Progress in Neurobiology 92 (2010) 345–369

Contents lists available at ScienceDirect

Progress in Neurobiology

journal homepage: www.elsevier.com/locate/pneurobio

# Toward a neurobiology of delusions

P.R. Corlett [a,*], J.R. Taylor [a], X.-J. Wang [b], P.C. Fletcher [c], J.H. Krystal [a]

[a] Department of Psychiatry, Yale University School of Medicine, Connecticut Mental Health Centre, Abraham Ribicoff Research Facility, 34 Park Street, New Haven, CT 06519, USA
[b] Department of Neurobiology and Kavli Institute of Neuroscience, Yale University School of Medicine, New Haven, CT, USA
[c] Department of Psychiatry, Brain Mapping Unit and Behavioural and Clinical Neurosciences Institute, University of Cambridge, School of Clinical Medicine, Cambridge,

ARTICLE INFO

*Article history:*
Received 11 February 2010
Received in revised form 6 May 2010
Accepted 8 June 2010

*Keywords:*
Delusions
Prediction
Error
Learning
Memory
Reconsolidation
Habit

*United Kingdom*

ABSTRACT

Delusions are the false and often incorrigible beliefs that can cause severe suffering in mental illness. We cannot yet explain them in terms of underlying neurobiological abnormalities. However, by drawing on recent advances in the biological, computational and psychological processes of reinforcement learning, memory, and perception it may be feasible to account for delusions in terms of cognition and brain function. The account focuses on a particular parameter, prediction error – the mismatch between expectation and experience – that provides a computational mechanism common to cortical hierarchies, fronto-striatal circuits and the amygdala as well as parietal cortices. We suggest that delusions result from aberrations in how brain circuits specify hierarchical predictions, and how they compute and respond to prediction errors. Defects in these fundamental brain mechanisms can vitiate perception, memory, bodily agency and social learning such that individuals with delusions experience an internal and external world that healthy individuals would find difficult to comprehend. The present model attempts to provide a framework through which we can build a mechanistic and translational understanding of these puzzling symptoms.



**Contents**



*Abbreviations:* AMPA receptors, α-amino-3-hydroxyl-5-methyl-4-isoxazole-propionate receptors; $D_1$ receptors, the $D_1$ subtype of dopamine receptors; $D_2$ receptors, the $D_2$ subtype of dopamine receptors; DLPFC, dorsolateral prefrontal cortex; fMRI, functional magnetic resonance imaging; NMDA receptors, N-methyl-D-aspartic acid; NAC, N-acetyl-cysteine; PFC, prefrontal cortex; OFC, orbitofrontal cortex; STS, superior temporal sulcus; TPJ, temporoparietal junction; VTA, ventral tegmental area.

* Corresponding author.
*E-mail address:* philip.corlett@yale.edu (P.R. Corlett).

0301-0082/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.pneurobio.2010.06.007

## 1. Introduction

Delusions are the extraordinary and tenacious false beliefs suffered by patients with various ailments ranging from schizophrenia (Schneider, 1959), to traumatic brain injury (Coltheart et al., 2007), Alzheimer's (Flint, 1991) and Parkinson's disease (Ravina et al., 2007), the ingestion of psychotogenic drugs (Corlett et al., 2009a) and, less frequently, autoimmune disorders such as Morvan's syndrome (Hudson et al., 2008) or potassium channel encephalopathy (Parthasarathi et al., 2006). Given this range of potential diagnoses, each with its own candidate neuropathology, it is perhaps unsurprising that we have not converged upon an agreed neurobiology of delusions. Delusions are particularly hard to study because of their insidious onset and tonic nature, their conceptual rather than behavioral basis (making them difficult to study using animal models), and the absence of a coherent theoretical model. We aim to address these issues in the current review by developing a translational model of delusion formation which we believe makes delusions tractable for animal modeling, amenable to investigation with functional neuroimaging and grounded within a theoretical framework that makes testable predictions.

Our task is made more difficult when one considers the range of odd beliefs from which people suffer; fears of persecution by clandestine forces (Melo et al., 2006); beliefs that televisions or newspapers are communicating a specific and personal message (Conrad, 1958b; Startup and Startup, 2005), the conviction that one's thoughts and movements are under the control of an external agent or are broadcast out loud (Schneider, 1959); an unrealistic belief in one's own fame or power (Karson, 1980; Kraeplin, 1902), that one is infested with parasites (Thiebierge, 1894) or deceased (Cotard, 1880), or the subject of a stranger's love (De Clerambault, 1942), or that family members have been replaced by imposters or even robots (Capgras and Reboul-Lachaux, 1923).

We take a cognitive neuropsychiatric approach to delusions. That is, the starting point is to review what we understand about the healthy functioning of a particular process, e.g. familiar face recognition, before extrapolating to the disease case, when face recognition fails and delusions of misidentification form (Halligan and David, 2001). This approach has proven successful for explaining certain delusions (Ellis and Young, 1990) but not yet for delusions in general. Perhaps this is because there are difficulties defining delusions as well as deciding what they have in common (if anything) with normal, healthy beliefs (Berrios, 1991; Delespaul and van Os, 2003; Jones, 2004; Owen et al., 2004). Beliefs are not easily accessible to the techniques of neuroscience which are more suited to representing states with clear experiential boundaries (Damasio, 2000; Knobel et al., 2008).

Furthermore, delusions are difficult to model in animals, given that they involve dysfunctions of what many consider uniquely human faculties like consciousness, language, reality monitoring and meta-cognition (Angrilli et al., 2009; Berrios, 1991; Moritz et al., 2006). Computational models of core cognitive functions (such as working memory) are being applied to gain insights into neural dysfunction in schizophrenia (Seamans and Yang, 2004; Winterer, 2006) and some are beginning to address the phenomenology of specific psychotic symptoms (Loh et al., 2007), however, these models have focused on circuit mechanisms within a local area (like prefrontal cortex), they are unable to capture the content of particular symptoms which involve information processing across large networks of interacting brain regions (Fuster, 2001).

There is a need for a testable conceptual model of delusions, one that is rooted in translational cognitive neuroscience. We, and others, propose that beliefs (both normal and abnormal) arise through a combination of innate or endowed processes, learning, experience and interaction with the world (Friston, 2010). Like other forms of information, beliefs are represented in the brain through the formation and strengthening of synaptic connections between neurons, for example causal beliefs may be mediated by a strengthening of the synaptic associations between pools of neurons representing a particular cause and their counterparts representing an associated effect (Dickinson, 2001; McLaren and Dickinson, 1990; Shanks, 2010). There are neural (and hence cognitive) limits set on the range of possible connections that can be made (Kandel, 1998). The strength of those connections is modifiable such that those conveying an adaptive advantage are strengthened and those that are disadvantageous are weakened (Hebb, 1949b; Thorndike, 1911).

This set of sculpted connections is used to predict subsequent states of the internal and external world and respond adaptively (Friston, 2005b); however, should that next state be surprising, novel or uncertain new learning is required (Schultz and Dickinson, 2000). Our premise is based upon the idea that the brain is an inference machine (Helmholtz, 1878/1971) and that delusions correspond to false inference. This inference is necessarily probabilistic and rests upon some representation of predictions (prediction error) and uncertainty (i.e. precision) about those predictions. Within this framework, we see delusions as maladaptive beliefs that misrepresent the world. They might arise through any number of perturbations within this scheme, from an unconstrained specification of the possible or lawful set of neural connections (Hoffman and Dobscha, 1989); providing the potential for bizarre beliefs to form (Hemsley and Garety, 1986a), to an adventitious and inappropriate reinforcement of particular neural connections (King et al., 1984; Shaner, 1999); engendering unexpected percepts, attentional capture and beliefs that deviate grossly from reality (Corlett et al., 2009a, 2007a; Fletcher and Frith, 2009). Impaired predictive mechanisms have been previously implicated in delusions of alien control; whereby the sufferer believes their movements are under the control of an external agent because of an inability to appropriately predict the sensory consequences of their actions (Frith et al., 2000b). We propose that this account generalizes from actions to numerous cognitive processes, that predictive learning and prediction errors are general mechanisms of brain function (Friston, 2005b; Schultz and Dickinson, 2000) and that aberrant predictions and prediction errors provide a unifying explanation for delusions with disparate contents.

A crucial distinction, which we will appeal to repeatedly, is between prediction errors per se and the precision or uncertainty about those errors. We will develop the argument that delusions (and their neurotransmitter basis) represent a failure to properly encode the precision of predictions and prediction errors; in other

words, a failure to optimize uncertainty about sensory information. Here, prediction errors encode information that remains to be explained by top-down predictions (Rao and Ballard, 1999). This distinction is important because it is easy to confuse the role of phasic dopaminergic discharges as encoding reward prediction error (Montague et al., 1996; Schultz et al., 1997), and the role of dopamine in modulating or optimizing the precision of prediction errors that may or may not be reward-related (Friston et al., 2009), for example by modulating the signal to noise response properties of neural units encoding prediction error. In what follows, we will assume that the pathophysiology of delusions involves a misrepresentation of salience, uncertainty, novelty or precision (mathematically precision is the inverse of uncertainty). Biologically, this corresponds to aberrant modulation of post-synaptic gain that, presumably, involves NMDA receptor function (Friston, 2010). This fits comfortably with the role of dopamine in controlling signal to noise and the numerous proposals that dopamine (at least in terms of its tonic discharge rates) encodes uncertainty or violation of expectations (Fiorillo et al., 2003; Preuschoff et al., 2006).

The challenge is to provide empirical data that test the hypothesis. Numerous investigators have accepted this challenge and, by sharing a set of common simplifying assumptions, we are beginning to develop an understanding of delusions in the brain. Here, we review this growing understanding, beginning with a set of principles which, we believe, are important in developing our understanding of the neurobiology of delusions.

## 2. Reductionist principles for a neuroscience of delusion

The four principles are as follows: Beliefs and memories share cognitive and neural mechanisms (1); learning memory and belief influence perception (2); affect impacts upon learning and memory and hence belief (3); our sense of self, agency, free will and beliefs about others are governed by the same simple neural learning mechanisms (4). By taking a reductionist approach, grounded in formal animal learning theory, computational and cognitive neuroscience we can begin to tackle the hard problems of belief, delusion, and the brain; problems often considered beyond the scope of neuroscience. Below, we consider the principles in more detail before discussing their implications for understanding the cognitive neuroscience of delusions.

### 2.1. Beliefs and memories share cognitive and neural underpinnings

Beliefs are notoriously difficult to define (Dennett, 1995), but generally refer to the attitude we have with regard to propositions about the world. Perhaps a pragmatic analysis might help. What functions do beliefs serve? Like memories, beliefs help us to organize incumbent information and coordinate adaptive responses (Dennett, 1995). In other words, though beliefs and memories are based on past experiences they are utilized to predict the future and respond accordingly (Corlett et al., 2009a). The most rigorous and formal definition of beliefs appeals to probability theory, and in particular Bayesian formulations (Bayes, 1763). This framework, which we use later, associates beliefs with probability distributions that are represented by the brain (Fiser et al., 2010). These comprise posterior beliefs that are conditioned upon sensory information and are constrained by prior beliefs. In the context of hierarchical Bayesian inference, the posterior belief (having seen the evidence) rests on empirical priors. Empirical priors are prior beliefs that are themselves optimized during hierarchical inference (Friston, 2005b). Assuming that the brain uses hierarchical inference to make predictions about the world, most of the beliefs it entertains can be regarded as empirical prior beliefs. From now on, we will refer to these as prior beliefs or priors and associate these with the source of top-down predictions that are used to form prediction errors. Some have equated beliefs with stimulus-response habits in experimental animals: the behaviors that track previously experienced contingencies but are insensitive to alterations in those contingencies (Eichenbaum and Bodkin, 2000). Indeed, in view of their tenacity and tendency to misrepresent true contingency, some have pointed out the similarities of beliefs to superstitious behaviors (Beck et al., 2007). Thus, beliefs, and therefore delusions, are regarded as representing adventitiously reinforced superstitions; predictions about the future that were formed accidentally and inappropriately but that nevertheless persist (Freeman et al., 2009; Shaner, 1999). Despite capturing aspects of belief phenomenology, these theories offer neither a mechanistic nor a neurobiological explanation of belief or delusion formation. This is what we seek here.

One compelling approach equates the process of human belief formation with Pavlovian conditioning. The same processes that drive animals to learn predictive associations between sensory stimuli and salient events (rewards or punishments) also contribute to the acquisition of beliefs in humans (Dickinson, 2001). Expectancy and experience, or, more specifically, mismatches between the two, are crucial for learning (Alloy and Tabachnik, 1984; Courville et al., 2006; Waldmann and Martignon, 1998). This mismatch, or prediction error, is central to formal associative learning theories, driving learning directly (Rescorla and Wagner, 1972) and indirectly, via the allocation of attention toward potentially explanatory cues (Pearce and Hall, 1980). However, there is also a tendency to focus on, and learn about, highly salient stimuli that consistently predict important consequences (Mackintosh, 1975). Under one account (Grossberg, 1982), the occurrence of an expected event that matches an active expectancy will amplify its representation in short-term memory, increasing the likelihood that it will be consolidated within long-term memory as well as the strength of this consolidation. By contrast, when an unexpected event violates the active expectancy, an orienting system is activated which resets short-term memory (dropping active expectancies) and engages an orienting response, permitting the acquisition of new explanatory associations. In essence, organisms learn associations between stimuli, events, thoughts and percepts to build an internal model of their environment (Sokolov, 1960; Tolman, 1932). This model is itself predictive and, whenever significant novelty is detected due to a mismatch between its predictions and actual experience it must be updated (Grossberg, 1982). In short, the allocation of attention toward appropriately salient events depends upon the optimization of the precision of top-down priors, relative to bottom-up evidence; both in sensory cortices [involving acetylcholine (Yu and Dayan, 2005)] and in fronto-striatal circuits [involving dopamine (Friston et al., 2009)].

This presents the organism with a challenge: to navigate the world successfully, we must sustain a set of prior beliefs (our internal model), sufficiently robust that we do not react reflexively and chaotically to any incoming sensory stimulus. At the same time, these beliefs (priors) must not be so immutable that our responses become fixed, stereotypical and insensitive to change (Corlett et al., 2009a). According to learning models of delusions, during the earliest phases of delusion formation aberrant novelty, salience or prediction error signals drive attention toward redundant or irrelevant environmental cues, the world seems to have changed, it feels strange and sinister, such signals and experiences provide an impetus for new learning which updates the world model inappropriately, manifest as a delusion (Corlett et al., 2009a, 2007a; Gray, 2004; Gray et al., 1991; Hemsley, 1994; Kapur, 2003). The insight relief that delusions bring engages strong memory consolidation, furthermore, they are deployed reflexively in response to similar aberrant experiences (Mishara and Corlett, 2009) and as such, they are rapidly rendered impervious to contradiction (Corlett et al., 2009a, see below).

*2.1.1. Neural instantiation of predictive learning and belief*

Midbrain dopamine neurons in substantia nigra (SN) and ventral tegmental area (VTA) code a reward prediction error (Montague et al., 1996; Schultz et al., 1997). When primates (Schultz et al., 1993; Waelti et al., 2001) and rodents (Takahashi et al., 2009) learn, activity in these neurons reflects a mismatch between expected and experienced reward that is redolent of the prediction error signal from formal learning theories (Waelti et al., 2001) and machine learning models (Montague et al., 1996; Sutton and Barto, 1998). However, recent studies have identified punishment prediction error signals (Matsumoto and Hikosaka, 2009) and mismatches between expected and experienced information (Bromberg-Martin and Hikosaka, 2009) in distinct anatomical populations of midbrain dopamine neurons, suggesting that these neurons and the circuits in which they are embedded are involved in the processing of salient events that will guide future adaptive behavior, for both positively and negatively valenced events (Hikosaka et al., 2008a). In human subjects, a circuit involving the midbrain and its projection sites in the striatum and prefrontal cortex signal prediction errors that guide causal learning (Corlett et al., 2004; Fletcher et al., 2001; Turner et al., 2004).

Prediction error driven learning and memory may represent a basic mode of brain function, referred to as predictive coding (Friston, 2005b, 2009; Schultz and Dickinson, 2000), that is, brains, component brain systems and even single neurons minimize uncertainty about incident information (either external or internal) by structurally or functionally embodying a prediction and responding to errors in the accuracy of the prediction (Fiorillo, 2008). Rapid excitatory and inhibitory neurotransmitters (glutamate and GABA) interact with slower neuromodulatory transmitters to instantiate this predictive coding scheme (Friston, 2005b, 2009), but the precise mechanism for computing prediction error signals remains poorly understood. Across successive levels of cortical hierarchies, top-down signaling from neurons in layers higher up the hierarchy confers expectancies, possibly through glutamatergic NMDA receptors but this is still not established empirically. Bottom-up inputs to a layer are signaled from the layer below through fast glutamatergic and GABAergic mechanisms. At a given level, any mismatch between expectancy and experience is transmitted up the cortical hierarchy to the level above via AMPA receptor signaling (Angelucci et al., 2002a,b; Friston, 2005b, 2009; Sherman and Guillery, 1998). Slower neuromodulatory transmitters, like dopamine, acetylcholine, serotonin and cannabinoids are engaged (Corlett et al., 2009a), mediating the post prediction error response by encoding the precision of or uncertainty associated with a particular prediction error (Friston, 2005c). Such uncertainty signals engage subsequent processing such as enhancing neural maintenance of working memory (Lavin et al., 2005) and modulating synaptic plasticity down the hierarchy thus tuning subsequent responses (Grace, 1991; Herrero et al., 2008). We shall refer this perspective on cortical processing, through feedforward signaling of sensory stimuli and feedback signaling of expectation and priors, as the Bayesian model.

According to this model, a prior belief is updated by prediction errors to provide a probabilistic prediction of expected inputs. Input probabilities are learnt at synapses by virtue of experience-dependent learning (Soltani and Wang, 2010), and read out at the level of neural activity populations (Ma et al., 2006). However, beliefs and priors are more than expectancies; strong prior beliefs can enhance, attenuate or vitiate sensed inputs sculpting them to conform to expectations (Jaspers, 1963). The power of prior expectancies can be observed in visual illusions, for example the hollow mask illusion in which a hollow mask is perceived as a convex face as a result of extended lifetime experience that faces are not concave but convex. Likewise strong neural priors can sculpt input signals so that they conform to expectancies (Rao and Ballard, 1999). Beliefs then, not only provide a mechanism through which current information is interpreted in light of the past; they involve an inductive inference that ensures experiences conform with expectancies (Clifford, 1877/1999). In associative learning, such behavioral inflexibility involves training in which expectancies are continuously confirmed (Adams and Dickinson, 1981). The representations and neural circuits controlling behavior gradually shift from more plastic goal-directed, knowledge-based frontal, temporal and ventral striatal regions of the brain toward more inflexible habitual behavior, decreased involvement of frontal cortices and a shift toward dorsal striatal circuits (Belin et al., 2009; Daw et al., 2005; Eichenbaum and Bodkin, 2000). This shift is marked by an increasing strength of the behavior even when the contingency no longer pertains or when the consequences of that behavior are no longer desired.

Whilst Bayesian models are often considered rational and optimal (Shanks, 2006), they have nevertheless been deployed to explain irrational processes such as the spread of panic and rumor within a crowd (which occurs rapidly in salient situations with few explanatory priors; Butts, 1998) and, more recently, a biophysically plausible model offers an explanation for base rate neglect in probabilistic decision making (Soltani and Wang, 2010). Essentially we advocate an explanation of delusions as a disruption to the normal Bayesian predictive mechanisms of the brain such that predictable and irrelevant events mismatch with expectancies and their salience demands new learning and explanation; a delusion represents an explanatory mechanism, an attempt to impose order on a disordered perceptual and cognitive world (McReynolds, 1960; Maher, 1974; Gray et al., 1991; Kapur, 2003; Corlett et al., 2007a,b, 2009a; Fletcher and Frith, 2009).

*2.1.2. Oscillation signatures of match and mismatch events*

In our introduction we alluded to the importance of dysfunctional neural circuits (rather than isolated regions) when considering the pathophysiological mechanisms underpinning delusions. That is, psychoses could be conceived as 'disconnection syndromes' (Friston and Frith, 1995). Inter- and intra-regional neural connections and disconnections are still poorly understood at the present time. One of the active research areas is the examination of the role of neural oscillations in inter-areal communication (Uhlhaas et al., 2008, 2006a; Uhlhaas and Singer, 2010). For example, oscillatory activity in the gamma frequency band (30–50 Hz) contributes to synchronizing populations of neurons in different brain regions, mediating the temporal structuring of neural activity necessary for sharing, transfer and storage of information (or learning) between these groups of coordinated cells or cell assemblies (Buzsaki, 2007). Such oscillations are thought to reflect the engagement of high-level cognitive processes such as attention (Joliot et al., 1994). A recent computational model of selective attention, consisting of a reciprocally connected loop between a sensory circuit and a high-level cognitive circuit, found that top-down signaling enhances gamma-band oscillatory coherence only when there is a match between the attended stimulus feature (expectation) and the actual stimulus feature (experience), and that this occurs exclusively in sensory neurons selective for the actual feature and in memory neurons (that are the source of top-down signaling) selective for the attended feature (Ardid et al., 2010).

Learning from the violation and confirmation of our expectancies can both be traced in oscillatory activity of recurrent neural circuits (Grossberg, 2009). Match-based learning captures the Hebbian notion of cell assemblies; collections of synaptically interconnected cells whose pre- and post-synaptic firing correlates and becomes mutually excitatory such that when a fraction of an input pattern is incident upon the assembly, the whole output is

realized (Hebb, 1949a). In human learners, gamma oscillations (measured using EEG) increase during acquisition of new associations, as does the coherence of oscillations in cortical regions representing the stimuli being associated (Miltner et al., 1999). Neural synchrony impacts on learning because synaptic plasticity depends on the timing of pre- and post-synaptic neural spikes (Bi and Poo, 2001).

But as we have observed, learning does not proceed by contiguity alone (Konorski, 1948). Cell assemblies also represent events that do not match our expectancies (O'Donnell, 2003). In terms of synaptic machinery, one type of mismatch-based learning, which is based on expected rewards, appears to be implemented in the mesocorticolimbic system through a tri-synaptic arrangement between pre and post-synaptic glutamatergic signaling with a modulatory role for the dopaminergic prediction error input from VTA (Pennartz et al., 2000; Schultz, 1998). Ensembles of neurons are defined by their membrane potential states; periods of very negative resting membrane potential or down states are periodically interrupted by a plateau depolarization or Up state (Haider et al., 2006; Ros et al., 2009; Sanchez-Vives and McCormick, 2000). Striatal up states are synchronized with those in frontal cortex (Goto and O'Donnell, 2001). Dopamine $D_2$ receptor signaling is associated with an instability of prefrontal representations (Seamans and Yang, 2004), providing an ensemble-level mechanism for surprise driven resetting of representations, search and new learning (Braver and Cohen, 1999; Grossberg, 1982). On the other hand, dopamine, acting through $D_1$ receptors and their interaction with NMDA channels facilitates the maintenance of Up states in target neurons (Cepeda and Levine, 1998; Wang and O'Donnell, 2001) and reinforces cell assemblies representing expected salient events (O'Donnell, 2003). In this scheme, the excessive $D_2$ signaling, impaired $D_1$ and impoverished NMDA signaling that comprise psychotic states would lead to a poor specification of prior expectancies and fronto-striatal cell assemblies comprised of cells representing merely coincident events and spurious associations.

But, how are predictions and prediction errors reflected more generally in the oscillatory signals of cortical hierarchies? While gamma oscillations are commonly enhanced under conditions that involve cognitive control, the top-down specification of priors may be reflected in beta-band (15–30 Hz) oscillations (Wang, in press). For instance, when recordings are made from the lateral intraparietal cortex and prefrontal cortex of behaving monkeys during a visual search task, the inter-areal coherence is enhanced in the beta-frequency band when the target and distractors are similar and visual search depends on top-down signaling, relative to when the target and distractors are dissimilar and target detection is based by feedforward perceptual 'pop-out' (Buschman and Miller, 2007).

