Review article

Comprehensive review: Computational modelling of schizophrenia

Vincent Valton[a,b,1], Liana Romaniuk[b], J. Douglas Steele[c], Stephen Lawrie[b], Peggy Seriès[a,*]

[a] Institute for Adaptive and Neural Computation, Edinburgh University, UK
[b] Division of Psychiatry, Royal Edinburgh Hospital, Edinburgh University, UK
[c] School of Medicine (Neuroscience), Dundee University, UK

ARTICLE INFO

Keywords:
Psychotic symptoms
Schizophrenia
Computational models
Computational psychiatry

ABSTRACT

Computational modelling has been used to address: (1) the variety of symptoms observed in schizophrenia using abstract models of behavior (e.g. Bayesian models – top-down descriptive models of psychopathology); (2) the causes of these symptoms using biologically realistic models involving abnormal neuromodulation and/or receptor imbalance (e.g. connectionist and neural networks – bottom-up realistic models of neural processes). These different levels of analysis have been used to answer different questions (i.e. understanding behavioral vs. neurobiological anomalies) about the nature of the disorder. As such, these computational studies have mostly supported diverging hypotheses of schizophrenia's pathophysiology, resulting in a literature that is not always expanding coherently. Some of these hypotheses are however ripe for revision using novel empirical evidence.

Here we present a review that first synthesizes the literature of computational modelling for schizophrenia and psychotic symptoms into categories supporting the dopamine, glutamate, GABA, dysconnection and Bayesian inference hypotheses respectively. Secondly, we compare model predictions against the accumulated empirical evidence and finally we identify specific hypotheses that have been left relatively under-investigated.

1. Background

Schizophrenia is a psychiatric disorder with a lifetime prevalence of 0.3–0.66% (Bhugra, 2005; van Os and Kapur, 2009). This condition manifests itself through a variety of symptoms across patients, classified into three distinct categories: positive, negative and cognitive symptoms. Positive symptoms refer to hallucinations (i.e. vivid perceptions of complex stimuli, such as hearing voices or seeing objects/people, in the absence of an external stimulus), and delusions (i.e. persistent false beliefs maintained despite being contradicted by reality or rational evidence and out of keeping with the individual's socio-cultural norms). Negative symptoms include flattened affect, social withdrawal, apathy, poverty of speech, and anhedonia. Cognitive deficits cover decreased memory performance, attentional and reasoning deficit, which is usually associated with an average IQ drop of about 10 points following the disease onset (Bhugra, 2005; Frith et al., 1991; Johnstone et al., 1991; McIntosh et al., 2005; van Os and Kapur, 2009). Schizophrenia is highly debilitating, leading to an average loss of 15–20 years of life expectancy when compared to the general population (Andrew et al., 2012; Mangalore and Knapp, 2007). It is argued that unhealthy lifestyles and increased suicidal rates (found to be about 12 times higher in schizophrenia; Caldwell and Gottesman, 1990) might account for this

general reduction in life expectancy (World Health Organisation, 1996).

Besides this devastating prospect for patients and their relatives, schizophrenia has also been found to generate a high economical burden on society (Knapp et al., 2004; Mangalore and Knapp, 2007; Serretti et al., 2009). Recently, the total societal cost of schizophrenia has been estimated to be around £6.7 to £11.8 billion per year for England alone (Andrew et al., 2012; Mangalore and Knapp, 2007). This is including direct treatment costs and indirect societal costs such as loss of employment. In fact, it has been estimated that around 80–93% of patients with schizophrenia remain unemployed, leading to large societal costs due to loss productivity (Andrew et al., 2012; Mangalore and Knapp, 2007). Lack of employment is argued to result largely from cognitive deficits, problems of attention and working memory (Insel, 2010). However, it is worth noting that negative symptoms, which include amotivation, anhedonia and apathy, are associated with social functioning impairments and as such, could also potentially contribute to the unemployment status observed in patients (Hoffmann and Kupper, 1997; Lysaker and Bell, 1995; Solinski et al., 1992; Suslow et al., 2000; Weinberg et al., 2009).

Unfortunately, there is currently no cure for schizophrenia, mainly due to a poor understanding of the causes and mechanisms of the

disorder. The best treatment to date consists of managing the occurrence of positive symptoms through a combination of anti-psychotic medications and psychosocial treatments. These treatments aims to minimize symptoms, potential risks to the patient or others (e.g. hallucinations/delusions leading to self-neglect or harm), and to avoid the relapse of psychosis. It is estimated that about 45% of patients recover after one or more episodes, 20% show a gradual worsening of symptoms and a final 35% exhibit a mix of remission with a worsening of some of the symptoms (relapsing-remitting; World Health Organisation, 1996).

