

Journal of Psychopharmacology
2016, Vol. 30(11) 1145–1155
© The Author(s) 2016



Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/0269881116650087
jop.sagepub.com



Prediction error, ketamine and psychosis: An updated model

Philip R Corlett¹, Garry D Honey² and Paul C Fletcher^{3,4}

Abstract

In 2007, we proposed an explanation of delusion formation as aberrant prediction error-driven associative learning. Further, we argued that the NMDA receptor antagonist ketamine provided a good model for this process. Subsequently, we validated the model in patients with psychosis, relating aberrant prediction error signals to delusion severity. During the ensuing period, we have developed these ideas, drawing on the simple principle that brains build a model of the world and refine it by minimising prediction errors, as well as using it to guide perceptual inferences. While previously we focused on the prediction error signal per se, an updated view takes into account its precision, as well as the precision of prior expectations. With this expanded perspective, we see several possible routes to psychotic symptoms – which may explain the heterogeneity of psychotic illness, as well as the fact that other drugs, with different pharmacological actions, can produce psychotomimetic effects. In this article, we review the basic principles of this model and highlight specific ways in which prediction errors can be perturbed, in particular considering the reliability and uncertainty of predictions. The expanded model explains hallucinations as perturbations of the uncertainty mediated balance between expectation and prediction error. Here, expectations dominate and create perceptions by suppressing or ignoring actual inputs. Negative symptoms may arise due to poor reliability of predictions in service of action. By mapping from biology to belief and perception, the account proffers new explanations of psychosis. However, challenges remain. We attempt to address some of these concerns and suggest future directions, incorporating other symptoms into the model, building towards better understanding of psychosis.

Keywords

Drug model, ketamine, psychosis, schizophrenia, delusions, hallucinations, computational psychiatry

Introduction

In 2007, in this journal, we outlined a theory of delusion formation expressed in terms of associative learning theory (Corlett et al., 2007a). It was not the first theory of delusions expressed in this framework (Gray et al., 1991; Miller, 1976), but it did implicate a specific psychological process (prediction error [PE]) and its neurochemical underpinnings in the genesis of delusions. We are very honoured to have been invited to revisit this article, and would like to take this opportunity to discuss the origins of the ideas, the key features of the theory, the evidence that has emerged to support and challenge it, and, importantly, the ways in which it has evolved in the context of a cognitive neuroscience field that has advanced rapidly and helped to shape its development.

We begin with a consideration of what attracted us to the central ideas outline in the original article and what precisely we were hypothesising by invoking PE to account for how delusional thinking might emerge.

Delusions and PE: The key ideas

We, like many, feel the need for a perspective on the key features of psychosis – delusions and hallucinations – that links the perceptual and inferential aspects of these experiences with the underlying biological mechanisms. Without such a linking mechanism, biological explanations of mental illness will remain incomplete. Computational cognitive neuroscience offers the potential to unite multiple levels of explanation

through deployment of computational models that can be plausibly related to activity of brain systems, the instantiation of cognitive processes and to high-level behaviour and experience (Corlett and Fletcher, 2014). Our early attempt to do this drew on a number of areas in the existing literature. In particular, we noted that, going back to Bleuler's (1908) earliest formulations (Miller, 1976), the delusions characteristic of schizophrenia occur against a background of strange and sometimes bizarre associations and experiences. This theme – linking delusional thinking to associative learning (Hartley, 1749/1976) – was developed more formally (Dickinson, 2001) in the context of the theoretically and empirically rich field of animal reinforcement learning, itself given formal foundations by advances in machine learning (Sutton and Barto, 1998) and artificial intelligence (Widrow and Hoff, 1960). Both of these fields conceptualised a fundamental role of the brain as identifying and updating

¹Department of Psychiatry, Yale University, New Haven, CT, USA

²Roche Switzerland, Basel, Switzerland

³Department of Psychiatry, University of Cambridge, Cambridge, UK

⁴Cambridgeshire and Peterborough NHS Foundation Trust, Cambridge, UK

Corresponding author:

Paul C Fletcher, Department of Psychiatry, University of Cambridge, UK.
Email: philip.corlett@yale.edu

statistical regularities (or associations), in effect to build an internal model of the world (Conant and Ashby, 1970).

Ensuing and increasingly complex and sophisticated perspectives on this idea have envisaged that this challenge is met through an iterative process of predicting and updating (Adams et al., 2013; Corlett et al., 2009a; Fletcher and Frith, 2009). Briefly, the brain uses prior knowledge (prior experience of associative relationships) to predict what its next input will be, and any mismatch between the prediction and what actually ensues is signalled as a PE. This is sometimes referred to as predictive coding or predictive processing (Clark, 2013). PE is a key drive to new learning in so far as it indicates incorrect predictions and hence a model that may need to be updated.

This simple formulation provided the framework that supported our initial consideration of PE in psychosis. Moreover, it offered us an operationally defined and quantifiable parameter that could readily be applied in analyses of neuroimaging data in humans (Corlett et al., 2004; Corlett and Fletcher, 2012; Corlett et al., 2007b; Fletcher et al., 2001; Murray et al., 2008; Turner et al., 2004). If PE signalling is altered, arising inappropriately or with an anomalous degree of strength or precision, then there would ensue, we argued, a compelling sense that one's existing model of the world was wrong, that something had changed. Perhaps there would be an enhanced tendency to see spurious associations between stimuli and events that, in reality, were not related. Perhaps too, there would arise a sense of a world that had changed, become more sinister and laden with meaning. And, having established a new set of associations, then these would form the framework dictating how attention might be diverted towards ensuing stimuli and shaping the inferences that arose. Put simply, an aberrant PE signal would lead to new learning, and new learning would engender new expectations that would themselves govern how the individual sampled and interpreted the world. In effect, this could lead to a developing change in which emerging beliefs created the evidence that supported them.

So it may be, we suggested, that delusions begin to emerge. Importantly, this explanatory framework, though far from comprehensive, offered the possibility of linking symptoms to brain processes because associative learning processes, including PE, had all been extensively studied in terms of their underlying neurobiology (Fletcher et al., 2001; Lavin et al., 2005; Schultz and Dickinson, 2000; Turner et al., 2004). Our attempts to establish this link focused on dopaminergic and glutamatergic systems and their interactions, given the evidence that these are critical to PE signalling (Lavin et al., 2005; Schultz and Dickinson, 2000). Since ketamine, an influential and compelling model of early psychosis (Krystal et al., 1994; Pomarol-Clotet et al., 2006), impacts both transmitter systems (Kegeles et al., 2000), this framework enabled a comprehensive account of the drug's psychotomimetic effects.

This early model, as we review below, has proven useful in contributing to subsequent ideas and research in our groups and beyond. Moreover, it has garnered support from a number of experimental approaches. But, like all models, it was necessarily a simplification. In particular, it focused primarily on the formation of simple associations between stimuli or between cause and effect, and it drew out some basic ideas of how a perturbed PE signal might disrupt association formation as well as perception and attention. But it is important to move beyond deterministic

first-order associations: the statistical regularities of the world exist at many levels and interact in complex ways. Just as the association between, for example, an apple and an apple tree is dependent on a higher level concept of seasons and weather, so the challenge faced by the brain is to represent associations in ways that are sensitive to context and to more remote second- and third-order associations. And within this framework, not only does the model become complex and intricate, but the circumstances under which it must be updated become ever more opaque. Consider seeing an apple tree with no apples: there are numerous reasons why this PE should not be a reason for updating our association between apples and apple trees, and the likelihood of modifying our belief about this association would depend on many factors. Computational models coming to grips with this complexity are producing insights that show great promise for informing and testing hypotheses about mechanisms by which psychosis – aberrant world modelling – may emerge (Adams et al., 2013; Corlett and Fletcher, 2014; Fletcher and Frith, 2009).

Thus, there were a number of important parameters that were not part of our model, and we consider these in the latter part of this article, showing how their inclusion enriches, without substantially changing, the basic ideas adumbrated in the original article. In particular, we consider how an aberrant PE signal, over time, could lead to adjustments in attention and perception, as well as the readiness to update one's model of the world in response to new information (Adams et al., 2013; Corlett and Fletcher, 2014; Fletcher and Frith, 2009). In this richer and more nuanced consideration, we find promising ways of extending the model to explain other key aspects of delusions, as well as important accompanying features such as hallucinations and negative symptoms (Adams et al., 2013; Corlett, 2015; Corlett and Fletcher, 2014; Fletcher and Frith, 2009).

Prior to this, we focus on how the basic theory, as originally outlined, has fared empirically and how it relates to other accounts of delusions (Coltheart and Davies, 2000) and to schizophrenia more broadly.

Empirical approaches to the PE model of delusions

In 2007, we made the case that ketamine infusion in healthy people provided a window on a hitherto experimentally challenging situation: the emergence of psychotic experience and belief. With the advent of early intervention approaches in psychosis and studies of the psychosis prodrome, it became possible to study patients closer to this illness phase. Furthermore, the continuum from healthy beliefs through to delusions has been increasingly appreciated. Studying attenuated psychotic symptoms (in otherwise healthy subjects) has proven a fruitful avenue of inquiry for testing the model. It should be noted that the basic experimental design that we have initially favoured, in provoking PE signal in order to characterise neural responses using functional neuroimaging, has been challenged (Griffiths et al., 2014). Having responded to this challenge (Corlett and Fletcher, 2015), we do not propose to revisit the argument here, but do note that the fundamental ideas underlying the model were applauded and the overall evidence in their favour was considered robust (Griffiths et al., 2014). There is a more fundamental challenge relating to whether a PE signal disruption is a

sufficient circumstance to engender delusions, and given that this relates to a long-standing dialogue among delusion theorists, we consider this in more detail below.

Before considering the challenges to the model and its shortcomings, it is reassuring to note that there is a good deal of support for the idea that psychosis is associated with altered PE signalling. Importantly, in a study of patients with first-episode psychosis, using an identical task, we observed a very similar pattern of PE responses in the right dorsolateral prefrontal cortex as that seen in healthy people administered ketamine (Corlett et al., 2006, 2007b). Crucially, these aberrant responses correlated with the severity of altered beliefs across subjects. In the same participants, using a different (reward-based) learning task, we observed a pattern of altered PE response that was highly comparable to controls (Murray et al., 2008). In both tasks, causal learning and reward PE signals in frontal, striatal and midbrain regions were inappropriately engaged. However, only rDLPFC PE during causal belief formation correlated with delusion scores in patients with first-episode psychosis. In healthy, non-psychotic people too, the degree to which their prefrontal PE response resembles that observed in patients with delusions correlates with the distress they feel with regard to their delusion-like ideas, though ventral striatal responses associated with the ideas themselves, irrespective of accompanying distress (Corlett and Fletcher, 2012). That is, if you hold your beliefs like a patient with psychosis, your right frontal PE response approximates that observed in patients with delusions (Corlett and Fletcher, 2012). These data came from our own work. Others have similarly observed aberrant PE signals in striatum, amygdala and frontal cortex in patients with psychotic illness that correlate with the severity of delusions (Gradin et al., 2011; Romaniuk et al., 2010; Schlagenhauf et al., 2009; Waltz et al., 2010).

In short, empirical data from patients with psychosis have consistently linked aberrant PE signal (measured with functional magnetic resonance imaging in various task contexts; electroencephalogram [EEG] and magnetoencephalogram correlates of PE have yet to be related to delusions) to the severity of delusions. However, there have been a number of theoretical and empirical challenges to the model that we now go on to discuss.

Shortcomings of the 2007 model

When we presented the model in 2007, it was an initial sketch of how ketamine might give rise to delusion-like ideas. While it gave us some elbow room to begin carving a more complete explanation of clinical delusions in terms of mind and brain function, it was by no means complete. It did not, for example, address one of the cardinal features of delusions: their fixity in the face of contradictory evidence. Indeed, it could be argued that in positing a model that could explain how beliefs are too readily updated, we were inherently failing to explain why the new beliefs themselves – the delusions – are actually tenacious and seemingly immune to new and contradictory evidence.

Moreover, the model had little to say directly about the content of delusions and in particular their focus on the social realm (the fact that they are often about other people and one's relationships to them). Nor did it deal with the affective charge of delusions. After all, why are delusional beliefs so deeply coloured by emotion and mood? Finally, delusions do not occur in isolation. They are frequently accompanied by hallucinations and often

co-occur with negative symptoms such as social withdrawal, apathy and self-neglect. We consider each of these in turn.

The main advance since 2007 involves an appreciation that models of associative learning might pertain not just to animal conditioning and human beliefs but also to perception. This allows us to address many of the earlier model's shortcomings. It entails conceiving of perception not as a passive process of sensory reception, but rather as active analysis by synthesis.

Our perceptual experience comprises not the actual sensory input but rather the most likely (based on prior experience) cause of this input. We do so by exploiting past regularities (encapsulated in our world model). Herman Von Helmholtz called this 'unconscious inference' and made the provocative claim that all perception was a form of controlled hallucination (given how reliant perception is upon these prior regularities rather than raw sense data). Pavlov also appreciated the deep connection between conditioning and perception: 'Evidently what the genius Helmholtz was referring to in unconscious inference, is the mechanism of the conditioned reflex' (Pavlov, 1928).

At the neuro-computational level, these ideas are realised in the hierarchical organisation of the brain. Prior expectations are realised top-down via NMDA and GABA signalling. Any mismatch between those priors and incoming information is signalled bottom-up via AMPA receptors. The impact of a given PE on future predictions is governed by its precision (or inverse variance). This computational motif is recapitulated across successive layers in a hierarchical manner – moving away from raw sensory data, the representations become increasingly complex, multifaceted (Mesulam, 1998) and perhaps distant from the immediate evidence of perception. The priors from the level above impact the signals from the level below. Different neuromodulators (dopamine, acetylcholine, serotonin, oxytocin) may implement the precision of priors and PEs in particular processing hierarchies.

Is altered PE signalling enough to explain delusions?

The PE account of delusions can be considered a one-factor account in that it argues that disruption of a single process (or collection of interrelated processes) may suffice to explain such beliefs. This raises a conflict with neuropsychological models asserting that two factors are necessary for delusions to form (a perceptual disruption and a belief evaluation deficit). We argue that actually this conflict is more apparent than real and arises because the PE account is pitched at a different level of explanation to that of such neuropsychological accounts.

The argument for there being two necessary factors (disruptions) in delusions is compelling. It is often related to monothematic delusions following brain damage, although it has been applied to patients with schizophrenia who have delusions. The main focus of the two-factor theory is Capgras delusion – the belief that a loved one has been replaced by an imposter – and it has been suggested that it can only arise when two things happen. First, a person fails to show the normal autonomic response to a person (leading to a lack of feeling of familiarity, even though there is a strong recognition). This would fit with the unlikely explanation that this person has been replaced, but such an explanation would only be accepted (i.e. the delusion would only form) if the sufferer also had a deficit in their ability to evaluate

and reject improbable beliefs. An advantage of this explanation is that it is based on standard neuropsychological methodology and draws on the existence of a single dissociation across the two factors (the experience of a lack of sense of familiarity and the emergence of the ensuing belief).

The two-factor theory implies a separation between perception and cognition. The essence of the predictive processing approach that underpins the PE model of delusions is that such separation, though it functions well at a descriptive level (some mental phenomena can meaningfully be described as beliefs and some as perceptions) does not require a separation at a deeper level. More specifically, it argues that both perceptions and beliefs are inferences (based upon an integration of upcoming data and existing prior expectation). Though they act at different levels within the hierarchy, the fundamental processes that underpin them and that may cause their perturbation may be common. In our model, hierarchy is key. Factor one (altered experience) could be specified lower in the hierarchy, and factor two (altered belief evaluation) higher up. But crucially, PE and its resolution unite them. Furthermore, enough disruption low down or high up can result in delusions.

It has been argued that the two-factor theory subsumes a PE model, but that the PE model alone is insufficient, since it only explains the abnormal experience but not the altered belief. We disagree with this perspective. While it is compelling to us that PE may offer a good explanation for abnormal experiences, we encourage the question of what the effects might be of a perturbed PE signal as we move up to higher, more abstract and complex levels of inference. There, we would see a disruption in the ability to evaluate and, where appropriate, reject, models of the world. In short, viewed at this deeper level, the same process that accounts for abnormal perception can also account for abnormal belief.

Simply put, the two explanations (two-factor and predictive processing) are cast at different explanatory levels. The two-factor theory is concerned with describing cognitive architectures. Predictive processing aims to unite brain, behavioural and phenomenological data for all delusions (neurological and those that occur in schizophrenia) and, as we argue presently, other psychotic symptoms such as hallucinations and amotivation.

The psychologist Kurt Lewin coined the aphorism 'there is nothing as practical as a good theory'. Since 2007, this expanded hierarchical model has been applied to explain other aspects of delusions, other psychotic symptoms (hallucinations and negative symptoms) and the psychotomimetic effects of other interventions (serotonergic hallucinogens, tetra-hydro-cannabinol [THC]). We now enumerate some of those advances.

Why do delusions persist?

One remarkable feature of delusional beliefs is their elasticity: they expand and morph to include new contradictory data. The person with Capgras, claiming their spouse is an imposter, might respond to other family members who greet the spouse warmly by saying, 'Of course they hug her [the imposter spouse] – they're in on it!' This is difficult to understand in at least two respects. First, the sufferer can often learn about other new things (they don't have an all-encompassing deficit in learning), so why do they seem unable to incorporate new and often strongly contradictory evidence into their beliefs rather than so

tenaciously holding the central delusional belief? Second, and related to this, the PE model of delusional emergence hypothesises an inappropriately enhanced tendency to develop new beliefs. *Prima facie*, this would surely militate against those beliefs being unshakeable.

In trying to answer these questions, we pursued two related lines of thought. The first concerns a disruption in the ways in which newly learned associations become updated. The idea here is that delusions differ from other beliefs in several ways that change their encoding and reconsolidation in memory. First, we suggest, their emergence occurs in response to a world that has become strange and mysterious. There is a puzzle to be solved. Compelling coincidences and seemingly significant events provoke a search for meaning, and the belief that eventually seems to resolve the ambiguity and uncertainty has a powerful function in relieving stress and anxiety. The fact that the belief is often unpleasant does not detract from this function, since it may be easier to bear and respond to a difficult certainty than a nameless and shapeless fear. Given their explanatory utility, they are rehearsed extensively. When delusions are questioned, bringing them to mind may actually serve to reinforce rather than to disrupt the memory (Corlett et al., 2010; Corlett et al., 2009b; Corlett et al., 2013). The idea here is that re-evocation of an association may, under certain circumstances (notably when PE signalling is inherently disrupted), strengthen a memory, even when it is not formally reinforced. We have modelled this process in humans with ketamine. By creating new associations (either appetitive or aversive) and then, a day after, reactivating them (in the absence of the original reinforcer) under ketamine, we observed that the memories became strengthened in comparison to the same procedure under placebo (Corlett et al., 2013). Indeed, the magnitude of this effect correlated with ketamine-induced psychosis and PE brain signal (Corlett et al., 2013). We replicated this memory-enhancing effect in rodents (Honsberger et al., 2015). Conditioned fear memories reactivated under ketamine are subsequently strengthened (Honsberger et al., 2015). This effect was blocked by ifenprodil infusion in the amygdala (a procedure commonly used to block memory destabilisation and updating in preclinical studies of reconsolidation; Honsberger et al., 2015).