Cortical areas have a well defined laminar structure and, in the neocortex, gamma-band oscillations are prominent in superficial layers (2/3) endowed with abundant horizontal connections (Binzegger et al., 2004; Buhl et al., 1998). In contrast, deeper layers (5/6) tend to display lower frequency beta-band oscillations (Wang, in press). Between reciprocally connected cortical areas, feedforward projections from a lower to a higher area originate in superficial layers of the lower area. Feedback connections begin in deep layers of the higher area and project to superficial layers of the lower area as well as subcortical structures. Thus beta oscillations, produced in the deep layers, may be especially involved in long distance signaling along feedback pathways. Top-down beta oscillations may encode the expectations that guide match-based learning and perception (Berke et al., 2008). Moreover, prior specifying, beta-frequency oscillatory feedback signals emanating from a 'cognitive area' project to superficial layers 2/3 in a 'sensory area', hence are well suited to modulating gamma oscillations that are locally generated in the superficial layers, in a context dependent manner (Wang, in press).

There are competing theories regarding the roles of different oscillatory bands in conveying neuronal predictions and prediction errors (Grossberg, 2009). For example, the relationship between high frequency gamma and lower frequency theta band oscillations in hippocampal neurons appears important for the recall of temporal sequences of events (Lisman and Buzsaki, 2008), this form of coding may be especially important in specifying predictions about the future (Lisman and Redish, 2009) and, if it is disrupted, prediction errors may result (Lisman and Grace, 2005); these aberrant errors may be propagated to target structures though inappropriate entrainment of oscillations between structures (Sirota et al., 2008). Furthermore, there are magnetoencephalography data suggesting that, during a face perception task in human subjects, higher-frequency gamma oscillations at lower levels of a neural hierarchy can entrain lower frequency (alpha-band) oscillations in regions higher up the hierarchy, which may represent accumulating prediction error for perceptual synthesis (Chen et al., 2010). Through non-linear coupling, gamma oscillations in the higher region increase, providing a mechanism through which ascending prediction errors are damped down or explained away (manifest as a decrease in alpha-band power; Chen et al., 2010). More data are clearly required. However, we can predict that in delusion-prone individuals, if predictions are poorly specified and errors signaled inappropriately, then low frequency oscillations, gamma oscillations and their interaction should be perturbed. Consistent with this prediction, highly schizotypal subjects have electrocortical responses to sensory stimulation in the gamma and beta frequency ranges that were slower to habituate following repeated presentation of the stimuli, indicative of maladaptive prior expectancies as well as aberrant prediction error responses (Vernon et al., 2005). Furthermore, patients with schizophrenia have reduced long-range phase synchrony in the beta-band during gestalt stimulus perception, perhaps indicative of aberrant prediction error. This aberrant signaling correlated with delusion severity across subjects (Uhlhaas et al., 2006a).

### 2.1.3. Delusions as aberrant neural learning

Excessive and inappropriate dopamine signaling is thought to render merely coincident events highly salient (Gray et al., 1991; Hemsley, 1994; Kapur, 2003), this may result from a dysfunction in glutamatergic and GABAergic signaling and thence, the regulation of dopamine signaling (Carlsson et al., 2001; Grace, 1991; Laruelle et al., 2003). Either directly or indirectly, this dysregulation leads to the inappropriate signaling of prediction error (Corlett et al., 2007a; Grace, 1991; Gray et al., 1991). Since prediction error may guide attention toward events that may explain the feeling of surprise or uncertainty (Pearce and Hall, 1980) and engage learning mechanisms (Rescorla and Wagner, 1972), we can see that such a disruption has could lead to altered attention, learning, and ultimately belief formation.

To consider the nature of this disruption in a little more detail, inappropriate prediction error signals could be conceived of as resulting from a change in the signal to noise properties of dopamine signaling (Grace, 1991; Miller, 1976; Spitzer, 1995); due to deficits in glutamatergic regulation of VTA dopamine neurons. Physiological noise is perceived by the system as real signal that engenders the cascade of events that a true prediction error would engage, namely a search for explanation and new learning. Ultimately, both of these possibilities; inappropriate prediction error and an altered signal to noise ratio of the dopamine system; are reflective of poor precision in the estimation of prediction error (Friston et al., 2009; Preuschoff et al., 2006), which will vitiate inference, biasing it toward misrepresenting inputs (be they sensory or neural). If persistent, this imprecision may ultimately

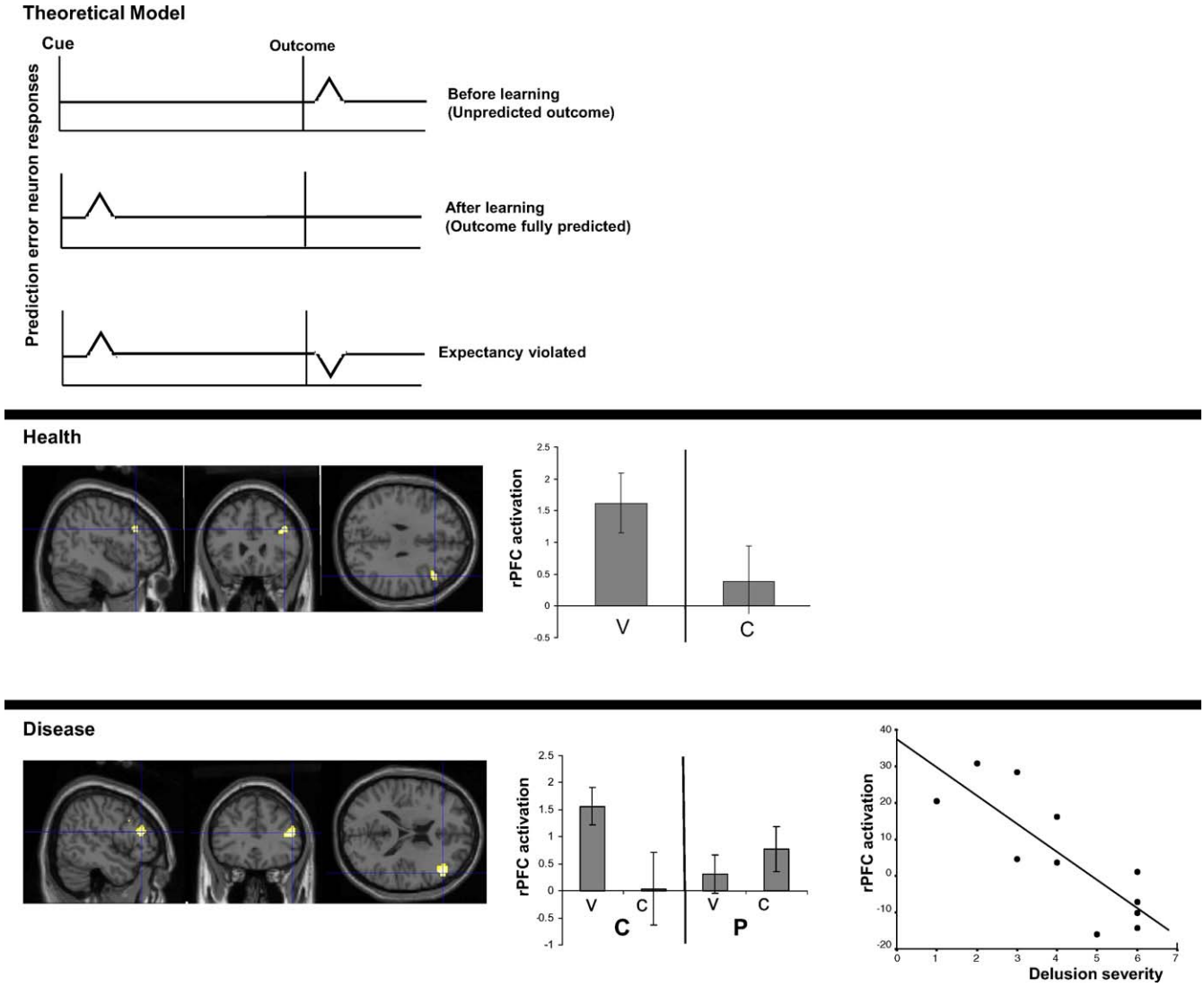


**Fig. 1.** Neural instantiation of predictive learning and belief. *Theoretical model*: Schematic of reward prediction error signals before learning, following learning and during extinction. *Health*: Right DLPFC prediction error response during casual learning in healthy subjects (Corlett et al., 2004) – V: violation of expectancy, C: confirmation of expectancy. *Disease*: Aberrant right frontal prediction error response in patients with first episode psychosis. The more profound the disruption, the more severe the delusions (Corlett et al., 2007b) – C: controls, P: patients with psychosis.

lead to the formation of a new explanatory prior, or delusion, that consolidates the misrepresentation allowing it to pervade the deluded individual's future perception and action (Jaspers, 1963). Aberrant mesocorticolimbic prediction error signals have been recorded during causal learning with functional neuroimaging in patients with schizophrenia and furthermore, the magnitude of those signal aberrations correlated with the severity of delusions across subjects (Corlett et al., 2007b) [see Fig. 1].

The relationship between conditioning and delusions has also been confirmed in the context of a reward learning task (Schlagenhauf et al., 2009) and an aversive conditioning task (Holt et al., 2008); in both cases, aberrant learning was related to the severity of delusional beliefs. It appears that the brain systems that govern normal causal belief formation are internally and inappropriately engaged when delusions form.

### 2.1.4. Multiple neural origins for prediction error and its dysfunction?

The computation of VTA prediction error signals involves the interplay between the basal ganglia and the prefrontal cortex (Soltani and Wang, 2010), especially the anterior cingulate cortex (Matsumoto and Hikosaka, 2007; Behrens et al., 2009) and the orbitofrontal cortex (Takahashi et al., 2009). Other studies point to hippocampus, specifically for signaling novelty in the form of mismatches between actual and expected information (i.e. prediction errors) which may then be transmitted to the VTA via the striatum (Lisman and Grace, 2005). This signaling of unexpected and salient events causes the organism to stop its ongoing behavior and search for explanatory cues (Gray et al., 1991). Patients with psychosis have increased regional cerebral blood flow (an indirect measure of neural activity) in CA1 and, those in whom this effect is most pronounced have the most severe delusions (Schobel et al., 2009). Likewise, individuals in the prodrome (the very earliest phases of psychosis) release more striatal dopamine than controls and again, the magnitude of that dopamine release correlates with the severity of delusion-like ideas (Howes et al., 2009). Contrary to the predictions of Gray and Kapur, this dopamine dysfunction has been observed not in the limbic striatum but in the associative striatum, a sub-region that is reciprocally connected with the dorsolateral prefrontal cortex (Haber, 2003; Haber et al., 2006). The latter is a part of the circuit engaged by prediction error driven learning and, moreover, shows aberrant responses in subjects experiencing disturbed percepts and odd beliefs (Corlett et al., 2004, 2006, 2007b). We discuss these observation in more detail below.

The rapidity with which reward prediction error signals are registered in VTA (of the order of milliseconds) may be incommensurate with the calculation of a reward prediction error (Redgrave and Gurney, 2006). Instead these signals could represent unexpected sensory events through cholinergic inputs from the pedunculopontine tegmentum (Dommett et al., 2005), or PPT, inputs which are combined with context representations from the prefrontal cortex and hippocampus as well as motor representations from the putamen in order to ascertain whether the organism or the environment was responsible for the unpredicted event. This agency account suggests that dysfunctions in dopamine signaling could explain both the sense of excessive agency for events in the world associated with paranoia (Kaney and Bentall, 1992) as well as the externalization of agency associated with delusions of passivity (Blakemore et al., 2002; Frith et al., 2000a). See below.

A further candidate site for prediction error dysfunction in psychosis is the habenula (Shepard et al., 2006). The habenula, in concert with the prefrontal cortex, is responsible for instantiating negative prediction error signals in the VTA; the dips below baseline firing that engage extinction learning; abandoning what we have previously learned in favor of a new prediction (Pan et al., 2008). A deficit in this signaling would raise baseline mesocorticolimbic dopamine levels (Lecourtier et al., 2008) and impair extinction learning (Holt et al., 2008; Waltz et al., 2007), perhaps explaining why deluded individuals stick with maladaptive and erroneous ideas (or corticostriatal cell assemblies) despite their demonstrable falsehood (Corlett et al., 2009b).

Bringing these observations together, it appears that the mesocorticolimbic dopamine system codes numerous types of expectation, their violation and the new learning that expectancy violation engenders; permitting adaptation to prevailing environmental contingencies (Schultz and Dickinson, 2000). When events that violate perceptual expectations are experienced, the hippocampal projection to the striatum engages a broader population of dopamine neurons in the VTA (Lodge and Grace, 2006a). Furthermore, the prefrontal cortex maintains higher level expectancies representing goals and the actions required to achieve those goals (Grace et al., 2007; Sesack and Grace, 2010) as well as reward values for sensory stimuli (Takahashi et al., 2009) and actions (Behrens et al., 2009). When events occur that violate those expectancies, PFC modulates the responses of active VTA dopamine neurons: engaging burst firing through its influence over the PPT (Lodge and Grace, 2006a,b), allowing updating of expectancies through new learning. Furthermore, PFC enables the quiescence of those same VTA neurons (through its influence on the habenula) when contingencies change and learning is extinguished (Hikosaka et al., 2008b; Pan et al., 2008). Reciprocal connections between VTA, PFC, striatum and hippocampus are involved in this updating process so that future expectancies conform to the prevailing environmental contingencies.

We predict that delusions are associated with a threefold disturbance in this circuitry: (i) Excessive hippocampal drive to VTA (via striatum) engaging a broader population of VTA dopamine neurons; (ii) Inappropriate engagement of PPT due to PFC dysfunction, instigating burst firing in that expanded pool of recruited neurons and (iii) Impaired habenula mediated inhibition of VTA dopamine neurons (which would normally instantiate extinction learning when an expected event fails to occur).

These three deficits would confer the cardinal characteristics of delusions, their bizarreness and tenacity: Bizarreness; due to the aberrant recruitment of VTA cells and their incorporation into cell assemblies which sculpt future expectancies; and tenacity; due to the failure of PFC to control the habenula, and hence coordinate the dips in VTA neuron firing below baseline that engage extinction learning when the predictions of the delusion are not borne out.

While this model begins to implicate aberrant learning processes in delusion formation, it does not address the range of different themes that form the content of delusions, nor does it fully explain the behaviors in which deluded individuals engage when confronted with evidence that challenges their belief (see below). In order to extend out explanation to encompass these characteristics, we discuss below what we consider to be key factors: the role of beliefs in instrumental conditioning (learning the relationships between our actions and their effects in the world) and the impact of repeated recall and rehearsal of that information on subsequent processing.

### 2.2. Learning, memory and belief alter perception

Perception is substantially constructive. That is, our expectancies (based on previous experience) contribute to what we currently perceive by sculpting sensory inputs (Bruner et al., 1949; Helmholtz, 1878/1971). The concepts and categories we have learned through experience can influence what we perceive, for example, if subjects are shown simple objects and asked to reproduce their colors, their responses are heavily influenced by the shape of the object (Goldstone, 1995). Motivational state, itself interacting with learning and memory (Berridge, 2007), can impact upon perceptual judgment; poorer children judge coins to be larger and heavier than do richer children (McCurdy, 1956). When presented with noisy, unstructured visual inputs, hungry subjects claim to see food objects (Atkinson and McClelland, 1948). The impact of motivation on bottom-up perceptual inputs may be mirrored in the mechanisms we use to imagine given percepts, a mechanism which, when inappropriately engaged may elicit hallucinations (Grossberg, 2000) and the impaired reality monitoring associated with delusional ideation (Simons et al., 2008)). For example, the spontaneous confabulations of patients with orbitofrontal lesions represent an excessive influence of past experience on current perception (Schnider, 2001) and delusional misidentification may reflect a failure to specify perceptual expectations such that known people or places lack a sense of familiarity (Fleminger, 1992).

#### 2.2.1. Neural mechanisms of the memory–perception cycle

Predictive coding and prediction error may be a basic mode of brain function (Friston, 2005b, 2009; Mesulam, 2008). This theory is best encapsulated by sensory cortices, in particular the visual cortex; whose anatomy recapitulates the idea of a hierarchically organized predictive coding system. Further up the neural hierarchy, more distal to the site of sensory input, approaching association cortices, the representations of sensory stimulation become more abstract (Mesulam, 2008). But the percept does not emerge as a consequence of a simple uni-directional progression up this hierarchy (Sperry, 1990). Rather the hierarchy is nested (Feinberg, 2000; Feinberg and Keenan, 2005) or enriched by interactions (feedback as well as feedforward) between its layers (Friston, 2005b, 2009). These interactions are instantiated by sparse and rapid feedforward AMPA and GABAergic signaling meeting feedback (possibly NMDA-mediated) signaling representing predicted inputs embodied in the layers above (Friston, 2005b). Any mismatch between expectancy and experience (signaled via AMPA receptors) can serve to update future priors. Dysinteractions within this Bayesian hierarchical arrangement may be responsible for the symptoms of psychosis (Corlett et al., 2009a; Fletcher and Frith, 2009). For example, in the absence of stable prior expectancies, certain perceptual illusions may not be perceived by patients with schizophrenia (Dakin et al., 2005; Dima et al., 2009; Emrich, 1989) nor individuals administered NMDA receptor antagonists (Phillips and Silverstein, 2003) perhaps indicative of a common underlying mechanism [although see Passie et al. (2003)

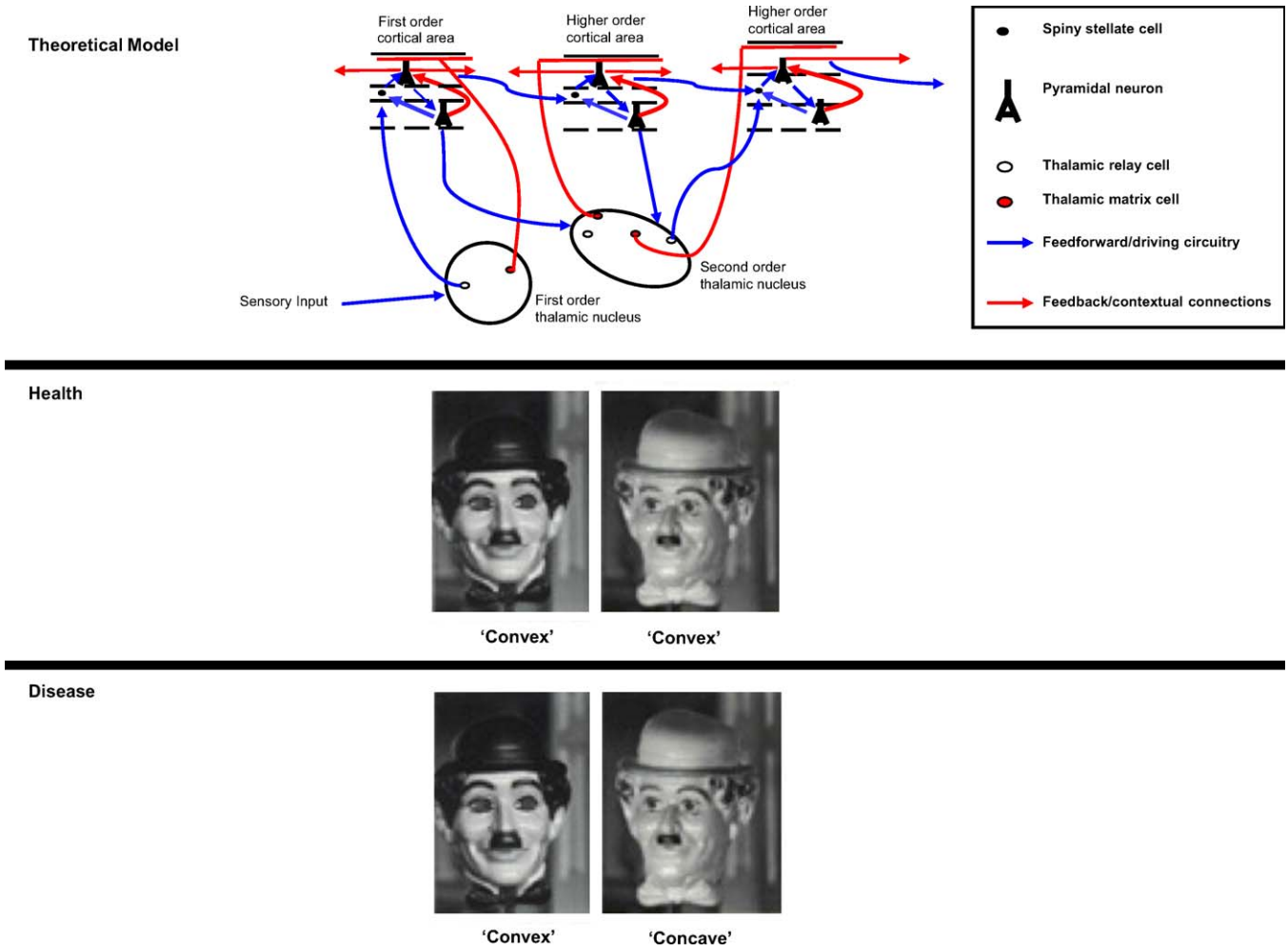


**Fig. 2.** Learning memory and belief alter perception. *Theoretical model*: Feedforward and feedback thalamocortical projections (adapted from http://wiki.tkk.fi/display/SYNB/Neocortex). *Health*: The rotating hollow mask is continuously perceived as convex due to our consistent experience of faces as convex. *Disease*: Individuals prone to or experiencing psychosis report the hollow mask as a hollow percept (Emrich, 1989).

for a dissociation between the effects of ketamine on a perceptual illusion and its psychotomimetic effects]. See Fig. 2.

The thalamus has also been strongly implicated in conscious perception. Thalamocortical circuits have intrinsic resonance in the gamma frequency range which is critical for conscious perception, prediction and learning (Steriade et al., 1991). GABAergic neurons in the basal ganglia projecting to the thalamus exert an inhibitory influence on thalamocortical neurons thus protecting the cortex from sensory overload (Sharp et al., 2001). Hyperactivity of dopamine or hypo-activity of glutamate in the striatum would compromise these protective mechanisms leading to excessive cortical stimulation and psychosis (Carlsson et al., 2001; Carlsson and Carlsson, 1990; Geyer and Vollenweider, 2008). Such a deficit could conceivably alter the sense of background and foreground that permeates normal perception (Conrad, 1958a). This could explain why other Gestalt principles, which involve grouping the perceptual field on the basis of learned environmental regularities (Fiser, 2009; Vickery and Jiang, 2009), are impaired by psychotomimetic drugs that alter dopaminergic and glutamatergic function (Kurylo and Gazes, 2008). Gestalt organizing principles are similarly disrupted in patients with schizophrenia (Silverstein et al., 2006; Uhlhaas and Mishara, 2007; Uhlhaas et al., 2006b).

Like other systems, thalamocortical circuits and their interaction with cortical information processing have been subject to a Bayesian analysis (Koechlin et al., 1996). According to this scheme, thalamocortical information represents the feedforward aspect (the information being represented) and cortico-cortical processing represents the prior expectancies, the operations to be performed on that information. Similar models have been developed to account for perception of coherent visual motion and mental rotation, as well as the predictive functions involved in enacting adaptive movements (Koechlin et al., 1996; Llinas and Roy, 2009).