1.1. Schizophrenia – diverging hypotheses

Several studies have identified neuroanatomical differences in patients (e.g. Kreczmanski et al., 2007; Lawrie et al., 2008; Seeman, 1994) as well as susceptible genes increasing the risk of developing psychiatric disorders (e.g. Chubb et al., 2008). However, while it is well established that genetic risk factors alone are not sufficient to account for the development of the disorder (Lawrie et al., 2008); it is widely accepted that an interaction between genetic (Berry et al., 2003; Bertolino and Blasi, 2009; Chubb et al., 2008) and environmental risk factors (i.e. stress, traumatic experiences, etc. Jones et al., 1994; McDonald and Murray, 2000; Mortensen et al., 1999) are necessary to lead to the emergence of schizophrenia. So far, research in the field has identified various differences between patients and healthy controls, which has led to divergent – although not mutually exclusive – hypotheses about the origins of the disorder. First, the Dopamine (DA) hypothesis was established through the observation of alleviated positive symptoms upon treatment with typical anti-psychotic drugs (APD), which block dopamine receptors D2 (D2r). Consistent with this hypothesis, subsequent imaging studies found elevated dopaminergic signalling (Meyer-Lindenberg et al., 2005; Murray et al., 2008; Waltz et al., 2009), elevated presynaptic striatal DA synthesis and release, and increased striatal D2 receptor densities (Howes and Kapur, 2009). More recently studies have also found deregulated D1 receptor densities in the pre-frontal regions of patients (Howes and Kapur, 2009). The second hypothesis, the glutamate (Glu) hypothesis emerged from the observation of induced psychosis in healthy subjects when exposed to psychoactive drugs, such as ketamine and phencyclidine (PCP), which acts primarily by blocking the glutamate binding sites of NMDA receptors (Corlett et al., 2007a; van Os and Kapur, 2009). Post-mortem studies also identified reduced glutamate levels in the pre-frontal areas of patients (Sherman et al., 1991). It is therefore plausible that reduced NMDA receptor densities or receptor hypo-function can account for the symptomatology observed in patients (Gilmour et al., 2012; Javitt and Zukin, 1991; Olney et al., 1999). The third hypothesis, the GABAergic hypothesis is supported by experimental studies reporting reduced cortical GABA, dysfunctional activity and reduced markers of inhibitory inter-neurons in the pre-frontal areas of patients (Lewis and Hashimoto, 2005; Nakazawa et al., 2012; Tanaka, 2008). Finally, the dysconnection hypothesis stemmed from several findings of reduced cortical volume, abnormal pre-frontal cortical folding, enlarged ventricles, abnormal synaptic connectivity (Harrison, 1999; Lawrie et al., 2008) and increased cortical activation during cognitive tasks (Manoach et al., 1999; Winterer and Weinberger, 2004). This increased activation is thought to be the result of a reduced synchrony or dysconnection between different cortical areas (Friston, 2005; Stephan et al., 2009), therefore requiring increased effort during completion of cognitive tasks.

Whilst the origins of the disorder are highly debated and led to diverging hypotheses of schizophrenia, researchers and practitioners alike tend to agree that until reliable biological markers are found – which can robustly and reliably predict the emergence of schizophrenia and its symptoms – the best course of action for current diagnostic purposes is to rely on clinical interviews and an interpretation of symptoms by trained professionals.

1.2. Categorical vs. dimensional diagnosis

In the absence of reliable biological markers, diagnosis of mental disorders is obtained from a clinical examination of the symptoms and behaviors expressed by patients (World Health Organisation, 1996). Using the diagnostic and statistical manual of mental disorders (DSM; DSM-IV-TR, 2000) or the international statistical classification of disease and related health problems (ICD-10 Chapter V; World Health Organisation, 2009), clinicians diagnose a patient's illness from the number of symptoms present and the duration of these symptoms. Specifically, for schizophrenia, the current DSM (DSM-5, 2013) criteria for diagnosis is met when two or more symptoms are present continuously for a period of one month or more, and had an impact on the patient's functioning for at least 6 months. The first symptom being either: delusions, hallucinations or disorganized speech, while the second symptom can be any negative or cognitive symptoms causing social or occupational dysfunction. That is, positive symptoms remain the predominant criteria necessary for the diagnosis of schizophrenia.

Recently however there has been an attempt to bridge the gap between categorical diagnoses based on the clinical consensus of symptoms and the identification of potential biological markers identified by neuroscience research (Insel et al., 2010). For example, the research domain criteria (RDoC), aims to develop a precision medicine approach (or personalized, i.e. that takes into account individual variability in genes, environment, and lifestyle for each person) to mental disorders based on behavioral and neurobiological markers (Cuthbert and Insel, 2013; Insel et al., 2010). More importantly, the RDoC proposes to cut across the typical categorical boundaries delineating current mental disorders and instead investigate the variations present in mental illness as belonging to a dimensional continuum (Cuthbert and Insel, 2013). For example, using the semi-structured Present State Examination (PSE), Strauss (1969) identified that psychotic experiences lie on a continuum of intensity in psychotic patients rather than being simply either present or absent. Using the PSE as a template, Peters et al. (1999) then developed the Peters delusion inventory, so as to measure non-clinical delusional ideation in the general population (Johns and van Os, 2001; Peters et al., 2004). Using this scale, delusions have since then been found to be present in the general population on multiple occasions (Corlett and Fletcher, 2012; Freeman et al., 2008; Schmack et al., 2013), albeit to a milder degree than those present in patients with schizophrenia. This growing body of evidence suggests that psychotic experiences might lie on a continuum (Allardyce et al., 2007; Corlett and Fletcher, 2012; David, 2010; Johns and van Os, 2001; Linscott and van Os, 2010; van Os et al., 2000). It is argued however that such a continuum would be impractical for clinical diagnosis (Lawrie et al., 2010). Similarly, a recent joint consortium between the American Psychological Association (APA), the National Institute for Mental Health (NIMH) and the World Health Organisation (WHO) agreed that while neurobiological parameters are of high importance for future diagnostic systems, according to the current state of knowledge, it seems more appropriate for use in research than for immediate clinical use (Insel et al., 2010).