The above, empirically supported but currently tentative explanation for how a belief begins to strengthen, even in the absence of objective evidence, can be considered alongside another important phenomenon relating to PE signal. Specifically, there is growing evidence that we are sensitive not just to the magnitude of PE but also to its variability. A high degree of variability can be encoded and leads in time to an adaptation of learning such that a given magnitude of a specific PE produces less updating (Diederer and Schultz, 2015; Preuschoff and Bossaerts, 2007; Preuschoff et al., 2008). In effect, we encode not just surprise or the unpredictability of a single event, but also keep a running tally of how likely we are to be able to predict the current environment, downregulating PE-dependent learning when our best predictions are unable to reduce PE. Thus, one can envisage that in the emergence of psychosis, a person is not just updating beliefs to suppress PE, but also, at a larger timescale, beginning to downregulate the importance of PE. This could lead to the possibility that early beliefs persist and can then remain relatively unchallenged as the person adapts learning rate (learning not to update). Again, this is

speculative, but it perhaps demonstrates the enhanced power to the PE model of taking into account dynamic, evolving effects and, importantly, suggests further experiments.

What may account for the characteristic contents of delusions?

Our original article was primarily concerned with describing a mechanism by which delusions might arise and had little to say about the characteristic content of delusional beliefs. Here, we outline some ideas that may fill this gap in relation to four characteristics of delusional thinking. First, delusions are highly personal in one sense, reflecting the fears and preoccupations of the individual, while at the same time, there are more generic and common aspects to them, and they draw more broadly on the contents of that individual's culture and era (Stompe et al., 2003). Second, delusions usually relate to agents and can often seem to reflect a fundamental alteration in the ability to attribute agency appropriately. Third, related to this, delusions are predominantly social (Bell, 2013; Fineberg and Corlett, 2016). They are often about people and their intentions or goals rather than merely physical entities. Thus, for example, a person may come to view an array of unusual coincidences and sinister experiences as arising from the actions of a persecutor (Kihlstrom and Hoyt, 1988). We note of course that there are some delusions with apparent positive content, and have recently argued that delusions may serve an adaptive function of facilitating continued engagement with an unpredictable environment (Fineberg and Corlett, 2016). Ultimately though, even grandiose delusions can be a source of distress and uncertainty (the person may feel themselves to be responsible for important events, including unpleasant ones, and may see themselves as vulnerable to envious and powerful enemies; Corlett et al., 2007a).

The first characteristic is, in one sense, very straightforward to understand. The delusion is seen as a person's hypothesis about the origins of their perceptual experiences. It is, as Coltheart et al. (2010) have observed, an abductive inference in which data are used to infer their underlying cause. Given that such inference relies on a person's best guess (Pierce, 1931–1958), then it follows that their own prior knowledge and expectation will necessarily determine the content of the emergent belief. And since their own expectations are, *inter alia*, socioculturally determined, there will be a strong overlap between those of the person and the time and culture they inhabit. The shifting nature of delusions across the decades testifies to this (Stompe et al., 2003).

A further idea that emerges from this model is that if the delusion arises to explain uncertainty, it seems feasible that certain features of our environment, being inherently more uncertain, may prove more likely to become the subject of delusional thoughts. Phenomena such as the intentions of others, their hidden goals, the import of their actions and their facial expressions may all be areas that are most likely to change in the face of altered PE (Corlett, 2015).

The question of agency in delusions, and psychosis more generally, is a very interesting one. While delusions are by no means uniquely about agents, they are nonetheless frequently preoccupied with the intentions and actions of other agents, and, in some cases, may explicitly entail a sense that another agent is exerting control over oneself (Frith, 2005a, 2005b). More broadly, other psychotic symptoms that can co-occur with delusions such as

hearing voices have also been attributed to a failure to attribute agency correctly such that one's own inner speech feels as though it has been externally generated (Ford, 2016; Ford and Mathalon, 2005; Ford et al., 2007). This has been framed in terms of a general source-monitoring deficit in schizophrenia (Keefe et al., 1999; Keefe and Kraus, 2009; Kraus et al., 2009).

In fact, there is a compelling body of theoretical literature relating prediction and expectation to the attribution and experience of agency. This comes from observations that even under normal circumstances, it is possible to produce a sense of agency in people who are not in fact the authors of an action and, conversely, to disavow agency for their own actions (Frith, 2005a, 2005b). Both of these phenomena can be produced by altering expectations and external cues, leading to the emerging view that sense of agency emerges from the integration of prior expectations with internal (e.g. proprioceptive) and external cues (Moore et al., 2011a; Moore and Fletcher, 2012). This has been very well-illustrated in the work of Daniel Wegner who shows that both expectations as well as external cues can profoundly alter the degree to which one sees oneself as the cause of events or attributes them to some external agent (Wegner and Wheatley, 1999).

In this sense, attribution of agency has the same status as other abductive inferences characterising delusional thinking (as discussed above), and perturbations within the system, as would be the case with disrupted PE signalling, could fundamentally alter the experience both of one's own agency and of that attributed to events in the outside world. Put in simple terms, one's own actions are characterised partly by the predictability of their consequences, so an action accompanied by an unpredictable consequence is perhaps more likely to originate externally (Frith, 2005a, 2005b).

There is some evidence for the above perspective. Ketamine also enhances intentional binding (the perceived compression in time between action and outcome for deeds for which we feel agency; Moore et al., 2011b). Intentional binding is likewise enhanced in patients with first-episode psychosis (Hauser et al., 2011; Voss et al., 2010). PEs have been implicated in intentional binding effects; binding effects are subject to Kamin blocking (prior learning of action–event associations can block the intentional binding effect). The blocking is weaker in individuals with higher schizotypy scores (Moore et al., 2011a). Furthermore, one's sense of agency for actions can be probed by considering forward modelling comparing predictions to feedback sensations and cancelling what was predicted to guide one's sense of ownership for thoughts and actions. Specifically, we are more likely to feel ourselves to be agents when the consequences of actions are predicted and more likely to assume external agency when they are unpredicted. It has been shown that for self-produced forces, we are likely to cancel out the sensory consequences (which are predicted) such that when trying to match an external force that we have just experienced on our finger by pressing down on the same finger, we overcompensate. The extra force exerted is thought to overcome the cancellation effect of our anticipation prior to performing an action (Shergill et al., 2003, 2005).

Patients with schizophrenia do not overcompensate; they are more accurate (Shergill et al., 2005). Also, more accurate force matching (without predictive over-compensation) correlates with delusion-like ideation in healthy people (Teufel et al., 2010). Taken together, these data confirm that PE-driven inferences are

central to a range of experiences. Ketamine and psychosis similarly perturb these PEs, and those perturbations relate to the severity of endogenous and ketamine induced psychotic symptoms (Moore et al., 2011b).

PEs have been invoked to explain the sense of agency for our actions and ownership for our bodies. Ketamine augments experience of the rubber-hand illusion – the spurious sense of ownership of a prop-hand if the hand is stroked at the same time as one's own hand (Morgan et al., 2011). People on ketamine get the illusion more strongly, and they experience it even in a control condition when the real and rubber hands are stroked asynchronously (Morgan et al., 2011). Patients with schizophrenia (Peled et al., 2003) and chronic ketamine abusers evince the same excessive experience of the illusion in the synchronous and asynchronous conditions (Tang et al., 2015). Activity in the right anterior insula cortex increases to the extent that individuals experience the illusion. Anil Seth and others have argued that the anterior insula is a key nexus for the PE-driven inferences that guide perceptions of bodily ownership and agency (Palmer et al., 2015; Seth, 2013; Seth et al., 2011).

In the remainder of the paper, we attempt to bring other symptoms – specifically, hallucinations and negative symptoms – into the explanatory fold. In so doing, we consider new dimensions of the theory (including its relationship to artificial intelligence and deep learning).

Predictive processing and hallucinations

Can predictive processing theory explain hallucinations? These have been related to aberrant PE signals in primary sensory cortices (Horga et al., 2014). Furthermore, prior theories of hallucinations can be cast in predictive processing terms. For example, auditory verbal hallucinations (AVH; 'voices') have been explained as aberrations of predictive forward models of inner speech (Ford, 2016; Ford and Mathalon, 2005; Ford et al., 2007). This corollary discharge explanation posits that thoughts and inner speech are prosecuted by means of an efferent copy of the motor acts that such inner speech would entail. This efferent copy from cerebellum to parietal cortex is used to cancel the sensorimotor consequences' actions (Blakemore et al., 1999). This mechanism may explain delusions of motor control in which a person experiences their own actions as arising from, and being controlled by, an external agent.

The same theory has been applied to inner speech (Feinberg, 1978) – that we predict the consequences of speaking in our heads (Frith, 2005a, 2005b). Any mismatch leads to the perception of alien agency for the inner speech – it is perceived as external. Accordingly, having subjects open their mouth wide, preventing pre-articulatory motions of the facial muscles, may well attenuate AVHs, perhaps because of effects on these motor predictions (Bick and Kinsbourne, 1987).

Wilkinson (2014) recently cast this model in predictive processing terms. The phenomenology of AVH comes about because the individual infers the best explanation for unexpected inner speech that reaches the threshold for awareness must be an individual talking inside one's head. Despite its intuitive appeal, corollary discharge or efferent copy theory has not fared so well empirically (Ford, 2016). Corollary discharge processing as measured by frontal cortex EEG signals during speech

production and the perception of perturbed versions of one's own speech is impaired in patients who hear voices, but it is likewise impaired in patients with schizophrenia who do not have AVH (Ford, 2016). Furthermore, the severity of corollary discharge impairments does not correlate with AVH severity across subjects (Ford, 2016). Finally, recent work suggests that rather than a failure of prediction that subtends aberrant PE, hallucinations may come about via an undue influence of priors on current processing (Teufel et al., 2015). Patients with an at-risk mental state are more likely to use prior visual information in making visual decisions (Teufel et al., 2015).

Taken together with observations of conditioned hallucinations (sensory experiences without stimuli that can be trained in the lab and to which patients with AVH are more sensitive; Kot and Serper, 2002), the empirical data suggest that hallucinations and delusions, whilst related, may be differently driven by the specification of prior expectations and how they shape subsequent processing. How can this be? In what follows, we focus on elaborations of the PE model that might explain the genesis of hallucinations, as well as the co-occurrence of positive and negative symptoms.

Reliability and uncertainty

To explain how hallucinations, delusions and perhaps even negative symptoms can co-occur despite being related to subtly different aspects of predictive processing, we turn to statistical learning theory, in particular how learning theories deal with attentional allocation. Many different, sometimes opposing, events can be salient. Importantly, both unpredicted events and those that are consistent predictors of important outcomes are potentially salient. These types of events (predictive and unpredicted) seem to have very different relationships to prediction. Statistical models of attentional allocation, unlike many formal learning theory models, allow for this. They simultaneously assess the relevance of stimuli for predicting outcomes (this is considered as their 'reliability'), and the relevance of their uncertainty, or failure to predict outcomes correctly for adjusting the predictions (Dayan et al., 2000). Considering the above, simple example of the apple tree, this might be a highly reliable predictor of the presence of apples, given that it does much better than other trees, but nonetheless a very uncertain one (given the importance of the seasons or weather).

Reliabilities lead to competition between stimuli for making the predictions (Mackintosh, 1975) and so are different from uncertainties in prediction which lead to competition for learning the predictions (Pearce and Hall, 1980). Reliabilities are the statistical account of which stimuli we deem important predictors, whereas uncertainties quantify how well those predictions are known. Reliability and uncertainty have different neurobiological and neurochemical mechanisms (Yu and Dayan, 2005). They have opposing relationships to PE. We argue that ketamine-induced positive and negative symptoms may relate to impaired reliability and enhanced uncertainty processing, respectively.

Both reliability and uncertainty affect attentional allocation. Organisms attend to reliable predictors of salient events (Anderson et al., 2011; Mackintosh, 1975), but on the other hand, PEs elicit surprise. Therefore, reliable events amass fewer PEs and should therefore be attended to less than unreliable predictors (Pearce and Hall, 1980). How can this be? Holland and Schiffrino

(2016) propose that the different ecological demands on attention in learning and action selection may explain the difference. Action decisions are optimised through a bias towards attending to reliable predictors of future states. On the other hand, learning is best served by focusing on the unknown. Pearce and Hall (1980) made this distinction by contrasting controlled versus automatic processing. The learning rate or associability of cues may not be equivalent to that for actions. Attention in learning may be driven by PEs. For action, predictions may dominate. For example, animals can attend to one element of a stimulus array for action guidance (based on its reliability) and another independent feature for learning (based on uncertainty). In the five-choice serial reaction time task, rats can be cued to which action to select. Degradation of those cues (by shortening the cues or decreasing their salience) can increase both errors of commission and omission. The more reliable predictors of which action to commit garner the most attention. On the other hand, those same cues can be learned as probabilistic predictors of food outcomes. Here, uncertain Pavlovian predictors accrue attention and are thus subsequently more associable: they are more readily learned about. In rodents, lesions of medial prefrontal cortex and parietal cortex doubly dissociate reliability-based from uncertainty-based attention (Holland and Schiffino, 2016).

We believe that this distinction may be helpful in reconciling the co-occurrence of delusions, hallucinations and negative symptoms in the same patients. These disparate (seemingly contradictory) symptoms may be differentially reliant on impairments in reliability processing (action selection) or uncertainty processing (learning).

Given the association between negative symptoms and goal-directed action selection, we might predict that patients with schizophrenia might show attenuated responses to cues that guide action selection (impaired reliability estimates). At the same time, patients with delusions might show spurious responses at the time of the outcome (enhanced uncertainty). Data from the monetary incentive delay task support this assertion. Negative symptoms in people with schizophrenia correlate with attenuated striatal responses to action eliciting cues that portend, for example, which action to select and its associated value (Waltz et al., 2010). On the other hand, positive symptoms (specifically delusions) correlate with aberrant PE signals in lateral prefrontal cortex (Corlett et al., 2007b; Waltz et al., 2010), medial prefrontal cortex (Schlagenhauf et al., 2009) and midbrain (Romaniuk et al., 2010) at the time of the outcome.

Predictive processing theory is cast in terms of the hierarchical arrangement of neural systems. Priors are specified top-down and PEs communicated bottom-up, but at each hierarchical level, there may be different relative precisions of predictions and PE (reliability and uncertainty, respectively). And these trade-offs may be different for visual, auditory, motor and other hierarchies. Thus, it may be that a global PE dysfunction (e.g. from disrupted excitatory inhibitory balance in the cortex; Bastos et al., 2012) may impact these hierarchies to different degrees, and to the extent that specific hierarchies (perceptual, motor) are disrupted, different symptoms (positive and negative) might arise.

A brain that is receiving noisy signals can become hungry for the priors that could possibly make sense of that noise and thus resolve uncertainty (Teufel et al., 2015). Thus, when signal from the lowest levels of sensory input are noisy, it may impose precise priors top-down higher in the hierarchy – weighting perception towards expectation (rather than input) in a listening attitude as

Arieti (1974) put it – which produces hallucinations (Hoffman, 2010). Sensory isolation can engender the same effect as bias towards prior expectations that engenders hallucinations (Corlett et al., 2009a).

Overwhelming unreliability of previous actions will likely produce a change in one's higher-level beliefs about the efficacy of one's action, perhaps leading to the conclusion that no behaviours will be effective at reducing uncertainty, so it is best not to act at all, producing negative symptoms (Corlett, 2015).

Other drug models

In 2007, we focused on predictive learning and ketamine. By expanding the model in terms of predictive processing and its role in perception and action, we were able to bring the psychotomimetic effects of other drugs into the explanatory fold. Serotonergic hallucinogens such as LSD induce visual hallucinations (Geyer and Vollenweider, 2008).

They do not, however, induce delusions (Young, 1974). In rats, they enhance glutamatergic responses to sensory stimuli in the locus coeruleus (Rasmussen and Aghajanian, 1986) and frontal cortex (Aghajanian and Marek, 1997), and may actually enhance NMDA signalling (Lambe and Aghajanian, 2006). Excessive AMPA signalling in the absence of NMDA impairment would lead to increased sensory noise in the context of normal priors. This is exactly the context in which we expect hallucinations to arise. The neural correlates of priors and PEs, top-down and bottom-up, have yet to be completely delineated. One theory of the default mode network, a brain circuit engaged when subjects are in a task-free mind-wandering state, is that it reflects PEs to be explained and the process of learned resolutions (Carhart-Harris and Friston, 2010). Rodent data support this idea (Berkes et al., 2011). Serotonergic hallucinogens increase default mode responses in human subjects in a manner that correlates with their psychotomimetic effects (Carhart-Harris et al., 2013). However, behavioural tasks that engage associative learning, perception and belief formation have yet to be examined in this context.

Cannabinoids such as Δ -9-THC also have psychotomimetic effects (D'Souza et al., 2004). The binocular depth inversion illusion (a stereoscopic effect thought to be driven by prior expectations about stimulus curvature) is attenuated by Δ -9-THC (Koethe et al., 2006; Semple et al., 2003). This weakening of top-down influences would suggest that hallucinations should not predominate under Δ -9-THC administration, and this appears to be the case. However, delusion-like ideas do occur (D'Souza et al., 2004).

Amphetamine elevates dopamine levels in the striatum in healthy volunteers and more so in individuals with schizophrenia (Laruelle et al., 2003). A single dose of amphetamine does not induce delusion-like ideas or hallucinations. Rather, elevated mood, grandiose ideas and hyperactivity are more characteristic (Jacobs and Silverstone, 1986). It also increases perceptual acuity of the whole visual field (Fillmore et al., 2005), unlike ketamine, which enhances the salience of discrete and apparently random objects, events and stimuli (Corlett et al., 2007a; Oye et al., 1992). We suggest that this pattern of psychopathology is due to increased precision of both priors and PEs through enhanced dopaminergic (Kegeles et al., 1999; Laruelle et al., 1995) and cholinergic function (Acquas and Fibiger, 1998).

In sum, considering the trade-off between prior experiences and current inputs via their relative precision opens a whole new explanatory scope for the model, both in terms of symptoms other than delusions and interventions other than ketamine (Corlett et al., 2009a).

The future: The emergence of computational psychiatry

It would be remiss not to acknowledge the fact that the PE model of delusions was formulated in the setting of exciting developments in the application of computational models to psychiatric questions, and arrogant to fail to acknowledge that computational psychiatry, though it has come into the spotlight fairly recently (Corlett and Fletcher, 2014; Friston et al., 2014; Montague et al., 2012), has a long history.

Some of the earliest work in artificial intelligence (AI) was rapidly used to explore the genesis of psychosis. For example, symbolic models of natural language were trained to implement simple responses to verbal questions. By altering the model parameters (e.g. its input–output mappings), paranoid responses could be elicited (Colby, 1960). Hoffman and Dobscha (1989) went beyond the language of thought analogy, implementing an artificial neural network model called a Hopfield network that could memorise inputs and give appropriate outputs. By pruning the allowable connections in this model, it produced spurious recall that Hoffman and Dobscha (1989) related to hallucinations and delusions. Before his untimely death, Ralph Hoffman combined Hopfield networks with a modular cognitive architecture to examine story learning and recall. When he increased PE signalling in the network, it began recalling spurious agents in the narratives it produced, inserting itself into those stories in a manner consistent with some delusions (Hoffman et al., 2011). Again, these networks were not hierarchical.