Inherent in all of these related schemes is the notion of a balance, between bottom-up and top-down (or feedforward and feedback) signaling. This balance is necessary in order to meet the afore-mentioned challenge of a system that is robust to noisy inputs (through reliance on empirically derived prior expectations) but is also flexibly responsive to new contexts and situations (through the capacity to alter priors on the basis of bottom-up signal). With this in mind, it is clear that, in addition to poorly specified predictions, excessively strong priors may be profoundly disruptive and psychotogenic (Corlett et al., 2009a). Perceptual associations between sensory modalities appear to be learned using mesocorticolimbic prediction error signals (den Ouden et al., 2010, 2009), which may explain the phenomenon of sensory conditioned hallucinations (Ellison, 1941; Seashore, 1895) whereby, learned associations between sensory stimuli (a tone predicts light stimulation for example) alter perception, such that

presentation of one stimulus (tone) induces experience of the other (light) even though the latter is not present. Learned associations can alter perception; Hallucination-prone individuals are more susceptible to experiencing sensory conditioned hallucinations (Kot and Serper, 2002). Likewise delusional beliefs can alter percepts such that they conform to the delusion (Jaspers, 1963). Excessively strong top-down predictions may explain the psychotogenic effects of LSD and sensory deprivation (Corlett et al., 2009a). Furthermore, individuals prone to abnormal experiences and beliefs are more susceptible to the Deese–Roediger–McDermott memory illusion whereby they claim to have experienced an event that was strongly expected but nevertheless did not occur (Corlett et al., 2009c). We predict that such expectation-based psychotic phenomena would be associated with inappropriate gamma and beta oscillations, reflective of inappropriate reverberatory activity in recurrent neural circuits and of pattern completion within Hebbian cell assemblies that are not relevant to the situation at hand.

### 2.3. Affect impacts upon learning, memory, perception and hence belief

The aberrant percepts that drive delusion formation often occur during periods of stress and are themselves anxiogenic (Keinan and Keinan, 1994). Furthermore, individuals with a low tolerance for ambiguity are more prone to paranormal beliefs and odd experiences (Houran and Houran, 1998). Some models posit a vicious circle in which fear and aberrant perception are mutually reinforcing and demand explanation, culminating in a delusion which then subtends future aberrant percepts and inappropriate fear (Lange et al., 1998; Pally, 2007). These models are descriptively compelling but are expressed largely at the higher cognitive level. We seek a more fundamental neural and cognitive explanation. Simply put, we argue that affectively charged uncertainty drives delusion formation, through establishment of predictive associations that, whilst maladaptive, represent attempts to render the world more predictable.

#### 2.3.1. Neural mechanisms of affective modulation

The uncertainty engendered by aberrations of experience is affectively charged (Vinogradov et al., 1992). Affective learning is also prediction error driven, involving a circuit incorporating the VTA, amygdala and hippocampus as well as the striatum and prefrontal cortex (Delgado et al., 2008b; Laviolette and Grace, 2006; Milad et al., 2007, 2004; Schiller et al., 2008). Dysfunctions within these nodes could engender fear in the wrong context, leading to maladaptive learning about the danger of adverse consequences. The top-down instantiation of extinction learning is particularly interesting in this respect; the dopaminergic and GABAergic mechanisms that override old fear learning with new extinction learning (Bissiere et al., 2003) may be impaired in schizophrenia (Holt et al., 2008). It is clear that paranoia could be accounted for parsimoniously by appealing to an inappropriate engagement of the brain's fear system and its persistence by an impairment of the brain's mechanisms of extinction.

The amygdala is crucial for fear learning in rodents and humans (Critchley et al., 2002; Morris et al., 2002). However, its role may not be limited to fear; the amygdala is involved in coding, processing and learning about salient stimuli (Balleine and Killcross, 2006; Paton et al., 2006). The link between fear and uncertainty is emphasized by theorists who posit that the amygdala is also engaged during conditions of uncertainty about biologically relevant stimuli that warrant vigilance (Sander et al., 2003; Whalen et al., 1998). For example, fearful faces represent ambiguous stimuli, since they signal the presence but not the source of threat (Whalen et al., 2001). Amygdala responses to appetitive and aversive events are modulated by predictability, being more marked when salient events are uncertain (Belova et al., 2007). In this respect, it is noteworthy that animals with lesions of the central nucleus of the amygdala do not allocate more attention to surprising events (Holland and Gallagher, 1993b).

Cholinergic interneurons in the substantia innominata/nucleus basalis and their projections to posterior parietal cortices are important for the surprise-induced enhancement of attention (Chiba et al., 1995; Bucci et al., 1998; Han et al., 1999). In humans, cues that predict aversive events engage both striatum (Delgado et al., 2008a) and amygdala (Schiller et al., 2008) but only the striatum codes aversive prediction error (Schiller et al., 2008), suggesting that the amygdala is involved in representing the salience of events learned as a consequence of prediction error signals transmitted from other regions. Aberrant prediction error responses in the midbrain or striatum could therefore encourage inappropriate assignment of significance to stimuli, thoughts and percepts (Kapur, 2003) which are then allocated attention in the amygdala (Laviolette and Grace, 2006) through changes in fronto-parietal spatial representations (Mohanty et al., 2009). These environmental contingencies are also subjected to strong consolidation through changes in synaptic strength in the rhinal and entorhinal cortices (Hikosaka et al., 2008a), hence, future encounters with similar cues will engender rapid and powerful predictions of aversive stimulation which would engage avoidance behaviors. Impairments in this system could then contribute to the maintenance of paranoia (Freeman et al., 2007; Moutoussis et al., 2007).

Uncertainty is a powerful and uncomfortable experience. A consequence of such perceived and unsettling lacking of control is that subjects strive to find consistent relationships. They consequently become prone to finding illusory patterns, seeing figures in noise, recognizing correlations between unrelated events, creating superstitious rituals and endorsing conspiracy beliefs (Whitson and Galinsky, 2008). We contend that these healthy coping mechanisms are magnified in individuals with psychosis, culminating in the formation of delusions. These 'filling in' processes may result from top-down influences of orbitofrontal cortex, which receives information from the each modality-specific cortical pathway specifying what a particular sensory object is (Rolls et al., 2008), for example; the inferior temporal cortex where object and face identity are encoded (Rolls, 2007) and the superior temporal sulcus where face expression and gesture are represented (Hasselmo et al., 1989a,b). Furthermore, the orbitofrontal cortex has inputs from the amygdala and the ventral tegmental area (Takahashi et al., 2009) which may drive its ability to learn affective value representations (Padoa-Schioppa and Assad, 2006) which appear to modulate perception in a top-down manner (de Araujo et al., 2005); when affectively charged external labels are applied to percepts, OFC responses bias cingulate and striatal responses in the direction of the label (Grabenhorst et al., 2008). Furthermore, damage to the OFC can result in spontaneous confabulation, a delusion-like disorder in which patients confuse ongoing reality with past experiences (Schnider, 2003). Thus, hyper-engagement of top-down attentional biases may contribute to the aberrant salience underpinning delusional beliefs (Kapur, 2003) as well as to their maintenance (Corlett et al., 2009a,b).

### 2.4. Simple synaptic learning and memory mechanisms of belief govern

#### 2.4.1. Our sense of self, agency and free will

Like beliefs, the self is difficult to define and multifaceted (Mishara, 2007). We will focus on one conception of self, that of an agent that is responsible for actions (Wegner, 2004). In this respect, excessive agency accounts of paranoia (Kaney and Bentall, 1992) may be enriched by a consideration of the phenomenon of

superstitious instrumental conditioning (Skinner, 1948), in which spurious associations are learned between an action and a salient outcome and the action persists despite there being no causal connection between it and the salient outcome. An excessively noisy dopamine system would be fertile grounds for superstitions, which are essentially delusional associations that are reinforced between merely coincident thoughts or actions and environmental events (Shaner, 1999). According to action reselection hypotheses of dopaminergic prediction error signals (Redgrave and Gurney, 2006), inappropriate dopaminergic prediction error signals would confer a spurious sense of agency for events.

Initial lesion studies suggested that hippocampal damage increased superstitious learning in experimental animals (Devenport, 1979). However, more extensive investigations implicated the parietal cortex in superstitious responding, suggesting that collateral damage to this region of cortex may have occurred when the hippocampus was aspirated (Mittleman et al., 1990). Elevated superstitious responding has been demonstrated in chronic ketamine users with delusion-like ideation and perceptual aberrations (Freeman et al., 2009) and patients with schizophrenia who have delusions (Roiser et al., 2009), although the rate of superstitious responding in (presumably non-delusional) control subjects was high in both of these studies.

Lesions of the parietal cortex grossly alter bodily perception and representation, for example, hemi-spatial neglect involves a failure to appreciate half of the body, external world and mental images (Bisiach and Luzzatti, 1978). Perhaps another function of the parietal cortex in instrumental learning involves keeping track of the sense of self as agent in the environment (Farrer et al., 2008). Wegner and others hypothesize that a sense of self agency may be learned through experience; having an intention to act very frequently precedes the action itself and this contiguity binds intentions with actions through associative learning (Glymour, 2004; Hume, 1739/2007; Wegner, 2004). This system can be fooled using subliminal prime events that alter the contiguity between actions and outcomes (Aarts et al., 2005) and furthermore, subjects judge the time between performing an action and producing an outcome as shorter when the action was intentional, a process of action-outcome binding (Moore et al., 2009). Schizophrenic patients with severe positive symptoms show a hyper-binding effect, an exaggerated binding between their actions and the outcomes they produce, consistent with a disturbed agency account of paranoia (Franck et al., 2005; Haggard et al., 2003). This process of learned intentionality has been modeled using Bayesian mechanisms; in essence, the task of inferring causal agency involves conditioning the evidence (whether the outcome occurred?) over the priors (was there an intention to act and would the outcome be consistent with the outcome performed? (Hendricks et al., 2007; Lau et al., 2007). Inappropriate engagement of this inference mechanism could account for excessive and inappropriate agency underpinning, for example, beliefs in telekinesis or telepathy, but what about delusions of passivity or external control?

The parietal cortex has also been implicated in passivity experiences through prediction error; in this case, the mismatch between expected and experienced consequences of movements (Schnell et al., 2008). Producing movements over which we feel a sense of agency also involves predictive learning and prediction error (Blakemore et al., 2002). Again, a Bayesian mechanism may underlie motor control; an internal predictive model of motor commands which is used to predict the sensory consequences of movements and compare them with the actual sensory feedback during movement execution (Wolpert et al., 1995; Wolpert and Miall, 1996). The cerebellum appears to store internal world models and compute discrepancies between predicted and experienced sensory consequences of actions (Blakemore et al., 2001). Event related functional MRI studies of the period before a movement show that activations changes in the cerebellum and PFC occur several seconds before movement onset and the degree of cerebellar activation correlates with that in prefrontal and inferior parietal cortices (Allen et al., 2005).

Internal 'forward' models use an efference copy of motor commands (Von Holst, 1954) to make a prediction about the sensory consequences of an action (Blakemore, 2003). This comparison can be used to cancel sensory effects of performing the action, compared with identical movements that are externally produced (Blakemore et al., 1999; Weiskrantz et al., 1971). An impairment in such a predictive system would result in a failure to attenuate the sensory consequences of self-produced actions, making them appear indistinguishable from externally generated sensations and engendering the inference that one's own movements were externally caused (Blakemore et al., 2002; Frith et al., 2000a). This theory provides an elegant explanation for why we cannot tickle ourselves, since we cancel the predicted sensory consequences of the action (Blakemore et al., 2000b). However, patients experiencing passivity phenomena and hallucinations, in whom sensory cancellation is presumed to be impaired, rate self generated stimulation as ticklish (Blakemore et al., 2000a). Impaired cancellation of efference copies has likewise been implicated in the pathophysiology of hallucinations; here internally generated speech is misperceived as externally generated due to this impairment in the cancellation of forward model predictions (Ford and Mathalon, 2005; Ford et al., 2007).

There are some rare patients who call the proposed model of passivity into question; subjects who have suffered haptic deafferentiation and therefore do not perceive sensory feedback from the actions they perform (Fourneret et al., 2002). Since a haptically deafferented subject does not suffer from delusions of passivity; some have argued that aberrant percepts of one's own action are not sufficient to explain passivity delusions; invoking a further belief evaluation dysfunction that is necessary for the delusional inference to occur (Coltheart, 2010). To clarify the prediction error based explanation of these phenomena; patients with passivity experiences do not use forward model predictions to cancel the predicted consequences of their movements so they experience the sensory consequences of their actions and therefore attribute the source of their actions externally. Haptically deafferented subjects should therefore be protected from passivity experiences; since such experiences do not depend on absence of feedback but on inappropriately large or unexpected feedback. It is this persistence and unexpected nature of aberrant prediction error that engages delusion formation.

Parietal cortex receives inputs from the cerebellar internal model (Ito, 1993), possibly combining them with a multisensory salience map of the external world and the motor plans necessary to approach or avoid salient features (Mohanty et al., 2009). Activity in the parietal operculum is also attenuated during self initiated movements compared with passive movements (Weiller et al., 1996) and during self-produced compared with external stimulation. Patients with lesion to the right hemisphere in white matter underlying the parietal operculum delusion that their limb belonged to their niece (Bottini et al., 2002).

Even healthy individuals can be tricked into accepting that a false hand belongs to their own body (Botvinick and Cohen, 1998). If subjects perceive the false hand being stimulated at the same time as they feel their own (occluded) hand receiving the same stimulation, they begin to feel that the false hand belongs to them, incorporating it into their body schema such that, when asked to estimate where their own hand is positioned, they point to a location closer to the false hand (Makin et al., 2008). Patients with schizophrenia are more susceptible to this illusion (Peled et al., 2003). It appears that the processes of multisensory integration

involved in judging ownership of a body part involve synaptic learning via associative Hebbian mechanisms, representing the confluence of seeing a hand stimulated and feeling a hand stimulated (Keysers et al., 2004). Furthermore, top-down attentional biases seem to influence the illusion (Tsakiris and Haggard, 2005). These biases again emerge through associative learning and are subject to the same formal rules, a surprising mismatch between the expected confluence of sensation and vision weakens the illusion. Likewise the illusion does not occur for a stick: people perceive rubber hand illusions more readily than rubber object illusions (Press et al., 2008). Physiological noise in the multisensory integration process that confers bodily ownership may engender mutated prior expectations about the body which bias subsequent perception, resulting in somatoparophrenias, delusions of body representation and agency (Vallar and Ronchi, 2009).

*2.4.2. Social learning and therefore our beliefs about others*

Social neuroscientists also appreciate the power of prediction error and predictive coding (Behrens et al., 2009; Kilner et al., 2007a,b; Lee, 2008a). Reinforcement learning circuits are engaged when human subjects make social value judgments and a further network of brain regions is engaged when subjects make judgments about the intentions of others – including the superior temporal sulcus/temporoparietal junction (STS/TPJ) (Behrens et al., 2009). These data build upon previous suggestions that associative principles like prediction error govern various social attribution processes (Miller, 1959). For example social attributions made about worker productivity are susceptible to associative learning phenomena like Kamin blocking (Cramer et al., 2002).

fMRI studies of prediction error driven reinforcement learning usually require participants to learn which of two stimuli to choose in order to win the most points (Pessiglione et al., 2006). In an extension to the standard paradigm, Behrens et al. gave subjects an additional source of information, the suggestion of a confederate who may or may not know the appropriate choice to make. Hence the subjects learned simultaneously whether to choose the blue or the green card and also whether they could trust the advice of the confederate. They were able to distinguish brain regions coding a mismatch between expected and experienced reward from brain regions coding a mismatch between expected and experienced truth. Intriguingly, these analyses revealed that adjacent but distinct regions of the anterior cingulate cortex coded reward and truth prediction error. The STS/TPJ also appeared to reflect social prediction errors about the truth of the confederate's advice (Behrens et al., 2008).

The analysis of social learning in terms of prediction error has recently bridged theories of both reinforcement learning and predictive coding. Building upon the empirical Bayes model of brain function, this approach combines the forward model of intentional motor control (Blakemore, 2003; Blakemore et al., 2001; Wolpert et al., 1995; Wolpert and Miall, 1996) with the observations of social prediction errors in STS (Behrens et al., 2008; Hampton et al., 2008) to explain the function of the brain's mirror neurons system through its direct link between action and observation (Kilner et al., 2007a,b). Here, the most likely cause of an observed action can be inferred by minimizing the prediction error across all levels in the cortical hierarchy that are engaged by that observation.

Observing, imagining, or in any way representing an action excites the motor program used to execute the same action (Jeannerod, 1994). Mirror Neurons discharge not only during action execution but also during action observation; they were identified in non-human primates, using neural recording, in area F5 and the inferior parietal lobule (Fogassi and Luppino, 2005; Gallese et al., 1996; Rizzolatti et al., 1996). Functional magnetic resonance imaging data have been used to infer the presence of mirror neurons in the human inferior parietal lobule (Chong et al., 2008) and inferior frontal gyrus (Kilner et al., 2009). However some have failed to find evidence of mirror neuron-like activations (Lingnau et al., 2009). Indeed, the spatial resolution of fMRI is such that it may be inappropriate to ascribe the response in a particular region to a specific population of cells. Furthermore, some have questioned the reified status of mirror neurons; that is, instead of being indivisible, they may simply reflect conditioning of an association between a motor program for an action and a visual representation of that action; learned by experience across the life course (Heyes, 2010). The present theory does not depend on the exact origin of mirror representations and, given that the regions in which mirror neurons have been identified with direct recording in non-human primates largely overlap with those regions that responded to action observation and execution in human subjects, we proceed by discussing the potential role of mirror neurons in human social cognition (Gallese et al., 2004).

Implicit in the description of mirror neurons is the idea that information is passed by forward connections from low-level representations of the movement kinematics to high-level representations of the intentions subtending the action. Observation of an action activates the STS, which in turn drives the inferior parietal lobule which drives the inferior frontal gyrus. Formally this is a recognition model that operates by the inversion of a generative model (Kilner et al., 2007a,b). A generative model will produce an estimate of the visual consequences of an executed action given the causes or goals of that action. By inverting the model it is possible to infer the cause or goal of an action given the visual input (Kilner et al., 2007a,b).

Again, bottom-up or top-down biases in this inference process would lead to gross misrepresentations of other's intentions. Those biases may arise due to aberrant prediction error signals, forging maladaptive social expectations manifest phenomenologically as intense feelings of social uncertainty and ultimately paranoia. More recently, it has emerged that beliefs about somebody's mental experience can influence how we perceive their physical attributes (Teufel et al., 2009). While the full connotations of this have yet to be explored, it seems that we may perceive someone's behavior depending on what we think that they are thinking.

## 3. The fixity of delusions

By inappropriately updating subject's priors, delusions are applied to all subsequent experiences (Conrad, 1958b; Mishara and Corlett, 2009). Why might this be? Indeed, if we are arguing that delusions form under the influence of inappropriate, uncertain and imprecise prediction error, why do delusions become so tenacious? Here we turn to a process that has received increasing empirical attention in recent years; memory reconsolidation (Misanin et al., 1968; Nader et al., 2000). We conceive of beliefs and delusions as a kind of memory (Eichenbaum and Bodkin, 2000), that is, a means through which past experiences and processing organize responses to current inputs. Memories serve a more dynamic function than simple storage; they can be recalled, returned to a labile state (Misanin et al., 1968; Nader et al., 2000), updated with new information (Estes, 1997) and strengthened (Lee, 2008b); a set of reconsolidation processes that appear to be engaged when unexpected events occur (Eisenhardt and Menzel, 2007). This updating process involves a streamlining or schematization of the representation (Stickgold and Walker, 2007). We have previously argued that, once delusions are formed, future prediction errors engage a reactivation, reconsolidation and strengthening of the delusion; rendering it impervious to

contradictory evidence; each time a delusion is deployed, it is reinforced further, conferring resistance to contradiction (Corlett et al., 2009b), rather like the formation of an instrumental habit with overtraining (Adams and Dickinson, 1981; Lee, 2008b; Stickgold and Walker, 2007). That is, when subsequent prediction errors occur, they are explicable in terms of the delusion and they serve to reinforce it, hence the paradoxical observation that challenging subjects' delusions can actually strengthen their conviction (Milton et al., 1978; Simpson and Done, 2002). Neurobiologically, this reconsolidation based strengthening would shift control of behavior toward the dorsal striatal habit system (see Fig. 3) and would manifest as immutable prior expectancies in Bayesian cortical hierarchies (Corlett et al., 2009a,b; Mishara and Corlett, 2009). Delusions may be maintained despite being fallacious through disruptions in fronto-striatal synaptic metaplasticity, a form of 'plasticity of plasticity' (Abraham and Bear, 1996) that allows old associations to be overridden by new learning. Metaplasticity can be restored with N-acetyl-cysteine (Moussawi et al., 2009), a drug which increases the availability of glutamate in extrasynaptic spaces by stimulating the cysteine-glutamate antiporter (Baker et al., 2008). This analysis of delusions, in terms of a shift away from computationally expensive prefrontal processing toward striatal habit (Daw et al., 2005; Mishara and Corlett, 2009) may also explain the waxing and waning of delusional conviction and the paradoxical double book-keeping; patients endorse particular delusions but do not act as if they truly believe them (Bleuler, 1908; Sass, 2004); such situations would transpire if the goal-directed system occasionally won the competition for control of behavior, a state of the system that can be engendered by enhancing plasticity in prefrontal brain regions (Hitchcott et al., 2007; Moussawi et al., 2009).

Here we draw upon advances in the cognitive neuroscience of addiction to make our case about delusions. Like delusions, aberrant prediction error accounts have been outlined for the generation of addictive behaviors (Lapish et al., 2006; Redish, 2004) as well as their maintenance as habits despite maladaptive consequences (Takahashi et al., 2008). We posit that the inappropriate prediction error that occurs in endogenous psychosis is internally generated [rather than a plastic response to drug consumption, although see Corlett et al. (2009a) for a review of drug induced psychoses] and that they track merely coincident environmental stimuli rather than cues that predict access to drug and drug induced hedonic states. However, maladaptive prediction error responses in addiction and psychosis may be indicative of a fronto-striatal system that is sensitized toward aberrant learning and may therefore explain the strong co-morbidity between drug abuse and psychosis (Kalayasiri et al., 2006; van Nimwegen et al., 2005).

Reactivating a delusion (perhaps having a patient engage with and ruminate upon it) may drive its representation into a labile state; providing a novel therapeutic window in which to intervene and destabilize the delusion. This approach has been taken previously with some success (Rubin, 1976), however, future well-controlled investigations are essential.

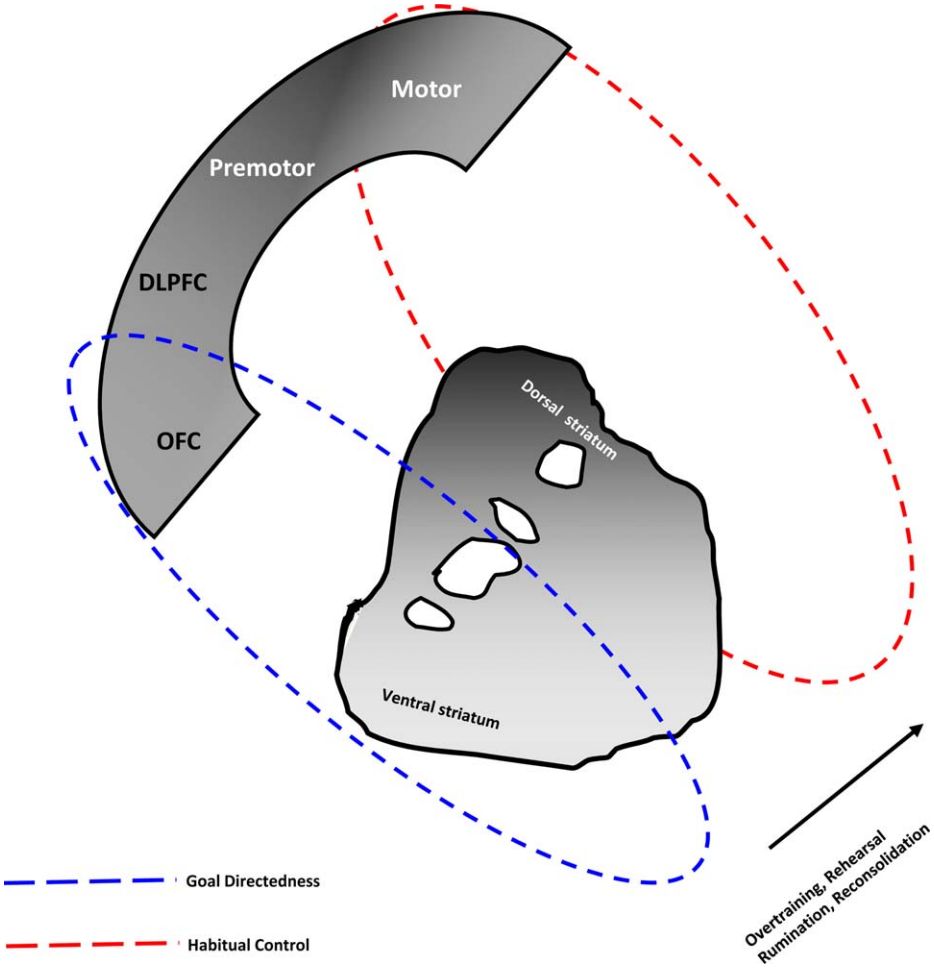


**Fig. 3.** Neural circuitry of goal directedness (knowledge) and habit (belief). With repetition, rumination and reconsolidation, the control of behavior shifts from flexible goal-directed ventral corticostriatal control toward control by the inflexible dorsal striatum and motor cortex.