1.3. Computational modelling

While experimental studies provide valuable information to understand the abnormal biological and cognitive processes in schizophrenia, experimental work alone is often limited by ethical, economic or practical factors. Recently, computational and mathematical models have shown to be very useful research tools for the exploration of neural computation, and understanding of the interaction between neural systems and functions (Montague et al., 2012, 2004). Specifically, Marr (1982) proposed that computational models may be used to investigate three distinct although complementary levels of analysis, namely the computational level (*"What"* does the brain compute, and *"why"*?), the algorithmic level (*"Which"* representations and algorithms

can describe these computations?) and the physical level (*"How"* are these algorithms implemented neurally?; Dayan and Abbott, 2005). These different levels of analyses typically lead to two categories of models: top-down and bottom-up models, which – although not incompatible – generally attempt to answer different scientific questions. That is, depending on the main scientific questions being investigated, the primary strength and predictive power of a model resides on their ability to accurately and realistically model the main variable of interest (e.g. neural dysfunction or behavioral deficit). Top-down models usually start from the computational level, for example, the behavioral phenomena and can remain only descriptive in nature (e.g. Bayesian models of perception; Colombo and Series, 2012), while bottom-up models start from the physical substrate and aim at acquiring a mechanistic understanding of how neural computations and processes are performed and give rise to behavior.

By integrating data from diverse experimental studies and levels of description, top-down and bottom-up models can offer a concise and formal description of a phenomenon, shed light on the underlying mechanisms and make predictions leading to novel experimental tests and hypotheses (Dayan and Durstewitz, 2016; Huys et al., 2011).

1.3.1. Computational psychiatry

Computational psychiatry is a young field in expansion at the intersection between computational neuroscience and psychiatry (Huys et al., 2011, 2016; Huys, 2013; Montague et al., 2012). This discipline builds on the initial effort in the 80s using connectionist models, but has also evolved to get closer to the physiological substrate and to more testable predictions (Huys, 2013; Montague et al., 2012). Although psychiatric disorders are characterized essentially by their high-level symptoms, following principles, computational models can help formalize symptoms and hypotheses to bridge the gap between neurobiology and psychiatry (Huys et al., 2011). That is, computational models are able to provide a normative framework to explicitly define and rigorously test competing hypotheses of mental disorders (Huys et al., 2011), while providing a link between different levels of descriptions (Huys, 2013).

For example, Maia and Frank (2011) illustrated how modelling using a deductive or abductive approach can lead to different predictions for psychiatry. Using the deductive approach scientists start from the premise of known neurobiological deficits observed in mental disorders, and implement these deficits in a computational model. The performance of the model is then compared to those of patients. If the model can account for the performance deficits observed in patients, it provides a plausible mechanistic account that bridges biological abnormalities to behavior or neural activity (Maia and Frank, 2011). The abductive approach, on the other hand, starts from the premise of a model of normal behavior and alter the model in multiple ways to generate distinct novel hypotheses of brain dysfunction. All these models are then fitted to the performance of patients to find which hypothesis (different models) accounts best for the patients' performance (Maia and Frank, 2011). The winning hypothesis can then be refined in an attempt to explain the deficits at lower levels of description, or used to devise new experimental tests that will precisely assay the dysfunction suggested by the winning hypothesis. However, this strategy assumes that all the competing hypotheses of dysfunction are tested simultaneously on the same dataset (Ahn and Busemeyer, 2016). Failure to test for all the competing hypotheses could result in conflicting research output, and misleading conclusions (Ahn and Busemeyer, 2016).

Computational techniques have been used not only as a tool to inform on the origins and mechanisms of a disorder, but also for diagnosis and prognostic of treatment efficacy (e.g. machine learning classification techniques – e.g. Chekroud et al., 2016). These latter techniques are promising and could potentially lead to automated classification and "model-based assays" used to diagnose mental disorders (Chekroud et al., 2016; Stephan and Mathys, 2014). In this review, however, we focus on the former type of models that attempt to understand the origins and mechanisms of mental disorders.

2. Questions, aims and methodology

The aim of this review is to synthesize the expanding literature of computational modelling of schizophrenia and psychotic symptoms and address the following questions:

- What predictions have computational models been able to achieve in terms of explaining the mechanisms of psychotic symptoms in schizophrenia?
- How well did these predictions hold up to the accumulated empirical evidence?
- Are there specific hypotheses that have been left relatively underinvestigated?

To extract an exhaustive bibliography of computational models in schizophrenia and psychosis, we used regular expressions to search through the PubMed and Web of Science databases using the following criteria:

Title and/or abstract including: (*"schizo*"* or *"psychos*"* or *"hallucin*"* or *"delusion*"*) and (*"neural?network*?* or *"comput* model*"* or *"model*"* or *"comput*"* or *"framework"*).

Exclusion criteria: We excluded papers that were not in English or peer-reviewed journals. Conference abstracts and animal models without computational modelling were discarded from this analysis, as well as computational models that were not designed to inform on the aetiology or mechanisms of the disorder (e.g. models developed for diagnostic purposes, data analysis or to identify medication interactions). This resulted in a list of more than 100 articles published between 1968 and 2016, comprising all levels of description of the psychopathology of schizophrenia (i.e. the "what?" "how?" and "why?" of computational levels of analysis; Dayan and Abbott, 2005; Marr, 1982).

For the sake of clarity, we have classified hypotheses into the distinct categories of dopamine, glutamate, GABAergic, disconnection and Bayesian inference, so that we could more readily compare similar models against each other, and contrast them against recently accumulated empirical evidence. However, this is not to say that these hypotheses are mutually exclusive and incompatible with each other. On the contrary, a disruption of one of these categories could result in a cascading chain of events (Lewis and Gonzalez-Burgos, 2006; Maia and Frank, 2016) so as to compensate for the dysfunction, leading to downstream up or down-regulation resulting in the symptoms we observe in patients. For example, while dopamine appears to be related to positive symptoms, the genesis of the dopaminergic dysfunction in schizophrenia may very well be the result of upstream glutamatergic and/or GABAergic deficits.