We and others have argued that hierarchical organisation is likely a key organisational principle for cortical processing. This principle is likewise captured in ‘deep learning’ (LeCun et al., 2015) – state of the art AI that involves hierarchical (or ‘deep’) architectures comprised of many Hopfield networks, separated by hidden layers, that learn representations of data without supervision and use reinforcement learning on these representations to guide action selection, beating humans at Atari games (Mnih et al., 2015) and Go (Silver et al., 2016).

In deep learning, each stage in the hierarchy learns to generate or reconstruct the activation patterns in the stage below. One such network, the Deep Boltzmann machine (DBM), utilises both feedforward and feedback processing (Salakhutdinov and Hinton, 2012), which better suits the recurrent processing in brain hierarchies we have described. In a DBM, each hidden layer receives input from a layer below (that conveys bottom-up information), and from a layer above that has learned predict the activity of the layer below. The hidden units learn latent variable representations of the input data. This means a deep network can synthesise representations of input data, even in the absence of such data (Yuille and Kersten, 2006).

Such network behaviour suggests these architectures might implement Helmholtz’s analysis by synthesis and may be particularly helpful in examining the genesis of hallucinations. Indeed, Reichert et al. (2013) have done just that, using a DBM to model the genesis of visual hallucinations in Charles Bonnet Syndrome,

a syndrome occurring in association with visual deficits (e.g. macular degeneration) in which illusory percepts and sometime complex hallucinations occur. As in Yu and Dayan (2002), acetylcholine can be used as a model parameter to set the balance in between feedforward and feedback flow of information in perception. In a version perhaps most relevant to our concerns, Dayan and Hinton describe a Helmholtz machine that uses acetylcholine to trade off priors and PEs and minimise free energy (Dayan et al., 1995; Hinton and Dayan, 1996). This formulation has much in common with the Kalman Filter approach to predictive learning that can embody statistical reliabilities and uncertainty and employs noradrenaline and acetylcholine to do as such (Dayan et al., 2000). These neurotransmitters may be opponent (Yu and Dayan, 2005) and have recently been implicated in the genesis of hallucinations (Collerton et al., 2005; Geddes et al., 2016). Kersten et al. (2004) further characterise a generative model as ‘strong’ if samples can be produced from it that consistently look like the data it learned from (Kersten et al., 2004). Reichert et al. (2013) were able to read out which particular stimuli were being hallucinated, which is a significant advance on earlier work with Hopfield networks.

Summary and conclusion

In conclusion, our 2007 model offered a rudimentary framework for thinking about how delusions may emerge, one that linked the experiences, via a cognitive model of associative learning, to neural processes. Here, we have shown how an extension of the model – one that brings in related parameters (precision, reliability, certainty) and a more dynamic view of how PE learning might change as PE signal evolves – offers new breadths of explanation. A key next step is to turn some of these insights into practical benefits for the patients who, along with their families and friends, may suffer greatly with their experiences. We believe that the elucidation of mechanisms by which these experiences arise is a necessary prelude to a comprehensive and precise diagnostic system, as well as to the development of individually targeted interventions.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Connecticut Mental Health Center (CMHC) and Connecticut State Department of Mental Health and Addiction Services (DMHAS). PRC was funded by an IMHRO/Janssen Rising Star Translational Research Award and CTSA Grant Number UL1 TR000142, R01MH067073 from the National Center for Research Resources (NCRR) and the National Center for Advancing Translational Science (NCATS), components of the National Institutes of Health (NIH), and NIH roadmap for Medical Research. The contents of this work are solely the responsibility of the authors and do not necessarily represent the official view of NIH or the CMHC/DMHAS. PCF is supported by the Wellcome Trust and the Bernard Wolfe Health Neuroscience Fund.

References

- Acquas E and Fibiger HC (1998) Dopaminergic regulation of striatal acetylcholine release: the critical role of acetylcholinesterase inhibition. *J Neurochem* 70: 1088–1093.

- Adams RA, Stephan KE, Brown HR, et al. (2013) The computational anatomy of psychosis. *Front Psychiatry* 4: 47.
- Aghajanian GK and Marek GJ (1997) Serotonin induces excitatory postsynaptic potentials in apical dendrites of neocortical pyramidal cells. *Neuropharmacology* 36: 589–599.
- Anderson BA, Laurent PA and Yantis S (2011) Value-driven attentional capture. *Proc Natl Acad Sci U S A* 108: 10367–10371.
- Arieti S (1974) An overview of schizophrenia from a predominantly psychological approach. *Am J Psychiatry* 131: 241–249.
- Bastos AM, Usrey WM, Adams RA, et al. (2012) Canonical microcircuits for predictive coding. *Neuron* 76: 695–711.
- Bell V (2013) A community of one: social cognition and auditory verbal hallucinations. *PLoS Biol* 11: e1001723.
- Berkes P, Orban G, Lengyel M, et al. (2011) Spontaneous cortical activity reveals hallmarks of an optimal internal model of the environment. *Science* 331: 83–87.
- Bick PA and Kinsbourne M (1987) Auditory hallucinations and subvocal speech in schizophrenic patients. *Am J Psychiatry* 144: 222–225.
- Blakemore SJ, Wolpert DM and Frith CD (1999) The cerebellum contributes to somatosensory cortical activity during self-produced tactile stimulation. *Neuroimage* 10: 448–459.
- Bleuler E (1908) Die Prognose der Dementia praecox (Schizophreniegruppe). *Allgem Zeit Psychiat Psychisch Gerich Med* 65: 436–464.
- Carhart-Harris RL and Friston KJ (2010) The default-mode, ego-functions and free-energy: a neurobiological account of Freudian ideas. *Brain* 133: 1265–1283.
- Carhart-Harris RL, Leech R, Erritzoe D, et al. (2013) Functional connectivity measures after psilocybin inform a novel hypothesis of early psychosis. *Schizophr Bull* 39: 1343–1351.
- Clark A (2013) Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav Brain Sci* 36: 181–204.
- Colby KM (1960) *Artificial Paranoia: A Computer Simulation of Paranoid Processes*. New York: Pergamon.
- Collerton D, Perry E and McKeith I (2005) Why people see things that are not there: a novel Perception and Attention Deficit model for recurrent complex visual hallucinations. *Behav Brain Sci* 28: 737–757.
- Coltheart M and Davies M (2000) *Pathologies of Belief*. Oxford: Blackwell.
- Coltheart M, Menzies P and Sutton J (2010) Abductive inference and delusional belief. *Cogn Neuropsychiatry* 15: 261–287.
- Conant RC and Ashby WR (1970) Every good regulator of a system must be a model of that system. *Int J Systems Sci* 1: 89–97.
- Corlett PR (2015) Answering some phenomenal challenges to the prediction error model of delusions. *World Psychiatry* 14: 181–183.
- Corlett PR, Aitken MR, Dickinson A, et al. (2004) Prediction error during retrospective reevaluation of causal associations in humans: fMRI evidence in favor of an associative model of learning. *Neuron* 44: 877–888.
- Corlett PR, Cambridge V, Gardner JM, et al. (2013) Ketamine effects on memory reconsolidation favor a learning model of delusions. *PLoS One* 8: e65088.
- Corlett PR and Fletcher PC (2012) The neurobiology of schizotypy: fronto-striatal prediction error signal correlates with delusion-like beliefs in healthy people. *Neuropsychologia* 50: 3612–3620.
- Corlett PR and Fletcher PC (2014) Computational psychiatry: a Rosetta Stone linking the brain to mental illness. *Lancet Psychiatry* 1: 399–402.
- Corlett PR and Fletcher PC (2015) Delusions and prediction error: clarifying the roles of behavioural and brain responses. *Cogn Neuropsychiatry* 20: 95–105.
- Corlett PR, Frith CD and Fletcher PC (2009a) From drugs to deprivation: a Bayesian framework for understanding models of psychosis. *Psychopharmacology (Berl)* 206: 515–530.
- Corlett PR, Honey GD, Aitken MR, et al. (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 PR, Honey GD and Fletcher PC (2007a) From prediction error to psychosis: ketamine as a pharmacological model of delusions. *J Psychopharmacol* 21: 238–252.
- Corlett PR, Krystal JH, Taylor JR, et al. (2009b) Why do delusions persist? *Front Hum Neurosci* 3: 12.
- Corlett PR, Murray GK, Honey GD, et al. (2007b) Disrupted prediction-error signal in psychosis: evidence for an associative account of delusions. *Brain* 130: 2387–2400.
- Corlett PR, Taylor JR, Wang XJ, et al. (2010) Toward a neurobiology of delusions. *Prog Neurobiol* 92: 345–369.
- Dayan P, Hinton GE, Neal RM, et al. (1995) The Helmholtz machine. *Neural Comput* 7: 889–904.
- Dayan P, Kakade S and Montague PR (2000) Learning and selective attention. *Nat Neurosci* 3: 1218–1223.
- Dickinson A (2001) The 28th Bartlett Memorial Lecture. Causal learning: an associative analysis. *Q J Exp Psychol B* 54: 3–25.
- Diederer KM and Schultz W (2015) Scaling prediction errors to reward variability benefits error-driven learning in humans. *J Neurophysiol* 114: 1628–1640.
- D'Souza DC, Perry E, MacDougall L, et al. (2004) The psychotomimetic effects of intravenous delta-9-tetrahydrocannabinol in healthy individuals: implications for psychosis. *Neuropsychopharmacology* 29: 1558–1572.
- Feinberg I (1978) Efference copy and corollary discharge: implications for thinking and its disorders. *Schizophr Bull* 4: 636–640.
- Fillmore MT, Rush CR and Abroms BD (2005) d-Amphetamine-induced enhancement of inhibitory mechanisms involved in visual search. *Exp Clin Psychopharmacol* 13: 200–208.
- Fineberg SK and Corlett PR (2016) The doxastic shear pin: delusions as errors of learning and memory. *Cogn Neuropsychiatry* 21: 73–89.
- Fletcher PC, Anderson JM, Shanks DR, et al. (2001) Responses of human frontal cortex to surprising events are predicted by formal associative learning theory. *Nat Neurosci* 4: 1043–1048.
- Fletcher PC and Frith CD (2009) Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia. *Nat Rev Neurosci* 10: 48–58.
- Ford JM (2016) Studying auditory verbal hallucinations using the RDoC framework. *Psychophysiology* 53: 298–304.
- Ford JM and Mathalon DH (2005) Corollary discharge dysfunction in schizophrenia: can it explain auditory hallucinations? *Int J Psychophysiol* 58: 179–189.
- Ford JM, Roach BJ, Faustman WO, et al. (2007) Synch before you speak: auditory hallucinations in schizophrenia. *Am J Psychiatry* 164: 458–466.
- Friston KJ, Stephan KE, Montague PR, et al. (2014) Computational psychiatry: the brain as a phantastic organ. *Lancet Psychiatry* 1: 148–158.
- Frith C (2005a) The neural basis of hallucinations and delusions. *C R Biol* 328: 169–175.
- Frith C (2005b) The self in action: lessons from delusions of control. *Conscious Cogn* 14: 752–770.
- Geddes MR, Tie Y, Gabrieli JD, et al. (2016) Altered functional connectivity in lesional peduncular hallucinosis with REM sleep behavior disorder. *Cortex* 74: 96–106.
- Geyer MA and Vollenweider FX (2008) Serotonin research: contributions to understanding psychoses. *Trends Pharmacol Sci* 29: 445–453.
- Gradin VB, Kumar P, Waiter G, et al. (2011) Expected value and prediction error abnormalities in depression and schizophrenia. *Brain* 134: 1751–1764.
- Gray JA, Feldon J, Rawlins JNP, et al. (1991) The neuropsychology of schizophrenia. *Behav Brain Sci* 14: 1–84.
- Griffiths O, Langdon R, Le Pelley ME, et al. (2014) Delusions and prediction error: re-examining the behavioural evidence for disrupted

- error signalling in delusion formation. *Cogn Neuropsychiatry* 19: 439–467.
- Hartley D (1749/1976) *Observations on Man, His Frame, His Duty, and His Expectations*. New York: Delmar.
- Hauser M, Moore JW, de Millas W, et al. (2011) Sense of agency is altered in patients with a putative psychotic prodrome. *Schizophr Res* 126: 20–27.
- Hinton GE and Dayan P (1996) Varieties of Helmholtz machine. *Neural Netw* 9: 1385–1403.
- Hoffman RE (2010) Revisiting Arieti's 'listening attitude' and hallucinated voices. *Schizophr Bull* 36: 440–442.
- Hoffman RE and Dobscha SK (1989) Cortical pruning and the development of schizophrenia: a computer model. *Schizophr Bull* 15: 477–490.
- Hoffman RE, Grasemann U, Gueorguieva R, et al. (2011) Using computational patients to evaluate illness mechanisms in schizophrenia. *Biol Psychiatry* 69: 997–1005.
- Holland PC and Schiffrino FL (2016) Mini-review: prediction errors, attention and associative learning. *Neurobiol Learn Mem* 131: 207–215.
- Honsberger MJ, Taylor JR and Corlett PR (2015) Memories reactivated under ketamine are subsequently stronger: a potential pre-clinical behavioral model of psychosis. *Schizophr Res* 164: 227–233.
- Horga G, Schatz KC, Abi-Dargham A, et al. (2014) Deficits in predictive coding underlie hallucinations in schizophrenia. *J Neurosci* 34: 8072–8082.
- Jacobs D and Silverstone T (1986) Dextroamphetamine-induced arousal in human subjects as a model for mania. *Psychol Med* 16: 323–329.
- Keefe RS, Arnold MC, Bayen UJ, et al. (1999) Source monitoring deficits in patients with schizophrenia; a multinomial modelling analysis. *Psychol Med* 29: 903–914.
- Keefe RS and Kraus MS (2009) Measuring memory-prediction errors and their consequences in youth at risk for schizophrenia. *Ann Acad Med Singapore* 38: 414–416.
- Kegeles LS, Abi-Dargham A, Zea-Ponce Y, et al. (2000) Modulation of amphetamine-induced striatal dopamine release by ketamine in humans: implications for schizophrenia. *Biol Psychiatry* 48: 627–640.
- Kegeles LS, Zea-Ponce Y, Abi-Dargham A, et al. (1999) Stability of [¹²³I]IBZM SPECT measurement of amphetamine-induced striatal dopamine release in humans. *Synapse* 31: 302–308.
- Kersten D, Mamassian P and Yuille A (2004) Object perception as Bayesian inference. *Annu Rev Psychol* 55: 271–304.
- Kihlstrom JF and Hoyt IP (1988) Hypnosis and the psychology of delusions. In: Oltmanns TF and Maher BA (eds) *Delusional Beliefs*. New York: John Wiley.
- Koethe D, Gerth CW, Neatby MA, et al. (2006) Disturbances of visual information processing in early states of psychosis and experimental delta-9-tetrahydrocannabinol altered states of consciousness. *Schizophr Res* 88: 142–150.
- Kot T and Serper M (2002) Increased susceptibility to auditory conditioning in hallucinating schizophrenic patients: a preliminary investigation. *J Nerv Ment Dis* 190: 282–288.
- Kraus MS, Keefe RS and Krishnan RK (2009) Memory-prediction errors and their consequences in schizophrenia. *Neuropsychol Rev* 19: 336–352.
- Krystal JH, Karper LP, Seibyl JP, et al. (1994) Subanesthetic effects of the noncompetitive NMDA antagonist, ketamine, in humans. Psychotomimetic, perceptual, cognitive, and neuroendocrine responses. *Arch Gen Psychiatry* 51: 199–214.
- Lambe EK and Aghajanian GK (2006) Hallucinogen-induced UP states in the brain slice of rat prefrontal cortex: role of glutamate spillover and NR2B-NMDA receptors. *Neuropsychopharmacology* 31: 1682–1689.
- Laruelle M, Abi-Dargham A, van Dyck CH, et al. (1995) SPECT imaging of striatal dopamine release after amphetamine challenge. *J Nucl Med* 36: 1182–1190.
- Laruelle M, Kegeles LS and Abi-Dargham A (2003) Glutamate, dopamine, and schizophrenia: from pathophysiology to treatment. *Ann N Y Acad Sci* 1003: 138–158.
- Lavin A, Nogueira L, Lapish CC, et al. (2005) Mesocortical dopamine neurons operate in distinct temporal domains using multimodal signaling. *J Neurosci* 25: 5013–5023.
- LeCun Y, Bengio Y and Hinton G (2015) Deep learning. *Nature* 521: 436–444.
- Mackintosh NJ (1975) A theory of attention: variations in the associability of stimuli with reinforcement. *Psychol Rev* 82: 276–298.
- Mesulam MM (1998) From sensation to cognition. *Brain* 121: 1013–1052.
- Miller R (1976) Schizophrenic psychology, associative learning and the role of forebrain dopamine. *Med Hypotheses* 2: 203–211.
- Mnih V, Kavukcuoglu K, Silver D, et al. (2015) Human-level control through deep reinforcement learning. *Nature* 518: 529–533.
- Montague PR, Dolan RJ, Friston KJ, et al. (2012) Computational psychiatry. *Trends Cogn Sci* 16: 72–80.
- Moore JW, Dickinson A and Fletcher PC (2011a) Sense of agency, associative learning, and schizotypy. *Conscious Cogn* 20: 792–800.
- Moore JW and Fletcher PC (2012) Sense of agency in health and disease: a review of cue integration approaches. *Conscious Cogn* 21: 59–68.
- Moore JW, Turner DC, Corlett PR, et al. (2011b) Ketamine administration in healthy volunteers reproduces aberrant agency experiences associated with schizophrenia. *Cogn Neuropsychiatry* 16: 264–381.
- Morgan HL, Turner DC, Corlett PR, et al. (2011) Exploring the impact of ketamine on the experience of illusory body ownership. *Biol Psychiatry* 69: 35–41.
- Murray GK, Corlett PR, Clark L, et al. (2008) Substantia nigra/ventral tegmental reward prediction error disruption in psychosis. *Mol Psychiatry* 13: 239, 267–276.
- Oye I, Paulsen O and 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.
- Palmer CJ, Seth AK and Hohwy J (2015) The felt presence of other minds: predictive processing, counterfactual predictions, and mentalising in autism. *Conscious Cogn* 36: 376–389.
- Pavlov IP (1928) Natural science and the brain. In: Pavlov IP (ed) *Lectures on Conditioned Reflexes*. New York, pp. 126–127.
- Pearce JM and 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 AB, et al. (2003) Somatosensory evoked potentials during a rubber-hand illusion in schizophrenia. *Schizophr Res* 64: 157–163.
- Pierce CS (1931–1958) *Collected Papers of Charles Sanders Peirce*. Cambridge, MA: Harvard University Press.
- Pomarol-Clotet E, Honey GD, Murray GK, et al. (2006) Psychological effects of ketamine in healthy volunteers. Phenomenological study. *Br J Psychiatry* 189: 173–179.
- Preuschoff K and Bossaerts P (2007) Adding prediction risk to the theory of reward learning. *Ann N Y Acad Sci* 1104: 135–146.
- Preuschoff K, Quartz SR and Bossaerts P (2008) Human insula activation reflects risk prediction errors as well as risk. *J Neurosci* 28: 2745–2752.
- Rasmussen K and Aghajanian GK (1986) Effect of hallucinogens on spontaneous and sensory-evoked locus coeruleus unit activity in the rat: reversal by selective 5-HT₂ antagonists. *Brain Res* 385: 395–400.
- Reichert DP, Series P and Storkey AJ (2013) Charles Bonnet syndrome: evidence for a generative model in the cortex? *PLoS Comput Biol* 9: e1003134.
- Romaniuk L, Honey GD, King JR, et al. (2010) Midbrain activation during Pavlovian conditioning and delusional symptoms in schizophrenia. *Arch Gen Psychiatry* 67: 1246–1254.

