action and Behavior: A Free-Energy Formulation. *Biological Cybernetics*, 102(3): 227-60. doi:10.1007/s00422-010-0364-z.

Friston, K., Shiner, T., Fitzgerald, T., Galea, J. M., Adams, R., Brown, H., Dolan, J. R., Moran, R., Stephan, K. E., and Bestmann, S. 2012. Dopamine, Affordance and Active Inference. *PLoS Computational Biology*, 8(1): e1002327. doi:10.1371/journal. pcbi.1002327.

Friston, K. 2005. Hallucinations and Perceptual Inference. *Behav Brain Sci*, 28(06): 764-66. doi:10.1017/S0140525X05290131.

Friston, K., and Kiebel, S. 2009. Predictive Coding under the Free-Energy Principle. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 364(1521): 1211-21. doi:10.1098/rstb.2008.0300.

Friston, K., Stephan, K., Li, B., and Daunizeau, J. 2010. Generalised Filtering. *Mathematical Problems in Engineering*: 1-35. doi:10.1155/2010/621670.

Frith, C., Rees, G., and Friston, K. 1998. Psychosis and the Experience of Self. Brain Systems Underlying Self-Monitoring. *Annals of the New York Academy of Sciences*, 843: 170-78.

Frith, C., and Friston, K. 2012. False Perceptions and False Beliefs: Understanding Schizophrenia. *Working Group on Neurosciences and the Human Person: New Perspectives on Human Activities, The Pontifical Academy of Sciences, 8-10 November 2012*. Casina Pio IV.

Garety, P. A., and Freeman, D . 1999. Cognitive Approaches to Delusions: A Critical Review of Theories and Evidence. *The British Journal of Clinical Psychology / the British Psychological Society*, 38(Pt 2) (June): 113-54.

Ginzburg, V. L. 1955. On the Theory of Superconductivity. *Il Nuovo Cimento*, 2(6): 1234-50. doi:10.1007/BF02731579.

Goff, D. C., and Coyle, J. T. 2001. The Emerging Role of Glutamate in the Pathophysiology and Treatment of Schizophrenia. *The American Journal of Psychiatry*, 158(9): 1367-77.

Goldman-Rakic, P. S., Castner, S. A., Svensson, T. H. Siever, L. J., and Williams, G. V. 2004. Targeting the Dopamine D1 Receptor in Schizophrenia: Insights for Cognitive Dysfunction. *Psychopharmacology*, 174(1): 3-16. doi:10.1007/s00213-004-1793-y.

Gonzalez-Burgos, G., and Lewis, D. A. 2012. NMDA Receptor Hypofunction, Parvalbumin-Positive Neurons, and Cortical Gamma Oscillations in Schizophrenia. *Schizophrenia Bulletin*, 38(5): 950-57. doi:10.1093/schbul/sbs010.

Gonzalez-Burgos, G., and Lewis, D. A. 2008. GABA Neurons and the Mechanisms of Network Oscillations: Implications for Understanding Cortical Dysfunction in Schizophrenia. *Schizophrenia Bulletin*, 34(5): 944-61. doi:10.1093/schbul/sbn070.

Gray, J. A., Feldon, J., Rawlins, J. N. P., Hemsley, D. R., and Smith, A. D. 1991. The Neuropsychology of Schizophrenia. *Behavioral and Brain Sciences*, 14(01): 1-20. doi:10.1017/S0140525X00065055.

Greenwood, T. A., Light, G. A., Swerdlow, N. R., Radant, A.D., and Braff, D. L. 2012. Association Analysis of 94 Candidate Genes and Schizophrenia-Related Endophenotypes. *PLoS ONE*, 7(1): e29630. doi:10.1371/journal.pone.0029630.

Gunduz-Bruce, H., Reinhart, R. M. G., Roach, B. J., Gueorguieva, R., Oliver, S., D'Souza, D. C., Ford, J. M., Krystal, J. H., and Mathalon, D. H. 2012. Glutamatergic Modulation of Auditory Information Processing in the Human Brain. *Biological Psychiatry*, 71(11): 969-77. doi:10.1016/j.biopsych.2011.09.031.

Haken, H. 1983. *Synergetics: Introduction and Advanced Topics*. 3rd ed. Berlin: Springer.

Hall, J., Romaniuk, L., McIntosh, A. M., Steele, J. D., Johnstone, E. C., and Lawrie, S. M. 2009. Associative Learning and the Genetics of Schizophrenia. *Trends in Neurosciences*, 32(6): 359-65. doi:10.1016/j.tins.2009.01.011.