3. The dopamine (D2) hypothesis

3.1. Experimental evidence

The dopamine hypothesis has been popular in the search for aetiological factors of schizophrenia. The hypothesis emerged from the discovery of first generation of anti-psychotic drugs (APD), which relieve patients from positive symptoms by blocking dopamine D2 receptors (D2r). Consistent with these findings, further support originated from the discovery of several psychotomimetic drugs (i.e. such as amphetamines) that can induce psychotic-like episodes in healthy individuals by increasing sub-cortical DA levels (Corlett et al., 2009a; Grace, 1991; Jentsch and Roth, 1999). Over the past decade, the dopamine hypothesis has been supported by various neuroimaging studies reporting increased pre-synaptic dopamine synthesis and storage in the

striatum of acutely psychotic patients (Fusar-Poli and Meyer-Lindenberg, 2013; Howes and Kapur, 2009; Howes and Murray, 2014). These dopamine levels were found to directly correlate with the degree of cognitive deficits and positive symptoms (Howes and Kapur, 2009; Howes and Murray, 2014). Additionally, increases in D2 dopamine receptors densities have been identified in the striatum of patients, together with reduced receptor densities in the thalamus and the anterior cingulate cortex (Howes and Kapur, 2009), although these effects appear to be relatively small (Howes and Murray, 2014). Recent reviews suggest that the influence of striatal D3 receptors in schizophrenia are not significant (Howes and Kapur, 2009; Howes and Murray, 2014), further supporting the role of D2 receptors in psychosis. Consistent with the DA hypothesis, many of the top genetic factors for developing schizophrenia involve genes directly interacting with the dopaminergic pathways (Frank, 2008; Hall et al., 2009; Howes and Kapur, 2009; Winterer and Weinberger, 2004; for a review see Howes et al., 2016). While it is likely that excessive D2r-activation is directly involved in psychosis, scientists are still attempting to link the molecular level anomalies to behavior and positive symptoms. One difficulty is that increased striatal dopaminergic D2r and decreased frontal D1r densities are found to more easily explain cognitive deficits and negative symptoms than positive symptoms (Maia and Frank, 2011).

3.2. Models and support

Within the computational literature supporting the dopamine hypothesis, we identified four main categories of models that support a deficit in dopaminergic transmission, namely: a decreased signal-to-noise ratio (SNR), inappropriate sensory gating, aberrant salience and abnormal reward prediction error (RPE).

3.2.1. Signal-to-noise ratio (SNR) models

Early computational models attempted to explain cognitive deficits in schizophrenia through a generalized decline of the signal-to-noise ratio (SNR) of cortical neurons. Specifically, in these models, DA was thought to function as a signal-to-noise enhancer that modulates neuronal activity by amplifying the neurons' signal while reducing distortions induced by cortical noise.

3.2.1.1. SNR in connectionist frameworks. In artificial neural networks (i.e. interconnected networks of simple processing units called neurons by analogy with the neural system), the signal-to-noise ratio can be altered by changing the neurons' gain or bias parameter (Aakerlund and Hemmingsen, 1998). This directly influences the activation pattern and the stochastic activity of the neurons in the system. SNR models traditionally focused on modelling cognitive symptoms and the performance of patients in tasks where they usually show deficits (e.g. Continuous Performance Task, Stroop Task, Rorschach inkblots, Wisconsin Card Sort Test (WCST), Facial Affect – Amos, 2000; Carter and Neufeld, 2007; Cohen and Servan-Schreiber, 1992, 1993; Jobe et al., 1994; Monchi et al., 2000; Peled and Geva, 2000). In these models, poor performance on cognitive tasks stems from working-memory deficits in units representing the prefrontal cortex (PFC), due to a low signal-to-noise ratio. Through a complete exploration of the parameter space from low to high gain modulation (i.e. hypo-dopaminergic to hyper-dopaminergic states), the models addressed the validity of different dopamine dysfunctions leading to the observed reduced performance. All these models reached the same conclusions, namely that prefrontal DA hypo-function was responsible for the deficient cognitive performance observed in patients (Amos, 2000; Carter and Neufeld, 2007; Cohen and Servan-Schreiber, 1992, 1993; Jobe et al., 1994; Monchi et al., 2000; Peled and Geva, 2000). With respect to working-memory, low DA levels are thought to result in a signal that is easily corrupted by internal cortical noise which in turn becomes incapable of transmitting and maintaining meaningful contextual information about the ongoing task (Cohen and Servan-

Schreiber, 1992, 1993). Another theory suggests that DA hypo-function results in a failure to update task relevant information into WM (Amos, 2000). A deficit in WM updating would then result in a failure to switch to new contextual information, and lead to perseverative behavior (Amos, 2000). Indeed, in switching-tasks such as the WCST where participants are required to infer a sorting rule that changes once it has been correctly acquired, patients are usually able to infer the initial rule but consistently fail to flexibly update it once it has been changed.