## 4. One or two factors?

There are competing accounts of delusions in cognitive neuropsychiatry (Coltheart et al., 2007; Freeman et al., 2002; Garety, 1991; Garety and Freeman, 1999; Gerrans, 2002; Kinderman and Bentall, 1997; McKay et al., 2007). Some argue that perceptual aberrations are all that is required for a delusion to form (Gerrans, 2002; Maher, 1974), others that delusions result from top-down reasoning impairments (Freeman et al., 2002; Garety, 1991; Garety and Freeman, 1999), others still posit some combination of both factors, a two-factor approach in which perceptual and reasoning abnormalities combine (Coltheart et al., 2007; McKay et al., 2007). The latter derive from observations that neurological patients with delusions often have two sites of damage; a lesion in a perceptual region (such as the fusiform face area) and an additional lesion in 'belief evaluation' regions, possibly in the right frontal cortex (Ramachandran and Blakeslee, 1998). The first damage engenders odd percepts and the second generates bizarre explanations.

Prediction error driven Bayesian models of delusions (Corlett et al., 2009a; Fletcher and Frith, 2009) subsume both factors into a single deficit in Bayesian inference; noise in predictive learning mechanisms engender inappropriate percepts which update future priors, leading to the formation and maintenance of delusions. Prediction error signals have been registered in right dorsolateral prefrontal cortex during causal learning (Corlett et al., 2004; Fletcher et al., 2001; Turner et al., 2004), psychotogenic drug administration and endogenous psychosis are associated with inappropriate responding in this region, the magnitude of which was predictive of delusion severity (Corlett et al., 2006, 2007b).

Two-factor theorists have recently equated the inappropriate prediction error signals that we reported in dorsolateral prefrontal cortex with their aberrant belief evaluation process or factor 2 (Coltheart, 2010). However, a single deficit in Bayesian inference is able to explain more of what we know about the interactions between perception and belief-based expectation, the neurobiology of the delusions that occur in schizophrenia and the maintenance of delusions in the face of contradictory evidence. That is, unlike two-factor theory, our model allows for dysfunctional prediction error to be calculated in PFC and imposed upon the rest of the brain or, alternatively for surprising perceptual inputs to arrive at PFC engaging surprise and demanding explanation. Both of these possibilities (bottom-up and top-down) are aberrations of a single factor; Bayesian inference.

We recognize the strong neurological evidence that perceptual aberration and delusional ideation are dissociable (Coltheart, 2010). However, we emphasize the potential consequences of prefrontal cortical damage alone (their factor 2) as well peripheral perceptual dysfunction (their factor 1); there are patients who suffer from delusion-like spontaneous confabulations following damage to ventromedial and lateral prefrontal cortex (Schnider, 2003; Turner et al., 2004) and at least one patient in whom peripheral sensations are perturbed (following damage to the brachial plexus) who has somatic delusions in the absence of any apparent structural damage and by extension any deficit in factor 2 (Ghoreishi, 2010).

In short, the present model suggests that inappropriate mismatches between expectancy and experience engender prediction error where there ought to be none, driving new and aberrant learning directly and through the allocation of attention toward irrelevant but potentially explanatory cues (Corlett et al., 2007a). This learning normally provides the basis for a variety of vital perceptual and cognitive functions that govern our interactions with the environment and other agents so when it malfunctions, gross misrepresentations of reality, delusions and perceptual aberrations, result.

## 5. A neurodevelopmental dimension?

Developmental studies suggest that children who go on to develop schizophrenia and therefore likely delusions (although not all patients with schizophrenia have delusions) have subtle neurological 'soft-signs' indicative of aberrant sensorimotor integration (Mohr et al., 1996). In healthy individuals, there are relationships between motor developmental milestones, structural integrity of the frontal cortex, striatum and cerebellum and executive cognitive function, associations which are not present in patients with schizophrenia (Ridler et al., 2006) suggesting impaired bootstrapping of cortical pathways into systems that can predict and respond to their inputs and thus, an impairment of adaptive interaction with the environment and other agents; individuals with impaired sensorimotor integration throughout development would learn impoverished or maladaptive prior expectancies about the world (Hemsley and Garety, 1986b).

Different homeobox genes are responsible for controlling the development and patterning of the frontal cortex (Tabares-Seisdedos and Rubenstein, 2009), midbrain dopamine neurons (Maxwell and Li, 2005), the striatum (Long et al., 2009), the amygdala (Tole et al., 2005) and cerebellum (Sillitoe et al., 2008). Some of these genes and their expression products have been associated with psychotic symptoms, for example; DLX1 expression is decreased in the thalamus of individuals with psychosis compared with those without a history of psychosis and matched healthy controls (Kromkamp et al., 2003). Likewise, the homeogene Engrailed 2 which controls cerebellar development is associated with schizophrenia (Gourion et al., 2004). Knocking out FGF17, a gene that controls the patterning and organization of frontal cortical development, leads to profound deficits in social interaction in mice, perhaps indicative of a relationship to paranoia (Scearce-Levie et al., 2008). Indeed, a human genetic association study revealed a link between the chromosome region where FGF17 is found (8p13) and delusional beliefs (Chiu et al., 2002). We acknowledge that we are speculating here and we appreciate the dangers of anthropomorphizing social behaviors in rodents; future work should address the validity of FGF-knockout as a model of paranoia by exploring other prediction error related processes in these animals; do they have a deficit in conditioned avoidance learning, for example? We believe that the different themes of delusional beliefs entertained by different subjects may have their origins in subtle developmental dysfunctions in the circuits we have outlined, biasing prediction error driven deficits in glutamatergic and dopaminergic processing toward a particular set of experiences and a specific explanatory belief. Normal variation in these same genetic loci may underpin individual differences in perceptual aberration as well as the themes and severity of delusion-like ideation in the healthy population.

## 6. Explaining delusion content

We now attempt to account for different kinds of delusion within this framework. While the scope of this section is by no means exhaustive, we believe that the range of delusions potentially accounted for within the framework is compelling (see Fig. 4).

### 6.1. Paranoia and delusions of reference

Referential delusions involve the belief that objects, events and agents in the environment are communicating specific and personal messages (Conrad, 1958b) ranging from the inanimate to animate, from newspapers, to television newsreaders (Startup

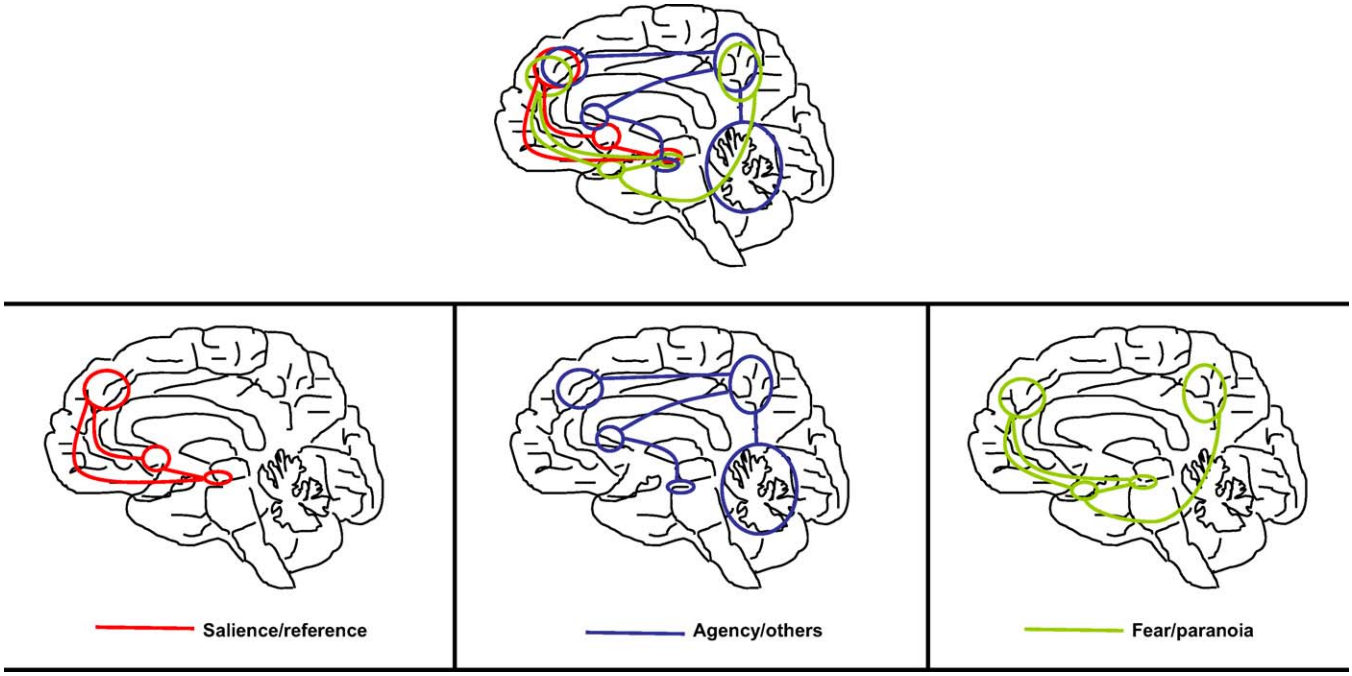


**Fig. 4.** Putative delusion circuits. *Salience/reference*: A circuit incorporating the midbrain dopaminergic nuclei, the associative striatum and frontal cortex. Aberrant prediction errors in midbrain update expectancies in the frontal cortex leading to aberrantly salient percepts. *Agency/others*: The midbrain, PFC, parietal cortex and cerebellum as well as the bimodal cells of the putamen. This circuit describes forward model predictions used to discern whether sensory stimulation was internally of externally generated. A breakdown in this predictive mechanism would manifest as hallucinatory tactile percepts and inferences of external control of intentional action. *Fear/paranoia*: A circuit incorporating the midbrain, amygdala, frontal and parietal cortices. Here, neutral or irrelevant stimuli, thoughts and percepts come to engender fear and anxiety. A dysfunction in fronto-parietal circuitry engenders inappropriate social predictions and maladaptive inferences about the intentions of others. *Interaction between circuits*: These circuits interact and likely mutually reinforce one another.

and Startup, 2005) and even the fictional television detective Columbo (Chadwick, 2007). The psychotomimetic drug ketamine transiently induces delusions of reference in healthy volunteers (Krystal et al., 1994; Oye et al., 1992; Pomarol-Clotet et al., 2006). It blocks NMDA receptors (thus impairing the specification of top-down prior expectancies) while at the same time enhancing bottom-up AMPA signaling (Jackson et al., 2004) and engages acetylcholine release (Sarter et al., 2005). Low, sub psychotic doses of the drug engage the right fronto-striatal prediction error signaling system in response to unsurprising and highly predictable events and the extent to which it does this showed a strong trend toward predicting the severity of heightened perception and delusional ideation (Corlett et al., 2006). We argue that delusions of reference form due to the attentional effects of aberrant prediction error (Pearce and Hall, 1980) mediated via surprise-induced acetylcholine release from the nucleus basalis of meynert (Bao et al., 2001; Holland and Gallagher, 1993a, 1999, 2006; Lee et al., 2005); subjects find their attention drawn toward irrelevant stimuli and events in the environment and impute personal meaning upon them, an experience that demands explanation, culminating in delusions of reference.

Paranoid ideation is associated with excessive fear or anxiety (Moutoussis et al., 2007). In the context of the present analysis, paranoia would result when aberrant prediction error in fronto-striatal learning systems engages the amygdala, engendering a feeling of fear and a state of hypervigilance. Relevant to this contention, delusions of reference and paranoid/persecutory ideation tend to co-occur in patients with delusions (Startup and Startup, 2005), that is, hypervigilance and the perception of meaning in irrelevant and innocuous events may engender paranoia, since uncertainty and unpredictability are inherently fear inducing (Vinogradov et al., 1992; Whitson and Galinsky, 2008). However, paranoid thoughts are commonly about other people (Melo et al., 2006) and, as such, they may involve a prediction error driven dysfunction in the social learning mechanisms that we use to infer the intentions of others localized to fronto-striatal and parietal circuits and the superior temporal sulcus/temporoparietal junction (Behrens et al., 2009, 2008).

Physiological noise in this system, as a result of NMDA receptor hypofunction (which would disturb the specification of priors), AMPA receptor hyperfunction (which would signal prediction error where there should be none) and elevated dopamine levels within the mirror neuron circuit would impair the sufferer's ability to use what they have learned about their own actions and intentions to make inferences about other agents (Kilner et al., 2007a,b). Those disturbances in predicting and learning the consequences of our own actions may also have their origins in a disruption in the extended fronto-striatal-parietal reinforcement learning circuit; as we outlined, the midbrain dopamine neurons implicated in the pathophysiology of schizophrenia (Murray et al., 2008) may report an error in prediction, which is then processed in a circuit incorporating the frontal cortex, striatum and hippocampus (Redgrave and Gurney, 2006) as well as parietal cortex (Mittleman et al., 1990). This signal may be used to discern whether the organisms' actions caused a particular outcome, or whether the outcome happened due to external events (Redgrave and Gurney, 2006), while hypofunctioning of this circuit would lead to a decreased sense of agency for one's own actions, perhaps most relevant to delusions of external control (see below), we posit that the hyper-engagement of this circuit could engender paranoia. That is, paranoid persecutory ideation is associated with superstitious biases in action-outcome learning (Kaney and Bentall, 1992). When playing rigged computer games paranoid individuals claimed to control both negative and positive outcomes when in fact there was no programmed contingency between their actions and the salient events.

Haggard et al. have reported an excessive binding between intentional actions and the outcomes they produce in patients with schizophrenia, however they did not relate this effect to delusions or paranoia in particular (Haggard et al., 2003). This maladaptive perception of contiguity between actions and outcomes would seem to offer an explanation for bizarre beliefs about telekinesis or enhanced predictive abilities, however in the context of the mirror neuron system account for computing and inferring the intentions of others (Kilner et al., 2007a,b), an individual who had learned spurious associations between their actions and salient environmental outcomes would also be expected to use those associations to infer the intentions of other agents. They would then ascribe supernatural abilities or excessively powerful status to individuals whom they encountered. In the context of prediction error induced amygdala responses, this inference would be affectively charged and result in a fear and distrust that is incommensurate with the current situation. This model makes some progress toward integrating neurobiology with psychodynamic explanations of paranoia (Kinderman and Bentall, 1996, 1997) in which attentional biases toward perceived threats are driven by mismatches between current self perceptions and how the patient believes they ought to be, focusing or projecting a threatening attributional bias onto external agents (Colby, 1977), patients with paranoia may attempt to avoid feelings of low self-esteem by attributing the cause of adverse experiences externally (Bentall et al., 2001). We believe that impairments in the brain's mirror neuron system and its ability to infer the intentions of others based on inverting its own predictions (Kilner et al., 2007a,b) may underpin these processes of inappropriate external projection of threat.

### 6.2. Delusions of motor passivity

"My fingers pick up the pen, but I don't control them. What they do is nothing to do with me" (Mellor, 1970). It appears that these odd beliefs result from an impairment in the cancellation of predicted sensory consequences of motor behaviors (Blakemore, 2003; Blakemore et al., 2003; Frith, 2005), involving a defect in the specification of motor predictions by the cerebellum which subsequently inappropriately engages parietal and frontal cortices (Frith, 2005; Schnell et al., 2008; Spence et al., 1997). An action produced without apparent forward model expectation is therefore ascribed to an external agent. A similar aberrant efference copy account has been made with respect to auditory hallucinations (Ford and Mathalon, 2005; Ford et al., 2007; Heinks-Maldonado et al., 2007).

However, some have criticized this model for failing to explain how patients with these delusions (and underlying brain pathology) can engage in any behavior at all. Having a sense of one's self as the source of our intentional actions may be essential for goal-directed instrumental learning (Glymour, 2004). This sense may be learned by the contiguous association between perceiving an intention to act, executing the motor program and encountering the consequences (Hume, 1739/2007; Wegner, 2004). Prediction errors due to physiological noise from dysregulated midbrain dopamine neurons projecting to prefrontal cortex could render those predictive associations unreliable (Corlett et al., 2007a). However, there is a less computationally intensive brain system that can control instrumental learning in the dorsolateral striatum. This system is said to mediate stimulus-response habits (Daw et al., 2005; McDonald and Hong, 2004; Reading et al., 1991; Tang et al., 2007; Tricomi et al., 2009; Yin et al., 2004). The information used to guide behavior in this system is insensitive to the current value of the outcome (Daw et al., 2005). Habitual organisms behave reflexively, emitting motor responses to environmental cues irrespective of their consequences (Adams and Dickinson, 1981). The dorsal striatal habit system is believed to govern compulsive drug seeking and taking (Belin et al., 2009). The goal-directed and habit systems are conceived of as competitors for the control of behavior – the system that is least uncertain about the appropriate behavior given the context may win that competition (Daw et al., 2005). Competition between them can be biased towards the habit system by extended behavioral training (Adams and Dickinson, 1981); boosting synaptic dopamine levels in the striatum (Nelson and Killcross, 2006), or modulating AMPA receptor function (Bespalov et al., 2007). Goal directedness can be rescued by restoring dopamine induced plasticity in the prefrontal cortex (Hitchcott et al., 2007). It is possible that the habit system wins the competition in individuals with delusions (see below). Passivity experiences may therefore be explained as instrumental actions controlled by the habit system in the context of a noisy and inaccurate goal-directed system.

### 6.3. Delusions of parasitosis

Individuals with delusional parasitosis are convinced that small animals such as insects or lice are living on or within their skin (Berrios, 1982, 1985). This particular symptom highlights the overlap between delusions and hallucinations, perceptions and beliefs which calls in to question the strict clinical distinction (Corlett et al., 2009a; Fletcher and Frith, 2009; Frith and Dolan, 2000). Striatal lesions (Huber et al., 2008), dopamine agonist medications (Charuvastra and Yaeger, 2006; Mitchell and Vierkant, 1991), cocaine (Mitchell and Vierkant, 1991; Siegel, 1978; Wallis, 1949) and amphetamine (Ellinwood, 1968; Ellinwood et al., 1974) abuse can all engender delusions of parasitosis. Indeed, chronic treatment with dopamine antagonists can induce behaviors indicative of parasitosis in experimental animals (Ellison, 1994). In human stroke patients, delusions of parasitosis often occur following lesions of right temporoparietal cortex, thalamus and putamen (Huber et al., 2008). Putamen strongly influences visuotactile perception (Graziano and Gross, 1993; Ladavas et al., 1998; Romo et al., 1995; Yoo et al., 2003), it contains bimodal cells with visual and tactile receptive fields, which help to encode the location of sensory stimuli mainly near the face. These cells project to parietal (ventral intra-parietal cortex), primary somatosensory and pre-motor cortices (Graziano and Gross, 1993; Ladavas et al., 1998).

We contend that sensations on the skin are a result of the same interaction between top-down and bottom-up mechanisms that we argue are crucial for visual perception. This is supported by the cutaneous rabbit illusion (Geldard and Sherrick, 1972) where simultaneous stimulation of two points on the skin gives rise to the percept of a rabbit 'hopping' between the two points; stimulation at a particular frequency is best explained by movement along a trajectory between the two points. There are Bayesian accounts of the illusion (Goldreich, 2007). Parasitosis may arise either due to bottom-up sensation that is normally ignored – for example a lack of adaptation of skin sensation over time or, alternatively, due to inappropriate top-down expectations – the power of cognition in cutaneous sensation is also underlined by contagious itch sensations experienced when subjects are exposed to conversations about insects on the skin (Heaven and McBrayer, 2000; Mitchell, 1995).

The same learning mechanisms that underpin the rubber hand illusion (Press et al., 2008) might also be involved in parasitosis; a deficit in Bayesian multisensory integration would lead aberrant prediction error, driving attention toward potentially explanatory cues and forging inappropriate visuotactile associations. These associations, between sensation and a particular spatial location, might be represented by bimodal cells in the striatum, forming a new prior, a top-down bias in attention to the skin which would

contribute to the maintenance of the delusion (Berrios, 1982; Corlett, 2009).

### 6.4. Delusions of misidentification

There are two main classes of misidentification delusion; Capgras; in which patients believe that their close family members have been replaced by imposters (Capgras and Reboul-Lachaux, 1923), and Fregoli; in which patients believe that strangers that they encounter are their relatives in disguise (Courbon and Fail, 1927). Additionally, some patients have misidentification of their own home either feeling it is unfamiliar (Feinberg and Keenan, 2005; Fleminger, 1992) or that the hospital in which they find themselves is really their house hundreds of miles away (Schnider, 2001). Two-factor models of these disorders assume a dual deficit, one in perception of affect, the other in belief evaluation (Coltheart et al., 2007). Instead, we argue that phenomenology of the percepts are such that bizarre beliefs are inevitable; surprising experiences demand surprising explanations (Kihlstrom and Hoyt, 1988). In our Bayesian, predictive learning scheme, Capgras results when patients experience an anomalous lack of affective responding when confronted with their relatives (Ellis and Young, 1990), the delusion constitutes a new prior driven by the experience, a means for explaining it away (Young, 2008). It is possible that the initial affective disturbance results from a failure to guide affect perception by prior experience, that is, just like sensory perception, emotions are predicted (Gilbert and Wilson, 2009); we have emotional priors, indeed, it is the prior expectancy of a familiar face combined with an emotional response (learned through experience) which breaks down in Capgras patients (Fleminger, 1992); fostering the misidentification of someone (or something) familiar as unfamiliar (Young, 2008). With the Fregoli delusion, it is a misplaced sense of familiarity (rather like a delusion of reference, specific to people) which guides patients to infer that people they do not know are actually their relatives in disguise.

In a meta-analysis of patients with delusional misidentification (Fregoli and Capgras delusions) about persons or objects, surveying 48 cases following neurological insult, Feinberg et al. found that the overwhelming majority had damage to the right hemisphere, commonly the frontal cortex. This observation is in line with our own work on prediction error during causal learning implicating a region of right dorsolateral prefrontal cortex in prediction error signaling (Corlett et al., 2004; Fletcher et al., 2001; Turner et al., 2004) and implicating it in delusion formation (Corlett et al., 2006, 2007b).

The laterality of damage that induces delusions seems replicable across studies of neurological patients with delusions (Devinsky, 2009). Spitzer and Walter (2003) speculate that this hemispheric bias can be explained by appealing to the different hemispheric modes of information processing (Kosslyn et al., 1992). Whereas the left hemisphere is characterized by smaller receptive fields resulting in focused, conjunctive coding, the right hemisphere is characterized by larger, overlapping receptive fields resulting in a coarse coding (Spitzer and Walter, 2003). In terms of Bayesian brain theory, receptive fields are related to the top-down specification of expected inputs (Rao and Ballard, 1999). Increasing dopamine levels may alter the signal to noise ratio of neurons, that is, it will increase the precision or certainty with which a prediction error is signaled (Friston et al., 2009) such that subjects respond to physiological noise as if it were meaningful signal (Grace, 1991). An increase in dopamine levels would serve to inappropriately increase confidence in noisy signals. It will therefore affect a system which relies on coarse coding, i.e. the right hemisphere, more prominently than a system which relies on conjunctive coding, i.e. the left hemisphere. That is, the right hemisphere is more susceptible to inappropriate optimization of prediction error because its predictions and prediction errors are inherently more noisy than the processing on the left hemisphere. Some speculate that, in response to right hemisphere error signals, the left hemisphere begins to construct explanations resulting in delusions (Devinsky, 2009), however the difficulty identifying and tracking delusions forming (Corlett et al., 2007a) means that this contention has not found empirical support.