Harrison, P. J., and Weinberger, D. R. 2005. Schizophrenia Genes, Gene Expression, and Neuropathology: On the Matter of Their Convergence. *Molecular Psychiatry*, 10(1): 40-68; image 5. doi:10.1038/sj.mp.4001558.

Harrison, P. J., Lewis, D. A., and Kleinman, J. E. 2011. Neuropathology of Schizophrenia. Weinberger D. R., and Harrison, P. J. *Schizophrenia*, BM, BCh, DM(Oxon) FRCPsych, Wiley-Blackwell: 372-92. http://onlinelibrary.wiley.com/doi/10.1002/9781444327298.ch18/summary.

Hinton, G. E., and van Camp, D. 1993. Keeping the Neural Networks Simple by Minimizing the Description Length of the Weights. *Proceedings of the Sixth Annual Conference on Computational Learning Theory*, 5-13. COLT '93. New York, NY, USA: ACM. doi:10.1145/168304.168306. http://doi.acm.org/10.1145/168304.168306.

Howes, O. D., and Kapur, S . 2009. The Dopamine Hypothesis of Schizophrenia: Version III--the Final Common Pathway. *Schizophrenia Bulletin*, 35(3): 549-62. doi:10.1093/schbul/sbp006.

Hutton, S. B., Crawford, T. J., Gibbins, H., Cuthbert, I., Barnes, T. R., Kennard, C., and Joyce, E. M. 2001. Short and Long Term Effects of Antipsychotic Medication on Smooth Pursuit Eye Tracking in Schizophrenia. *Psychopharmacology*, 157(3): 284-91. doi:10.1007/s002130100803.

Hutton, S. B., Crawford, T. J., Puri, B. K., Duncan, L. J., Chapman, M., Kennard, C., Barnes, T. R., and Joyce, E. M. 1998. Smooth Pursuit and Saccadic Abnormalities in First-Episode Schizophrenia. *Psychological Medicine*, 28(3): 685-92.

Jabben, N., Arts, B., van Os, J., and Krabbendam, J. 2010. "Neurocognitive Functioning as Intermediary Phenotype and Predictor of Psychosocial Functioning across the Psychosis Continuum: Studies in Schizophrenia and Bipolar Disorder. *The Journal of Clinical Psychiatry*, 71(6): 764-74. doi:10.4088/JCP.08m04837yel.

Johnstone, E. C., Owens, D. G., Frith, C. D. and Crow, T. J. 1987. The Relative Stability of Positive and Negative Features in Chronic Schizophrenia. *The British Journal of Psychiatry*, 150(1): 60-64. doi:10.1192/bjp.150.1.60.

Kantrowitz, J.T., and Javitt, D. C. 2010. N-Methyl-D-Aspartate (NMDA) Receptor Dysfunction or Dysregulation: The Final Common Pathway on the Road to Schizophrenia? *Brain Research Bulletin*, 83(3-4): 108-21. doi:10.1016/j.brainresbull.2010.04.006.

Kapur, S. 2003. Psychosis as a State of Aberrant Salience: A Framework Linking Biology, Phenomenology, and Pharmacology in Schizophrenia. *The American Journal of Psychiatry*, 160(1): 13-23.

Keane, B. P., Silverstein, S. M., Wang, Y., and Papathomas, T. V. 2013. Reduced Depth Inversion Illusions in Schizophrenia Are State-Specific and Occur for Multiple Ob-

ject Types and Viewing Conditions. *Journal of Abnormal Psychology*, 122(2): 506-12. doi:10.1037/a0032110.

Kessler, R. M., Woodward, N. D., Riccardi, P., Li, R., Ansari, M. S., Anderson, S., Dawant, B., Zald, D., and Meltzer, H. Y. 2009. Dopamine D2 Receptor Levels in Striatum, Thalamus, Substantia Nigra, Limbic Regions, and Cortex in Schizophrenic Subjects. *Biological Psychiatry*, 65(12): 1024-31. doi:10.1016/j.biopsych.2008.12.029.

Kiebel, S., Daunizeau, J., and Friston, K . 2009. Perception and Hierarchical Dynamics. *Frontiers in Neuroinformatics*, 3: 20. doi:10.3389/neuro.11.020.2009.

Krystal, J. H., Karper, L. P., Seibyl, J. P., Freeman, G. K., Delaney, R., Bremner, J. D., Heninger, G. R., Bowers Jr, M. B., and Charney, D. S. 1994. Subanesthetic Effects of the Noncompetitive NMDA Antagonist, Ketamine, in Humans. Psychotomimetic, Perceptual, Cognitive, and Neuroendocrine Responses. *Archives of General Psychiatry*, 51(3): 199-214.