A possible criticism of these models is that they can only account for poor cognitive performance following a hypo-dopaminergic state (low SNR). While frontal hypo-dopaminergia is consistent with neuroimaging findings in schizophrenia (Howes and Kapur, 2009), it has been shown experimentally that weak or excessive frontal D1r activation also lead to poor working-memory performance (Vijayraghavan et al., 2007). In the SNR models, however, increasing the gain of neuronal units so as to model a hyper-dopaminergic state leads to a high SNR, which would result in an improvement rather than a deterioration of cognitive performance. Such models thus fail to account for working-memory deficits following frontal hyper-dopaminergia.

3.2.1.2. SNR in attractor networks (cortical stability). Hopfield attractor networks (Hopfield, 1982, 1984) are recurrent artificial neural network of binary units used to model memory storage and retrieval. Such models have also been used to model patients' behavior by adding SNR perturbations to the network. During training, these networks can be made to store specific patterns of activation (memories) by updating the weights of connections between neuronal units. After training, the network can recover an entire memory from a degraded or partial memory input by gradually flowing into the closest pattern of activation (attractor). All the attractors learnt by that network (memories) collectively form the attractor landscape. Such models have usually been used to explain the occurrence of spurious memories (hallucinations; Chen, 1994, 1995; Rolls et al., 2008) or specific aspects of positive symptoms such as the perseverance of delusions (Rolls et al., 2008).

In early models, Spitzer (1995) argued that a hyper-dopaminergic state in cortical networks results in a high SNR, leading to strongly anchored activation of memories encoding high-level constructs such as ideas, concepts and meanings (Spitzer, 1995). Consistent with this hypothesis, Rolls et al. (2008) argued that the perseverance of delusions could be explained by the depth of the basins of attractions in the attractor landscape of the network, where again attractors would correspond to ideas, meanings or an interpretation of the environment. That is, the depth of the basins of attraction would prevent unlearning or switching to new attractors (new ideas or interpretation), leading to a perseverance and an inability to adapt to novel cues from the environment (Rolls et al., 2008). In Hopfield networks, the SNR is modulated by changing the temperature parameter of the neurons, which in turn alters their firing probabilities. A low SNR leads to the inability for the network to recover learnt memories due to a high amount of noise. A high SNR instead results in recurring patterns of activation, irrespective of the original input, or spurious memories, analogous to delusional thoughts or hallucinations (Chen, 1994, 1995). When studying the whole spectrum of temperature changes, Chen (1994, 1995) predicted an inverted-U response profile, whereby intermediate temperatures induced normal behavior and memory retrieval, high temperature resulted in parasitic foci/spurious attractors analogous to hallucinations and delusions (positive symptoms), while low temperature impeded memory retrieval (i.e. cognitive deficits). Interestingly, the inverted-U response profile in working-memory performance was later validated experimentally by electrophysiological recordings of primates' PFC neurons during working-memory tasks (Cools and D'Esposito, 2011; Vijayraghavan et al., 2007).

Recent implementations of attractor networks have reached a high level of biological and physiological detail using integrate-and-fire spiking neurons together with realistic AMPA, GABA, NMDA and DA

pathways (with D1r *vs.* D2r mediated SNR; Rolls et al., 2008). In these studies, GABAergic interneurons inhibit the activity of excitatory neurons that are not encoding the current memory so as to keep the activated memory pattern stable, while NMDA receptors modulate the stochastic firing probabilities of the pyramidal cells. DA modulates the SNR by stabilizing the firing patterns of NMDA and GABA activity, whereby a D1-dominated state increases excitatory and inhibitory activity leading to deeper basins of attraction, while D2-dominated states flatten the energy landscape and facilitates jumps from one attractor to the other. The reduction of excitatory (NMDA) and inhibitory (GABA) activity leads to an impossibility for the network to keep the firing patterns stable, resulting in random jumps between attractors. These random jumps have been argued to be responsible for the positive symptoms and cognitive deficits observed in schizophrenia (Loh et al., 2007; Rolls et al., 2008).

These models make precise and valuable neurophysiological predictions regarding the global inhibitory and excitatory activity of the cortical networks in patients *vs.* that of healthy controls. It would be extremely valuable to be able to test these experimentally. However, such predictions are difficult to test using present neuroscientific tools (typical neuroimaging tools simply do not have the resolution required to monitor the activity of excitatory and inhibitory neurons, while invasive multi-electrode recording can only be used serendipitously in patients undergoing epilepsy surgery). The models presented above suggest that, at the cognitive level, attractors encode high-level constructs such as ideas, meaning, concepts or interpretations. These predictions are also challenging to test experimentally, as high-level constructs are likely to be encoded over a wide array of sparsely interconnected neurons (Huth et al., 2016). However, it may be possible to test some of these neurophysiological predictions and the effect of DA manipulation in vitro using multi-electrode recording of neural cell-cultures derived from patients' induced pluripotent stem cells (Brennand et al., 2011).

3.2.2. Sensory gating models

Sensory gating was the earliest theory of abnormal dopamine function in schizophrenia implemented using computational models (Callaway and Naghdi, 1982; Carr and Wale, 1986). This theory postulates that the brain has to gate relevant information to working-memory and filter-out irrelevant stimuli from all modalities. This mechanism enables subjects to flexibly adapt their behavior to the demands of particular tasks, favouring the processing of task-relevant information over other sources of competing information. This process, also known as cognitive control (Cohen et al., 1996), is thought to be automatic. In these models, the sensory gating process works by preventing task-irrelevant stimuli to access working memory, while maintaining the integrity of task-relevant information against distractors and is assumed to be related to DA signalling (Cohen et al., 1996). Biologically, the gating of relevant information is thought to occur through the simultaneous phasic burst of DA neurons when relevant stimuli are presented, while tonic DA is though to be responsible for the maintenance and protection of working-memory (Tretter and Albus, 2007). In schizophrenia, sensory gating would be disrupted due to inappropriate phasic and tonic dopaminergic signalling, leading to incorrect updates (intrusion of irrelevant stimuli) and maintenance of information (perseveratory behavior). This would finally lead to deficits in attention and cognition (Grace, 1991).