When considering delusions of misidentification of neurological origin, it seems puzzling that damage in the same region could be associated with both an increase and a decrease in perceived familiarity. Two-factor theorists would suggest that this is parsimoniously explained by ascribing the right frontal cortex the function of belief evaluation (Coltheart, 2010; Coltheart et al., 2007). However, we found that right frontal prediction error signal during causal learning was also related to ketamine induced perceptual aberrations (Corlett et al., 2006) and, furthermore, a study of individuals with lesions in the right dorsolateral prefrontal cortex suggested that lesion patients attended to and learned about irrelevant stimulus dimensions during a reward learning task (Hornak et al., 2004). It is possible that damage or dysfunction in prefrontal cortex could, paradoxically elevate activation in the remaining neurons since, in healthy individuals they provide a brake on subcortical dopamine nuclei through glutamatergic (Grace, 1991; Laruelle et al., 2003) and GABAergic mechanisms (Carlsson et al., 2001). Consequently, either due to a release from inhibition or an alteration of signal to noise properties, dopamine neurons projecting back from VTA to prefrontal cortex would increase in burst firing (Jackson et al., 2004) inducing rapid and random post-synaptic potentials in remaining functional cortical neurons (Lavin et al., 2005).

### 6.5. Cotard delusion

Perhaps one of the most bizarre delusions is the sufferer believing that they have died (Cotard, 1880), associated with claims that parts of them have "rotted away" or "disappeared" (Gerrans, 2002). It is possible that the same impoverished habitual mechanisms of instrumental action are engaged (see above) and the subject infers that the intentional agent that they were has disappeared, that is, the Cotard delusion may be a special case of passivity. Additionally, Some hypothesize that Capgras patients fail to recognize family members due a disconnection between face recognition units in the fusiform face area and the ascription of emotional meaning in the limbic system, therefore, patients with Cotard delusion may have no connection at all between sensation and affective processing (Gerrans, 2002; Ramachandran and Blakeslee, 1998). In this analysis, Cotard delusion is the converse of paranoia, instead of heightened and inappropriate emotional intensity it is a failure to ascribe emotional significance to any event (Gerrans, 2002; Ramachandran and Blakeslee, 1998). Such a lack of emotional engagement with experiences would be surprising, engendering prediction error and sculpting the erroneous conclusion that the patient had died. Again, affective prediction fails, but instead of the rather specific effect in Capgras, it is a generalized failure in predicting the affective qualities of all sensory inputs. Like Capgras and Fregoli, this may involve a dysfunction in orbitofrontal cortex specifying top-down emotional predictions (Rolls and Grabenhorst, 2008). The delusion has been reported in a case study following right temporoparietal and bilateral frontal damage (Young et al., 1992), it also occurs in schizophrenia (Coltheart et al., 2007). This delusion involves both a deficit in affective forecasting (by the orbitofrontal cortex and amygdala), as well as (potentially) a deficit in motor forecasting (and thus sensory cancellation), with a diminished sense of self and emotional disengagement, the patient concludes that he/she is dead.

## 7. Why that odd belief? Individual differences in delusion susceptibility

While some psychotic patients get paranoid, others experience passivity, others still have multiple bizarre delusions. We posit a single factor, prediction error dysfunction for delusion formation and maintenance (Corlett et al., 2009a, 2007a; Corlett et al., 2009a; Fletcher and Frith, 2009). We have recently applied this single factor account to explain the range of phenomenological effects of pharmacologically distinct psychotomimetic drugs from dopamine agonist amphetamines, to NMDA antagonists, cannabinoids and serotonergic hallucinogens (Corlett et al., 2009a). We believe the same explanation may be possible for the individual differences in susceptibility to different delusional themes observed in patients with schizophrenia.

Schizophrenia is a heritable but heterogeneous mental illness; its genetic inheritance appears to involve multiple genes of small effect (Tabares-Seisdedos and Rubenstein, 2009) or alternatively multiple rare genetic variants each with a large impact (Walsh et al., 2008). However, common to many of the identified risk genes for schizophrenia is a role in associative learning, prediction error signaling and NMDA receptor dependent synaptic plasticity (Hall et al., 2009; Stephan et al., 2006; Walsh et al., 2008). Some of the genes implicated in prediction error driven learning (Frank et al., 2007; Heyser et al., 2000) increase the risk for schizophrenia; the COMT val/met polymorphism may enhance maladaptive feedback between frontal cortex and subcortical dopamine neurons and is associated with risk for schizophrenia, aberrant salience and delusions (Bilder et al., 2004). In addition, PP1R1b, the gene coding for neostriatal signaling nexus DARPP-32 which integrates midbrain dopamine inputs with cortical glutamatergic signaling has been associated with prediction error driven learning (Frank et al., 2007; Heyser et al., 2000) fronto-striatal structure and function as well as risk for schizophrenia (Meyer-Lindenberg et al., 2007). Variation in the function of these genes may explain inter-subject variability in susceptibility to delusions following psychotomimetic drug administration (Corlett et al., 2009a, 2007a; Svenningsson et al., 2003).

However, different delusional themes are characteristic following the administration of different psychotomimetics; paranoia is more intense following cannabis administration (D'Souza et al., 2009) whereas ketamine engenders delusions of reference (Krystal et al., 1994; Oye et al., 1992; Pomarol-Clotet et al., 2006), although the two themes are by no means mutually exclusive (Startup and Startup, 2005). We believe that a second genetic insult may confer susceptibility to particular kinds of delusion in schizophrenia, an insult involving disrupted cortical patterning and how the developing cortex interacts with environmental inputs in forming and maintaining cortical hierarchies (Sur and Rubenstein, 2005). Although this appears to be a two-factor theory, when we consider how Bayesian hierarchies like the brain develop into prediction engines (Friston, 2005b) through interactions between neural circuitry and incoming stimulation (Sur and Rubenstein, 2005), delusions really involve a singular dysfunction in predictive learning (i.e. an interaction between the two deficits which leads to (dys)interactions between poorly specified top-down predictions and noisy feedforward inputs; inducing aberrant and imprecise prediction errors (Corlett et al., 2009a; Fletcher and Frith, 2009). The genes for building cortical hierarchies may also engender prediction error dysfunction irrespective of dopaminergic/glutamatergic 'prediction error' risk gene status and furthermore, the two insults may interact to produce more severe or varied delusions in the same patient.

Are there any empirical data to support of our contention that delusions with different themes are mediated by distinct (but overlapping) neural circuits? Patients with delusions secondary to neurological damage often have lesions in right frontal cortex but, according to two-factor theories, the theme of the belief is conferred by damage to a second structure; for example the fusiform face area in Capgras delusion. Patients suffering from dementia with Lewy bodies experience delusions (Nagahama et al., 2010, 2007) like Capgras (Hirono and Cummings, 1999). Nagahama et al. used factor analysis to classify psychotic symptoms in dementia with Lewy bodies. They found that hallucinations, misidentification experiences and delusions were independent symptom domains (Nagahama et al., 2007). More recently they replicated this factor structure in an independent group of patients and assessed the neural correlates of those factors by regressing factor scores onto resting state neuroimaging data across subjects (Nagahama et al., 2010).

Patients suffering from misidentification had hypo-perfusion in left hippocampus, insula, inferior frontal gyrus and nucleus accumbens compared to patients without those symptoms. Individuals who had visual hallucinations of person or a feeling of presence had hypo-perfusion in bilateral parietal and left ventral occipital gyrus. Patients with persecutory delusions showed significant hyperactivity in right cingulate sulcus, bilateral middle frontal gyri, right inferior frontal gyrus, left medial superior frontal gyrus and left middle frontopolar gyrus. These distinct circuits tend to support our predicted delusion circuits (see Fig. 4); that is, paranoia involves a frontal hyperactivity; delusions that potentially involve hyper salience of own body representations (e.g. hallucinations of people and feeling of presence) involve a parietal dysfunction and reduplications of person and place involve a predictive memory impairment; impaired familiarity processing and fronto-hippocampal as well as fronto-striatal dysfunction.

Lewy bodies appear to accumulate in the space between bands of cortex; occupied by afferent or efferent connections with different cortical sites or with subcortical regions, that is, they have a laminar distribution (Armstrong et al., 2001). Depending on which layer, they preferentially influence the feedforward (prediction error specifying) connections originating in laminae I–III and terminating in granular lamina IV of the adjacent lobe (Armstrong et al., 2001). Alternatively, Lewy bodies may accumulate in the feedback fibers (responsible for specifying prior expectations and attentional modulation) which originate in laminae V and VI (and to some extent III) and terminate in lamina I (De Lacoste and White, 1993). Why the feedforward and feedback pathways of one particular circuit would be more sensitive to Lewy body inclusions than another circuit (conferring a particular delusion content) has yet to be determined, however, the disconnection that they engender within particular circuits is consistent with the putative disconnections invoked to explain the symptoms of schizophrenia (Friston, 2005a; Friston and Frith, 1995).

Finally, we turn to a rare but intriguing phenomenon, Folie a Deux (Lasegue and Falret, 1877), to evaluate our proposal that delusional themes are mediated by inherited biological processes. Folie a Deux (FD) is a psychotic disorder shared between two sufferers; an 'inducer' who initially develops the belief and the 'induced', an apparently otherwise healthy individual who comes to share the delusional belief. All kinds of rare delusional contents can be transmitted, e.g. Cotard, Capgras, Fregoli (Wolff and McKenzie, 1994).

FD commonly occurs in persons who live close together, the delusion perhaps being transmitted through social learning processes. Additionally, if both patients are related, they may share the same genetically driven illness or predisposition. Monozygotic twins can share the same delusional themes (Lazarus, 1986); however, since they often share both genetic and environmental exposure, it is difficult to discern the unique contributions made by genes and environment. Scharfetter

attempted to dissect these contributions by identifying dyads in whom there was no consanguinity (e.g. husband and wife) then evaluating the risk for schizophrenia in each respective family. Incidence in both inducer and induced was very high (6.5–26.2%, compared with 1% population incidence), suggesting that a general predisposition toward delusions was necessary for accepting someone else's aberrant belief (Scharfetter, 1970). Future empirical research should investigate the personality, cognitive and neural functions of related and unrelated FD dyads to ascertain the roles of specific neural circuits in instantiating particular delusional beliefs.

## 8. Testing the hypothesis

Our sketch of the emerging neurobiology of delusional beliefs makes a number of testable predictions which will assess the validity of the venture:

(1) We have argued that delusions arise and are maintained due to aberrations of glutamatergic synaptic plasticity, specifically chronically elevated synaptic glutamate which renders inappropriate salience and learning that engenders a limit on metaplasticity. Given its effectiveness against cocaine induced deficits in metaplasticity (Moussawi et al., 2009), we predict that N-acetylcysteine should be an effective treatment for delusions.
(2) Patients with delusional parasitosis and delusions of passivity should be more susceptible to the rubber hand illusion because of the dysinteraction between the bimodal cells in their striatum and parietal and cerebellar circuits responsible for coding top-down, motor expectancies and cancelling the sensory consequences of actions.
(3) Paranoia should be associated with prediction error dysfunction in mesocorticolimbic regions as well as the mirror neuron circuit, especially the superior temporal sulcus region involved in learning to infer the intentions of other agents (Behrens et al., 2008).
(4) Given a large enough sample and phenomenologically rigorous assessment it should be possible to test our aetiological hypothesis about homeobox genes, development and the specification of priors; we predict that subtle variation in the gene coding for pax6 will alter amygdala development and therefore confer a risk for paranoia (Tole et al., 2005); Engrailed 2 will be associated with an increased likelihood of cerebellar dysfunction (Sillitoe et al., 2008) and as such will confer risk for passivity delusions.
(5) Reconsolidation processes should be enhanced in individuals with intractable delusions – engaging and challenging their belief should increase its severity but treatments that block reconsolidation (such as the alpha adrenergic receptor antagonist, propanalol) should ameliorate delusions (if they have been actively engaged).
(6) Physical interventions that target reactivated representations of delusions should also have therapeutic benefits (Rubin, 1976), for example, it may be possible to disrupt the reconsolidation of delusions with transcranial magnetic stimulation (TMS) (Corlett, 2009). Based on the observed relationship between DLPFC dysfunction and delusional ideation (Corlett et al., 2006, 2007b) as well as the role of DLPFC in controlled memory retrieval and updating (Fletcher and Henson, 2001) we suggest that specifically targeting that region with TMS following memory engagement may prove beneficial.
(7) Individuals with high positive schizotypy or treated with psychotomimetic drugs should demonstrate aberrant prediction error signaling and therefore form learned habits more rapidly than controls.
(8) If delusions are learned habits, then pharmacological interventions that restore goal directedness should be effective therapeutically; for example, antagonizing AMPA receptors (Bespalov et al., 2007), boosting PFC dopamine levels (Hitchcott et al., 2007) and attenuating striatal dopamine (Nelson and Killcross, 2006) should favor plasticity and goal directedness. The dopamine partial agonist Aripiprazole combines both antagonism of elevated striatal dopamine and an elevation of attenuated prefrontal dopamine and may specifically target aberrant prediction error signaling in midbrain dopamine neurons (Hamamura and Harada, 2007). It may be particularly effective against cognitive habits like delusions.

In order to complete a revolution of translation, having been inspired by the role of prediction error in associative learning in infrahuman species to develop our account of delusions, we should use invasive preclinical neuroscientific approaches in combination with associative learning phenomena to model delusion formation and maintenance in experimental animals. There are a number of potential opportunities here; combining acute psychotomimetic pharmacological models (to recapitulate putative neurobiological mechanisms of psychosis) with associative learning tasks that are sensitive to prediction error (to model delusion formation) or habit learning and memory reconsolidation (to model delusion maintenance).

## 9. Conclusion

We have outlined an account of delusional beliefs based on the tenets of animal learning theory and hierarchical Bayesian inference. We apply those tenets not only to explain dysfunctions in Pavlovian predictive learning (Corlett et al., 2006, 2007b) and instrumental conditioning (Freeman et al., 2009; Murray et al., 2008; Roiser et al., 2009; Schlagenhauf et al., 2009), but also to account for the perceptual, affective and social disruptions that attend delusions (Bentall et al., 2001; Maher, 1974; Vinogradov et al., 1992).

In deluded individuals, the ability to use learned information to constrain current experience is impaired resulting aberrations of sensory and affective perception as well as cognition (Gray et al., 1991; Hemsley, 1994). Delusions may arise as an explanation for these odd happenings and they engage new learning (Kapur, 2003; Maher, 1974; McGhie and Chapman, 1961). They bring such relief that they are stamped into memory and become a new explanatory scheme for the sufferer (Jaspers, 1963), that is, delusions are elastic; they encompass new experiences and maintain a certain consistency of the world for the patient. In terms of the Bayesian model we outlined, delusions become the sufferer's new priors and they are used to predict and explain future experiences. We believe that the same prediction error driven learning mechanisms can account for the fixity of delusional beliefs, since, now, when subsequent physiological noise elicits a reactivation of the delusion, it is reinforced and reconsolidated more strongly (Corlett et al., 2009b). These hypotheses are readily testable in individuals suffering endogenous delusions, in healthy subjects exposed to psychotomimetic model psychoses and in preclinical models by focusing on the framework for translational cognitive neuroscience provided by formal associative learning theory, hierarchical Bayesian learning, predictive coding and information theory – the concept that intersects all of these is surprise or prediction error (Friston, 2010) and our model implicates aberrant prediction in the pathophysiology of delusions.

We have applied this model to various different kinds of delusions, examining beliefs that result from neurological damage as well as those that result from ingestion of psychotomimetic compounds and those that occur in schizophrenia, we feel, with

some success. However, Brendan Maher, Emeritus Professor of the Psychology of Personality at Harvard, astutely aligned delusions with scientific theories (Maher, 1988), suggesting that scientists, like individuals with delusions, were extremely resistant to giving up their preferred theories even in the face damningly negative evidence. Like scientists, deluded individuals are confronted by surprising data which they explain away by abductive inference, generating hypotheses that explain away the surprise (Coltheart et al., 2010). Scientists (some of them at least) will engage in deductive inference to test the validity of their conclusions; whilst patients with delusions may not engage in this process (Miller, 1976), showing a bias against disconfirmatory evidence (Woodward et al., 2006). Furthermore, inductive inference, that is, reasoning from the specific to the general, has been invoked to explain the influence of prior experience over current perception (Barlow, 1990). We propose that the inductive process, reasoning beyond the data, may provide a mechanism through which delusions are maintained and pervade future experiences (Jaspers, 1963). Whilst the theory outlined in the present piece is our preferred explanation of delusions, we hope that we engender discussion, debate and investigation. As Maher says of science and psychosis: "*Puzzles demand an explanation; the search for an explanation begins and continues until one has been devised*". We hope that this article might encourage others to join the search.

## References


Aarts, H., Custers, R., Wegner, D.M., 2005. On the inference of personal authorship: enhancing experienced agency by priming effect information. Conscious. Cogn. 14, 439–458.

Abraham, W.C., Bear, M.F., 1996. Metaplasticity: the plasticity of synaptic plasticity. Trends Neurosci. 19, 126–130.

Adams, C.D., Dickinson, A, 1981. Actions and habits: variations in associative representations during instrumental learning. In: Spear, N.E., Miller, R.R. (Eds.), Information Processing in Animals: Memory Mechanisms. Erlbaum, New Jersey.

Allen, G., McColl, R., Barnard, H., Ringe, W.K., Fleckenstein, J., Cullum, C.M., 2005. Magnetic resonance imaging of cerebellar-prefrontal and cerebellar-parietal functional connectivity. Neuroimage 28, 39–48.

Alloy, L.B., Tabachnik, N., 1984. Assessment of covariation by humans and animals: the joint influence of prior expectations and current situational information. Psychol. Rev. 91, 112–149.

Angelucci, A., Levitt, J.B., Lund, J.S., 2002a. Anatomical origins of the classical receptive field and modulatory surround field of single neurons in macaque visual cortical area V1. Prog. Brain Res. 136, 373–388.

Angelucci, A., Levitt, J.B., Walton, E.J., Hupe, J.M., Bullier, J., Lund, J.S., 2002b. Circuits for local and global signal integration in primary visual cortex. J. Neurosci. 22, 8633–8646.

Angrilli, A., Spironelli, C., Elbert, T., Crow, T.J., Marano, G., Stegagno, L., 2009. Schizophrenia as failure of left hemispheric dominance for the phonological component of language. PLoS One 4, e4507.

Ardid, S., Wang, X.J., Gomez-Cabrero, D., Compte, A., 2010. Reconciling coherent oscillation with modulation of irregular spiking activity in selective attention: gamma-range synchronization between sensory and executive cortical areas. J. Neurosci. 30, 2856–2870.

Armstrong, R.A., Cairns, N.J., Lantos, P.L., 2001. What does the study of the spatial patterns of pathological lesions tell us about the pathogenesis of neurodegenerative disorders? Neuropathology 21, 1–12.

Atkinson, J.W., McClelland, D.C., 1948. The projective expression of needs. II. The effect of different intensities of the hunger drive on Thematic Apperception. J. Exp. Psychol. 38, 643–658.

Baker, D.A., Madayag, A., Kristiansen, L.V., Meador-Woodruff, J.H., Haroutunian, V., Raju, I., 2008. Contribution of cystine-glutamate antiporters to the psychotomimetic effects of phencyclidine. Neuropsychopharmacology 33, 1760–1772.

Balleine, B.W., Killcross, S., 2006. Parallel incentive processing: an integrated view of amygdala function. Trends Neurosci. 29, 272–279.

Bao, S., Chan, V.T., Merzenich, M.M., 2001. Cortical remodelling induced by activity of ventral tegmental dopamine neurons. Nature 412, 79–83.

Barlow, H., 1990. Conditions for versatile learning, Helmholtz's unconscious inference, and the task of perception. Vision Res. 30, 1561–1571.

Bayes, T., 1763. An essay towards solving a problem in the doctrine of chances. Philos. Trans. R. Soc. Lond. 53, 370–418.

Beck, J., Beck, J., Forstmeier, W., 2007. Superstition and belief as inevitable by-products of an adaptive learning strategy. Hum. Nat. 18, 35.

Behrens, T.E., Hunt, L.T., Rushworth, M.F., 2009. The computation of social behavior. Science 324 (5931), 1160–1164.

Behrens, T.E., Hunt, L.T., Woolrich, M.W., Rushworth, M.F., 2008. Associative learning of social value. Nature 456, 245–249.

Belin, D., Jonkman, S., Dickinson, A., Robbins, T.W., Everitt, B.J., 2009. Parallel and interactive learning processes within the basal ganglia: relevance for the understanding of addiction. Behav. Brain Res. 199, 89–102.

Belova, M.A., Paton, J.J., Morrison, S.E., Salzman, C.D., 2007. Expectation modulates neural responses to pleasant and aversive stimuli in primate amygdala. Neuron 55, 970–984.

Bentall, R.P., Corcoran, R., Howard, R., Blackwood, N., Kinderman, P., 2001. Persecutory delusions: a review and theoretical integration. Clin. Psychol. Rev. 21, 1143–1192.

Berke, J.D., Hetrick, V., Breck, J., Greene, R.W., 2008. Transient 23–30 Hz oscillations in mouse hippocampus during exploration of novel environments. Hippocampus 18, 519–529.

Berridge, K.C., 2007. The debate over dopamine's role in reward: the case for incentive salience. Psychopharmacology (Berl.) 191, 391–431.

Berrios, G.E., 1982. Tactile hallucinations: conceptual and historical aspects. J. Neurol. Neurosurg. Psychiatry 45, 285–293.

Berrios, G.E., 1985. Delusional parasitosis and physical disease. Compr. Psychiatry 26, 395–403.

Berrios, G.E., 1991. Delusions as "wrong beliefs": a conceptual history. Br. J. Psychiatry Suppl. 6–13.

Bi, G., Poo, M., 2001. Synaptic modification by correlated activity: Hebb's postulate revisited. Annu. Rev. Neurosci. 24, 139–166.

Bespalov, A.Y., Harich, S., Jongen-Relo, A.L., van Gaalen, M.M., Gross, G., 2007. AMPA receptor antagonists reverse effects of extended habit training on signaled food approach responding in rats. Psychopharmacology (Berl.) 195, 11–18.

Bilder, R.M., Volavka, J., Lachman, H.M., Grace, A.A., 2004. The catechol-O-methyltransferase polymorphism: relations to the tonic-phasic dopamine hypothesis and neuropsychiatric phenotypes. Neuropsychopharmacology 29, 1943–1961.

Binzegger, T., Douglas, R.J., Martin, K.A., 2004. A quantitative map of the circuit of cat primary visual cortex. J. Neurosci. 24, 8441–8453.

Bisiach, E., Luzzatti, C., 1978. Unilateral neglect of representational space. Cortex 14, 129–133.

Bissiere, S., Humeau, Y., Luthi, A., 2003. Dopamine gates LTP induction in lateral amygdala by suppressing feedforward inhibition. Nat. Neurosci. 6, 587–592.

Blakemore, S.J., 2003. Deluding the motor system. Conscious. Cogn. 12, 647–655.

Blakemore, S.J., Frith, C.D., Wolpert, D.M., 1999. Spatio-temporal prediction modulates the perception of self-produced stimuli. J. Cogn. Neurosci. 11, 551–559.

Blakemore, S.J., Frith, C.D., Wolpert, D.M., 2001. The cerebellum is involved in predicting the sensory consequences of action. Neuroreport 12, 1879–1884.