Laruelle, M., Abi-Dargham, A., van Dyck, C. H., Gil, R., D'Souza, C. D., Erdos, J., McCance, E., et al. 1996. Single Photon Emission Computerized Tomography Imaging of Amphetamine-Induced Dopamine Release in Drug-Free Schizophrenic Subjects. *Proceedings of the National Academy of Sciences of the United States of America*, 93(17): 9235-40.

Leung, S., Croft, R. J., Baldeweg, T., and Nathan, P. J. 2007. Acute Dopamine D(1) and D(2) Receptor Stimulation Does Not Modulate Mismatch Negativity (MMN) in Healthy Human Subjects. *Psychopharmacology*, 194(4): 443-51. doi:10.1007/s00213-007-0865-1.

Liddle, P. F. 1987. The Symptoms of Chronic Schizophrenia. A Re-Examination of the Positive-Negative Dichotomy. *The British Journal of Psychiatry: The Journal of Mental Science*, 151(August): 145-51.

Lisman, J. E., Coyle, J. T., Green, R. W., Javitt, D. C., Benes, F. M., Heckers, S., and Grace, A. A. 2008. Circuit-Based Framework for Understanding Neurotransmitter and Risk Gene Interactions in Schizophrenia. *Trends in Neurosciences*, 31(5): 234-42. doi:10.1016/j.tins.2008.02.005.

Lochner, J. P. A., and Burger, J. F. 1961. Form of the Loudness Function in the Presence of Masking Noise. *The Journal of the Acoustical Society of America*, 33(12): 1705-7. doi:10.1121/1.1908548.

Luthringer, R., Rinaudo, G., Toussaint, M., Bailey, P., Muller, G., Muzet, A., and Macher, J. 1999. Electroencephalographic Characterization of Brain Dopaminergic Stimulation by Apomorphine in Healthy Volunteers. *Neuropsychobiology*, 39(1): 49-56. doi:26560.

Maher, B. A. 1974. Delusional Thinking and Perceptual Disorder. *Journal of Individual Psychology*, 30(1): 98-113.

Mockler, D., Riordan, J., and Sharma, T. 1997. Memory and Intellectual Deficits Do Not Decline with Age in Schizophrenia. *Schizophrenia Research*, 26(1): 1-7. doi:10.1016/S0920-9964(97)00031-5.

Montgomery, J. M., and Madison, D. V. 2004. Discrete Synaptic States Define a Major Mechanism of Synapse Plasticity. *Trends in Neurosciences*, 27(12): 744-50. doi:10.1016/j.tins.2004.10.006.

Morgan, H. L., Turner, D. C. Corlett, P. R., Absalom, A. R., Adapa, R., Arana, F. S., Pigott, J., et al. 2011. Exploring the Impact of Ketamine on the Experience of Illusory Body Ownership. *Biological Psychiatry*, 69(1): 35-41. doi:10.1016/j.biopsych.2010.07.032.

Morrison, P. D. 2012. 5: The Search for Madness. *Schizophrenia: The Final Frontier—A Festschrift for Robin M. Murray*, 71-85. Hove: Psychology Press. http://books.google.co.uk/books?hl=en&lr=&id=A49Fm_9zfBAC&oi=fnd&pg=PA71&ots=fb6brPVa5S&sig=1XnMwcDROLDSEtxuJnhJNBiqqrg.

Moutoussis, M., Bentall, R. P., El-Deredy, W., and Dayan, P. 2011. Bayesian Modelling of Jumping-to-Conclusions Bias in Delusional Patients. *Cognitive Neuropsychiatry*, 16(5): 422-47. doi:10.1080/13546805.2010.548678.

Mumford, D. 1992. On the Computational Architecture of the Neocortex. II. The Role of Cortico-Cortical Loops. *Biological Cybernetics*, 66(3): 241-51.

O'Driscoll, G. A., and Callahan, B. L. 2008. Smooth Pursuit in Schizophrenia: A Meta-Analytic Review of Research since 1993. *Brain and Cognition*, 68(3): 359-70. doi:10.1016/j.bandc.2008.08.023.

O'Tuathaigh, C. M. P., Salum, C., Young, A. M. J., Pickering, A. D., Joseph, M. H., and Moran, P. M. 2003. The Effect of Amphetamine on Kamin Blocking and Overshadowing. *Behavioural Pharmacology*, 14(4): 315-22. doi:10.1097/01.fbp.0000080416.18561.3e.

Olney, J. W., and Farber, N. B. 1995. Glutamate Receptor Dysfunction and Schizophrenia. *Archives of General Psychiatry*, 52(12): 998-1007.