Gating models traditionally used connectionist frameworks to reproduce the performance of healthy controls or the perseveratory behavior of patients at the WCST and CPT, CPT-X tasks (Braver et al., 1999; Braver and Cohen, 1999). In these models, the DA signal exerts a top-down influence on behavior by gating task-relevant information, allowing the update, maintenance and protection against distracting stimuli (Braver et al., 1999; Braver and Cohen, 1999). Such models of working-memory gating converged to similar conclusions, namely, that DA hypo-function was most likely to be responsible for the cognitive

deficits observed in schizophrenia (Braver et al., 1999; Braver and Cohen, 1999). Additionally, a number of descriptive models (i.e. not formalized using computational simulations; Javanbakht, 2005, 2006) also concurred with earlier models, also suggesting that a DA hypo-function would lead to positive symptoms due to weakened top-down behavioral control (Javanbakht, 2005, 2006). Finally, using a connectionist framework of facial affect recognition, Carter and Neufeld (2007) attempted to address a question that is often neglected in the literature: Why are patients with schizophrenia constantly found to exhibit reaction-time deficits in cognitive tasks? In this model, inefficient gating of information led to an overflow of incoming stimuli, resulting in additional processing for task completion. The increased amount of processing leads to a reaction-time escalation, as observed in patients (Carter and Neufeld, 2007). It is worth noting however that these patients were receiving medication, which provides an alternative explanation for the increased processing times. Importantly, increased reaction-time is not specific to schizophrenia; it has also been observed in other psychiatric conditions such as major depressive disorder (Austin et al., 2001).

3.2.3. Aberrant salience model

The aberrant 'motivational salience' hypothesis has its origins in a recent interpretation of the role of DA as signalling rewards associated to stimuli so as to guide behavior (Berridge, 1998; Wise, 1978). An aberrant 'motivational salience' is an incorrect assignment of motivational salience to innocuous stimuli, where DA acts as an indicator of motivation, desire, or attention attributed to a stimulus (Howes and Kapur, 2009; Kapur, 2003). The theory of incentive or 'motivational salience' was first used to explain drug addiction, where inappropriate rewards for drug intake gradually increase the motivational drive to relapse and repeat behavior (Berridge, 1998; Redish et al., 2008; Torregrossa et al., 2011). In schizophrenia, scientists have posited that an aberrant DA signalling would result in incorrect stimulus-reinforcer associations, attributing inappropriate salience to innocuous stimuli (Abboud et al., 2016; Anticevic et al., 2011; Gray et al., 1991; Howes and Kapur, 2009; Kapur, 2003; Roiser et al., 2009, 2013). This inappropriate salience attribution is hypothesized to lead to an increase and perseverance of delusional thinking, even in the face of opposing evidence (Anticevic and Corlett, 2012; Corlett et al., 2009b; Howes and Kapur, 2009; Kapur, 2003). Recent behavioral and neuroimaging experiments appear to confirm the link between aberrant salience, DA signalling and the strength of delusions in schizophrenia patients (Roiser et al., 2009; Romaniuk et al., 2010) and patients at ultra-high risk (UHR) of psychosis (Roiser et al., 2013), but not in patients with long-standing treatment-refractory persistent delusions (Abboud et al., 2016).

Grasemann et al. (2009) and Hoffman et al. (2011) adapted the aberrant salience framework using a connectionist model of story learning and recall to study thought disorder (delusions and derailments). This model mimics the multiple stages of syntax processing, where in each processing stage, artificial neural networks are trained to recall chains of words and sentences to reproduce a previously learnt story from a partial original input. The model is trained to learn the sequences of words and sentences through back-propagation. Excessive DA signalling during learning (termed 'hyperlearning' by Grasemann et al., 2009; Hoffman et al., 2011), was modelled by increasing the learning rate of the last 500 training cycles of the model. This manipulation was argued to be consistent with the aberrant salience hypothesis. That is, since increased DA transmission would lead to an aberrant assignment of salience, it should eventually result in excessive learning. The authors also implemented various alternative mechanisms such as working memory dysconnection (loss of synaptic connections) and hypo-dopaminergic states (as in the sensory gating models presented above) by altering the gain and bias of the response curve of neurons. When comparing the performance of each model to that of controls and schizophrenia patients, only the hyperlearning and

dysconnection models provided satisfactory fit to the data. However, the hyperlearning achieved the best fit to the experimental data. The hyperlearning model could account for derailments from the original story through a confusion between the characters of different stories ('agent-slotting errors') leading to delusion-like ideas (Grasemann et al., 2009; Hoffman et al., 2011). Specifically, these studies suggest that fixed delusions could stem from contaminated memories (i.e. due to misappropriated agents/characters between stories). However, it is difficult to verify whether 'agent-slotting errors' genuinely lead to false beliefs (delusions) as the authors argue. That is, a falsely reconstructed story within the model can stem from an incorrect recombination of memories during recall, but this appears to be in contrast with the idea that delusions are false beliefs strongly anchored in memory. However, it is interesting that out of all types of story recall errors that were possible, agent slotting errors were the most frequent, as is observed experimentally in the subgroup of patients exhibiting delusions.