Blakemore, S.J., Oakley, D.A., Frith, C.D., 2003. Delusions of alien control in the normal brain. Neuropsychologia 41, 1058–1067.

Blakemore, S.J., Smith, J., Steel, R., Johnstone, C.E., Frith, C.D., 2000a. The perception of self-produced sensory stimuli in patients with auditory hallucinations and passivity experiences: evidence for a breakdown in self-monitoring. Psychol. Med. 30, 1131–1139.

Blakemore, S.J., Wolpert, D., Frith, C., 2000b. Why can't you tickle yourself? Neuroreport 11, R11–R16.

Blakemore, S.J., Wolpert, D.M., Frith, C.D., 2002. Abnormalities in the awareness of action. Trends Cogn. Sci. 6, 237–242.

Bleuler, E., 1908. Die Prognose der Dementia praecox (Schizophreniegruppe). Allgemeine Zeitschrift für Psychiatrie und psychischgerichtliche Medizin 65, 436–464.

Bottini, G., Bisiach, E., Sterzi, R., Vallar, G., 2002. Feeling touches in someone else's hand. Neuroreport 13, 249–252.

Botvinick, M., Cohen, J., 1998. Rubber hands 'feel' touch that eyes see. Nature 391, 756.

Braver, T.S., Cohen, J.D., 1999. Dopamine, cognitive control, and schizophrenia: the gating model. Prog. Brain Res. 121, 327–349.

Bromberg-Martin, E.S., Hikosaka, O., 2009. Midbrain dopamine neurons signal preference for advance information about upcoming rewards. Neuron 63, 119–126.

Bruner, J., Bruner, J., Postman, L., 1949. Perception, cognition, and behavior. J. Pers. 18, 14.

Bucci, D.J., Holland, P.C., et al., 1998. Removal of cholinergic input to rat posterior parietal cortex disrupts incremental processing of conditioned stimuli. J. Neurosci. 18 (19), 8038–8046.

Buhl, E.H, Tamas, G., Fisahn, A., 1998. Cholinergic activation and tonic excitation induce persistent gamma oscillations in mouse somatosensory cortex in vitro. J. Physiol. 513 (Part 1), 117–126.

Buschman, T.J., Miller, E.K., 2007. Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. Science 315, 1860–1862.

Butts, C., 1998. A Bayesian model of panic in belief. Computat. Math. Org. Theor. 4 (4), 373–404.

Buzsaki, G., 2007. The structure of consciousness. Nature 446, 267.

Capgras, J., Reboul-Lachaux, J., 1923. L'illusion des "soises" dans un delire systematise. Bull. Soc. Clin. Med. Mentale 11, 6–16.

Carlsson, A., Waters, N., Holm-Waters, S., Tedroff, J., Nilsson, M., Carlsson, M.L., 2001. Interactions between monoamines, glutamate, and GABA in schizophrenia: new evidence. Annu. Rev. Pharmacol. Toxicol. 41, 237–260.

Carlsson, M., Carlsson, A., 1990. Schizophrenia: a subcortical neurotransmitter imbalance syndrome? Schizophr. Bull. 16, 425–432.

Cepeda, C., Levine, M.S., 1998. Dopamine and N-methyl-D-aspartate receptor interactions in the neostriatum. Dev. Neurosci. 20, 1–18.

Chadwick, P.K., 2007. Peer-professional first-person account: schizophrenia from the inside—phenomenology and the integration of causes and meanings. Schizophr. Bull. 33, 166–173.

Charuvastra, A., Yaeger, D., 2006. Tactile hallucinations associated with therapeutic doses of bupropion in 2 patients. J. Clin. Psychiatry 67, 1820–1821.
Chiba, A.A., Bucci, D.J., et al., 1995. Basal forebrain cholinergic lesions disrupt increments but not decrements in conditioned stimulus processing. J. Neurosci. 15 (11), 7315–7322.
Chiu, Y.F., McGrath, J.A., Thornquist, M.H., Wolyniec, P.S., Nestadt, G., Swartz, K.L., Lasseter, V.K., Liang, K.Y., Pulver, A.E., 2002. Genetic heterogeneity in schizophrenia. II: Conditional analyses of affected schizophrenia sibling pairs provide evidence for an interaction between markers on chromosome 8p and 14q. Mol. Psychiatry 7, 658–664.
Chong, T.T., Cunnington, R., Williams, M.A., Kanwisher, N., Mattingley, J.B., 2008. fMRI adaptation reveals mirror neurons in human inferior parietal cortex. Curr. Biol. 18, 1576–1580.
Clifford, W.K., 1877/1999. The ethics of belief. In: Madigan, T. (Ed.), The ethics of belief and other essays. Prometheus, Amherst, MA, pp. 70–96.
Colby, K.M., 1977. Appraisal of four psychological theories of paranoid phenomena. J. Abnorm. Psychol. 86, 54–59.
Coltheart, M., 2010. The neuropsychology of delusions. Ann. N. Y. Acad. Sci. 1191, 16–26.
Coltheart, M., Langdon, R., McKay, R., 2007. Schizophrenia and monothematic delusions. Schizophr. Bull. 33, 642–647.
Coltheart, M., Menzies, P., Sutton, J., 2010. Abductive inference and delusional belief. Cogn. Neuropsychiatry 15, 261–287.
Conrad, K., 1958a. Die Beginnende Schizophrenie. G. Thieme, Stuttgart.
Conrad, K., 1958b. Die BeginnendeSchizophrenie. G. Thieme, Stuttgart.
Corlett, P.R., Aitken, M.R., Dickinson, A., Shanks, D.R., Honey, G.D., Honey, R.A., Robbins, T.W., Bullmore, E.T., Fletcher, P.C., 2004. Prediction error during retrospective revaluation of causal associations in humans: fMRI evidence in favor of an associative model of learning. Neuron 44, 877–888.
Corlett, P.R., Frith, C.D., Fletcher, P.C., 2009a. From drugs to deprivation: a Bayesian framework for understanding models of psychosis. Psychopharmacology (Berl.) 206 (4), 515–530.
Corlett, P.R., Krystal, J.H., Taylor, J.R., Fletcher, P.C., 2009b. Why do delusions persist? Front. Hum. Neurosci. 3, 12.
Corlett, P.R., Simons, J.S., Pigott, J.S., Gardner, J.M., Murray, G.K., Krystal, J.H., Fletcher, P.C., 2009c. Illusions and delusions: relating experimentally-induced false memories to anomalous experiences and ideas. Front. Behav. Neurosci. 3, 53.
Corlett, P.R., Honey, G.D., Aitken, M.R., Dickinson, A., Shanks, D.R., Absalom, A.R., Lee, M., Pomarol-Clotet, E., Murray, G.K., McKenna, P.J., Robbins, T.W., Bullmore, E.T., Fletcher, P.C., 2006. Frontal responses during learning predict vulnerability to the psychotogenic effects of ketamine: linking cognition, brain activity, and psychosis. Arch. Gen. Psychiatry 63, 611–621.
Corlett, P.R., Honey, G.D., Fletcher, P.C., 2007a. From prediction error to psychosis: ketamine as a pharmacological model of delusions. J. Psychopharmacol. 21, 238–252.
Corlett, P.R., Murray, G.K., Honey, G.D., Aitken, M.R., Shanks, D.R., Robbins, T.W., Bullmore, E.T., Dickinson, A., Fletcher, P.C., 2007b. Disrupted prediction-error signal in psychosis: evidence for an associative account of delusions. Brain 130, 2387–2400.
Cotard, J., 1880. Du délire hypocondriaque dans une forme grave de la melancolie anxieuse. Memoire lu à la Société médicopsychologique dans la Séance du 28 Juin 1880. Ann. Medico-Psychol. Med. 168–174.
Courbon, P., Fail, G., 1927. Syndrome d' ''illusion de Frégoli'' et schizophrénie. Bull. Soc. Clin. Méd. Mentale 15, 121–124.
Courville, A.C., Daw, N.D., Touretzky, D.S., 2006. Bayesian theories of conditioning in a changing world. Trends Cogn. Sci. 10, 294–300.
Cramer, R.E., Weiss, R.F., William, R., Reid, S., Nieri, L., Manning-Ryan, B., 2002. Human agency and associative learning: Pavlovian principles govern social process in causal relationship detection. Q. J. Exp. Psychol. B 55, 241–266.
Critchley, H.D., Mathias, C.J., Dolan, R.J., 2002. Fear conditioning in humans: the influence of awareness and autonomic arousal on functional neuroanatomy. Neuron 33, 653–663.
D'Souza, D.C., Sewell, R.A., Ranganathan, M., 2009. Cannabis and psychosis/schizophrenia: human studies. Eur. Arch. Psychiatry Clin. Neurosci. 259, 413–431.
Dakin, S., Carlin, P., Hemsley, D., 2005. Weak suppression of visual context in chronic schizophrenia. Curr. Biol. 15, R822–R824.
Damasio, A., 2000. Thinking about belief: concluding remarks. In: Schacter, D.L., Scarry, E. (Eds.), Memory, Brain and Belief. Harvard University Press, Cambridge Massachussetts/London, England.
Daw, N.D., Niv, Y., Dayan, P., 2005. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nat. Neurosci. 8, 1704–1711.
de Araujo, I.E., Rolls, E.T., et al., 2005. Cognitive modulation of olfactory processing. Neuron 46 (4), 671–679.
De Clerambault, G.G., 1942. Les Psychoses Passionelles. Oeuvere Psychiatrique. Presses Universities, de France, Paris, pp. 331, 337–339, 357, 408.
De Lacoste, M.C., White 3rd, C.L., 1993. The role of cortical connectivity in Alzheimer's disease pathogenesis: a review and model system. Neurobiol. Aging 14, 1–16.
Delespaul, P., van Os, J., 2003. Jaspers was right after all—delusions are distinct from normal beliefs. Against. Br. J. Psychiatry 183, 286.
Delgado, M.R., Li, J., Schiller, D., Phelps, E.A., 2008a. The role of the striatum in aversive learning and aversive prediction errors. Philos. Trans. R. Soc. Lond. B Biol. Sci. 363, 3787–3800.
Delgado, M.R., Nearing, K.I., Ledoux, J.E., Phelps, E.A., 2008b. Neural circuitry underlying the regulation of conditioned fear and its relation to extinction. Neuron 59, 829–838.
den Ouden, H.E., Daunizeau, J., Roiser, J., Friston, K.J., Stephan, K.E., 2010. Striatal prediction error modulates cortical coupling. J. Neurosci. 30, 3210–3219.
den Ouden, H.E., Friston, K.J., Daw, N.D., McIntosh, A.R., Stephan, K.E., 2009. A dual role for prediction error in associative learning. Cereb. Cortex 19, 1175–1185.
Dennett, D., 1995. Do animals have beliefs. In: Roitblat, H.L., Meyer, J.A. (Eds.), Comparative Approaches to Cognitive Science. MIT Press, Cambrudge Massachusetts; London, England.
Devenport, L.D., 1979. Superstitious bar pressing in hippocampal and septal rats. Science 205, 721–723.
Devinsky, O., 2009. Delusional misidentifications and duplications: right brain lesions, left brain delusions. Neurology 72, 80–87.
Dickinson, A., 2001. The 28th Bartlett Memorial Lecture. Causal learning: an associative analysis. Q. J. Exp. Psychol. B 54, 3–25.
Dima, D., Roiser, J.P., Dietrich, D.E., Bonnemann, C., Lanfermann, H., Emrich, H.M., Dillo, W., 2009. Understanding why patients with schizophrenia do not perceive the hollow-mask illusion using dynamic causal modelling. Neuroimage 46, 1180–1186.
Dommett, E., Coizet, V., Blaha, C.D., Martindale, J., Lefebvre, V., Walton, N., Mayhew, J.E., Overton, P.G., Redgrave, P., 2005. How visual stimuli activate dopaminergic neurons at short latency. Science 307, 1476–1479.
Eichenbaum, H., Bodkin, J.A., 2000. Belief and knowledge as distinct forms of memory. In: Schacter, D.L., Scarry, E. (Eds.), Memory Brain and Belief. Harvard University Press, Cambridge, Massachusetts/London England.
Eisenhardt, D., Menzel, R., 2007. Extinction learning, reconsolidation and the internal reinforcement hypothesis. Neurobiol. Learn. Mem. 87, 167–173.
Ellinwood Jr., E.H., 1968. Amphetamine psychosis. II. Theoretical implications. Int. J. Neuropsychiatry 4, 45–54.
Ellinwood Jr., E.H., Sudilovsky, A., Nelson, L.M., 1974. Behavior and EEG analysis of chronic amphetamine effect. Biol. Psychiatry 8, 169–176.
Ellis, H.D., Young, A.W., 1990. Accounting for delusional misidentifications. Br. J. Psychiatry 157, 239–248.
Ellison, D.G., 1941. Hallucinations produced by sensory conditioning. J. Exp. Psychol. 28, 1–20.
Ellison, G., 1994. Stimulant-induced psychosis, the dopamine theory of schizophrenia, and the habenula. Brain Res. Brain Res. Rev. 19, 223–239.
Emrich, H.M., 1989. A three-component-system hypothesis of psychosis. Impairment of binocular depth inversion as an indicator of a functional dysequilibrium. Br. J. Psychiatry Suppl. (5) 37–39.
Estes, W.K., 1997. Processes of memory loss, recovery, and distortion. Psychol. Rev. 104, 148–169.
Farrer, C., Frey, S.H., Van Horn, J.D., Tunik, E., Turk, D., Inati, S., Grafton, S.T., 2008. The angular gyrus computes action awareness representations. Cereb. Cortex 18, 254–261.
Feinberg, T.E., 2000. The nested heirarchy of consciousness: a neurobiological soultion to the problem of mental unity. Neurocase 6, 75–81.
Feinberg, T.E., Keenan, J.P., 2005. Where in the brain is the self? Conscious. Cogn. 14, 661–678.
Fiorillo, C.D., 2008. Towards a general theory of neural computation based on prediction by single neurons. PLoS One 3, e3298.
Fiorillo, C.D., Tobler, P.N., Schultz, W., 2003. Discrete coding of reward probability and uncertainty by dopamine neurons. Science 299, 1898–1902.
Fiser, J., 2009. Perceptual learning and representational learning in humans and animals. Learn. Behav. 37, 141–153.
Fiser, J., Berkes, P., Orban, G., Lengyel, M., 2010. Statistically optimal perception and learning: from behavior to neural representations. Trends Cogn. Sci. 14, 119–130.
Fleminger, S., 1992. Seeing is believing: the role of 'preconscious' perceptual processing in delusional misidentification. Br. J. Psychiatry 160, 293–303.
Fletcher, P.C., Anderson, J.M., Shanks, D.R., Honey, R., Carpenter, T.A., Donovan, T., Papadakis, N., Bullmore, E.T., 2001. Responses of human frontal cortex to surprising events are predicted by formal associative learning theory. Nat. Neurosci. 4, 1043–1048.
Fletcher, P.C., Frith, C.D., 2009. Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia. Nat. Rev. Neurosci. 10, 48–58.
Fletcher, P.C., Henson, R.N., 2001. Frontal lobes and human memory: insights from functional neuroimaging. Brain 124, 849–881.
Flint, A.J., 1991. Delusions in dementia: a review. J. Neuropsychiatry Clin. Neurosci. 3, 121–130.
Fogassi, L., Luppino, G., 2005. Motor functions of the parietal lobe. Curr. Opin. Neurobiol. 15, 626–631.
Ford, J.M., Mathalon, D.H., 2005. Corollary discharge dysfunction in schizophrenia: can it explain auditory hallucinations? Int. J. Psychophysiol. 58, 179–189.
Ford, J.M., Roach, B.J., Faustman, W.O., Mathalon, D.H., 2007. Synch before you speak: auditory hallucinations in schizophrenia. Am. J. Psychiatry 164, 458–466.
Fourneret, P., Paillard, J., Lamarre, Y., Cole, J., Jeannerod, M., 2002. Lack of conscious recognition of one's own actions in a haptically deafferented patient. Neuroreport 13, 541–547.
Franck, N., Posada, A., Pichon, S., Haggard, P., 2005. Altered subjective time of events in schizophrenia. J. Nerv. Ment. Dis. 193, 350–353.

Frank, M.J., Moustafa, A.A., Haughey, H.M., Curran, T., Hutchison, K.E., 2007. Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. Proc. Natl. Acad. Sci. U.S.A. 104, 16311–16316.
Freeman, D., Garety, P.A., Kuipers, E., Fowler, D., Bebbington, P.E., 2002. A cognitive model of persecutory delusions. Br. J. Clin. Psychol. 41, 331–347.
Freeman, D., Garety, P.A., Kuipers, E., Fowler, D., Bebbington, P.E., Dunn, G., 2007. Acting on persecutory delusions: the importance of safety seeking. Behav. Res. Ther. 45, 89–99.
Freeman, T.P., Morgan, C.J., Klaassen, E., Das, R.K., Stefanovic, A., Brandner, B., Curran, H.V., 2009. Superstitious conditioning as a model of delusion formation following chronic but not acute ketamine in humans. Psychopharmacology (Berl.) 206 (4), 563–573.
Friston, K., 2005a. Disconnection and cognitive dysmetria in schizophrenia. Am. J. Psychiatry 162, 429–432.
Friston, K., 2005b. A theory of cortical responses. Philos. Trans. R. Soc. Lond. B Biol. Sci. 360, 815–836.
Friston, K., 2009. The free-energy principle: a rough guide to the brain? Trends Cogn. Sci. 13, 293–301.
Friston, K., 2010. The free-energy principle: a unified brain theory? Nat. Rev. Neurosci. 11, 127–138.
Friston, K.J., 2005c. Hallucinations and perceptual inherence. Behav. Brain Sci. 28, 764–766.
Friston, K.J., Daunizeau, J., Kiebel, S.J., 2009. Reinforcement learning or active inference? PLoS One 4, e6421.
Friston, K.J., Frith, C.D., 1995. Schizophrenia: a disconnection syndrome? Clin. Neurosci. 3, 89–97.
Frith, C., 2005. The neural basis of hallucinations and delusions. C. R. Biol. 328, 169–175.
Frith, C.D., Blakemore, S., Wolpert, D.M., 2000a. Explaining the symptoms of schizophrenia: abnormalities in the awareness of action. Brain Res. Brain Res. Rev. 31, 357–363.
Frith, C.D., Blakemore, S.J., Wolpert, D.M., 2000b. Abnormalities in the awareness and control of action. Philos. Trans. R. Soc. Lond. B Biol. Sci. 355, 1771–1788.
Frith, C.D., Dolan, R.J., 2000. The role of memory in the delusions associated with schizophrenia. In: Schacher, D., Scarry, E. (Eds.), Memory, Brain and Belief. Harvard University Press.
Fuster, J.M., 2001. The prefrontal cortex—an update: time is of the essence. Neuron 30, 319–333.
Gallese, V., Fadiga, L., Fogassi, L., Rizzolatti, G., 1996. Action recognition in the premotor cortex. Brain 119 (Part 2), 593–609.
Gallese, V., Keysers, C., Rizzolatti, G., 2004. A unifying view of the basis of social cognition. Trends Cogn. Sci. 8, 396–403.
Garety, P., 1991. Reasoning and delusions. Br. J. Psychiatry Suppl. 14–18.
Garety, P.A., Freeman, D., 1999. Cognitive approaches to delusions: a critical review of theories and evidence. Br. J. Clin. Psychol. 38 (Part 2), 113–154.
Geldard, F.A., Sherrick, C.E., 1972. The cutaneous "rabbit": a perceptual illusion. Science 178, 178–179.
Gerrans, P., 2002. A one-stage explanation of the Cotard delusion. Philos. Psychiatry Psychol. 9, 47–53.
Geyer, M.A., Vollenweider, F.X., 2008. Serotonin research: contributions to understanding psychoses. Trends Pharmacol. Sci. 29, 445–453.
Ghoreishi, A., 2010. A somatic type delusional disorder secondary to peripheral neuropathy: a case report. Psychiatria Danubina 20, 85–87.
Gilbert, D.T., Wilson, T.D., 2009. Why the brain talks to itself: sources of error in emotional prediction. Philos. Trans. R. Soc. Lond. B Biol. Sci. 364, 1335–1341.
Glymour, C., 2004. We believe in freedom of the will so we can learn [Comment on Wegner Precis of the Illusion of Conscious Will]. Behav. Brain Sci. 27, 661–662.
Goldreich, D., 2007. A Bayesian perceptual model replicates the cutaneous rabbit and other tactile spatiotemporal illusions. PLoS One 2, e333.
Goldstone, R., 1995. Effects of categorization on color perception. Psychol. Sci. 6, 298–304.
Goto, Y., O'Donnell, P., 2001. Network synchrony in the nucleus accumbens in vivo. J. Neurosci. 21, 4498–4504.
Gourion, D., Leroy, S., Bourdel, M.C., Goldberger, C., Poirier, M.F., Olie, J.P., Krebs, M.O., 2004. Cerebellum development and schizophrenia: an association study of the human homeogene Engrailed 2. Psychiatry Res. 126, 93–98.
Grabenhorst, F., Rolls, E.T., Bilderbeck, A., 2008. How cognition modulates affective responses to taste and flavor: top-down influences on the orbitofrontal and pregenual cingulate cortices. Cereb. Cortex 18, 1549–1559.
Grace, A.A., 1991. Phasic versus tonic dopamine release and the modulation of dopamine system responsivity: a hypothesis for the etiology of schizophrenia. Neuroscience 41, 1–24.
Grace, A.A., Floresco, S.B., Goto, Y., Lodge, D.J., 2007. Regulation of firing of dopaminergic neurons and control of goal-directed behaviors. Trends Neurosci. 30, 220–227.
Gray, J.A., 2004. On biology, phenomenology, and pharmacology in schizophrenia. Am. J. Psychiatry 161, 377 author reply 377–378.
Gray, J.A., Feldon, J., Rawlins, J.N.P., Hemsley, D., Smith, A.D., 1991. The neuropsychology of schizophrenia. Behav. Brain Sci. 14, 1–84.
Graziano, M.S., Gross, C.G., 1993. A bimodal map of space: somatosensory receptive fields in the macaque putamen with corresponding visual receptive fields. Exp. Brain Res. 97, 96–109.
Grossberg, S., 1982. Processing of expected and unexpected events during conditioning and attention: a psychophysiological theory. Psychol. Rev. 89, 529–572.
Grossberg, S., 2000. How hallucinations may arise from brain mechanisms of learning, attention, and volition. J. Int. Neuropsychol. Soc. 6, 583–592.
Grossberg, S., 2009. Cortical and subcortical predictive dynamics and learning during perception, cognition, emotion and action. Philos. Trans. R. Soc. Lond. B Biol. Sci. 364, 1223–1234.
Haber, S.N., 2003. The primate basal ganglia: parallel and integrative networks. J. Chem. Neuroanat. 26, 317–330.
Haber, S.N., Kim, K.S., Mailly, P., Calzavara, R., 2006. Reward-related cortical inputs define a large striatal region in primates that interface with associative cortical connections, providing a substrate for incentive-based learning. J. Neurosci. 26, 8368–8376.
Haggard, P., Martin, F., Taylor-Clarke, M., Jeannerod, M., Franck, N., 2003. Awareness of action in schizophrenia. Neuroreport 14, 1081–1085.
Haider, B., Duque, A., Hasenstaub, A.R., McCormick, D.A., 2006. Neocortical network activity in vivo is generated through a dynamic balance of excitation and inhibition. J. Neurosci. 26, 4535–4545.
Hall, J., Romaniuk, L., McIntosh, A.M., Steele, J.D., Johnstone, E.C., Lawrie, S.M., 2009. Associative learning and the genetics of schizophrenia. Trends Neurosci. 32, 359–365.
Halligan, P.W., David, A.S., 2001. Cognitive neuropsychiatry: towards a scientific psychopathology. Nat. Rev. Neurosci. 2, 209–215.
Hamamura, T., Harada, T., 2007. Unique pharmacological profile of aripiprazole as the phasic component buster. Psychopharmacology (Berl.) 191, 741–743.
Hampton, A.N., Bossaerts, P., O'Doherty, J.P., 2008. Neural correlates of mentalizing-related computations during strategic interactions in humans. Proc. Natl. Acad. Sci. U.S.A. 105, 6741–6746.
Han, J.S., Holland, P.C., et al., 1999. Disconnection of the amygdala central nucleus and substantia innominata/nucleus basalis disrupts increments in conditioned stimulus processing in rats. Behav. Neurosci. 113 (1), 143–151.
Hasselmo, M.E., Rolls, E.T., Baylis, G.C., 1989a. The role of expression and identity in the face-selective responses of neurons in the temporal visual cortex of the monkey. Behav. Brain Res. 32, 203–218.
Hasselmo, M.E., Rolls, E.T., Baylis, G.C., Nalwa, V., 1989b. Object-centered encoding by face-selective neurons in the cortex in the superior temporal sulcus of the monkey. Exp. Brain Res. 75, 417–429.
Heaven, L., McBrayer, D., 2000. External motivators of self-touching behavior. Percept. Mot. Skills 90, 338–342.
Hebb, D.O., 1949a. The Organization of Behavior. John Wiley.
Hebb, D.O., 1949b. The Organization of Behaviour: a Neuropsychological Theory. Wiley, New York.
Heinks-Maldonado, T.H., Mathalon, D.H., Houde, J.F., Gray, M., Faustman, W.O., Ford, J.M., 2007. Relationship of imprecise corollary discharge in schizophrenia to auditory hallucinations. Arch. Gen. Psychiatry 64, 286–296.
Helmholtz, H.von., 1878/1971. The facts of perception. In: Kahl, R. (Ed.), Selected Writings of Herman von Helmholtz. Weslyan University Press.
Hemsley, D.R., 1994. Perceptual and cognitive abnormalities as the basis for schizophrenic symptoms. In: David, A.S., Cutting, J.C. (Eds.), The Neuropsychology of Schizophrenia. Laurence Erlbaum Associates, Hove, UK, pp. 97–118.
Hemsley, D.R., Garety, P.A., 1986a. The formation and maintenance of delusions: a Bayesian analysis. Br. J. Psychiatry 149, 51–56.
Hemsley, D.R., Garety, P.A., 1986b. The formation of maintenance of delusions: a Bayesian analysis. Br. J. Psychiatry 149, 51–56.
Hendricks, K.V., Wiggers, P., Jonker, C.M., Haselager, W.F., 2007. Towards a computational model of the self-attribution of agency. In: Oliver, P., Kray, C. (Eds.), The Artificial Intelligence and Simulation of Behaviour Annual Convention. pp. 350–356.
Herrero, J.L., Roberts, M.J., Delicato, L.S., Gieselmann, M.A., Dayan, P., Thiele, A., 2008. Acetylcholine contributes through muscarinic receptors to attentional modulation in V1. Nature 454, 1110–1114.
Heyes, C., 2010. Mesmerising mirror neurons. Neuroimage 51, 789–791.
Heyser, C.J., Fienberg, A.A., Greengard, P., Gold, L.H., 2000. DARPP-32 knockout mice exhibit impaired reversal learning in a discriminated operant task. Brain Res. 867, 122–130.
Hikosaka, O., Bromberg-Martin, E., Hong, S., Matsumoto, M., 2008a. New insights on the subcortical representation of reward. Curr. Opin. Neurobiol. 18, 203–208.
Hikosaka, O., Sesack, S.R., Lecourtier, L., Shepard, P.D., 2008b. Habenula: crossroad between the basal ganglia and the limbic system. J. Neurosci. 28, 11825–11829.
Hirono, N., Cummings, J.L., 1999. Neuropsychiatric aspects of dementia with Lewy bodies. Curr. Psychiatry Rep. 1, 85–92.
Hitchcott, P.K., Quinn, J.J., Taylor, J.R., 2007. Bidirectional modulation of goal-directed actions by prefrontal cortical dopamine. Cereb. Cortex 17, 2820–2827.
Hoffman, R.E., Dobscha, S.K., 1989. Cortical pruning and the development of schizophrenia: a computer model. Schizophr. Bull. 15, 477–490.
Holland, P.C., Gallagher, M., 1993a. Amygdala central nucleus lesions disrupt increments, but not decrements, in conditioned stimulus processing. Behav. Neurosci. 107, 246–253.
Holland, P.C., Gallagher, M., 1993b. Effects of amygdala central nucleus lesions on blocking and unblocking. Behav. Neurosci. 107, 235–245.
Holland, P.C., Gallagher, M., 1999. Amygdala circuitry in attentional and representational processes. Trends Cogn. Sci. 3, 65–73.
Holland, P.C., Gallagher, M., 2006. Different roles for amygdala central nucleus and substantia innominata in the surprise-induced enhancement of learning. J. Neurosci. 26, 3791–3797.
Holt, D.J., Lebron-Milad, K., Milad, M.R., Rauch, S.L., Pitman, R.K., Orr, S.P., Cassidy, B.S., Walsh, J.P., Goff, D.C., 2008. Extinction memory is impaired in schizophrenia. Biol. Psychiatry 65 (6), 455–463.
Hornak, J., O'Doherty, J., Bramham, J., Rolls, E.T., Morris, R.G., Bullock, P.R., Polkey, C.E., 2004. Reward-related reversal learning after surgical excisions in orbito-