Oranje, B., Gispen-de Wied, C. C. , Westenberg H. G. M., Kemner, C. Verbaten, M. N., and Kahn, R. S. 2004. Increasing Dopaminergic Activity: Effects of L-Dopa and Bromocriptine on Human Sensory Gating. *Journal of Psychopharmacology (Oxford, England)*, 18(3): 388-94. doi:10.1177/026988110401800310.

Oranje, B., Gispen-de Wied, C. C., Verbaten, M. N., and Kahn, R. S. 2002. Modulating Sensory Gating in Healthy Volunteers: The Effects of Ketamine and Haloperidol. *Biological Psychiatry*, 52(9): 887-95.

Passafaro, M., Piëch, V., and Sheng, M. 2001. Subunit-Specific Temporal and Spatial Patterns of AMPA Receptor Exocytosis in Hippocampal Neurons. *Nature Neuroscience*, 4(9): 917-26. doi:10.1038/nn0901-917.

Passie, T., Karst M., Borsutzky, M., Wiese, B., Emrich, H. M., and Schneider, U. 2003. Effects of Different Subanaesthetic Doses of (S)-Ketamine on Psychopathology and Binocular Depth Inversion in Man. *Journal of Psychopharmacology (Oxford, England)*, 17(1): 51-56.

Phillips, W. A., and Silverstein, S. M. 2003. Convergence of Biological and Psychological Perspectives on Cognitive Coordination in Schizophrenia. *The Behavioral and Brain Sciences*, 26(1): 65-82; discussion 82-137.

Radant, A. D., Bowdle, T. A., Cowley, D. S., Kharasch, E. D., and Roy-Byrne, P. P. 1998. Does Ketamine-Mediated N-Methyl-D-Aspartate Receptor Antagonism Cause Schizophrenia-like Oculomotor Abnormalities? *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, 19(5): 434-44. doi:10.1016/S0893-133X(98)00030-X.

Rao, R. P., and Ballard, D. H. 1999. Predictive Coding in the Visual Cortex: A Functional Interpretation of Some Extra-Classical Receptive-Field Effects. *Nature Neuroscience*, 2(1): 79-87. doi:10.1038/4580.

Razoux, F., Garcia, R., and Léna, I. 2007. Ketamine, at a Dose That Disrupts Motor Behavior and Latent Inhibition, Enhances Prefrontal Cortex Synaptic Efficacy and Glutamate Release in the Nucleus Accumbens. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, 32(3): 719-27. doi:10.1038/sj.npp.1301057.

Reilly, J. L., Lencer, R., Bishop, J. R., Keedy, S., and Sweeney, J. A. 2008. Pharmacological Treatment Effects on Eye Movement Control. *Brain and Cognition*, 68(3): 415-35. doi:10.1016/j.bandc.2008.08.026.

Reith, J., Benkelfat, C., Sherwin, A., Yasuhara, Y., Kuwabara., H., Andermann, F., Bachneff, S., et al. 1994. Elevated Dopa Decarboxylase Activity in Living Brain of Patients with Psychosis. *Proceedings of the National Academy of Sciences of the United States of America*, 91(24): 11651-54.

Richards, A. M. 1968. Monaural Loudness Functions under Masking. *The Journal of the Acoustical Society of America*, 44(2): 599-605.

Salinas, E., and Their, P. 2000. Gain Modulation: A Major Computational Principle of the Central Nervous System. *Neuron*, 27(1): 15-21.

Schmidt, A., Bachmann, R., Kometer, M., Csomor, P. A, Stephan, K. E., Seifritz, E, and Vollenweider, F. X. 2012. Mismatch Negativity Encoding of Prediction Errors Predicts S-Ketamine-Induced Cognitive Impairments. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, 37(4): 865-75. doi:10.1038/npp.2011.261.

Shergill, S. S., Bays, P. M., Frith, C. D., and Wolpert, D. M. 2003. Two Eyes for an Eye: The Neuroscience of Force Escalation. *Science (New York, N.Y.)*, 301(5630): 187. doi:10.1126/science.1085327.

Shergill, S. S., Samson, G., Bays, P. M., Frith, C. D, and Wolpert, D. M. 2005. Evidence for Sensory Prediction Deficits in Schizophrenia. *The American Journal of Psychiatry*, 162(12): 2384-86. doi:10.1176/appi.ajp.162.12.2384.

Shi, W.-X., and Zhang, X.-X. 2003. Dendritic Glutamate-Induced Bursting in the Prefrontal Cortex: Further Characterization and Effects of Phencyclidine. *The Journal of Pharmacology and Experimental Therapeutics*, 305(2): 680-87. doi:10.1124/jpet.102.046359.