- Salakhutdinov R and Hinton G (2012) An efficient learning procedure for deep Boltzmann machines. *Neural Comput* 24: 1967–2006.
- Schlagenhauf F, Sterzer P, Schmack K, et al. (2009) Reward feedback alterations in unmedicated schizophrenia patients: relevance for delusions. *Biol Psychiatry* 65: 1032–1039.
- Schultz W and Dickinson A (2000) Neuronal coding of prediction errors. *Annu Rev Neurosci* 23: 473–500.
- Semple DM, Ramsden F and McIntosh AM (2003) Reduced binocular depth inversion in regular cannabis users. *Pharmacol Biochem Behav* 75: 789–793.
- Seth AK (2013) Interoceptive inference, emotion, and the embodied self. *Trends Cogn Sci* 17: 565–573.
- Seth AK, Suzuki K and Critchley HD (2011) An interoceptive predictive coding model of conscious presence. *Front Psychol* 2: 395.
- Shergill SS, Bays PM, Frith CD, et al. (2003) Two eyes for an eye: the neuroscience of force escalation. *Science* 301: 187.
- Shergill SS, Samson G, Bays PM, et al. (2005) Evidence for sensory prediction deficits in schizophrenia. *Am J Psychiatry* 162: 2384–2386.
- Silver D, Huang A, Maddison CJ, et al. (2016) Mastering the game of Go with deep neural networks and tree search. *Nature* 529: 484–489.
- Stompe T, Ortwein-Swoboda G, Ritter K, et al. (2003) Old wine in new bottles? Stability and plasticity of the contents of schizophrenic delusions. *Psychopathology* 36: 6–12.
- Sutton RS and Barto AG (1998) *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press.
- Tang J, Morgan HL, Liao Y, et al. (2015) Chronic administration of ketamine mimics the perturbed sense of body ownership associated with schizophrenia. *Psychopharmacology (Berl)* 232: 1515–1526.
- Teufel C, Kingdon A, Ingram JN, et al. (2010) Deficits in sensory prediction are related to delusional ideation in healthy individuals. *Neuropsychologia* 48: 4169–4172.
- Teufel C, Subramaniam N, Doblér V, et al. (2015) Shift toward prior knowledge confers a perceptual advantage in early psychosis and psychosis-prone healthy individuals. *Proc Natl Acad Sci U S A* 112: 13401–13406.
- Turner DC, Aitken MR, Shanks DR, et al. (2004) The role of the lateral frontal cortex in causal associative learning: exploring preventative and super-learning. *Cereb Cortex* 14: 872–880.
- Voss M, Moore J, Hauser M, et al. (2010) Altered awareness of action in schizophrenia: a specific deficit in predicting action consequences. *Brain* 133: 3104–3112.
- Waltz JA, Schweitzer JB, Ross TJ, et al. (2010) Abnormal responses to monetary outcomes in cortex, but not in the basal ganglia, in schizophrenia. *Neuropsychopharmacology* 35: 2427–2439.
- Wegner DM and Wheatley T (1999) Apparent mental causation. Sources of the experience of will. *Am Psychol* 54: 480–492.
- Widrow B and Hoff ME Jr (1960) Adaptive switching circuits. *IRE Wescon Conv Rec* 4: 96–104.
- Wilkinson S (2014) Accounting for the phenomenology and varieties of auditory verbal hallucination within a predictive processing framework. *Conscious Cogn* 30: 142–155.
- Young BG (1974) A phenomenological comparison of LSD and schizophrenic states. *Br J Psychiatry* 124: 64–74.
- Yu AJ and Dayan P (2002) Acetylcholine in cortical inference. *Neural Netw* 15: 719–730.
- Yu AJ and Dayan P (2005) Uncertainty, neuromodulation, and attention. *Neuron* 46: 681–692.
- Yuille A and Kersten D (2006) Vision as Bayesian inference: analysis by synthesis? *Trends Cogn Sci* 10: 301–308.

From prediction error to psychosis: ketamine as a pharmacological model of delusions

Journal of Psychopharmacology
21(3) (2007) 238–252
© 2007 British Association
for Psychopharmacology
ISSN 0269-8811
SAGE Publications Ltd,
Los Angeles, London,
New Delhi and Singapore
10.1177/0269881107077716

P. R. Corlett *Brain Mapping Unit, Department of Psychiatry, University of Cambridge, School of Clinical Medicine, Addenbrooke's Hospital, Cambridge, UK.*

G. D. Honey *Brain Mapping Unit, Department of Psychiatry, University of Cambridge, School of Clinical Medicine, Addenbrooke's Hospital, Cambridge, UK.*

P. C. Fletcher *Brain Mapping Unit, Department of Psychiatry, University of Cambridge, School of Clinical Medicine, Addenbrooke's Hospital, Cambridge, UK.*

Abstract

Recent cognitive neuropsychiatric models of psychosis emphasize the role of attentional disturbances and inappropriate incentive learning in the development of delusions. These models highlight a pre-psychotic period in which the patient experiences perceptual and attentional disruptions. Irrelevant details and numerous associations between stimuli, thoughts and percepts are imbued with inappropriate significance and the attempt to rationalize and account for these bizarre experiences results in the formation of delusions. The present paper discusses delusion formation in terms of basic associative learning processes. Such processes are driven by prediction error signals. Prediction error refers to mismatches between

an organism's expectation in a given environment and what actually happens and it is signalled by both dopaminergic and glutamatergic mechanisms. Disruption of these neurobiological systems may underlie delusion formation. We review similarities between acute psychosis and the psychotic state induced by the NMDA receptor antagonist drug ketamine, which impacts upon both dopaminergic and glutamatergic function. We conclude by suggesting that ketamine may provide an appropriate model to investigate the formative stages of symptom evolution in schizophrenia, and thereby provide a window into the earliest and otherwise inaccessible aspects of the disease process.

Introduction

In this paper, we review a cognitive neuropsychiatric account of delusion formation that highlights the transition from disrupted visual and auditory perception, through attentional capture to delusional ideation. We suggest that this transition arises from inappropriate prediction error signaling. Prediction errors are a mismatch between expectation and occurrence and are used by organisms as teaching signals. In general, we try to minimize the error and thus improve our understanding of, and ability to predict, the environment (Dickinson, 2001; Rescorla and Wagner, 1972; Schultz and Dickinson, 2000). According to prediction error models of delusion formation, the experience of mismatch when there is none drives an individual to invent bizarre causal structures to explain away their experiences, these are manifest clinically as delusions (Gray, 1993, 1998b, 2004; Gray *et al.*, 1995; Hemsley, 1992, 1993, 1994, 2005a, b).

We consider evidence relating to these models, taken from autobiographical accounts and structured interviews of schizophrenic patients. This is corroborated by experimental evidence showing that impaired associative learning may have a role in symptom formation. We consider this evidence in light of the extensive animal research implicating the mesolimbic dopaminergic system, and more recently, glutamatergic mechanisms in prediction error signaling together with functional imaging data in humans also supporting the involvement of frontostriatal regions in prediction error processing.

Having set out the evidence supporting abnormal prediction error dependent learning in delusion formation, and speculated on its neurophysiological basis, we consider the extent to which ketamine may provide a means by which to explore the nature of delusions. We suggest that a ketamine-induced disruption of fronto-striatal dopamine/glutamate function leads to characteristic psychopathology via aberrant prediction error-based causal learning.

We begin by reviewing the fundamental tenets of associative learning theory and considering the impact of their disruption on learning, attention and belief formation.

Prediction error, associative learning and psychosis

Formal associative learning theories posit that prediction errors are used by organisms as teaching signals (Dickinson, 2001; Rescorla and Wagner, 1972; Schultz and Dickinson, 2000). These signals have both direct and indirect consequences for learning (see Fig. 1). By decreasing the magnitude of these errors, the organism improves its ability to predict relationships in its environment and thus adaptively increase its contact with rewards and decrease its contact with punishments (Dickinson, 2001; Rescorla and Wagner, 1972; Schultz and Dickinson, 2000). In terms of predicting the environment, minimization of prediction errors directly strengthens the associative relationship between a predictive cue and a rewarding outcome (Dickinson, 2001; Rescorla and Wagner, 1972; Schultz and Dickinson, 2000). In addition, prediction error signals influence learning indirectly, by altering the attention allocated to stimuli (greater attention is assigned to stimuli that have occurred in unpredictable environments). The more attention paid to a stimulus, the more readily the organism will associate it with a particular outcome in the environment (Fiorillo *et al.*, 2003, 2005; Grossberg, 1982; Pearce and Hall, 1980).

In brief, this view of associative learning intimately links prediction error signal, association formation and attentional allocation. We next consider the neuroscientific basis for this linkage before exploring the possible consequences of its disruption as a precipitant of psychotic symptoms.

Dopamine neurons in the ventral-tegmental area (VTA) of the macaque have been shown to code a reward prediction error. Their firing patterns are consistent with those predicted by formal associative learning theory (Rescorla and Wagner, 1972; Waelti *et al.*, 2001). Extracellular recordings from midbrain dopamine neurons reveal initial phasic activity in response to unpredicted reward delivery (Hollerman and Schultz, 1998; Ljungberg *et al.*, 1992; Romo and Schultz, 1990; Schultz, 1998a, b; Schultz *et al.*, 1997). The neurons gradually lose this response as rewards become predicted (Hollerman and Schultz, 1998; Hollerman *et al.*, 1998; Schultz *et al.*, 1993a, b, 1997). As the organism learns that certain stimuli predict the delivery of reward, those stimuli, rather than the reward itself, begin to evoke this activity (Schultz *et al.*, 1993a, b, 1997). Further work has shown that the dopamine neuron response to predictive stimuli is governed by the occurrence of a reward prediction error, rather than simply by the presence a stimulus-reward association (Waelti *et al.*, 2001). This phasic signal coding prediction error appears to be accompanied, and possibly complemented, by a tonic dopaminergic signal coding uncertainty (Fiorillo *et al.*, 2003, 2005; Schultz *et al.*, 1993a, b, 1997).

The neuroscientific evidence implicating dopamine in both prediction error and uncertainty concurs with formal learning theories suggesting that a 'learned-uncertainty' about stimuli is important to allocation of attention (Dayan *et al.*, 2000; Grossberg, 1982; Pearce and Hall, 1980; Yu and Dayan, 2002; Yu and Dayan, 2005). That is, stimuli associated with a prediction error on one trial (i.e. their

relationship with reward is uncertain) tend to receive more attention and hence, more is learned about them on the following trial (Grossberg, 1982; Mackintosh, 1975; Pearce and Hall, 1980). This may be mediated by an interaction between phasic prediction error firing and ramping tonic activity in dopamine neurons (Fiorillo *et al.*, 2003, 2005). Alternatively, the forebrain cholinergic system may code the allocation of attention (Chiba *et al.*, 1995; Everitt and Robbins, 1997; Sarter and Bruno, 1999; Yu and Dayan, 2002, 2005). Intriguingly, there is some evidence that acetylcholine release and cortical representation of stimuli may be sculpted by dopaminergic prediction error signals from the VTA (Bao *et al.*, 2001, 2003).

Functional neuroimaging studies of human subjects have provided further support for the importance of fronto-striatal systems in reward-based (Berns *et al.*, 2001; Dreher *et al.*, 2005; McClure *et al.*, 2003a; O'Doherty *et al.*, 2003; Tobler *et al.*, 2005) and non reward-based (Aron *et al.*, 2004; Corlett *et al.*, 2004; Rodriguez *et al.*, 2005) associative learning. Prediction errors are also important for more general cognitive control: that is, the process by which thoughts, plans and behaviour are organised in pursuit of a desired goal (Montague *et al.*, 2004). Influential models of cognitive control propose that the simple associations between predictive cues, behavioural responses and desired outcomes are used to learn and control the sequence of actions required to achieve a goal (Montague *et al.*, 2004). The prefrontal cortex is believed to maintain a representation of the current goals of the organism (Miller and Cohen, 2001), as well as playing a key role in the acquisition of the conditional associations that are used to guide behaviour towards that goal, 'the rules of the game' (Fuster, 1985; Miller and Cohen, 2001). For truly flexible behaviour, the goal representation must be changed or updated in light of novel information. This may be achieved by prediction error signals from the VTA (Braver *et al.*, 1999; Braver and Cohen, 1999; O'Reilly *et al.*, 1999). In the absence of a phasic signal from the VTA, the prefrontal cortex maintains its representation of the current goal, however, when afferent stimuli induce a phasic dopamine response from VTA neurons, the prefrontal gate is 'opened', allowing updating of the information maintained and hence the associations that are driving goal-directed behaviour. Goal directed behaviour fails when the gate is opened by behaviourally irrelevant stimuli, inducing distractibility.

There is therefore theoretical and empirical support for the relationship between prediction error, learning, attention and control, as well as for the importance of the mesolimbic dopamine system in these processes. In the next section we examine the possibility that deficits in this system might be related to the formation of delusions through abnormal prediction error firing. Such an abnormality might lead to the allocation of attention to inappropriate stimuli and the formation of inappropriate associations between stimuli, thoughts and percepts.

Associative learning deficits as a basis for delusions in schizophrenia

The link between dopamine, associative learning and schizophrenic symptoms was first proposed by Miller, who outlined how

an endogenous dopamine disruption in the striatum could lead to symptoms such as delusions and thought disorder via disrupted associative learning. Miller suggested that delusions could be explained as a lowering of the level of significance required to accept a conclusion (Miller, 1976). Based on rodent lesion work, the basal ganglia were implicated in the learning of associations between stimuli and outcomes in the environment (Mitcham and Thomas, 1972) via a dopamine dependent process (Fibiger *et al.*, 1974). Miller's contention was that dopamine overactivity in the basal ganglia lowered the threshold for concluding that an association existed between two entities (e.g. external stimuli and events). Delusions then, are erroneous conclusions that two unrelated stimuli or events are actually related (Miller, 1976). The theory was later elaborated to encompass attentional disturbances, due to disrupted dopamine firing, leading to the allocation of attention to irrelevant stimuli (Miller, 1989, 1993). A consequence of such attentional capture is the perception of (inappropriate) relatedness between the

stimuli that capture attention and other stimuli and events that the patient experiences (Miller, 1989, 1993).

Hemsley (1992, 1993, 1994, 1996, 1998, 2005a, b), Gray (1993, 1995, 1998a, b, 2004) and Gray *et al.* (1995) emphasized the role of attentional and perceptual disturbance in the development of psychotic symptoms. These models were influenced by Broadbent's attentional filter hypothesis, which posited an attentional 'pigeonholing' mechanism that scheduled behaviorally relevant stimuli over irrelevant stimuli using expectancies and context (Broadbent, 1958). This led to a renewed interest in the interactions between perceptual and cognitive abnormalities in schizophrenia (Arieti, 1955, 1974; Berze, 1914; Conrad, 1958; Maher, 1974, 1988; Matussek, 1952). Central to the Gray/Hemsley model is a comparator which brings together 'the current state of the organism's perceptual world with a predicted state' (Gray, 1993). If a mismatch occurs between expected and actual perception, then the current motor program is interrupted and attention is allocated to the stimuli in question. There

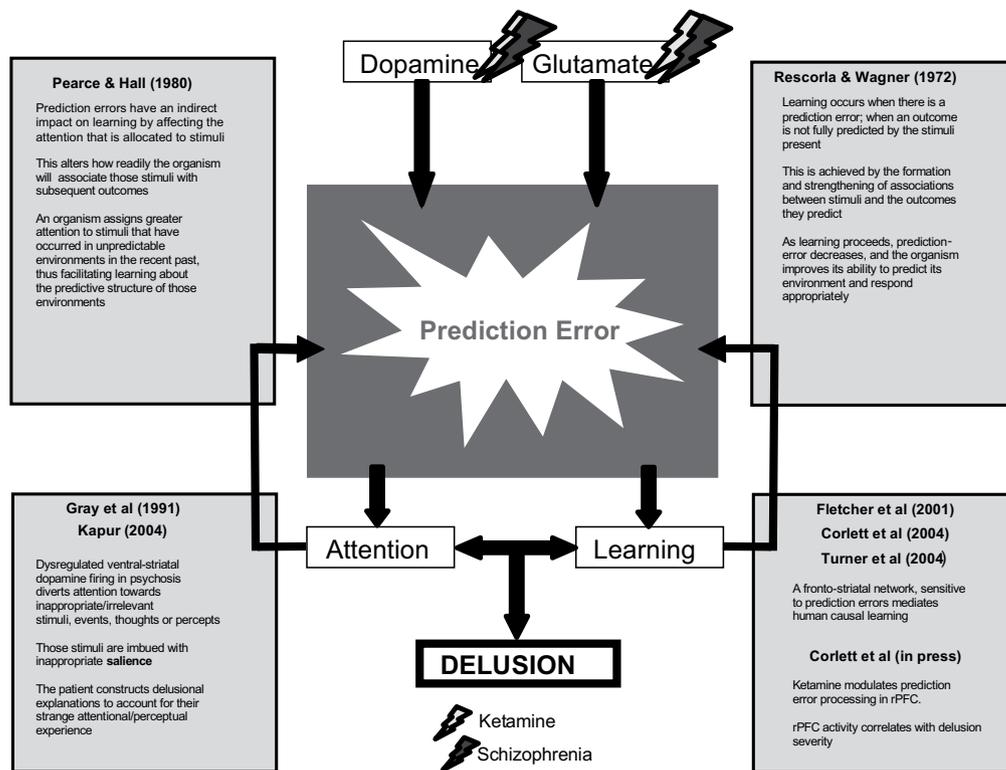


Figure 1 From prediction errors to delusions. The mismatch between expectancy and actual occurrence gives rise to a prediction error. These signals are used by organisms to guide behaviour; they drive learning and the allocation of attention to important stimuli in the environment. Noise in the system that generates prediction errors may be responsible for some of the symptoms of psychosis, notably delusions. Neurochemically these mismatch signals are coded by phasic dopamine activity in the midbrain which is under the regulatory influence of glutamate from the prefrontal cortex. We propose that ketamine provides a useful model psychosis in healthy volunteers (because of its impact upon both glutamatergic and dopaminergic function). The impact of ketamine upon prediction error processing may be assessed with sensitive neuroimaging techniques such as functional magnetic resonance imaging (fMRI). fMRI studies have identified a frontostriatal system, sensitive to prediction errors during causal learning. Investigating the impact of ketamine upon brain responses to error-driven causal learning and relating that impact to ketamine induced psychopathology ketamine provides a truer understanding of psychosis at the levels of symptoms, cognition and the brain.

are clear similarities between the output of this comparator and the attentional consequences of prediction errors under formal learning theory (Grossberg, 1982; Pearce and Hall, 1980).