frontal or dorsolateral prefrontal cortex in humans. J. Cogn. Neurosci. 16, 463–478.
Houran, J., Houran, J., 1998. Preliminary study of tolerance of ambiguity of individuals reporting paranormal experiences. Psychol. Rep. 82, 183.
Howes, O.D., Montgomery, A.J., Asselin, M.C., Murray, R.M., Valli, I., Tabraham, P., Bramon-Bosch, E., Valmaggia, L., Johns, L., Broome, M., McGuire, P.K., Grasby, P.M., 2009. Elevated striatal dopamine function linked to prodromal signs of schizophrenia. Arch. Gen. Psychiatry 66, 13–20.
Huber, M., Karner, M., Kirchler, E., Lepping, P., Freudenmann, R.W., 2008. Striatal lesions in delusional parasitosis revealed by magnetic resonance imaging. Prog. Neuropsychopharmacol. Biol. Psychiatry 32, 1967–1971.
Hudson, L.A., Rollins, Y.D., Anderson, C.A., Johnston-Brooks, C., Tyler, K.L., Filley, C.M., 2008. Reduplicative paramnesia in Morvan's syndrome. J. Neurol. Sci. 267, 154–157.
Hume, D., 1739/2007. A Treatise of Human Nature. Oxford University Press, Oxford.
Ito, M., 1993. Movement and thought: identical control mechanisms by the cerebellum. Trends Neurosci. 16, 448–450 discussion 453–444.
Jackson, M.E., Homayoun, H., Moghaddam, B., 2004. NMDA receptor hypofunction produces concomitant firing rate potentiation and burst activity reduction in the prefrontal cortex. Proc. Natl. Acad. Sci. U.S.A. 101, 8467–8472.
Jaspers, K., 1963. General Psychopathology. Manchester University Press.
Jeannerod, M., 1994. The hand and the object: the role of posterior parietal cortex in forming motor representations. Can. J. Physiol. Pharmacol. 72, 535–541.
Joliot, M., Ribary, U., Llinas, R., 1994. Human oscillatory brain activity near 40 Hz coexists with cognitive temporal binding. Proc. Natl. Acad. Sci. U.S.A. 91, 11748–11751.
Jones, H., 2004. Defining delusion. Br. J. Psychiatry 185, 354–355.
Kalayasiri, R., Sughondhabirom, A., Gueorguieva, R., Coric, V., Lynch, W.J., Morgan, P.T., Cubells, J.F., Malison, R.T., 2006. Self-reported paranoia during laboratory "binge" cocaine self-administration in humans. Pharmacol. Biochem. Behav. 83, 249–256.
Kandel, E.R., 1998. A new intellectual framework for psychiatry. Am. J. Psychiatry 155, 457–469.
Kaney, S., Bentall, R.P., 1992. Persecutory delusions and the self-serving bias. Evidence from a contingency judgment task. J. Nerv. Ment. Dis. 180, 773–780.
Kapur, S., 2003. Psychosis as a state of aberrant salience: a framework linking biology, phenomenology, and pharmacology in schizophrenia. Am. J. Psychiatry 160, 13–23.
Karson, C.N., 1980. A new look at delusions of grandeur. Compr. Psychiatry 21, 62–69.
Keinan, G., Keinan, G., 1994. Effects of stress and tolerance of ambiguity on magical thinking. J. Pers. Soc. Psychol. 67, 48.
Keysers, C., Wicker, B., Gazzola, V., Anton, J.L., Fogassi, L., Gallese, V., 2004. A touching sight: SII/PV activation during the observation and experience of touch. Neuron 42, 335–346.
Kihlstrom, J.F., Hoyt, I.P., 1988. Hypnosis and the psychology of delusions. In: Oltmanns, T.F., Maher, B.A. (Eds.), Delusional Beliefs. Wiley, New York.
Kilner, J.M., Friston, K.J., Frith, C.D., 2007a. The mirror-neuron system: a Bayesian perspective. Neuroreport 18, 619–623.
Kilner, J.M., Friston, K.J., Frith, C.D., 2007b. Predictive coding: an account of the mirror neuron system. Cogn. Process. 8, 159–166.
Kilner, J.M., Neal, A., Weiskopf, N., Friston, K.J., Frith, C.D., 2009. Evidence of mirror neurons in human inferior frontal gyrus. J. Neurosci. 29, 10153–10159.
Kinderman, P., Bentall, R.P., 1996. Self-discrepancies and persecutory delusions: evidence for a model of paranoid ideation. J. Abnorm. Psychol. 105, 106–113.
Kinderman, P., Bentall, R.P., 1997. Causal attributions in paranoia and depression: internal, personal, and situational attributions for negative events. J. Abnorm. Psychol. 106, 341–345.
King, R., Barchas, J.D., Huberman, B.A., 1984. Chaotic behavior in dopamine neurodynamics. Proc. Natl. Acad. Sci. U.S.A. 81, 1244–1247.
Knobel, A., Heinz, A., Voss, M., 2008. Imaging the deluded brain. Eur. Arch. Psychiatry Clin. Neurosci. 258 (Suppl. 5), 76–80.
Koechlin, E., Anton, J.L., Burnod, Y., 1996. Dynamical computational properties of local cortical networks for visual and motor processing: a bayesian framework. J. Physiol. Paris 90, 257–262.
Konorski, J., 1948. Conditioned Reflexes and Neuron Organization. Cambridge University Press.
Kosslyn, S.M., Chabris, C.F., Marsolek, C.J., Koenig, O., 1992. Categorical versus coordinate spatial relations: computational analyses and computer simulations. J. Exp. Psychol. Hum. Percept. Perform. 18, 562–577.
Kot, T., Serper, M., 2002. Increased susceptibility to auditory conditioning in hallucinating schizophrenic patients: a preliminary investigation. J. Nerv. Ment. Dis. 190, 282–288.
Kraeplin, E., 1902. Clinical Psychiatry. MacMillan, New York.
Kromkamp, M., Uylings, H.B., Smidt, M.P., Hellemons, A.J., Burbach, J.P., Kahn, R.S., 2003. Decreased thalamic expression of the homeobox gene DLX1 in psychosis. Arch. Gen. Psychiatry 60, 869–874.
Krystal, J.H., Karper, L.P., Seibyl, J.P., Freeman, G.K., Delaney, R., Bremner, J.D., Heninger, G.R., Bowers Jr., M.B., Charney, D.S., 1994. Subanesthetic effects of the noncompetitive NMDA antagonist, ketamine, in humans. Psychotomimetic, perceptual, cognitive, and neuroendocrine responses. Arch. Gen. Psychiatry 51, 199–214.
Kurylo, D.D., Gazes, Y., 2008. Effects of Ketamine on perceptual grouping in rats. Physiol. Behav. 95, 152–156.
Ladavas, E., Zeloni, G., Farne, A., 1998. Visual peripersonal space centred on the face in humans. Brain 121 (Part 12), 2317–2326.
Lange, R., Lange, R., Houran, J., 1998. Delusions of the paranormal: a haunting question of perception. J. Nerv. Ment. Dis. 186, 637.
Lapish, C.C., Seamans, J.K., Chandler, L.J., 2006. Glutamate-dopamine cotransmission and reward processing in addiction. Alcohol Clin. Exp. Res. 30, 1451–1465.
Laruelle, M., Kegeles, L.S., Abi-Dargham, A., 2003. Glutamate, dopamine, and schizophrenia: from pathophysiology to treatment. Ann. N. Y. Acad. Sci. 1003, 138–158.
Lasegue, C., Falret, J., 1877. La folie a deux (ou folie communiquee). Ann. Med. Psychol. 18, 321–355.
Lau, H.C., Rogers, R.D., Passingham, R.E., 2007. Manipulating the experienced onset of intention after action execution. J. Cogn. Neurosci. 19, 81–90.
Lavin, A., Nogueira, L., Lapish, C.C., Wightman, R.M., Phillips, P.E., Seamans, J.K., 2005. Mesocortical dopamine neurons operate in distinct temporal domains using multimodal signaling. J. Neurosci. 25, 5013–5023.
Laviolette, S.R., Grace, A.A., 2006. The roles of cannabinoid and dopamine receptor systems in neural emotional learning circuits: implications for schizophrenia and addiction. Cell. Mol. Life Sci. 63, 1597–1613.
Lazarus, A., 1986. Folie a deux in identical twins: interaction of nature and nurture. Br. J. Psychiatry 148, 324–326.
Lecourtier, L., Defrancesco, A., Moghaddam, B., 2008. Differential tonic influence of lateral habenula on prefrontal cortex and nucleus accumbens dopamine release. Eur. J. Neurosci. 27, 1755–1762.
Lee, D., 2008a. Game theory and neural basis of social decision making. Nat. Neurosci. 11, 404–409.
Lee, H.J., Groshek, F., Petrovich, G.D., Cantalini, J.P., Gallagher, M., Holland, P.C., 2005. Role of amygdalo-nigral circuitry in conditioning of a visual stimulus paired with food. J. Neurosci. 25, 3881–3888.
Lee, J.L., 2008b. Memory reconsolidation mediates the strengthening of memories by additional learning. Nat. Neurosci. 11, 1264–1266.
Lingnau, A., Gesierich, B., Caramazza, A., 2009. Asymmetric fMRI adaptation reveals no evidence for mirror neurons in humans. Proc. Natl. Acad. Sci. U.S.A. 106, 9925–9930.
Lisman, J., Buzsaki, G., 2008. A neural coding scheme formed by the combined function of gamma and theta oscillations. Schizophr. Bull. 34, 974–980.
Lisman, J., Redish, A.D., 2009. Prediction, sequences and the hippocampus. Philos. Trans. R. Soc. Lond. B Biol. Sci. 364, 1193–1201.
Lisman, J.E., Grace, A.A., 2005. The hippocampal-VTA loop: controlling the entry of information into long-term memory. Neuron 46, 703–713.
Llinas, R.R., Roy, S., 2009. The 'prediction imperative' as the basis for self-awareness. Philos. Trans. R. Soc. Lond. B Biol. Sci. 364, 1301–1307.
Lodge, D.J., Grace, A.A., 2006a. The hippocampus modulates dopamine neuron responsivity by regulating the intensity of phasic neuron activation. Neuropsychopharmacology 31, 1356–1361.
Lodge, D.J., Grace, A.A., 2006b. The laterodorsal tegmentum is essential for burst firing of ventral tegmental area dopamine neurons. Proc. Natl. Acad. Sci. U.S.A. 103, 5167–5172.
Loh, M., Rolls, E.T., Deco, G., 2007. A dynamical systems hypothesis of schizophrenia. PLoS Comput. Biol. 3, e228.
Long, J.E., Swan, C., Liang, W.S., Cobos, I., Potter, G.B., Rubenstein, J.L., 2009. Dlx1&2 and Mash1 transcription factors control striatal patterning and differentiation through parallel and overlapping pathways. J. Comp. Neurol. 512, 556–572.
Ma, W.J., Beck, J.M., Latham, P.E., Pouget, A., 2006. Bayesian inference with probabilistic population codes. Nat. Neurosci. 9, 1432–1438.
Mackintosh, N.J., 1975. A theory of attention: variations in the associability of stimuli with reinforcement. Psychol. Rev. 82 .
Maher, B.A., 1974. Delusional thinking and perceptual disorder. J. Individ. Psychol. 30 (1), 98–113.
Maher, B.A., 1988. Delusions as Normal Theories. Wiley, New York.
Makin, T.R., Holmes, N.P., Ehrsson, H.H., 2008. On the other hand: dummy hands and peripersonal space. Behav. Brain Res. 191, 1–10.
Matsumoto, M., Hikosaka, O., 2007. Lateral habenula as a source of negative reward signals in dopamine neurons. Nature 447 (7148), 1111–1115.
Matsumoto, M., Hikosaka, O., 2009. Two types of dopamine neuron distinctly convey positive and negative motivational signals. Nature 459, 837–841.
Maxwell, S.L., Li, M., 2005. Midbrain dopaminergic development in vivo and in vitro from embryonic stem cells. J. Anat. 207, 209–218.
McCurdy, H.G., 1956. Coin perception studies and the concept of schemata. Psychol. Rev. 63, 160–168.
McDonald, R.J., Hong, N.S., 2004. A dissociation of dorso-lateral striatum and amygdala function on the same stimulus-response habit task. Neuroscience 124, 507–513.
McGhie, A., Chapman, J., 1961. Disorders of attention and perception in early schizophrenia. Br. J. Med. Psychol. 34, 103–116.
McKay, R., Langdon, R., Coltheart, M., 2007. Models of misbelief: integrating motivational and deficit theories of delusions. Conscious. Cogn. 16, 932–941.
McLaren, I.P., Dickinson, A., 1990. The conditioning connection. Philos. Trans. R. Soc. Lond. B Biol. Sci. 329, 179–186.
McReynolds, P., 1960. Anxiety, perception and schizophrenia. In: Jackson, D.D. (Ed.), The etiology of schizophrenia. Basic Books, New York, pp. 248–292.
Mellor, C.S., 1970. First rank symptoms of schizophrenia. I. The frequnncy in schizophrenics on admission to hospital. II. Differences between individual first rank symptoms. Br. J. Psychiatry 117, 15–23.
Melo, S.S., Taylor, J.L., Bentall, R.P., 2006. 'Poor me' versus 'bad me' paranoia and the instability of persecutory ideation. Psychol. Psychother. 79, 271–287.
Mesulam, M., 2008. Representation, inference, and transcendent encoding in neurocognitive networks of the human brain. Ann. Neurol. 64, 367–378.