Earlier studies used connectionist frameworks to describe how the aberrant salience hypothesis might lead to cognitive deficits as well as negative symptoms (Grossberg, 1999, 2000). Particularly, such models were interested in investigating how symptoms might arise from impaired amygdala circuits and abnormal arousal levels in patients (Grossberg, 1999, 2000). In these studies, the arousal level was assumed to be driven by dopamine and to follow an inverted-U response profile. Specifically, DA release was postulated to drive the amygdala circuits, where hypo-dopaminergic or hyper-dopaminergic activation would lead to a reduced top-down control resulting in an inability to block incentive stimuli (Grossberg, 1999, 2000). These models were solely descriptive however, and were not tested using simulations, making it difficult to draw testable predictions.

3.2.4. Prediction error

The reward prediction error (RPE) hypothesis is the most recent interpretation of dopamine function. The theory dates back to the 60s when Sokolov (1960) proposed that our internal representation of the environment should be updated as a function of a mismatch between the predicted and actual stimuli (Schmajuk, 2005). This theory was later supported by clinical studies in animals and humans revealing that the midbrain dopaminergic signal was consistent with the expected reward signal of the temporal difference (TD) learning algorithm (Schultz et al., 1997). Reward prediction-error models do not necessarily contradict sensory-gating and signal-to-noise models discussed in earlier paragraphs. These models focused almost exclusively on modelling prefrontal cortices and predate the finding of associative learning through dopaminergic reward prediction-error signalling. This explains why such earlier models did not discriminate direct and indirect dopamine pathways (D1r vs. D2r) or tonic vs. phasic activity of dopaminergic neurons. Interestingly, the predictions made by these early computational studies of DA function (SNR, attractors, sensory gating), which mostly suggest that cognitive deficits stem from low prefrontal dopamine (D1r) activation, are still valid. Recent studies (e.g. aberrant salience and prediction error) aim at uncovering different phenomena: the role of dopamine in the basal ganglia and learning and how impaired learning lead to the emergence of positive symptoms and cognitive deficits. In associative learning experiments, the DA signal originating from the ventral tegmental area (VTA) is found to be similar to the reward prediction error signal used to drive learning in the TD-learning algorithm (Schultz et al., 1997; Smith et al., 2005). The DA signal is interpreted as the biological substrate of the reward prediction error, where an expected outcome leads to tonic DA release, unexpected positive outcome leads to phasic DA release and unexpected negative outcome are represented by dips of DA release below the tonic baseline (lack of expected reward; Grace, 1991). Consistent with these findings, Smith et al. (2003, 2004, 2007) successfully modelled patients' cognitive deficits in associative learning tasks by modelling aberrant DA reward prediction-error, which disrupts learning. The simulations successfully matched the behavioral performance of rodents in

experimental studies using amphetamines and anti-psychotics as pharmacological models of schizophrenia (Smith et al., 2003, 2004, 2007). Recent computational models from Frank and colleagues (Frank, 2008; Frank and Claus, 2006; Maia and Frank, 2011; O'Reilly and Frank, 2006; Waltz et al., 2007) also provide a very detailed mechanistic account of the direct and indirect pathways of the basal ganglia and how these pathways interact with frontal cortices. These cortico-basal-thalamo-cortical models have been able to provide a detailed account of motor and cognitive deficits in patients with Parkinson's disease (Frank et al., 2004; Maia and Frank, 2011; Moustafa et al., 2008a,b). Investigating the indirect and direct pathways modulated by D2r and D1r (indirect/NoGo and direct/Go pathways) could lead to novel predictions regarding D2r vs. D1r mediated cognitive deficits in schizophrenia (i.e. impairment in positive as opposed to negative reinforcers; Frank, 2008).

Using associative learning and functional magnetic resonance imaging (fMRI), multiple studies have identified strong distortions in the expected reward prediction-error signal of patients (Corlett et al., 2007b; Gradin et al., 2011; Murray et al., 2008; Roiser et al., 2009, 2013). Interestingly, the distortion magnitude of the prediction error signal was highly predictive of positive symptom severity (delusions; Corlett et al., 2007b; Gradin et al., 2011; Murray et al., 2008; Roiser et al., 2009, 2013; Romaniuk et al., 2010). These findings led to suggest that the RPE hypothesis is consistent with the aberrant salience hypothesis. That is, delusions might stem from faulty PE that fails to discriminate between logical, rational or adaptive associations in the environment such that patients would attend to stimuli they should normally ignore (Frank, 2008).

Interestingly, the original interpretation of the function of the midbrain dopaminergic signal as a RPE driving learning has recently given place to a more refined view, distinguishing different types of DA signals and functions (Bromberg-Martin et al., 2010; Grace, 2016; Schultz, 2016). Particularly, it appears that the DA response to stimuli consists of two distinct signals, a saliency signal and a valuation signal, that operate sequentially on a narrow timescale (Bromberg-Martin et al., 2010; Schultz, 2016). The first signal – the saliency signal – operates just after the stimulus presentation and responds positively (phasic bursts) to both rewarding and punishing events. It is thought to encode a measure of attention or stimulus 'salience' and is modulated by the novelty, intensity or physical characteristics of the stimulus to bring attention to potentially important and relevant information (Schultz, 2016). The second signal – the valuation signal – is the RPE, responding positively (phasic activity) to unexpected positive rewards, neutrally (tonic activity) to expected rewards and negatively (dips in tonic activity) to punishments. The valuation signal drives appetitive and avoidance conditioning, leading to the motivational drive to engage and approach rewarding stimuli (Bromberg-Martin et al., 2010; Schultz, 2016). These two signals are complementary and appear to originate from two different populations of dopaminergic neurons within the SNc (Bromberg-Martin et al., 2010; Grace, 2016). Interestingly, this means that both aberrant motivational valuation and aberrant attentional salience could be present concurrently. Particularly, due to the signals' close temporal proximity, it is believed (Schultz, 2016) that any overall increase in DA transmission could lead to an overlapping of the saliency and value signal, resulting in false valuation, and aberrant motivational drive. Grace (2016) proposes a detailed circuitry of these networks, delineating the cortical and subcortical mechanisms and glutamatergic dysfunctions that could lead to such aberrant DA signalling in schizophrenia.