In short, it is suggested that inappropriate mismatch signals (i.e. prediction errors) are ultimately responsible for the perceptual aberrations, capture of attention and perception of inappropriate causal relationships that are characteristic of psychosis and may be prelude to delusions (Gray, 1993, 1995, 1998a, b, 2004; Gray *et al.*, 1995; Hemsley, 1992, 1993, 1994, 1996, 1998, 2005a, b). The violation of expectancies in healthy individuals leads subjects to engage in causal reasoning to relieve the feeling of uncertainty and unpredictability about their environment (Einhorn and Hogarth, 1986). The consequences of an inappropriate prediction error may be the same as those of an appropriate one; the formation of causal associative relationships and the allocation of attention to potentially explanatory stimuli (Dickinson, 2001; Pearce and Hall, 1980). Stimuli that may enter into associative relationships include external environmental events as well as patients' internal cognitive and affective operations. This view has its antecedents in earlier models: Schneider, for example, suggested: 'Meaningful connections are created between temporarily coincident external impressions, external impression with the patient's present condition, a perception with thoughts which happened to be present, or events and recollections happening to occur in consciousness at about the same time.' (Schneider, 1930).

In an attempt to account for the therapeutic effects of antipsychotic medications, Kapur appeals to the concept of salience (Beninger, 1988; Kapur, 2003, 2004; Kapur *et al.*, 2005). Motivational salience describes a quality possessed by stimuli that makes those stimuli capable of capturing attention and driving goal-directed behaviour (Berridge and Robinson, 1998). According to the salience hypothesis, stimuli are attributed inappropriate salience due to aberrant dopamine firing in the ventral striatum. Again, delusions arise via a disrupted dopamine-driven learning mechanism which progresses from perceptual and attentional aberrations to delusional ideation. Kapur's model is related therefore to the idea that delusions arise from aberrant perception- and attention-related dopamine firing. Indeed, recent theoretical models have made explicit the links between prediction error, uncertainty and motivational salience. McClure describes a prediction error mediated mechanism whereby stimuli are attributed motivational salience according to stimulus unpredictability (McClure *et al.*, 2003b) and parallels can be drawn between the concept of motivational salience and those of attentional salience and associability also, according to formal learning theory, driven by prediction error (Grossberg, 1982; Pearce and Hall, 1980). It is possible that dopamine activity imbues behaviorally irrelevant stimuli with motivational salience via inappropriate prediction error signalling.

How, ultimately, do these fairly low level changes culminate in the rich and complex set of beliefs that may characterize a delusional system? Any attempt to extend the model in this regard is necessarily speculative. Maher (Maher, 1974, 1988) outlined an attributional account of delusion formation that may explain this transition. He terms aberrant perceptual experiences 'surprises', and posits they are the result of a discrepancy between expectation

and experience in the environment (identical to the prediction errors, central to formal learning theory and to associative accounts of delusion formation). Surprises are intense and pervasive experiences and as such are attributed personal significance [akin to the misattribution of salience hypothesis (Beninger, 1988; Kapur, 2003, 2004; Kapur *et al.*, 2005)].

We may further speculate on the emergence of delusional beliefs in relation to other characteristic features: Firstly, their content is crucially related to the individual's personal fears, needs, or security (Reed, 1972). The particular explanation will be coloured by aspects of the patient's past and present experience as well as cultural factors (Kihlstrom and Hoyt, 1988; Maher, 1974, 1988; Reed, 1972). Aberrant perceptual experiences may be anxiogenic in the same way as unpredictable events (Mineka and Kihlstrom, 1978). Thus anomalous events may be unpleasant and demand explanation. When humans make causal attributions they tend to suffer from a benefactance bias, such that they internalise the cause of positive events and externally attribute negatively valenced events (Greenwald, 1980; Kaney and Bentall, 1992). Hence a psychotic individual seeking an explanation for their unpleasant anomalous experiences will look to the environment outside them. Moreover, people tend to attribute causal significance to the most salient aspects of the perceptual field at the time the event actually occurred (Taylor, 1978). In the terms of associative theories, aberrant prediction error signals might randomly increase the attentional salience of aspects of the perceptual field, leading subjects to attribute inappropriate importance to irrelevant environmental features (Beninger, 1988; Kapur, 2003, 2004; Kapur *et al.*, 2005; Gray, 1993, 1995, 1998a, b, 2004; Gray *et al.*, 1995; Hemsley, 1992, 1993, 1994, 1996, 1998, 2005a, b).

Delusions also tend to be fixed: unshakeable in the face of evidence that appears to contradict them. If it is true that the delusional belief accounts for unpredictable and therefore anxiety-provoking experience, then it is possible that its emergence is accompanied by relief from anxiety. This outcome diminishes the person's subsequent motivation to question his or her original conclusions and increases resistance to contrary information. This theme is also represented in Miller's (1993) associative learning based account of psychosis. He argues that arriving at a causal explanation that accounts for aberrant experiences is so rewarding/relieving that it is accompanied by a surge of dopamine (Miller, 1993). Dopamine impacts upon the consolidation of memories (Dalley *et al.*, 2005) and as such, a delusional conclusion, formed under conditions of dopamine hyperactivity, is 'stamped-in' to long term memory, rendering it relatively impervious to disconfirmatory evidence.

The account outlined thus far deals best with referential delusions (for example delusions of grandiosity and paranoia). It is possible that inappropriate prediction errors (or noise) in systems other than the mesocorticolimbic dopamine system may underpin other positive symptoms for example delusions of passivity (Blakemore *et al.*, 2002; Frith *et al.*, 2000).

In summary, models of delusion formation emphasise inappropriate prediction error signalling as a basis for the inappropriate perception, attention, association and significance of stimuli, thoughts and percepts. In order to reconcile these unexpected experiences,

patients engage in delusional reasoning, in a similar manner to healthy individuals when their expectancies are violated (Dickinson, 2001; Einhorn and Hogarth, 1986; Hemsley, 1992, 1993, 1994). In the next section we consider the evidence that people with schizophrenia do indeed show associative learning abnormalities.

Evidence for an associative learning models of delusion formation in schizophrenia

The model of delusion formation outlined above would predict that the early stages of psychosis development would be accompanied by a period of transition, in which abnormal sensory perceptions, are experienced for the first time, and, with repeated occurrence, become ultimately becoming rationalized and integrated into the individual's understanding and representation of the environment. These abnormal perceptions and the aberrant explanations patients generate in order to account for them, can be seen as the early crystallization of the symptoms of psychosis (Hemsley, 1992, 1993, 1994). Thus information from informal and structured clinical interviews, taken at the very early stages of illness, are important in identifying the role of aberrant associations in patients' accounts of their symptoms. McGhie and Chapman carried out such interviews and recorded a number of early phenomena that seem to be consistent with the model (McGhie and Chapman, 1961). First, they remarked that, during the prodromal period, patients appear to lose their ability to direct attention focally and voluntarily. Instead, attention is diverted, involuntarily to the diverse stimuli in the environment that do not usually capture the attentional focus (McGhie and Chapman, 1961). One patient described this as follows: 'It's as if someone had turned up the volume . . . I noticed it most with background noises . . . noises that are always around but you don't notice them. Now they seem to be just as loud as and sometimes louder than the main noises that are going on' (McGhie and Chapman, 1961). Freedman and Chapman (1973) found that compared to non-schizophrenic psychiatric control subjects, patients with schizophrenia more frequently reported changes in their ability to concentrate, as a result of their inability to focus attention selectively on relevant stimuli and thoughts, and to screen out or ignore irrelevant stimuli and thoughts (Freedman and Chapman, 1973).

Also in keeping with the model, patients at the early stages of illness report perceptual changes. These changes may involve a heightening of sensory awareness, particularly in the visual and auditory modalities (McGhie and Chapman, 1961) and are perhaps an extension of the loss of selective attention outlined above. These perceptions, and alterations in their attentional status, may mediate a change in subjective reality (McGhie and Chapman, 1961). Schizophrenics more often than non-schizophrenics report changes, in the intensity of visual and auditory perception of real, identifiable external stimuli (Freedman and Chapman, 1973). Chapman (1966) suggests that patients' perceptual disruptions reflect an impaired capacity to select relevant stimuli from the diffuse mass of incident stimuli, based on previous learning and experience (Chapman, 1966). This is a theme echoed in Matussek's account of delusional perception (Matussek, 1952), in which patient's perceptual disruptions are interpreted as a

'loosening of the perceptual context'. Stimuli are perceived independently, rather than a part of a coherent scene, suggesting an inability to bring previous experience to bear upon perception (Matussek, 1952). The effect of this misperception of the whole scene as isolated stimuli is the attribution of delusional significance to those stimuli that happen to capture attention (Matussek, 1952) (see below).

Associative learning in schizophrenia

The data on attentional and perceptual disruption support a general intrusion of irrelevant stimuli into conscious experience. However, associative theories of psychosis require that the attentional disturbance should have consequences for learning. One paradigm in which this possibility can be tested is latent inhibition (Lubow and Moore, 1959). Latent inhibition (LI) is an adaptive learning process, used by organisms to prevent responding to behaviourally irrelevant stimuli. It is manifest as retardation of learning of relationships between a stimulus and an outcome if that stimulus has been experienced previously, without preceding the outcome (Lubow, 1965). According to formal learning theories, latently inhibited stimuli lose associability, they are less attentionally salient (Pearce and Hall, 1980).

According to associative theories of psychosis, LI should be impaired in acutely psychotic patients because they deploy inappropriate attention to behaviourally irrelevant stimuli (Gray, 1993; Hemsley, 1993). Latent inhibition is indeed impaired in acute schizophrenia, in that schizophrenics do not suffer any decrement in learning following non-reinforced pre-exposure (Baruch *et al.*, 1988; Bender *et al.*, 2001; Gal *et al.*, 2005; Gray 1993, 1998a, b, 2004; Gray and Snowden, 2005; Lubow and Gewirtz, 1995; Vaitl *et al.*, 2002; Yogeve *et al.*, 2004; Young *et al.*, 2005). This effect is abolished by disease chronicity and medication status, those receiving neuroleptic treatment does not show an attenuation of latent inhibition (Alves and Silva, 2001; Baruch *et al.*, 1988; Gal *et al.*, 2005). Another conditioning phenomenon that (in accordance with associative theories) is impaired in psychosis is conditioned blocking (Alves and Silva, 2001; Baruch *et al.*, 1988; Gal *et al.*, 2005; Jones *et al.*, 1992; Martins Serra *et al.*, 2001; Moran *et al.*, 2003; Oades *et al.*, 1992, 1996). Blocking occurs when nothing is learned about a novel stimulus when it is paired with a familiar stimulus that already fully predicts an outcome (Kamin, 1969). Blocking is disrupted in schizophrenia, in that patients learn an association between the novel stimulus and the outcome. According to associative theories, this is due to inappropriate prediction error signaling, leading to association formation between the irrelevant stimulus and the outcome (Escobar *et al.*, 2002).

Cognitive control in schizophrenia

Reasoning involves the purposeful manipulation of relevant details from previously acquired stored information. Patients report that the logical sequence of their ideas is replaced by sequences of merely associated thoughts, and the reasoning process therefore became increasingly concrete (McGhie and Chapman, 1961). Arieti suggests that reasoning is disrupted in schizophrenia due to 'a lack

of the inhibition of peripheral ideas necessary for effective abstraction' (Arieti, 1955). Similarly McKellar highlights the importance of the inability to inhibit associated but irrelevant ideas in patients with schizophrenia (McKellar, 1957). Empirical tests of these reasoning disruptions include sorting objects into related categories. Schizophrenic patients tend to be over-inclusive (Cameron, 1939). That is, instead of sorting on the basis of size, shape or colour, patients tend to ignore these main variables and concentrate on some insignificant features, for example scratches or irregularities on the surface of the object. Over-inclusiveness could be considered a cognitive manifestation of the attentional and learning disruptions induced by inappropriate prediction errors. Subjects can be over-inclusive due to attentional interference by distracting stimulus features or by forming amalgamations of items on the basis of weak relations between them. Prediction error signals could provide a mechanism for both of these possibilities (see Fig. 1), since a mismatch signal increases attention, in a search for explanatory stimuli (Pearce and Hall, 1980) and mismatch signals drive learning directly by strengthening associations (Dickinson, 2001; Rescorla and Wagner, 1972).

Reasoning may be considered one kind of cognitive control. Schizophrenic patients tend to be impaired upon most tasks of cognitive control (Barch, 2005; Braver *et al.*, 1999; Cohen and Servan-Schreiber, 1992; Goldman-Rakic, 1994; Weinberger and Gallhofer, 1997). Above we discussed a model of cognitive control that uses the same VTA dopamine signals that drive associative learning to gate information flow to the prefrontal cortex during tasks that require cognitive control (Braver *et al.*, 1999; Braver and Cohen, 1999; O'Reilly *et al.*, 1999). It has been proposed that the poor performance of schizophrenic patients on such tasks is a result of an impaired prefrontal gating mechanism (Braver *et al.*, 1999; Braver and Cohen, 1999). If the prefrontal gate is opened by behaviourally irrelevant stimuli the patient will be distracted by those stimuli and their task performance will be disrupted.

In brief, therefore, data from schizophrenic patients suggest that there are indeed abnormalities in perception, attention, associative learning and cognitive control, all of which are compatible with the model of delusion formation outlined above. Support the involvement of inappropriate prediction error signals symptom formation. However, such information is necessarily indirect and identifying a population who are currently in this very early phase of their illness is difficult. Consequently, this model has not been tested directly. In the next section, we suggest a way of exploring the psychological prelude to delusion formation using a drug model (ketamine). In this setting we are able to control and manipulate psychopathology and to relate it to psychological processing more directly.

Ketamine as a drug model for delusion formation

The evidence reviewed so far in support of an associative learning deficit model of psychosis is necessarily indirect, since more specific evidence would require investigation of these deficits during the prodromal phase, during which inappropriate associations begin to be linked and formed into a delusional framework. As we

have noted earlier, this formative period of the illness is intrinsically difficult to investigate, since patients lack insight into the inappropriate associations formed, and will typically not have been identified to psychiatric services at this stage. Given this situation, a model of the processes theoretically predicted to be impaired would be particularly useful, affording an opportunity to investigate and manipulate the process of associative learning and its impact on psychotic experience. The requirements of this model would be to induce an acute, transient state of psychotic experience in healthy volunteers, ideally via neurobiological mechanisms theoretically implicated in associative learning. Ketamine, an NMDA receptor antagonist, may be an appropriate candidate as a model to interrogate the association between impairments in associative learning and the experience of phenomena which are qualitatively similar to those seen in patients with schizophrenia. Here we discuss the psychotomimetic effects of ketamine in terms of a causal associative learning analysis.

Attentional and perceptual effects of ketamine

Ketamine appears to disrupt the intensity and integrity of the sensory experience (Krystal *et al.*, 1994). In the auditory domain, the effect is one of hyper-acuity. Background noises become unusually loud (Krystal *et al.*, 1994; Oye *et al.*, 1992; Vollenweider *et al.*, 1997a, b). Subjects become preoccupied by certain unimportant sounds like the ticking of a clock (Oye *et al.*, 1992). Shifting attention away from these unusually loud sounds requires conscious effort (Oye *et al.*, 1992). Visual phenomena are also common (Krystal *et al.*, 1994; Oye *et al.*, 1992; Vollenweider *et al.*, 1997a, b). Subjects become preoccupied with certain objects (Oye *et al.*, 1992), spatially distant, weak or insignificant stimuli are perceived as disproportionately salient (Krystal *et al.*, 1994). Colours within the focus of attention seem more vivid than usual, those outside the focus of attention are dulled (Krystal *et al.*, 1994). These visual effects are associated with strange meanings of the surroundings, confusion and difficulty in directing/focusing attention (Vollenweider *et al.*, 1997a); one subject reported seeing the 'shadow' of a person falling past a fourth floor window (Newcomer *et al.*, 1999). This is an example of how perceptual aberration may lead to delusional ideation (see above).

Ketamine and psychopathology

The perceptual aberrations induced by ketamine are remarkably similar to those described by schizophrenic patients early in their illness (Freedman and Chapman, 1973; Freedman, 1974; McGhie and Chapman, 1961), that is subjects report that their attention is drawn to irrelevant or background stimuli and that those stimuli are imbued with significance. The alterations in perceptual experience are consistent with an aberrant prediction error account, whereby noise in the brain system that generates prediction errors would drive attention towards irrelevant stimuli. However, ketamine does not increase subjects' tendency to externalize when guessing the source of ambiguous material (Honey *et al.*, 2005), unlike in schizophrenia, where patients tend to attribute ambiguous material to external sources, particularly those experiencing delusions of

control (Keefe *et al.*, 1999). The increased propensity of subjects under ketamine to internalize ambiguous information may well relate to the specific symptomatology that ketamine induces; subjects report paranoid ideas and ideas of reference on ketamine (Bowdle *et al.*, 1998; Krystal *et al.*, 1994, 1998). Therefore an increased tendency to internalize ambiguous information may precipitate the self-referentiality of the delusion. How episodic memory formation and source attribution relate to prediction error processing remains unclear, however, recent brain imaging data relate dopaminergic midbrain firing at encoding to successful subsequent retrieval of memoranda (Schott *et al.*, 2004, 2006). Noise in this system, induced by drug or disease, may well impact upon memory performance. How this relates to source attribution, especially under ambiguous situations, remains to be clarified experimentally.

Ketamine, cognitive control and learning

The impact of ketamine on tasks of cognitive control is well characterized. Ketamine impairs working memory (Adler *et al.*, 1998; Honey *et al.*, 2003, 2004; Morgan *et al.*, 2004; Rowland, 2005; Rowland *et al.*, 2005) and sustained attention (Passie *et al.*, 2005; Umbricht *et al.*, 2000). The pattern of deficits appears to reflect those of the psychotic patient fairly closely (Krystal *et al.*, 2000; Umbricht *et al.*, 2000). Krystal and colleagues assayed the effects of ketamine on the Wisconsin card sort task, a classic measure of cognitive control and executive attention (Krystal *et al.*, 2000) They used a novel 2-day design in which they could assess the impact of ketamine on the acquisition of rules and, on a separate occasion the implementation of those rules. Ketamine impaired the acquisition but not the implementation of abstract stimulus-response rules; it also increased distractibility, and psychotic symptoms (Krystal *et al.*, 2000). Although Krystal and colleagues do not report any relationship between distractibility, task performance and psychopathology in their subjects, it may be that the effects of ketamine on cognitive control may be mediated by the generation of inappropriate prediction error signals resulting in excessive malleability of the goal representation, distractibility and reduced cognitive control. Such cognitive disruptions are unexpected and distressing and may drive referential delusional reasoning.