Meyer-Lindenberg, A., Straub, R.E., Lipska, B.K., Verchinski, B.A., Goldberg, T., Callicott, J.H., Egan, M.F., Huffaker, S.S., Mattay, V.S., Kolachana, B., Kleinman, J.E., Weinberger, D.R., 2007. Genetic evidence implicating DARPP-32 in human frontostriatal structure, function, and cognition. J. Clin. Invest. 117, 672–682.
Milad, M.R., Quirk, G.J., Pitman, R.K., Orr, S.P., Fischl, B., Rauch, S.L., 2007. A role for the human dorsal anterior cingulate cortex in fear expression. Biol. Psychiatry 62, 1191–1194.
Milad, M.R., Vidal-Gonzalez, I., Quirk, G.J., 2004. Electrical stimulation of medial prefrontal cortex reduces conditioned fear in a temporally specific manner. Behav. Neurosci. 118, 389–394.
Miller, N.E., 1959. Liberalization of basic S-R concepts: extensions to conflict behaviour, motivation and social learning. In: Kock, S. (Ed.), Psychology: A Study of a Science. McGraw-Hill, New York, pp. 196–292.
Miller, R., 1976. Schizophrenic psychology, associative learning and the role of forebrain dopamine. Med. Hypotheses 2, 203–211.
Miltner, W.H., Braun, C., Arnold, M., Witte, H., Taub, E., 1999. Coherence of gamma-band EEG activity as a basis for associative learning. Nature 397, 434–436.
Milton, F., Patwa, V.K., Hafner, R.J., 1978. Confrontation vs. belief modification in persistently deluded patients. Br. J. Med. Psychol. 51, 127–130.
Misanin, J.R., Miller, R.R., Lewis, D.J., 1968. Retrograde amnesia produced by electroconvulsive shock after reactivation of a consolidated memory trace. Science 160, 554–555.
Mishara, A.L., 2007. Is minimal self preserved in schizophrenia? A subcomponents view. Conscious. Cogn. 16, 715–721.
Mishara, A.L., Corlett, P.R., 2009. Are delusions biologically adaptive? Salvaging the doxastic shear pin. Behav. Brain Sci. 32, 530–531.
Mitchell, C.W., 1995. Effects of subliminally presented auditory suggestions of itching on scratching behavior. Percept. Mot. Skills 80, 87–96.
Mitchell, J., Vierkant, A.D., 1991. Delusions and hallucinations of cocaine abusers and paranoid schizophrenics: a comparative study. J. Psychol. 125, 301–310.
Mittleman, G., Whishaw, I.Q., Jones, G.H., Koch, M., Robbins, T.W., 1990. Cortical, hippocampal, and striatal mediation of schedule-induced behaviors. Behav. Neurosci. 104, 399–409.
Mohanty, A., Egner, T., Monti, J.M., Mesulam, M.M., 2009. Search for a threatening target triggers limbic guidance of spatial attention. J. Neurosci. 29, 10563–10572.
Mohr, F., Hubmann, W., Cohen, R., Bender, W., Haslacher, C., Honicke, S., Schlenker, R., Wahlheim, C., Werther, P., 1996. Neurological soft signs in schizophrenia: assessment and correlates. Eur. Arch. Psychiatry Clin. Neurosci. 246, 240–248.
Montague, P.R., Dayan, P., Sejnowski, T.J., 1996. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. J. Neurosci. 16, 1936–1947.
Moore, J.W., Wegner, D.M., Haggard, P., 2009. Modulating the sense of agency with external cues. Conscious. Cogn. 18 (4), 1056–1064.
Moritz, S., Woodward, T.S., Chen, E., 2006. Investigation of metamemory dysfunctions in first-episode schizophrenia. Schizophr. Res. 81, 247–252.
Morris, J.S., deBonis, M., Dolan, R.J., 2002. Human amygdala responses to fearful eyes. Neuroimage 17, 214–222.
Moussawi, K., Pacchioni, A., Moran, M., Olive, M.F., Gass, J.T., Lavin, A., Kalivas, P.W., 2009. N-acetylcysteine reverses cocaine-induced metaplasticity. Nat. Neurosci. 12, 182–189.
Moutoussis, M., Williams, J., Dayan, P., Bentall, R.P., 2007. Persecutory delusions and the conditioned avoidance paradigm: towards an integration of the psychology and biology of paranoia. Cogn. Neuropsychiatry 12, 495–510.
Murray, G.K., Corlett, P.R., Clark, L., Pessiglione, M., Blackwell, A.D., Honey, G., Jones, P.B., Bullmore, E.T., Robbins, T.W., Fletcher, P.C., 2008. Substantia nigra/ventral tegmental reward prediction error disruption in psychosis. Mol. Psychiatry 13 (239), 267–276.
Nader, K., Schafe, G.E., Le Doux, J.E., 2000. Fear memories require protein synthesis in the amygdala for reconsolidation after retrieval. Nature 406, 722–726.
Nagahama, Y., Okina, T., Suzuki, N., Matsuda, M., 2010. Neural correlates of psychotic symptoms in dementia with Lewy bodies. Brain 133 (Pt 2), 557–567.
Nagahama, Y., Okina, T., Suzuki, N., Matsuda, M., Fukao, K., Murai, T., 2007. Classification of psychotic symptoms in dementia with Lewy bodies. Am. J. Geriatr. Psychiatry 15, 961–967.
Nelson, A., Killcross, S., 2006. Amphetamine exposure enhances habit formation. J. Neurosci. 26, 3805–3812.
O'Donnell, P., 2003. Dopamine gating of forebrain neural ensembles. Eur. J. Neurosci. 17, 429–435.
Owen, G., Harland, R., Antonova, E., Broome, M., 2004. Jaspers' concept of primary delusion. Br. J. Psychiatry 185, 77–78.
Oye, I., Paulsen, O., Maurset, A., 1992. Effects of ketamine on sensory perception: evidence for a role of N-methyl-D-aspartate receptors. J. Pharmacol. Exp. Ther. 260, 1209–1213.
Padoa-Schioppa, C., Assad, J.A., 2006. Neurons in the orbitofrontal cortex encode economic value. Nature 441 (7090), 223–226.
Pally, R., 2007. The predicting brain: unconscious repetition, conscious reflection and therapeutic change. Int. J. Psychoanal. 88, 861–881.
Pan, W.X., Schmidt, R., Wickens, J.R., Hyland, B.I., 2008. Tripartite mechanism of extinction suggested by dopamine neuron activity and temporal difference model. J. Neurosci. 28, 9619–9631.
Parthasarathi, U.D., Harrower, T., Tempest, M., Hodges, J.R., Walsh, C., McKenna, P.J., Fletcher, P.C., 2006. Psychiatric presentation of voltage-gated potassium channel antibody-associated encephalopathy. Case report. Br. J. Psychiatry 189, 182–183.
Passie, T., Karst, M., Borsutzky, M., Wiese, B., Emrich, H.M., Schneider, U., 2003. Effects of different subanaesthetic doses of (S)-ketamine on psychopathology and binocular depth inversion in man. J. Psychopharmacol. 17, 51–56.
Paton, J.J., Belova, M.A., Morrison, S.E., Salzman, C.D., 2006. The primate amygdala represents the positive and negative value of visual stimuli during learning. Nature 439, 865–870.
Pearce, J.M., Hall, G., 1980. A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. Psychol. Rev. 87, 532–552.
Peled, A., Pressman, A., Geva, A.B., Modai, I., 2003. Somatosensory evoked potentials during a rubber-hand illusion in schizophrenia. Schizophr. Res. 64, 157–163.
Pennartz, C.M., McNaughton, B.L., Mulder, A.B., 2000. The glutamate hypothesis of reinforcement learning. Prog. Brain Res. 126, 231–253.
Pessiglione, M., Seymour, B., Flandin, G., Dolan, R.J., Frith, C.D., 2006. Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. Nature 442, 1042–1045.
Phillips, W.A., Silverstein, S.M., 2003. Convergence of biological and psychological perspectives on cognitive coordination in schizophrenia. Behav. Brain Sci. 26, 65–82 discussion 82–137.
Pomarol-Clotet, E., Honey, G.D., Murray, G.K., Corlett, P.R., Absalom, A.R., Lee, M., McKenna, P.J., Bullmore, E.T., Fletcher, P.C., 2006. Psychological effects of ketamine in healthy volunteers. Phenomenological study. Br. J. Psychiatry 189, 173–179.
Press, C., Heyes, C., Haggard, P., Eimer, M., 2008. Visuotactile learning and body representation: an ERP study with rubber hands and rubber objects. J. Cogn. Neurosci. 20, 312–323.
Preuschoff, K., Bossaerts, P., Quartz, S.R., 2006. Neural differentiation of expected reward and risk in human subcortical structures. Neuron 51, 381–390.
Ramachandran, V., Blakeslee, S., 1998. Phantoms in the Brain: Probing the Mysteries of the Human Mind. William Morrow, New York.
Rao, R.P., Ballard, D.H., 1999. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nat. Neurosci. 2, 79–87.
Ravina, B., Marder, K., Fernandez, H.H., Friedman, J.H., McDonald, W., Murphy, D., Aarsland, D., Babcock, D., Cummings, J., Endicott, J., Factor, S., Galpern, W., Lees, A., Marsh, L., Stacy, M., Gwinn-Hardy, K., Voon, V., Goetz, C., 2007. Diagnostic criteria for psychosis in Parkinson's disease: report of an NINDS, NIMH work group. Mov. Disord. 22, 1061–1068.
Reading, P.J., Dunnett, S.B., Robbins, T.W., 1991. Dissociable roles of the ventral, medial and lateral striatum on the acquisition and performance of a complex visual stimulus-response habit. Behav. Brain Res. 45, 147–161.
Redgrave, P., Gurney, K., 2006. The short-latency dopamine signal: a role in discovering novel actions? Nat. Rev. Neurosci. 7, 967–975.
Redish, A.D., 2004. Addiction as a computational process gone awry. Science 306, 1944–1947.
Rescorla, R.A., Wagner, A.R., 1972. A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and non-reinforcement. In: Black, A.H., Prokasy, W.F. (Eds.), Classical Conditioning II: Current Research and Theory. Appleton-Century-Crofts, New York.
Ridler, K., Veijola, J.M., Tanskanen, P., Miettunen, J., Chitnis, X., Suckling, J., Murray, G.K., Haapea, M., Jones, P.B., Isohanni, M.K., Bullmore, E.T., 2006. Fronto-cerebellar systems are associated with infant motor and adult executive functions in healthy adults but not in schizophrenia. Proc. Natl. Acad. Sci. U.S.A. 103, 15651–15656.
Rizzolatti, G., Fadiga, L., Gallese, V., Fogassi, L., 1996. Premotor cortex and the recognition of motor actions. Brain Res. Cogn. Brain Res. 3, 131–141.
Roiser, J.P., Stephan, K.E., den Ouden, H.E., Barnes, T.R., Friston, K.J., Joyce, E.M., 2009. Do patients with schizophrenia exhibit aberrant salience? Psychol. Med. 39, 199–209.
Rolls, E.T., 2007. The representation of information about faces in the temporal and frontal lobes. Neuropsychologia 45, 124–143.
Rolls, E.T., Grabenhorst, F., 2008. The orbitofrontal cortex and beyond: from affect to decision-making. Prog. Neurobiol. 86, 216–244.
Rolls, E.T., Grabenhorst, F., Margot, C., da Silva, M.A., Velazco, M.I., 2008. Selective attention to affective value alters how the brain processes olfactory stimuli. J. Cogn. Neurosci. 20, 1815–1826.
Romo, R., Merchant, H., Ruiz, S., Crespo, P., Zainos, A., 1995. Neuronal activity of primate putamen during categorical perception of somaesthetic stimuli. Neuroreport 6, 1013–1017.
Ros, H., Sachdev, R.N., Yu, Y., Sestan, N., McCormick, D.A., 2009. Neocortical networks entrain neuronal circuits in cerebellar cortex. J. Neurosci. 29, 10309–10320.
Rubin, R.D., 1976. Clinical use of retrograde amnesia produced by electroconvulsive shock. A conditioning hypothesis. Can. Psychiatr. Assoc. J. 21, 87–90.
Sanchez-Vives, M.V., McCormick, D.A., 2000. Cellular and network mechanisms of rhythmic recurrent activity in neocortex. Nat. Neurosci. 3, 1027–1034.
Sander, D., Grafman, J., Zalla, T., 2003. The human amygdala: an evolved system for relevance detection. Rev. Neurosci. 14, 303–316.
Sarter, M., Nelson, C.L., Bruno, J.P., 2005. Cortical cholinergic transmission and cortical information processing in schizophrenia. Schizophr. Bull. 31, 117–138.
Sass, L.A., 2004. Some reflections on the (analytic) philosophical approach to delusion. Philos. Psychiatry Psychol. 11, 71–80.
Scearce-Levie, K., Roberson, E.D., Gerstein, H., Cholfin, J.A., Mandiyan, V.S., Shah, N.M., Rubenstein, J.L., Mucke, L., 2008. Abnormal social behaviors in mice lacking Fgf17. Genes Brain. Behav. 7, 344–354.

Scharfetter, C., 1970. On the hereditary aspects of symbiontic psychoses. A contribution towards the understanding of the schizophrenia-like psychoses. Psychiatr. Clin. (Basel) 3, 145–152.
Schiller, D., Levy, I., Niv, Y., LeDoux, J.E., Phelps, E.A., 2008. From fear to safety and back: reversal of fear in the human brain. J. Neurosci. 28, 11517–11525.
Schlagenhauf, F., Sterzer, P., Schmack, K., Ballmaier, M., Rapp, M., Wrase, J., Juckel, G., Gallinat, J., Heinz, A., 2009. Reward feedback alterations in unmedicated schizophrenia patients: relevance for delusions. Biol. Psychiatry 65, 1032–1039.
Schneider, K., 1959. Clinical Psychopathology. Grune & Stratton, New York.
Schnell, K., Heekeren, K., Daumann, J., Schnell, T., Schnitker, R., Moller-Hartmann, W., Gouzoulis-Mayfrank, E., 2008. Correlation of passivity symptoms and dysfunctional visuomotor action monitoring in psychosis. Brain 131, 2783–2797.
Schnider, A., 2001. Spontaneous confabulation, reality monitoring, and the limbic system—a review. Brain Res. Brain Res. Rev. 36, 150–160.
Schnider, A., 2003. Spontaneous confabulation and the adaptation of thought to ongoing reality. Nat. Rev. Neurosci. 4, 662–671.
Schobel, S.A., Lewandowski, N.M., Corcoran, C.M., Moore, H., Brown, T., Malaspina, D., Small, S.A., 2009. Differential targeting of the CA1 subfield of the hippocampal formation by schizophrenia and related psychotic disorders. Arch. Gen. Psychiatry 66, 938–946.
Schultz, W., 1998. Predictive reward signal of dopamine neurons. J. Neurophysiol. 80, 1–27.
Schultz, W., Apicella, P., Ljungberg, T., 1993. Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. J. Neurosci. 13, 900–913.
Schultz, W., Dayan, P., Montague, P.R., 1997. A neural substrate of prediction and reward. Science 275, 1593–1599.
Schultz, W., Dickinson, A., 2000. Neuronal coding of prediction errors. Annu. Rev. Neurosci. 23, 473–500.
Seamans, J.K., Yang, C.R., 2004. The principal features and mechanisms of dopamine modulation in the prefrontal cortex. Prog. Neurobiol. 74, 1–58.
Seashore, C.E., 1895. Measurements of illusions and hallucinations in normal life. Stud. Yale Psychol. Lab. 3, 1–67.
Sesack, S.R., Grace, A.A., 2010. Cortico-Basal Ganglia reward network: microcircuitry. Neuropsychopharmacology 35, 27–47.
Shaner, A., 1999. Delusions, superstitious conditioning and chaotic dopamine neurodynamics. Med. Hypotheses 52, 119–123.
Shanks, D.R., 2006. Bayesian associative learning. Trends. Cogn. Sci. 10 (11), 477–478.
Shanks, D.R., 2010. Learning: from association to cognition. Annu. Rev. Psychol. 61, 273–301.
Sharp, F.R., Tomitaka, M., Bernaudin, M., Tomitaka, S., 2001. Psychosis: pathological activation of limbic thalamocortical circuits by psychomimetics and schizophrenia? Trends Neurosci. 24, 330–334.
Shepard, P.D., Holcomb, H.H., Gold, J.M., 2006. Schizophrenia in translation: the presence of absence: habenular regulation of dopamine neurons and the encoding of negative outcomes. Schizophr. Bull. 32, 417–421.
Sherman, S.M., Guillery, R.W., 1998. On the actions that one nerve cell can have on another: distinguishing "drivers" from "modulators". Proc. Natl. Acad. Sci. U.S.A. 95, 7121–7126.
Siegel, R.K., 1978. Cocaine hallucinations. Am. J. Psychiatry 135, 309–314.
Sillitoe, R.V., Stephen, D., Lao, Z., Joyner, A.L., 2008. Engrailed homeobox genes determine the organization of Purkinje cell sagittal stripe gene expression in the adult cerebellum. J. Neurosci. 28, 12150–12162.
Silverstein, S., Uhlhaas, P.J., Essex, B., Halpin, S., Schall, U., Carr, V., 2006. Perceptual organization in first episode schizophrenia and ultra-high-risk states. Schizophr. Res. 83, 41–52.
Simons, J.S., Henson, R.N., Gilbert, S.J., Fletcher, P.C., 2008. Separable forms of reality monitoring supported by anterior prefrontal cortex. J. Cogn. Neurosci. 20, 447–457.
Simpson, J., Done, D.J., 2002. Elasticity and confabulation in schizophrenic delusions. Psychol. Med. 32, 451–458.
Sirota, A., Montgomery, S., Fujisawa, S., Isomura, Y., Zugaro, M., Buzsaki, G., 2008. Entrainment of neocortical neurons and gamma oscillations by the hippocampal theta rhythm. Neuron 60, 683–697.
Skinner, B.F., 1948. "Superstition" in the pigeon. J. Exp. Psychol. 38, 168–172.
Sokolov, E.N., 1960. Neuronal Models and the Orienting Reflex. Josiah Macy Jr Foundation, New York.
Soltani, A., Wang, X.J., 2010. Synaptic computation underlying probabilistic inference. Nat. Neurosci. 13 (1), 112–119.
Spence, S.A., Brooks, D.J., Hirsch, S.R., Liddle, P.F., Meehan, J., Grasby, P.M., 1997. A PET study of voluntary movement in schizophrenic patients experiencing passivity phenomena (delusions of alien control). Brain 120 (Part 11), 1997–2011.
Sperry, R.W., 1990. Forebrain commisurectomy and consciuos awareness. In: Trevarthen, C. (Ed.), Brain Circuits and the Mind. Cambridge University Press, New York.
Spitzer, M., 1995. A neurocomputational approach to delusions. Compr. Psychiatry 36, 83–105.
Spitzer, M., Walter, H., 2003. The cognitive neuroscience of agency in schizophrenia. In: David, A., Kircher, T. (Eds.), The Self in Neuroscience and Psychiatry. Cambridge University Press, Cambridge, pp. 436–444.
Startup, M., Startup, S., 2005. On two kinds of delusion of reference. Psychiatry Res. 137, 87–92.
Stephan, K.E., Baldeweg, T., Friston, K.J., 2006. Synaptic plasticity and dysconnection in schizophrenia. Biol. Psychiatry 59, 929–939.
Steriade, M., Dossi, R.C., Pare, D., Oakson, G., 1991. Fast oscillations (20–40 Hz) in thalamocortical systems and their potentiation by mesopontine cholinergic nuclei in the cat. Proc. Natl. Acad. Sci. U.S.A. 88, 4396–4400.
Stickgold, R., Walker, M.P., 2007. Sleep-dependent memory consolidation and reconsolidation. Sleep Med. 8, 331–343.
Sur, M., Rubenstein, J.L., 2005. Patterning and plasticity of the cerebral cortex. Science 310, 805–810.
Sutton, R.S., Barto, A.G., 1998. Reinforcement Learning: An Introduction. MIT Press.
Svenningsson, P., Tzavara, E.T., Carruthers, R., Rachleff, I., Wattler, S., Nehls, M., McKinzie, D.L., Fienberg, A.A., Nomikos, G.G., Greengard, P., 2003. Diverse psychotomimetics act through a common signaling pathway. Science 302, 1412–1415.
Tabares-Seisdedos, R., Rubenstein, J.L., 2009. Chromosome 8p as a potential hub for developmental neuropsychiatric disorders: implications for schizophrenia, autism and cancer. Mol. Psychiatry 14, 563–589.
Takahashi, Y., Schoenbaum, G., Niv, Y., 2008. Silencing the critics: understanding the effects of cocaine sensitization on dorsolateral and ventral striatum in the context of an actor/critic model. Front. Neurosci. 2, 86–99.
Takahashi, Y.K., Roesch, M.R., Stalnaker, T.A., Haney, R.Z., Calu, D.J., Taylor, A.R., Burke, K.A., Schoenbaum, G., 2009. The orbitofrontal cortex and ventral tegmental area are necessary for learning from unexpected outcomes. Neuron 62 (2), 269–280.
Tang, C., Pawlak, A.P., Prokopenko, V., West, M.O., 2007. Changes in activity of the striatum during formation of a motor habit. Eur. J. Neurosci. 25, 1212–1227.
Teufel, C., Alexis, D.M., Todd, H., Lawrance-Owen, A.J., Clayton, N.S., Davis, G., 2009. Social cognition modulates the sensory coding of observed gaze direction. Curr. Biol. 19, 1274–1277.
Thiebierge, G., 1894. Les acaraphobes. Ann. Dermatol. Syphiligraphie 3, 730–736.
Thorndike, E.L., 1911. Animal Intelligence. MacMillan, New York.
Tole, S., Remedios, R., Saha, B., Stoykova, A., 2005. Selective requirement of Pax6, but not Emx2, in the specification and development of several nuclei of the amygdaloid complex. J. Neurosci. 25, 2753–2760.
Tolman, E.C., 1932. Purposive Behaviour in Animals and Men. Century, New York.
Tricomi, E., Balleine, B.W., O'Doherty, J.P., 2009. A specific role for posterior dorsolateral striatum in human habit learning. Eur. J. Neurosci. 29, 2225–2232.
Tsakiris, M., Haggard, P., 2005. The rubber hand illusion revisited: visuotactile integration and self-attribution. J. Exp. Psychol. Hum. Percept. Perform. 31, 80–91.
Turner, D.C., Aitken, M.R., Shanks, D.R., Sahakian, B.J., Robbins, T.W., Schwarzbauer, C., Fletcher, P.C., 2004. The role of the lateral frontal cortex in causal associative learning: exploring preventative and super-learning. Cereb. Cortex 14, 872–880.
Uhlhaas, P.J., Haenschel, C., Nikolic, D., Singer, W., 2008. The role of oscillations and synchrony in cortical networks and their putative relevance for the pathophysiology of schizophrenia. Schizophr. Bull. 34, 927–943.
Uhlhaas, P.J., Linden, D.E., Singer, W., Haenschel, C., Lindner, M., Maurer, K., Rodriguez, E., 2006a. Dysfunctional long-range coordination of neural activity during Gestalt perception in schizophrenia. J. Neurosci. 26, 8168–8175.
Uhlhaas, P.J., Mishara, A.L., 2007. Perceptual anomalies in schizophrenia: integrating phenomenology and cognitive neuroscience. Schizophr. Bull. 33, 142–156.
Uhlhaas, P.J., Phillips, W.A., Mitchell, G., Silverstein, S.M., 2006b. Perceptual grouping in disorganized schizophrenia. Psychiatry Res. 145, 105–117.
Uhlhaas, P.J., Singer, W., 2010. Abnormal neural oscillations and synchrony in schizophrenia. Nat. Rev. Neurosci. 11, 100–113.
Vallar, G., Ronchi, R., 2009. Somatoparaphrenia: a body delusion. A review of the neuropsychological literature. Exp. Brain Res. 192, 533–551.
van Nimwegen, L., de Haan, L., van Beveren, N., van den Brink, W., Linszen, D., 2005. Adolescence, schizophrenia and drug abuse: a window of vulnerability. Acta Psychiatr. Scand. Suppl. 35–42.
Vernon, D., Haenschel, C., Dwivedi, P., Gruzelier, J., 2005. Slow habituation of induced gamma and beta oscillations in association with unreality experiences in schizotypy. Int. J. Psychophysiol. 56, 15–24.
Vickery, T.J., Jiang, Y.V., 2009. Associative grouping: perceptual grouping of shapes by association. Atten. Percept. Psychophys. 71, 896–909.
Vinogradov, S., King, R.J., Huberman, B.A., 1992. An associationist model of the paranoid process: application of phase transitions in spreading activation networks. Psychiatry 55, 79–94.
Von Holst, E., 1954. Relations between the central nervous system and the peripheral organs. Br. J. Anim. Behav. 2, 89–94.
Waelti, P., Dickinson, A., Schultz, W., 2001. Dopamine responses comply with basic assumptions of formal learning theory. Nature 412, 43–48.
Waldmann, M.R., Martignon, L., 1998. A Bayesian network model of causal learning. In: Gernsbacher, M.A., Derry, S.J. (Eds.), Proceedings of the Twentieth Annual Conference of the Cognitive Science Society. Earlbaum, Mahwah, NJ, pp. 1102–1107.
Wallis, G.G., 1949. A case of hallucinosis due to cocaine. J. R. Nav. Med. Serv. 35, 112.
Walsh, T., McClellan, J.M., McCarthy, S.E., Addington, A.M., Pierce, S.B., Cooper, G.M., Nord, A.S., Kusenda, M., Malhotra, D., Bhandari, A., Stray, S.M., Rippey, C.F., Roccanova, P., Makarov, V., Lakshmi, B., Findling, R.L., Sikich, L., Stromberg, T., Merriman, B., Gogtay, N., Butler, P., Eckstrand, K., Noory, L., Gochman, P., Long, R., Chen, Z., Davis, S., Baker, C., Eichler, E.E., Meltzer, P.S., Nelson, S.F., Singleton, A.B., Lee, M.K., Rapoport, J.L., King, M.C., Sebat, J., 2008. Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. Science 320, 539–543.

Waltz, J.A., Frank, M.J., Robinson, B.M., Gold, J.M., 2007. Selective reinforcement learning deficits in schizophrenia support predictions from computational models of striatal-cortical dysfunction. Biol. Psychiatry 62, 756–764.
Wang, J., O'Donnell, P., 2001. D(1) dopamine receptors potentiate nmda-mediated excitability increase in layer V prefrontal cortical pyramidal neurons. Cereb. Cortex 11, 452–462.
Wang, X.J. Neurophysiological and computational principles of cortical rhythms in cognition. Physiol. Rev., in press.
Wegner, D.M., 2004. Precis of the illusion of conscious will. Behav. Brain Sci. 27, 649–659 discussion 659–692.
Weiller, C., Juptner, M., Fellows, S., Rijntjes, M., Leonhardt, G., Kiebel, S., Muller, S., Diener, H.C., Thilmann, A.F., 1996. Brain representation of active and passive movements. Neuroimage 4, 105–110.
Weiskrantz, L., Elliott, J., Darlington, C., 1971. Preliminary observations on tickling oneself. Nature 230, 598–599.
Whalen, P.J., Rauch, S.L., Etcoff, N.L., McInerney, S.C., Lee, M.B., Jenike, M.A., 1998. Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. J. Neurosci. 18, 411–418.
Whalen, P.J., Shin, L.M., McInerney, S.C., Fischer, H., Wright, C.I., Rauch, S.L., 2001. A functional MRI study of human amygdala responses to facial expressions of fear versus anger. Emotion 1, 70–83.
Whitson, J.A., Galinsky, A.D., 2008. Lacking control increases illusory pattern perception. Science 322, 115–117.
Winterer, G., 2006. Cortical microcircuits in schizophrenia—the dopamine hypothesis revisited. Pharmacopsychiatry 39 (Suppl. 1), S68–S71.
Wolff, G., McKenzie, K., 1994. Capgras, Fregoli and Cotard's syndromes and Koro in folie a deux. Br. J. Psychiatry 165, 842.
Wolpert, D.M., Ghahramani, Z., Jordan, M.I., 1995. An internal model for sensorimotor integration. Science 269, 1880–1882.
Wolpert, D.M., Miall, R.C., 1996. Forward models for physiological motor control. Neural Netw. 9, 1265–1279.
Woodward, T.S., Moritz, S., Chen, E.Y., 2006. The contribution of a cognitive bias against disconfirmatory evidence (BADE) to delusions: a study in an Asian sample with first episode schizophrenia spectrum disorders. Schizophr. Res. 83, 297–298.
Yin, H.H., Knowlton, B.J., Balleine, B.W., 2004. Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. Eur. J. Neurosci. 19, 181–189.
Yoo, S.S., Freeman, D.K., McCarthy 3rd, J.J., Jolesz, F.A., 2003. Neural substrates of tactile imagery: a functional MRI study. Neuroreport 14, 581–585.
Young, A.W., Robertson, I.H., Hellawell, D.J., de Pauw, K.W., Pentland, B., 1992. Cotard delusion after brain injury. Psychol. Med. 22, 799–804.
Young, G., 2008. Capgras delusion: an interactionist model. Conscious. Cogn. 17, 863–876.
Yu, A.J., Dayan, P., 2005. Uncertainty, neuromodulation, and attention. Neuron 46, 681–692.