Silverstein, S. M., and Keane, B. P. 2011. Perceptual Organization Impairment in Schizophrenia and Associated Brain Mechanisms: Review of Research from 2005 to 2010. *Schizophrenia Bulletin*, 37(4): 690-99. doi:10.1093/schbul/sbr052.

Simpson, E. H., Kellendonk, C., and Kandel, E. 2010. A Possible Role for the Striatum in the Pathogenesis of the Cognitive Symptoms of Schizophrenia. *Neuron*, 65(5): 585-96. doi:10.1016/j.neuron.2010.02.014.

Speechley, W. J., Whitman, J. C., and Woodward, T. S. 2010. The Contribution of Hypersalience to the 'Jumping to Conclusions' Bias Associated with Delusions in Schizophrenia. *Journal of Psychiatry & Neuroscience: JPN*, 35(1): 7-17.

Stephan, K. E., Baldeweg T., and Friston, K. 2006. Synaptic Plasticity and Dysconnection in Schizophrenia. *Biological Psychiatry*, 59(10): 929-39. doi:10.1016/j.biopsych. 2005.10.005.

Stephan, K. E., Friston, K., and Frith, C. D. 2009. Dysconnection in Schizophrenia: From Abnormal Synaptic Plasticity to Failures of Self-Monitoring. *Schizophrenia Bulletin*, 35(3): 509-27. doi:10.1093/schbul/sbn176.

Stone, J. M., Erlandsson, K., Arstad, E., Squassante, L., Teneggi, V. Bressan, R. A., Krystal, J. H., Ell, P. J., and Pilowsky, L. S. 2008. Relationship between Ketamine-Induced Psychotic Symptoms and NMDA Receptor Occupancy: A [(123)I]CNS-1261 SPET Study. *Psychopharmacology*. 197(3): 401-8. doi:10.1007/s00213-007-1047-x.

Sweeney, J. A., Haas, G. L., Li, S., and Weiden, P. J. 1994. Selective Effects of Antipsychotic Medications on Eye-Tracking Performance in Schizophrenia. *Psychiatry Research*, 54(2): 185-98.

Teufel, C., Kingdon, A., Ingram, J. N., Wolpert, D. M., and Fletcher, P. C. 2010. Deficits in Sensory Prediction Are Related to Delusional Ideation in Healthy Individuals. *Neuropsychologia*, 48(14): 4169-72. doi:10.1016/j.neuropsychologia.2010.10.024.

Uhlhaas, P. J, Millard, I., Muetzelfeldt, Curran, H. V., and Morgan, C. J. A. 2007. Perceptual Organization in Ketamine Users: Preliminary Evidence of Deficits on Night of Drug Use but Not 3 Days Later. *Journal of Psychopharmacology (Oxford, England)*, 21(3): 347-52. doi:10.1177/0269881107077739.

Umbricht, D., Schmid, L., Koller, R., Vollenweider, F. X., Hell, D., and Javitt, D. C . 2000. Ketamine-Induced Deficits in Auditory and Visual Context-Dependent Processing in Healthy Volunteers: Implications for Models of Cognitive Deficits in Schizophrenia. *Archives of General Psychiatry*, 57(12): 1139-47.

Weiler, M. A., Thaker, G. K., Lahti, A. C., and Tamminga, C. A. 2000. Ketamine Effects on Eye Movements. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, 23(6): 645-53. doi:10.1016/S0893-133X(00)00156-1.

Williams, L. E., Ramachandran, V. S., Hubbard, E. M., Braff, D. L, and Light, G. A. 2010. Superior Size-Weight Illusion Performance in Patients with Schizophrenia: Evidence for Deficits in Forward Models. *Schizophrenia Research*, 121(1-3): 101-6. doi:10.1016/j.schres.2009.10.021.

Winterer, G., and McCarley, R. W. 2011. Electrophysiology of Schizophrenia. Weinberger, D. R., and Harrison, P. J. *Schizophrenia*, (3rd ed.): 311-33. Oxford: Wiley-Blackwell.
http://onlinelibrary.wiley.com/doi/10.1002/9781444327298.ch15/summary.

World Health Organization. 1992. *ICD-10: International Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines*. Geneva: WHO.

Young, A. M. J., Moran, P. M., and Joseph, M. H. 2005. The Role of Dopamine in Conditioning and Latent Inhibition: What, When, Where and How? *Neuroscience and Biobehavioral Reviews*, 29(6): 963-76. doi:10.1016/j.neubiorev.2005.02.004.