The impact of ketamine on latent inhibition in human subjects remains to be ascertained. However the effect of ketamine on attentional measures has been explored (Abel *et al.*, 2003; Kreitschmann-Andermahr *et al.*, 2001; Oranje *et al.*, 2000; Umbricht *et al.*, 2002; Umbricht *et al.*, 2000) for example; mismatch negativity (MMN, an event related potential generated in response to outlier or oddball stimuli embedded within trains of predictable stimuli). Ketamine disrupts the MMN signal in healthy volunteers in a pattern redolent of the disruption in schizophrenic patients (Umbricht *et al.*, 2000); furthermore, subjects' MMN response in the absence of drug correlates with their experience of positive symptoms on ketamine (Umbricht *et al.*, 2002). Under the scheme outlined in this review, prediction errors should have some consequence for learning, hence inappropriate errors should lead to maladaptive or erroneous learning. The MMN signal is generated by unexpected events (and so can be considered a prediction error), however, the impact of MMN

disruption on subsequent behaviour is unknown and a specific relationship with delusions has not been demonstrated.

In recent work from our own laboratory we demonstrate that ketamine does indeed disrupt prediction error processing in the context of a causal learning task (Corlett *et al.*, 2006). Across a series of studies we have demonstrated a relationship between prediction errors generated by expectancy violation during causal learning and BOLD response in right lateral prefrontal cortex (Corlett *et al.*, 2004; Fletcher *et al.*, 2001; Turner *et al.*, 2004). In a placebo controlled study we found that a low dose of ketamine attenuated the response to expectancy violation in this region whilst augmenting the response to unsurprising, predictable events. Outside of the scanner we increased the dose of ketamine to induce the psychopathology associated with ketamine. Across subjects, sensitivity to expectancy violation on placebo was predictive of perceptual aberration and attentional capture [as measured by the CADSS (Bremner *et al.*, 1998)] as well as ideas and delusions of reference [as measured by the PSE (Wing *et al.*, 1974)]. We believe that these data provide objective evidence in favour of the account of delusion formation described presently.

In the preceding section we have outlined some of the effects of ketamine and how they relate to the model under examination. At an acute dose, in healthy volunteers, ketamine induces psychopathological phenomena redolent of those described by psychotic patients reflecting retrospectively on the very early stages of their illness (Freedman and Chapman, 1973; Freedman, 1974; McGhie and Chapman, 1961). These accounts have formed the basis of theoretical considerations of delusion formation based on disrupted associative learning which emphasise disturbed volitional control of attention, reasoning and associative learning, culminating in the construction of a delusional system to account for such bizarre experiences (Beninger, 1988; Kapur, 2003, 2004; Kapur *et al.*, 2005; Gray, 1993, 1995, 1998a, b, 2004; Gray *et al.*, 1995; Hemsley, 1992, 1993, 1994, 1996, 1998, 2005a, b). It is possible that noise in a system responsible for signaling mismatches between expectancy and outcome (prediction errors) may provide a parsimonious mechanistic account for these phenomena in disease and under ketamine (Corlett *et al.*, 2006). However, the majority of experimental data on the neurobiology of mismatch signals implicate dopamine as the key neurotransmitter (Fiorillo *et al.*, 2003, 2005; Hollerman and Schultz, 1998; Hollerman *et al.*, 1998; Schultz, 1998a, b, 1999, 2001, 2002; Schultz *et al.*, 1997, 1998, 2000; Schultz and Dickinson, 2000; Waelti *et al.*, 2001). We have argued that ketamine (an NMDA receptor antagonist) provides one experimental tool to transiently induce a state redolent of early psychosis. The next section will attempt to address the apparent mismatch; How can an NMDA receptor antagonist, that perturbs glutamate function, induce psychosis via a dopamine dependent mechanism?

Fronto-striatal effects of ketamine on glutamatergic and dopaminergic function

We have thus far described how the phenomenology associated with both schizophrenia and ketamine share overlapping features, and we have argued that both states may be characterized by a

breakdown of association formation. We next consider the neurobiological aspects of ketamine-induced psychopathology, and the relevance of this to the associative learning account of psychosis.

Ketamine and dopamine

As reviewed earlier, prediction error signaling and stimulus-reward learning primarily implicates dopamine as the major neurotransmitter (Fiorillo *et al.*, 2003, 2005; Hollerman and Schultz, 1998; Hollerman *et al.*, 1998; Schultz 1998a, b, 1999, 2001, 2002; Schultz *et al.*, 1997, 1998, 2000; Schultz and Dickinson, 2000; Waelti *et al.*, 2001). Similarly, the dopamine hypothesis is the dominant pathophysiological model of psychosis, based upon the neurochemical actions of antipsychotics (Anden *et al.*, 1970; Carlsson and Lindqvist, 1963; Creese *et al.*, 1976; Nyback and Sedvall, 1970; Seeman *et al.*, 1976) as well as the psychotomimetic effects of amphetamines (Angrist and Gershon, 1970; Bell, 1965, 1973; Connell, 1958; Young and Scoville, 1938). Whilst the primary action of ketamine is the blockade of the glutamatergic NMDA receptor, ketamine has important direct and indirect actions on both glutamate and dopamine, indeed it has been argued that the cognitive and behavioral effects of ketamine might be attributable to stimulation of D₂ receptors rather than to the blockade of NMDA receptors (Kapur and Seeman, 2001).

The locomotor stimulating effects of NMDA receptor antagonists like PCP and MK-801 in experimental animals were indicative of an influence on dopaminergic neurotransmission (Carlsson and Carlsson, 1990; Whitton *et al.*, 1992). Acute treatment with these compounds increases dopamine release in the striatum, nucleus accumbens and prefrontal cortex of experimental animals (Mathe *et al.*, 1998; Moghaddam *et al.*, 1997; Sitges *et al.*, 2000; Wedzony *et al.*, 1994) and enhances the firing rate of dopamine neurons in the midbrain (Freeman and Bunney, 1984; Murase *et al.*, 1993; Svensson *et al.*, 1998; Zhang *et al.*, 1992). In addition, the locomotor response produced by NMDA receptor antagonists can be diminished by catecholamine depletion (Maj *et al.*, 1991; Willins *et al.*, 1993) or by dopamine receptor antagonists administered either systemically or directly into the striatum (Corbett *et al.*, 1995; Hoffman, 1992; Willins *et al.*, 1993).

In human subjects, positron emission tomography imaging of 11C-raclopride binding following ketamine administration revealed that ketamine induces striatal dopamine release (Smith *et al.*, 1998) and the magnitude of this release correlated with the intensity of ketamine induced psychosis (Breier *et al.*, 1998; Vollenweider *et al.*, 2000). Using a novel experimental design, Kegeles and colleagues demonstrated that ketamine enhanced the striatal dopamine release induced by amphetamine administration, suggesting that the psychotomimetic effects of NMDA receptor antagonism may be due to a disruption in the glutamatergic control of dopamine function (Kegeles *et al.*, 2000).

Ketamine and glutamate

The cognitive and behavioural effects of ketamine are not solely attributable to its effect at the D₂ receptor. Typical antipsychotics, which afford high D₂ receptor blockade, do not reverse the

psychotomimetic effects of ketamine (Krystal *et al.*, 1999) suggesting a dopamine independent mechanism contributing at least in part to the effects of ketamine. Both the psychopathological (Anand *et al.*, 2000) and cognitive (Krystal *et al.*, 2005a) effects of ketamine have been reported to be blocked by compounds reducing glutamatergic transmission.

Models of associative learning implicate glutamatergic transmission via NMDA and AMPA receptors in long term potentiation (For review see (Nicoll, 2003) In animal models, applications of compounds that disrupt dopamine, NMDA and AMPA receptors in the nucleus accumbens result in dissociable behavioural effects on associative learning (Di Ciano *et al.*, 2001). Using an autoshaping task, where stimuli predictive of rewards usurp the motivational qualities of the reward and elicit approach behaviour in the animal (Brown and Jenkins, 1968), DiCiano and colleagues demonstrated that dopaminergic and NMDA (but not AMPA) receptor dysfunction disrupted acquisition of Pavlovian approach behaviour. Infusion of the NMDA receptor antagonist AP5 into the nucleus accumbens core blocks spatial learning in a maze task (Maldonado-Irizarry and Kelley, 1995; Smith-Roe and Kelley, 2000) as well as acquisition of an instrumental response (lever-pressing) for food reward (Kelley *et al.*, 1997), however, once the response was acquired, NMDA receptor antagonist infusions were without impact upon behavioural performance, suggesting an impact upon learning but not execution of the response. The data suggest that dopamine and glutamate interaction is critical in triggering intracellular transductional and transcriptional mechanisms that lead to long term changes in gene expression, synaptic plasticity and ultimately behaviour (Dalley *et al.*, 2005; Floresco *et al.*, 2001; Kelley and Berridge, 2002; Scott *et al.*, 2002). Disruptions to this level of interaction may account for the longer term maintenance of delusions across psychotic episodes (Miller, 1993) as well as their elaboration and fixity despite overwhelming contradictory evidence (Miller, 1993). However, those roles are tangential to the critical process at hand; prediction error, and its disruption in early psychotic states.

Ketamine and dopamine-glutamate interaction

Recent developments in *in vivo* voltammetric measurements of cortical and subcortical dopamine and glutamate in experimental animals may provide some indication of the interaction between subcortical dopamine and cortical glutamatergic transmission. Lavin and co-workers recorded a rapid glutamate signal in the PFC in response to VTA stimulation (Lavin *et al.*, 2005). This suggests a reinterpretation of the consequences to prefrontal function of phasic, prediction error related dopaminergic firing in the VTA. The prefrontal cortex responds to such an error with a slowly increasing level of extracellular dopamine that rises to a plateau and then gradually returns to baseline. This is also the case when motivationally salient events occur, irrespective of their valence (Ahn and Phillips, 1999; Del Arco and Mora, 2000; Feenstra and Botterblom, 1996; Feenstra *et al.*, 1995, 2000; Finlay and Zigmond, 1997; Finlay *et al.*, 1995; Taber and Fibiger, 1997; Watanabe *et al.*, 1997). The firing of VTA codes a salience/prediction error signal in accordance with single-unit recording data, yet this signal may be transmitted via glutamate co-released from VTA terminals in the PFC (Lavin *et al.*,

2005). This hypothesis frees DA from having to encode both fast (in the VTA) and slow (in the PFC) signals and transfers fast signalling to the glutamate system that is exquisitely suited to produce transient changes in neural activity. The firing of VTA neurons in response to an unexpected outcome may release glutamate in the PFC, which can evoke a persistent activity state. It is possible that the PFC maintains the observation that the environment was different than expected and attempts to reconcile that information with what happens to the organism subsequently. Thus, when an animal enters an environment rich with unexpected rewards, DA may maintain a state of cognitive attention, lasting many minutes, such that the organism learns to predict the rewards and they are no longer unexpected (Lavin *et al.*, 2005).

Based on the brain imaging data that implicate frontostriatal systems in causal learning (Corlett *et al.*, 2004; Fletcher *et al.*, 2001; Turner *et al.*, 2004), we propose that ketamine induces inappropriate subcortical prediction error signals in human subjects (Corlett *et al.*, 2006). This would precipitate the engagement of prefrontal mechanisms, leading to the allocation of attention to potentially explanatory (although irrelevant) stimuli, the formation of inappropriate associations between stimuli and a disruption of goal-directed behaviour. Such cognitive disruptions lead to the deployment of higher-level, metacognitive strategies to account for subject's experiences. Given the anxiogenic nature of odd cognitive experiences and the perception of inappropriate relatedness and importance of various entities, subjects construct bizarre delusional accounts that are coloured by their personal knowledge of the world and causal interactions within it (Kilhlstrom and Hoyt, 1988). The perceptual, cognitive and pathological sequelae of endogenous and NMDA antagonist induced psychosis may be due to disruptions in dopamine and glutamate signals (see Fig. 1).

Relevance of other psychotomimetic drugs

There are of course other drugs that induce a transient psychotic state in healthy volunteers; for example amphetamine (Angrist and Gershon, 1970; Bell, 1965, 1973; Connell, 1958; Young and Scoville, 1938), lysergic acid diethylamide (LSD) (Osmond, 1957), psilocybin (Hasler *et al.*, 2004; Vollenweider *et al.*, 1998) and cannabis (D'Souza *et al.*, 2004). It is important to consider how the present framework might apply to these agents.

Thus far we have discussed how ketamine may impact upon prediction error processing across a network of brain regions and thus induce psychotic symptoms. We must also consider cellular and intracellular processes. Svenningsson and colleagues recently demonstrated that PCP (a more potent analogue of ketamine), LSD, cannabis and amphetamine all act via a common intracellular signaling pathway (Andersson *et al.*, 2005; Svenningsson *et al.*, 2003). This signaling cascade proceeds via the phosphorylation of Dopamine and cAMP regulated phosphoprotein (molecular weight = 32 kDa) or DARPP-32. DARPP-32 is highly concentrated in neostriatum (the caudate, putamen and nucleus accumbens). The neurons that contain DARPP-32 are the only efferent pathway conveying information out of the neostriatum to the cortex. Furthermore, the excitability of DARPP-32 containing neurons is

modulated by dopaminergic neurons that project from the VTA to the neostriatum. The prefrontal cortex also sends glutamatergic afferents back onto DARPP-32 rich neurons in the neostriatum (for review see Greengard, 2001).

Svenningsson and colleagues demonstrated the importance of DARPP-32 in the action of psychotomimetic drugs using a mutant mouse in which DARPP-32 function had been knocked out. The psychotomimetic effects of LSD, psilocybin, PCP, cannabis and amphetamine were all attenuated in this animal relative to wild-type (Andersson *et al.*, 2005; Svenningsson *et al.*, 2003). All of these drugs do have some affinity for dopamine receptors but they also bind to specific receptor systems; PCP to NMDA receptors; LSD and psilocybin to serotonergic receptors; Cannabis to endocannabinoid receptors and amphetamine to dopamine receptors. It is the intracellular cascade of biochemical events that mediate their psychotomimetic effects. It is likely that ketamine is also a potent modulator of DARPP-32 given its impact upon dopamine and glutamate function and its homology with PCP.

DARPP-32 knockout (KO) mice also exhibit behavioural deficits in learning about food rewards. In a reversal learning paradigm in which mice are trained to respond on a certain contingency during an acquisition phase and then trained on a reversed contingency in a subsequent training phase, DARPP-32 KO mice showed retarded reversal learning relative to wild-type animals. This effect was interpreted by the authors as relevant to associative learning theory and specifically to prediction error, since DARPP-32 animals were unable to learn rapidly from inappropriate responses and switch to an appropriate behavioural response (Heyser *et al.*, 2000).

Taken together, these data suggest that all psychotomimetic drugs mediate their effects via an intracellular signaling pathway that is critical for reward learning and adaptive responding, and that may code and respond to the prediction errors intracellularly.

Conclusions and future work

In summary, this paper has outlined a prediction error model of psychosis that attempts to explain the perceptual, attentional and cognitive disruptions characteristic of the earliest phases of psychosis (see Fig. 1). There is evidence to suggest that prediction error signaling is mediated by both dopamine and glutamate (Lavin *et al.*, 2005). The relevance of ketamine to this model is clear, given its influence on both glutamatergic and dopaminergic function (Aalto *et al.*, 2005; Breier *et al.*, 1998; Krystal *et al.*, 2005b; Moghaddam *et al.*, 1997; Smith *et al.*, 1998; Vollenweider *et al.*, 2000). However, the hypothesised relation between disrupted glutamate and dopamine function, attentional disruption and aberrant associative learning requires validation from experimental work in animal models (rat and non-human primate) as well as human subjects administered NMDA receptor antagonists and suffering endogenous psychosis. For example, the impact of systemic NMDA receptor antagonists on the putative salience signals of Schultz (Fiorillo *et al.*, 2003, 2005; Hollerman and Schultz, 1998; Hollerman *et al.*, 1998; Schultz, 1998a, b, 1999, 2000, 2001, 2002; Schultz *et al.*, 1997, 1998, 2000; Waelti *et al.*, 2001) and Lavin and colleagues (Lavin *et al.*, 2005) would provide further support for the model.

Assessing the effects of other psychotomimetic drugs, such as cannabis and psilocybin on prediction error signals in animals and humans might provide some insight upon the generality of the proposed framework. Functional polymorphisms in the gene that codes for DARPP-32 have been identified, they impact upon frontostriatal structure and function and increase risk of schizophrenia (Meyer-Lindenberg *et al.*, 2007). Such genetic markers of DARPP-32 function might permit *in vivo* investigation of the relationships between the function of this protein, the effects of psychotomimetic drugs and prediction error processing in humans.

As outlined above, acetylcholine modulates the attention allocated to environmental events (Chiba *et al.*, 1995; Everitt and Robbins, 1997; Sarter and Bruno, 1999; Yu and Dayan, 2002, 2005), possibly under the influence of prediction error signals from the midbrain (Bao *et al.*, 2001, 2003) in accordance with formal associative learning theories (Grossberg, 1982; Pearce and Hall, 1980). Ketamine enhances cortical acetylcholine release in experimental animals (Nelson *et al.*, 2002), potentially via the sensitization of mesolimbic dopamine system (see Sarter *et al.*, 2005 for review). Recent pathophysiological theories of schizophrenia postulate disturbances in both cholinergic and dopaminergic neurotransmission and synaptic plasticity as a result of NMDA receptor hypofunction (Friston, 2005; Stephan *et al.*, 2006). However, these theories ascribe separable consequences to cholinergic and dopaminergic dysfunction; impaired emotional learning due to dopaminergic dysfunction and impaired perceptual learning as a consequence of cholinergic impairments. Emotional learning deficits are held to underlie 'the disintegrative and autistic aspects of schizophrenic symptoms' whilst perceptual learning deficits underpin hallucinations (Stephan *et al.*, 2006). The present thesis argues against such strict pathophysiological separation on the basis that attentional processes (driven by an interaction between dopaminergic prediction error signals and cholinergic attentional modulation) may also be important in associative causal learning and hence delusion formation. Future work should interrogate the relationship between human associative learning and attention and its pharmacological basis using paradigms from associative learning theory.

The impact of ketamine on cognition and the relationship between this impact and the symptoms induced by ketamine needs to be further explored. For example, the model of delusion formation described would benefit from a fuller treatment of the interaction between aberrant prediction error, surprise and attribution. We have recently deployed functional imaging techniques that provide more sensitive, multidimensional measures of associative learning in human subjects than their predictive behaviour. Our findings support the present thesis, demonstrating a relationship between NMDA receptor hypofunction, causal learning, perceptual aberration and delusions (Corlett *et al.*, 2006). Future studies should explore the interaction between subjects' prediction error processing during causal learning and their source monitoring and cognitive control abilities. By demonstrating such interactions, we can inform upon and enrich the present model of delusion formation. By using fMRI in conjunction with pharmacological models in this way, it may be possible to identify functional brain markers that predict the onset of symptoms which may be applied to populations at high risk of psychosis.

References

- Aalto S, Ihalaainen J, Hirvonen J, Kajander J, Scheinin H, Tanila H, Nagren K, Vilkmann H, Gustafsson L L, Syvalahti E, Hietala J (2005) Cortical glutamate-dopamine interaction and ketamine-induced psychotic symptoms in man. *Psychopharmacology (Berl)* 182: 375–383
- Abel K M, Allin M P, Hemsley D R, Geyer M A (2003) Low dose ketamine increases prepulse inhibition in healthy men. *Neuropharmacology* 44: 729–737
- Adler C M, Goldberg T E, Malhotra A K, Pickar D, Breier A (1998) Effects of ketamine on thought disorder, working memory, and semantic memory in healthy volunteers. *Biol Psychiatry* 43: 811–816
- Ahn S, Phillips A G (1999) Dopaminergic correlates of sensory-specific satiety in the medial prefrontal cortex and nucleus accumbens of the rat. *J Neurosci* 19: RC29
- Alves C R, Silva M T (2001) Facilitation of latent inhibition by the atypical antipsychotic risperidone. *Pharmacol Biochem Behav* 68: 503–506
- Anand A, Charney D S, Oren D A, Berman R M, Hu X S, Cappiello A, Krystal J H (2000) Attenuation of the neuropsychiatric effects of ketamine with lamotrigine: support for hyperglutamatergic effects of N-methyl-D-aspartate receptor antagonists. *Arch Gen Psychiatry* 57: 270–276
- Anden N E, Butcher S G, Corrodi H, Fuxe K, Ungerstedt U (1970) Receptor activity and turnover of dopamine and noradrenaline after neuroleptics. *Eur J Pharmacol* 11: 303–314
- Andersson M, Usiello A, Borgkvist A, Pozzi L, Dominguez C, Fienberg A A, Svenningsson P, Fredholm B B, Borrelli E, Greengard P, Fisone G (2005) Cannabinoid action depends on phosphorylation of dopamine- and cAMP-regulated phosphoprotein of 32 kDa at the protein kinase A site in striatal projection neurons. *J Neurosci* 25: 8432–8438
- Angrist B M, Gershon S (1970) The phenomenology of experimentally induced amphetamine psychosis—preliminary observations. *Biol Psychiatry* 2: 95–107
- Arieti S (1955) *Interpretation of Schizophrenia*. NY: Robert Brunner
- Arieti S (1974) An overview of schizophrenia from a predominantly psychological approach. *Am J Psychiatry* 131: 241–249
- Aron A R, Shohamy D, Clark J, Myers C, Gluck M A, Poldrack R A (2004) Human midbrain sensitivity to cognitive feedback and uncertainty during classification learning. *J Neurophysiol* 92: 1144–1152
- Bao S, Chan V T, Merzenich M M (2001) Cortical remodelling induced by activity of ventral tegmental dopamine neurons. *Nature* 412: 79–83
- Bao S, Chan V T, Zhang L I, Merzenich M M (2003) Suppression of cortical representation through backward conditioning. *Proc Natl Acad Sci U S A* 100: 1405–1408
- Barch D M (2005) What can research on schizophrenia tell us about the cognitive neuroscience of working memory? *Neuroscience*
- Baruch I, Hemsley D R, Gray J A (1988) Differential performance of acute and chronic schizophrenics in a latent inhibition task. *J Nerv Ment Dis* 176: 598–606
- Bell D S (1965) Comparison of amphetamine psychosis and schizophrenia. *Br J Psychiatry* 111: 701–707
- Bell D S (1973) The experimental reproduction of amphetamine psychosis. *Arch Gen Psychiatry* 29: 35–40
- Bender S, Muller B, Oades R D, Sartory G (2001) Conditioned blocking and schizophrenia: a replication and study of the role of symptoms, age, onset-age of psychosis and illness-duration. *Schizophr Res* 49: 157–170
- Beninger R J (1988) *The Slow Therapeutic Action of Antipsychotic Drugs: A Possible Mechanism Involving the Role of Dopamine in Incentive Learning*. Karger, Basel
- Berns G S, McClure S M, Pagnoni G, Montague P R (2001) Predictability modulates human brain response to reward. *J Neurosci* 21: 2793–2798

- Berridge K C, Robinson T E (1998) What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Res Brain Res Rev* 28: 309–369
- Berze J (1914) Die primäre Insuffizienz der psychischen Aktivität. Ihr Wesen, ihre Erscheinungen und ihre Bedeutung als Grundstörungen der Dementia Praecox und der hypophrenen überhaupt. Leipzig: Franz Deuticke
- Blakemore S J, Wolpert D M, Frith C D (2002) Abnormalities in the awareness of action. *Trends Cogn Sci* 6: 237–242
- Bowdle T A, Radant A D, Cowley D S, Kharasch E D, Strassman R J, Roy-Byrne P P (1998) Psychedelic effects of ketamine in healthy volunteers: relationship to steady-state plasma concentrations. *Anesthesiology* 88: 82–88
- Braver T S, Barch D M, Cohen J D (1999) Cognition and control in schizophrenia: a computational model of dopamine and prefrontal function. *Biol Psychiatry* 46: 312–328
- Braver T S, Cohen J D (1999) Dopamine, cognitive control, and schizophrenia: the gating model. *Prog Brain Res* 121: 327–349
- Breier A, Adler C M, Weisenfeld N, Su T P, Elman I, Picken L, Malhotra A K, Pickar D (1998) Effects of NMDA antagonism on striatal dopamine release in healthy subjects: application of a novel PET approach. *Synapse* 29: 142–147
- Bremner J D, Krystal J H, Putnam F W, Southwick S M, Marmar C, Charney D S, Mazure C M (1998) Measurement of dissociative states with the Clinician-Administered Dissociative States Scale (CADSS). *J Trauma Stress* 11: 125–136
- Broadbent D (1958) Perception and Communication. London: Pergamon
- Brown P L, Jenkins H M (1968) Auto-shaping of the pigeon's key-peck. *J Exp Anal Behav* 11: 1–8
- Cameron N (1939) Schizophrenic thinking in a problem-solving situation. *Journal of Mental Science* 85: 1012–1035
- Carlsson A, Lindqvist M (1963) Effect of chlorpromazine or haloperidol on formation of 3-methoxytyramine and normetanephrine in mouse brain. *Acta Pharmacol Toxicol (Copenh)* 20: 140–144
- Carlsson M, Carlsson A (1990) Interactions between glutamatergic and monoaminergic systems within the basal ganglia—implications for schizophrenia and Parkinson's disease. *Trends Neurosci* 13: 272–276
- Chapman J (1966) The early symptoms of schizophrenia. *Br J Psychiatry* 112: 225–251
- Chiba A A, Buccini D J, Holland P C, Gallagher M (1995) Basal forebrain cholinergic lesions disrupt increments but not decrements in conditioned stimulus processing. *J Neurosci* 15: 7315–7322
- Cohen J D, Servan-Schreiber D (1992) Context, cortex, and dopamine: a connectionist approach to behavior and biology in schizophrenia. *Psychol Rev* 99: 45–77
- Connell P (1958) Amphetamine Psychosis. London: Chapman & Hall
- Conrad K (1958) Die Beginnende Schizophrenie. Stuttgart: G. Thieme
- Corbett R, Camacho F, Woods A T, Kerman L L, Fishkin R J, Brooks K, Dunn R W (1995) Antipsychotic agents antagonize non-competitive N-methyl-D-aspartate antagonist-induced behaviors. *Psychopharmacology (Berl)* 120: 67–74
- 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 reevaluation of causal associations in humans: fMRI evidence in favor of an associative model of learning. *Neuron* 44: 877–888
- Corlett P R, Honey G D, Aitken M R F, 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
- Creese I, Burt D R, Snyder S H (1976) Dopamine receptor binding predicts clinical and pharmacological potencies of antischizophrenic drugs. *Science* 192: 481–483
- D'Souza D C, Perry E, MacDougall L, Ammerman Y, Cooper T, Wu Y T, Braley G, Gueorguieva R, Krystal J H (2004) The psychotomimetic effects of intravenous delta-9-tetrahydrocannabinol in healthy individuals: implications for psychosis. *Neuropsychopharmacology* 29: 1558–1572
- Dalley J W, Laane K, Theobald D E, Armstrong H C, Corlett P R, Chudasama Y, Robbins T W (2005) Time-limited modulation of appetitive Pavlovian memory by D1 and NMDA receptors in the nucleus accumbens. *Proc Natl Acad Sci USA* 102: 6189–6194
- Dayan P, Kakade S, Montague P R (2000) Learning and selective attention. *Nat Neurosci* 3(Suppl): 1218–1223
- Del Arco A, Mora F (2000) Endogenous dopamine potentiates the effects of glutamate on extracellular GABA in the prefrontal cortex of the freely moving rat. *Brain Res Bull* 53: 339–345
- Di Ciano P, Cardinal R N, Cowell R A, Little S J, Everitt B J (2001) Differential involvement of NMDA, AMPA/kainate, and dopamine receptors in the nucleus accumbens core in the acquisition and performance of pavlovian approach behavior. *J Neurosci* 21: 9471–9477
- Dickinson A (2001) The 28th Bartlett Memorial Lecture. Causal learning: an associative analysis. *Q J Exp Psychol B* 54: 3–25
- Dreher J C, Kohn P, Berman K F (2005) Neural coding of distinct statistical properties of reward information in humans. *Cereb Cortex*
- Einhorn H J, Hogarth R M (1986) Judging probable cause. *Psychological Bulletin* 99: 3–19
- Escobar M, Oberling P, Miller R R (2002) Associative deficit accounts of disrupted latent inhibition and blocking in schizophrenia. *Neurosci Biobehav Rev* 26: 203–216
- Everitt B J, Robbins T W (1997) Central cholinergic systems and cognition. *Annu Rev Psychol* 48: 649–684
- Feenstra M G, Botterblom M H (1996) Rapid sampling of extracellular dopamine in the rat prefrontal cortex during food consumption, handling and exposure to novelty. *Brain Res* 742: 17–24
- Feenstra M G, Botterblom M H, Mastenbroek S (2000) Dopamine and noradrenaline efflux in the prefrontal cortex in the light and dark period: effects of novelty and handling and comparison to the nucleus accumbens. *Neuroscience* 100: 741–748
- Feenstra M G, van der Weij W, Botterblom M H (1995) Concentration-dependent dual action of locally applied N-methyl-D-aspartate on extracellular dopamine in the rat prefrontal cortex in vivo. *Neurosci Lett* 201: 175–178
- Fibiger H C, Phillips A G, Zis A P (1974) Deficits in instrumental responding after 6-hydroxydopamine lesions of the nigro-neostriatal dopaminergic projection. *Pharmacol Biochem Behav* 2: 87–96
- Finlay J M, Zigmond M J (1997) The effects of stress on central dopaminergic neurons: possible clinical implications. *Neurochem Res* 22: 1387–1394
- Finlay J M, Zigmond M J, Abercrombie E D (1995) Increased dopamine and norepinephrine release in medial prefrontal cortex induced by acute and chronic stress: effects of diazepam. *Neuroscience* 64: 619–628
- Fiorillo C D, Tobler P N, Schultz W (2003) Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* 299: 1898–1902
- Fiorillo C D, Tobler P N, Schultz W (2005) Evidence that the delay-period activity of dopamine neurons corresponds to reward uncertainty rather than backpropagating TD errors. *Behav Brain Funct* 1: 7
- 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

- Floresco S B, Blaha C D, Yang C R, Phillips A G (2001) Dopamine D1 and NMDA receptors mediate potentiation of basolateral amygdala-evoked firing of nucleus accumbens neurons. *J Neurosci* 21: 6370–6376
- Freedman B, Chapman L J (1973) Early subjective experience in schizophrenic episodes. *J Abnorm Psychol* 82: 46–54
- Freedman B J (1974) The subjective experience of perceptual and cognitive disturbances in schizophrenia. A review of autobiographical accounts. *Arch Gen Psychiatry* 30: 333–340
- Freeman A S, Bunney B S (1984) The effects of phencyclidine and N-allylnormetazocine on midbrain dopamine neuronal activity. *Eur J Pharmacol* 104: 287–293
- Friston K (2005) Disconnection and cognitive dysmetria in schizophrenia. *Am J Psychiatry* 162: 429–432
- Frith C D, Blakemore S J, Wolpert D M (2000) Abnormalities in the awareness and control of action. *Philos Trans R Soc Lond B Biol Sci* 355: 1771–1788
- Fuster J M (1985) The prefrontal cortex, mediator of cross-temporal contingencies. *Human Neurobiology* 4: 169–179
- Gal G, Schiller D, Weiner I (2005) Latent inhibition is disrupted by nucleus accumbens shell lesion but is abnormally persistent following entire nucleus accumbens lesion: The neural site controlling the expression and disruption of the stimulus preexposure effect. *Behav Brain Res* 162: 246–255
- Goldman-Rakic P S (1994) Working memory dysfunction in schizophrenia. *J Neuropsychiatry Clin Neurosci* 6: 348–357
- Gray J A (1993) Consciousness, schizophrenia and scientific theory. *Ciba Found Symp* 174: 263–273; discussion 273–281
- Gray J A (1995) Dopamine release in the nucleus accumbens: the perspective from aberrations of consciousness in schizophrenia. *Neuropsychologia* 33: 1143–1153
- Gray J A (1998a) Abnormal contents of consciousness: the transition from automatic to controlled processing. *Adv Neurol* 77: 195–208; discussion 208–211
- Gray J A (1998b) Integrating schizophrenia. *Schizophr Bull* 24: 249–266
- Gray J A (2004) On biology, phenomenology, and pharmacology in schizophrenia. *Am J Psychiatry* 161: 377; author reply 377–378
- Gray J A, Joseph M H, Hemsley D R, Young A M, Warburton E C, Boulenguez P, Grigoryan G A, Peters S L, Rawlins J N P, Taib C-T, Yee B K, Cassaday H, Weiner I, Gal G, Gusak O, Joel D, Shadach E, Shalev U, Tarrasch R, Feldon J (1995) The role of mesolimbic dopaminergic and retrohippocampal afferents to the nucleus accumbens in latent inhibition: implications for schizophrenia. *Behav Brain Res* 71: 19–31
- Gray N S, Snowden R J (2005) The relevance of irrelevance to schizophrenia. *Neurosci Biobehav Rev* 29: 989–999
- Greengard P (2001) The neurobiology of slow synaptic transmission. *Science* 294: 1024–1030
- Greenwald A G (1980) The totalitarian ego. *American Psychologist* 35: 603–618
- Grossberg S (1982) Processing of expected and unexpected events during conditioning and attention: a psychophysiological theory. *Psychol Rev* 89: 529–572
- Hasler F, Grimberg U, Benz M A, Huber T, Vollenweider F X (2004) Acute psychological and physiological effects of psilocybin in healthy humans: a double-blind, placebo-controlled dose-effect study. *Psychopharmacology (Berl)* 172: 145–156
- Hemsley D R (1992) Cognitive abnormalities and schizophrenic symptoms. *Psychol Med* 22: 839–842
- Hemsley D R (1993) A simple (or simplistic?) cognitive model for schizophrenia. *Behav Res Ther* 31: 633–645
- Hemsley D R (1994) A cognitive model for schizophrenia and its possible neural basis. *Acta Psychiatr Scand Suppl* 384: 80–86
- Hemsley D R (1996) Schizophrenia. A cognitive model and its implications for psychological intervention. *Behav Modif* 20: 139–169
- Hemsley D R (1998) The disruption of the ‘sense of self’ in schizophrenia: potential links with disturbances of information processing. *Br J Med Psychol* 71(Pt 2): 115–124
- Hemsley D R (2005a) The development of a cognitive model of schizophrenia: placing it in context. *Neurosci Biobehav Rev* 29: 977–988
- Hemsley D R (2005b) The schizophrenic experience: taken out of context? *Schizophr Bull* 31: 43–53
- 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
- Hoffman D C (1992) Typical and atypical neuroleptics antagonize MK-801-induced locomotion and stereotypy in rats. *J Neural Transm Gen Sect* 89: 1–10
- Hollerman J R, Schultz W (1998) Dopamine neurons report an error in the temporal prediction of reward during learning. *Nat Neurosci* 1: 304–309
- Hollerman J R, Tremblay L, Schultz W (1998) Influence of reward expectation on behavior-related neuronal activity in primate striatum. *J Neurophysiol* 80: 947–963
- Honey G D, O’Loughlin C, Turner D C, Pomarol-Clotet E, Corlett P R, Fletcher P C (2005) The effects of a subpsychotic dose of ketamine on recognition and source memory for agency: implications for pharmacological modelling of core symptoms of schizophrenia. *Neuropsychopharmacology*
- Honey R A, Turner D C, Honey G D, Sharar S R, Kumaran D, Pomarol-Clotet E, McKenna P, Sahakian B J, Robbins T W, Fletcher P C (2003) Subdissociative dose ketamine produces a deficit in manipulation but not maintenance of the contents of working memory. *Neuropsychopharmacology* 28: 2037–2044
- Honey R A, Honey G D, O’Loughlin C, Sharar S R, Kumaran D, Bullmore E T, Menon D K, Donovan T, Lupson V C, Bisbrown-Chippendale R, Fletcher P C (2004) Acute ketamine administration alters the brain responses to executive demands in a verbal working memory task: an fMRI study. *Neuropsychopharmacology* 29: 1203–1214
- Jones S H, Gray J A, Hemsley D R (1992) Loss of the Kamin blocking effect in acute but not chronic schizophrenics. *Biol Psychiatry* 32: 739–755
- Kamin L J (1969) Predictability, surprise, attention and conditioning. In Church R, Campbell B (eds), *Punishment and Aversive Behavior*. New York: Appleton-Century-Crofts, 279–296
- 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
- Kapur S (2004) How antipsychotics become anti-“psychotic”—from dopamine to salience to psychosis. *Trends Pharmacol Sci* 25: 402–406
- Kapur S, Mizrahi R, Li M (2005) From dopamine to salience to psychosis—linking biology, pharmacology and phenomenology of psychosis. *Schizophr Res* 79: 59–68
- Kapur S, Seeman P (2001) Ketamine has equal affinity for NMDA receptors and the high-affinity state of the dopamine D2 receptor. *Biol Psychiatry* 49: 954–957
- Keefe R S, Arnold M C, Bayen U J, Harvey P D (1999) Source monitoring deficits in patients with schizophrenia; a multinomial modelling analysis. *Psychol Med* 29: 903–914
- Kegeles L S, Abi-Dargham A, Zea-Ponce Y, Rodenhiser-Hill J, Mann J J, Van Heertum R L, Cooper T B, Carlsson A, Laruelle M (2000) Modulation of amphetamine-induced striatal dopamine release by

- ketamine in humans: implications for schizophrenia. *Biol Psychiatry* 48: 627–640
- Kelley A E, Berridge K C (2002) The neuroscience of natural rewards: relevance to addictive drugs. *J Neurosci* 22: 3306–3311
- Kelley A E, Smith-Roe S L, Holahan M R (1997) Response-reinforcement learning is dependent on N-methyl-D-aspartate receptor activation in the nucleus accumbens core. *Proc Natl Acad Sci USA* 94: 12174–12179
- Kihlstrom J F, Hoyt I P (1988) Hypnosis and the psychology of delusions. New York: Wiley
- Kreitschmann-Andermahr I, Rosburg T, Demme U, Gaser E, Nowak H, Sauer H (2001) Effect of ketamine on the neuromagnetic mismatch field in healthy humans. *Brain Res Cogn Brain Res* 12: 109–116
- Krystal J H, Karper L P, Seibyl J P, Freeman G K, Delaney R, Bremner J D, Heninger G R, Bowers M B, Jr., 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
- Krystal J H, Karper L P, Bennett A, D'Souza D C, Abi-Dargham A, Morrissey K, Abi-Saab D, Bremner J D, Bowers M B, Jr., Suckow R F, Stetson P, Heninger G R, Charney D S (1998) Interactive effects of subanesthetic ketamine and subhypnotic lorazepam in humans. *Psychopharmacology (Berl)* 135: 213–229
- Krystal J H, D'Souza D C, Karper L P, Bennett A, Abi-Dargham A, Abi-Saab D, Cassello K, Bowers M B, Jr., Vegso S, Heninger G R, Charney D S (1999) Interactive effects of subanesthetic ketamine and haloperidol in healthy humans. *Psychopharmacology (Berl)* 145: 193–204
- Krystal J H, Bennett A, Abi-Saab D, Belger A, Karper LP, D'Souza DC, Lipschitz D, Abi-Dargham A, Charney DS (2000) Dissociation of ketamine effects on rule acquisition and rule implementation: possible relevance to NMDA receptor contributions to executive cognitive functions. *Biol Psychiatry* 47: 137–143
- Krystal J H, Abi-Saab W, Perry E, D'Souza DC, Liu N, Gueorguieva R, McDougall L, Hunsberger T, Belger A, Levine L, Breier A (2005a) Preliminary evidence of attenuation of the disruptive effects of the NMDA glutamate receptor antagonist, ketamine, on working memory by pretreatment with the group II metabotropic glutamate receptor agonist, LY354740, in healthy human subjects. *Psychopharmacology (Berl)* 179: 303–309
- Krystal J H, Perry E B, Jr, Gueorguieva R, Belger A, Madonick S H, Abi-Dargham A, Cooper T B, Macdougall L, Abi-Saab W, D'Souza D C (2005b) Comparative and interactive human psychopharmacologic effects of ketamine and amphetamine: implications for glutamatergic and dopaminergic model psychoses and cognitive function. *Arch Gen Psychiatry* 62: 985–994
- 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
- Ljungberg T, Apicella P, Schultz W (1992) Responses of monkey dopamine neurons during learning of behavioral reactions. *J Neurophysiol* 67: 145–163
- Lubow R E (1965) Latent inhibition: effects of frequency of nonreinforced preexposure of the CS. *J Comp Physiol Psychol* 60: 454–457
- Lubow R E, Gewirtz J C (1995) Latent inhibition in humans: data, theory, and implications for schizophrenia. *Psychol Bull* 117: 87–103
- Lubow R E, Moore A U (1959) Latent inhibition: the effect of nonreinforced pre-exposure to the conditional stimulus. *J Comp Physiol Psychol* 52: 415–419
- Mackintosh N J (1975) A theory of attention: Variations in associability of stimuli with reinforcement. *Psychological Review* 82: 276–298
- Maher B A (1974) Delusional thinking and perceptual disorder. *J Individ Psychol* 30: 98–113
- Maher B A (1988) Delusions as Normal Theories. New York: Wiley
- Maj J, Rogoz Z, Skuza G (1991) Locomotor hyperactivity induced by MK-801 in rats. *Pol J Pharmacol Pharm* 43: 449–458
- Maldonado-Irizarry C S, Kelley A E (1995) Excitatory amino acid receptors within nucleus accumbens subregions differentially mediate spatial learning in the rat. *Behav Pharmacol* 6: 527–539
- Martins Serra A, Jones S H, Toone B, Gray J A (2001) Impaired associative learning in chronic schizophrenics and their first-degree relatives: a study of latent inhibition and the Kamin blocking effect. *Schizophr Res* 48: 273–289
- Mathe J M, Nomikos G G, Schilström B, Svensson T H (1998) Non-NMDA excitatory amino acid receptors in the ventral tegmental area mediate systemic dizocilpine (MK-801) induced hyperlocomotion and dopamine release in the nucleus accumbens. *J Neurosci Res* 51: 583–592
- Matusek P (1952) Studies in delusional perception. In Cutting J, Shepherd M (eds), *The Clinical Roots of the Schizophrenia Concept. Translations of Seminal European Contributions on Schizophrenia*. Cambridge: Cambridge University Press, 89–103
- McClure S M, Berns G S, Montague P R (2003a) Temporal prediction errors in a passive learning task activate human striatum. *Neuron* 38: 339–346
- McClure S M, Daw N D, Montague P R (2003b) A computational substrate for incentive salience. *Trends Neurosci* 26: 423–428
- McGhie A, Chapman J (1961) Disorders of attention and perception in early schizophrenia. *Br J Med Psychol* 34: 103–116
- McKellar P (1957) *Imagination and Thinking*. London: Cohen & West
- 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, Weinberg D R (2007) Genetic evidence implicating DARPP-32 in human frontostriatal structure, function, and cognition. *J Clin Invest* 117: 672–682
- Miller E K, Cohen J D (2001) An integrative theory of prefrontal cortex function. *Annu Rev Neurosci* 24: 167–202
- Miller R (1976) Schizophrenic psychology, associative learning and the role of forebrain dopamine. *Med Hypotheses* 2: 203–211
- Miller R (1989) Hyperactivity of associations in psychosis. *Aust N Z J Psychiatry* 23: 241–248
- Miller R (1993) Striatal dopamine in reward and attention: a system for understanding the symptomatology of acute schizophrenia and mania. *Int Rev Neurobiol* 35: 161–278
- Mineka S, Kihlstrom J F (1978) Unpredictable and uncontrollable events: a new perspective on experimental neurosis. *J Abnorm Psychol* 87: 256–271
- Mitcham J C, Thomas R K, Jr (1972) Effects of substantia nigra and caudate nucleus lesions on avoidance learning in rats. *J Comp Physiol Psychol* 81: 101–107
- Moghaddam B, Adams B, Verma A, Daly D (1997) Activation of glutamatergic neurotransmission by ketamine: a novel step in the pathway from NMDA receptor blockade to dopaminergic and cognitive disruptions associated with the prefrontal cortex. *J Neurosci* 17: 2921–2927
- Montague P R, Hyman S E, Cohen J D (2004) Computational roles for dopamine in behavioural control. *Nature* 431: 760–767
- Moran P M, Al-Uzri M M, Watson J, Reveley M A (2003) Reduced Kamin blocking in non paranoid schizophrenia: associations with schizotypy. *J Psychiatr Res* 37: 155–163
- Morgan C J, Mofeez A, Brandner B, Bromley L, Curran H V (2004) Acute effects of ketamine on memory systems and psychotic symptoms in healthy volunteers. *Neuropsychopharmacology* 29: 208–218
- Murase S, Mathe J M, Grenhoff J, Svensson T H (1993) Effects of dizocilpine (MK-801) on rat midbrain dopamine cell activity: differential actions on firing pattern related to anatomical localization. *J Neural Transm Gen Sect* 91: 13–25

- Nelson C L, Burk J A, Bruno J P, Sarter M (2002) Effects of acute and repeated systemic administration of ketamine on prefrontal acetylcholine release and sustained attention performance in rats. *Psychopharmacology (Berl)* 161: 168–179
- Newcomer J W, Farber N B, Jevtovic-Todorovic V, Selke G, Melson A K, Hershey T, Craft S, Olney J W (1999) Ketamine-induced NMDA receptor hypofunction as a model of memory impairment and psychosis. *Neuropsychopharmacology* 20: 106–118
- Nicoll R A (2003) Expression mechanisms underlying LTP: a post-synaptic view. *Philos Trans R Soc Lond B Biol Sci* 358: 721–726
- Nyback H, Sedvall G (1970) Further studies on the accumulation and disappearance of catecholamines formed from tyrosine-14C in mouse brain. Effect of some phenothiazine analogues. *Eur J Pharmacol* 10: 193–205
- O'Doherty J P, Dayan P, Friston K, Critchley H, Dolan R J (2003) Temporal difference models and reward-related learning in the human brain. *Neuron* 38: 329–337
- O'Reilly R C, Braver T S, Cohen J D (1999) A biologically-based neural network model of working memory. In Shah P, Miyake A (eds), *Models of Working Memory*. Cambridge: Cambridge University Press, 375–411
- Oades R D, Bunk D, Eggers C (1992) Paranoid schizophrenics may not use irrelevant signals. The use of measures of blocking and of urinary dopamine. *Acta Paedopsychiatr* 55: 183–184
- Oades R D, Zimmermann B, Eggers C (1996) Conditioned blocking in patients with paranoid, non-paranoid psychosis or obsessive compulsive disorder: associations with symptoms, personality and monoamine metabolism. *J Psychiatr Res* 30: 369–390
- Oranje B, van Berckel B N, Kemner C, van Ree J M, Kahn R S, Verbaten M N (2000) The effects of a sub-anaesthetic dose of ketamine on human selective attention. *Neuropsychopharmacology* 22: 293–302
- Osmond H (1957) A review of the clinical effects of psychotomimetic agents. *Ann N Y Acad Sci* 66: 418–434
- 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
- Passie T, Karst M, Wiese B, Emrich H M, Schneider U (2005) Effects of different subanesthetic doses of (S)-ketamine on neuropsychology, psychopathology, and state of consciousness in man. *Neuropsychobiology* 51: 226–233
- 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
- Reed G F (1972) *The Psychology of Anomalous Experience: A Cognitive Approach*. London: Hutchinson
- 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*. New York: Appleton-Century-Crofts
- Rodriguez P F, Aron A R, Poldrack R A (2005) Ventral-striatal/nucleus-accumbens sensitivity to prediction errors during classification learning. *Hum Brain Mapp*
- Romo R, Schultz W (1990) Dopamine neurons of the monkey midbrain: contingencies of responses to active touch during self-initiated arm movements. *J Neurophysiol* 63: 592–606
- Rowland L M (2005) Subanesthetic ketamine: how it alters physiology and behavior in humans. *Aviat Space Environ Med* 76: C52–C58
- Rowland L M, Astur R S, Jung R E, Bustillo J R, Lauriello J, Yeo R A (2005) Selective cognitive impairments associated with NMDA receptor blockade in humans. *Neuropsychopharmacology* 30: 633–639
- Sarter M, Bruno J P (1999) Abnormal regulation of corticopetal cholinergic neurons and impaired information processing in neuropsychiatric disorders. *Trends Neurosci* 22: 67–74
- Sarter M, Nelson C L, Bruno J P (2005) Cortical cholinergic transmission and cortical information processing in schizophrenia. *Schizophr Bull* 31: 117–138
- Schneider C (1930) *Die Psychologie der Schizophrenen*. Leipzig: Germany Thieme
- Schott B H, Sellner D B, Lauer C J, Habib R, Frey J U, Guderian S, Heinze H J, Duzel E (2004) Activation of midbrain structures by associative novelty and the formation of explicit memory in humans. *Learn Mem* 11: 383–387
- Schott B H, Seidenbecher C I, Fenker D B, Lauer C J, Bunzeck N, Bernstein H G, Tischmeyer W, Gundelfinger E D, Heinze H J, Duzel E (2006) The dopaminergic midbrain participates in human episodic memory formation: evidence from genetic imaging. *J Neurosci* 26: 1407–1417
- Schultz W, Apicella P, Ljungberg T (1993a) 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, Apicella P, Ljungberg T, Romo R, Scarnati E (1993b) Reward-related activity in the monkey striatum and substantia nigra. *Prog Brain Res* 99: 227–235
- Schultz W, Dayan P, Montague P R (1997) A neural substrate of prediction and reward. *Science* 275: 1593–1599
- Schultz W, Tremblay L, Hollerman J R (1998) Reward prediction in primate basal ganglia and frontal cortex. *Neuropharmacology* 37: 421–429
- Schultz W (1998a) The phasic reward signal of primate dopamine neurons. *Adv Pharmacol* 42: 686–690
- Schultz W (1998b) Predictive reward signal of dopamine neurons. *J Neurophysiol* 80: 1–27
- Schultz W (1999) The reward signal of midbrain dopamine neurons. *News Physiol Sci* 14: 249–255
- Schultz W, Dickinson A (2000) Neuronal coding of prediction errors. *Annu Rev Neurosci* 23: 473–500
- Schultz W, Tremblay L, Hollerman J R (2000) Reward processing in primate orbitofrontal cortex and basal ganglia. *Cereb Cortex* 10: 272–284
- Schultz W (2001) Reward signaling by dopamine neurons. *Neuroscientist* 7: 293–302
- Schultz W (2002) Getting formal with dopamine and reward. *Neuron* 36: 241–263
- Scott L, Kruse M S, Forssberg H, Brismar H, Greengard P, Aperia A (2002) Selective up-regulation of dopamine D1 receptors in dendritic spines by NMDA receptor activation. *Proc Natl Acad Sci USA* 99: 1661–1664
- Seeman P, Chau-Wong M, Tedesco J, Wong K (1976) Dopamine receptors in human and calf brains, using [3H]apomorphine and an antipsychotic drug. *Proc Natl Acad Sci USA* 73: 4354–4358
- Sitges M, Nekrassov V, Guarneros A (2000) Simultaneous action of MK-801 (dizclopine) on dopamine, glutamate, aspartate and GABA release from striatum isolated nerve endings. *Brain Res* 854: 48–56
- Smith-Roe S L, Kelley A E (2000) Coincident activation of NMDA and dopamine D1 receptors within the nucleus accumbens core is required for appetitive instrumental learning. *J Neurosci* 20: 7737–7742
- Smith G S, Schloesser R, Brodie J D, Dewey SL, Logan J, Vitkun SA, Simkowitz P, Hurley A, Cooper T, Volkow ND, Cancro R (1998) Glutamate modulation of dopamine measured in vivo with positron emission tomography (PET) and 11C-raclopride in normal human subjects. *Neuropsychopharmacology* 18: 18–25
- Stephan K E, Baldeweg T, Friston K J (2006) Synaptic plasticity and dysconnection in schizophrenia. *Biol Psychiatry*
- Svenningsson P, Zavara 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

- Svensson T H, Mathe J M, Nomikos G G, Schilström B (1998) Role of excitatory amino acids in the ventral tegmental area for central actions of non-competitive NMDA-receptor antagonists and nicotine. *Amino Acids* 14: 51–56
- Taber M T, Fibiger H C (1997) Activation of the mesocortical dopamine system by feeding: lack of a selective response to stress. *Neuroscience* 77: 295–298
- Taylor S E, Fiske S T (1978) *Saliency, Attention and Attribution: Top of the Head Phenomena*, Vol 11. San Diego: CA: Academic Press
- Tobler P N, O'Doherty J P, Dolan R J, Schultz W (2005) Human neural learning depends on reward prediction errors in the blocking paradigm. *J Neurophysiol*
- 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
- Umbrecht D, Koller R, Vollenweider F X, Schmid L (2002) Mismatch negativity predicts psychotic experiences induced by NMDA receptor antagonist in healthy volunteers. *Biol Psychiatry* 51: 400–406
- Umbrecht D, Schmid L, Koller R, Vollenweider F X, Hell D, Javitt D C (2000) Ketamine-induced deficits in auditory and visual context-dependent processing in healthy volunteers: implications for models of cognitive deficits in schizophrenia. *Arch Gen Psychiatry* 57: 1139–1147
- Vaitl D, Lipp O, Bauer U, Schuler G, Stark R, Zimmermann M, Kirsch P (2002) Latent inhibition and schizophrenia: Pavlovian conditioning of autonomic responses. *Schizophr Res* 55: 147–158
- Vollenweider F X, Leenders K L, Oye I, Hell D, Angst J (1997a) Differential psychopathology and patterns of cerebral glucose utilisation produced by (S)- and (R)-ketamine in healthy volunteers using positron emission tomography (PET). *Eur Neuropsychopharmacol* 7: 25–38
- Vollenweider F X, Leenders K L, Scharfetter C, Antonini A, Maguire P, Missimer J, Angst J (1997b) Metabolic hyperfrontality and psychopathology in the ketamine model of psychosis using positron emission tomography (PET) and [18F] fluorodeoxyglucose (FDG). *Eur Neuropsychopharmacol* 7: 9–24
- Vollenweider F X, Vollenweider-Scherpenhuyzen M F, Babler A, Vogel H, Hell D (1998) Psilocybin induces schizophrenia-like psychosis in humans via a serotonin-2 agonist action. *Neuroreport* 9: 3897–3902
- Vollenweider F X, Vontobel P, Oye I, Hell D, Leenders K L (2000) Effects of (S)-ketamine on striatal dopamine: a [¹¹C]raclopride PET study of a model psychosis in humans. *J Psychiatr Res* 34: 35–43
- Waelti P, Dickinson A, Schultz W (2001) Dopamine responses comply with basic assumptions of formal learning theory. *Nature* 412: 43–48
- Watanabe M, Kodama T, Hikosaka K (1997) Increase of extracellular dopamine in primate prefrontal cortex during a working memory task. *J Neurophysiol* 78: 2795–2798
- Wedzony K, Golembiowska K, Zazula M (1994) Differential effects of CGP 37849 and MK-801, competitive and noncompetitive NMDA antagonists, with respect to the modulation of sensorimotor gating and dopamine outflow in the prefrontal cortex of rats. *Naunyn Schmiedeberg Arch Pharmacol* 350: 555–562
- Weinberger D R, Gallhofer B (1997) Cognitive function in schizophrenia. *Int Clin Psychopharmacol* 12(Suppl 4): S29–S36
- Whitton P S, Biggs C S, Pearce B R, Fowler L J (1992) Regional effects of MK-801 on dopamine and its metabolites studied by in vivo microdialysis. *Neurosci Lett* 142: 5–8
- Willins D L, Narayanan S, Wallace L J, Uretsky N J (1993) The role of dopamine and AMPA/kainate receptors in the nucleus accumbens in the hypermotility response to MK801. *Pharmacol Biochem Behav* 46: 881–887
- Wing J K, Cooper, J E, Sartorius N (1974) *The Measurement and Classification of Psychiatric Symptoms*. Cambridge: Cambridge University Press
- Yogev H, Sirota P, Gutman Y, Hadar U (2004) Latent inhibition and overswitching in schizophrenia. *Schizophr Bull* 30: 713–726
- Young A M, Kumari V, Mehrotra R, Hemsley DR, Andrew C, Sharma T, Williams SC, Gray JA (2005) Disruption of learned irrelevance in acute schizophrenia in a novel continuous within-subject paradigm suitable for fMRI. *Behav Brain Res* 156: 277–288
- Young D, Scoville W B (1938) Paranoid psychosis in narcolepsy and the possible danger of benzedrine treatment. *The Medical Clinics of North America* 22: 637–646
- Yu A J, Dayan P (2002) Acetylcholine in cortical inference. *Neural Netw* 15: 719–730
- Yu A J, Dayan P (2005) Uncertainty, neuromodulation, and attention. *Neuron* 46: 681–692
- Zhang J, Chiodo L A, Freeman A S (1992) Electrophysiological effects of MK-801 on rat nigrostriatal and mesoaccumbal dopaminergic neurons. *Brain Res* 590: 153–163