4. The glutamate hypothesis

4.1. Experimental evidence

The glutamate hypothesis refers to the theory that glutamatergic signalling might be disrupted in schizophrenia. N-methyl-D-aspartate

(NMDA) receptors are glutamatergic receptors known to be essential for synaptic plasticity and learning through the stabilization of synaptic connections (long term potentiation (LTP); Kandel et al., 2013). Consistent with this hypothesis, increases in the expression of NMDA receptors of subtype NR2D were identified in the prefrontal regions of patients with schizophrenia (Akbarian et al., 1996 – NR2D mediated NMDAr are considered "hyperexcitable", and the increase observed in schizophrenia is believed to be a compensatory response from reduced prefrontal activity). Secondly, psychotomimetic drugs such as phencyclidine (PCP) and ketamine that block NMDA receptors (NMDAr antagonists) lead to negative symptoms, cognitive deficits and delusion-like ideation in healthy individuals (Javitt, 1987; Javitt and Zukin, 1991; Jentsch and Roth, 1999). As a result, ketamine has been widely used as a pharmacological model of schizophrenia (Anticevic et al., 2012; Corlett et al., 2011, 2013; Honey et al., 2006; Javitt, 1987; Javitt and Zukin, 1991; Moore et al., 2011). This led to the widely accepted hypothesis that schizophrenia patients might suffer from deficient NMDA receptors (NMDA receptor hypo-function; Honey et al., 2006; Javitt, 1987; Javitt and Zukin, 1991; Jentsch and Roth, 1999). More recently, it has been shown that genetically engineered NRG1 mice (NRG1 encodes the neuregulin protein, essential to NMDA receptor maturation) displayed abnormal behaviors reminiscent to those schizophrenia patients: abnormal social interactions, increased anxiety, abnormal levels of DA release and hypersensitivity to amphetamines, all of which can be reversed with anti-psychotics (Powell et al., 2009).

4.2. Models and support

While DA is widely accepted as playing a major role in psychotic symptoms, the glutamate hypothesis remains a strong potential candidate to explain the aetiology of schizophrenia as a whole. One reason for this is that the glutamate hypothesis can account for a wider range of symptoms, inducing positive, cognitive and negative symptoms when using ketamine or PCP in healthy controls (Javitt, 1987; Javitt and Zukin, 1991; Jentsch and Roth, 1999). However, it is worth noting that no pharmacological treatment affecting glutamate has been found to be effective to date in schizophrenia (Papanastasiou et al., 2013).

The glutamate hypothesis is relatively recent in comparison to the DA hypothesis, and as a result fewer computational models have been developed to assay its validity. Such models (e.g. Murray et al., 2012) consist mostly of biophysical models using integrate-and-fire neural networks that simulate memory or working-memory storage and retrieval through attractor networks. These networks provide realistic simulations of the interactions between excitatory and inhibitory (E/I) activity in cortical areas relevant to the task being modelled (e.g. hippocampus and/or prefrontal cortex). That is, making a number of assumptions regarding the topology of the network (e.g. Mexican hat connectivity), these models can predict the E/I balance within cortical areas that is necessary for memory storage and retrieval. While such realistic neural networks provide very detailed predictions at the biophysical level, these predictions are difficult to validate experimentally. In fact, as mentioned earlier, most of the data and measurements acquired in schizophrenia comes from neuroimaging or behavioral experiments, and are thus difficult to relate to predictions regarding precise neural activity.

4.2.1. NMDA receptor hypofunction – cortical stability

Models supporting the glutamate hypothesis usually explain cognitive deficits and/or negative symptoms through NMDA receptors hypofunction in the PFC (Hsu et al., 2008; Murray et al., 2012; Wang, 2006) or through a combination of NMDA receptors hypo-function in the PFC and the hippocampus (Diwadkar et al., 2008; Siekmeier et al., 2007). Wang (2006) simulated prefrontal networks of working-memory using integrate-and-fire neural networks. In this model, pyramidal cells (excitatory) and inhibitory interneurons were differently modulated by NMDA receptors. Wang (2006) then tested whether such biophysically

realistic attractor networks can simulate the sustained activity of PFC neurons observed during delayed-response tasks in primates. The author found that realistic models of WM maintenance can be instantiated by attractor networks, but that a precise E/I balance is critical in order to filter out distracting stimuli (Murray et al., 2012; Wang, 2006). A second class of models also addressed NMDA receptor hypo-function in the hippocampus (Diwadkar et al., 2008; Siekmeier et al., 2007). For example, Siekmeier et al. (2007) used a connectionist model of the hippocampus to simulate associative learning and context-dependent retrieval of verbal stimuli. The model predicted that NMDA receptor hypo-function in the hippocampus would result in poor memory retrieval (Siekmeier et al., 2007). Interestingly, the authors argue that a hyper-dopaminergic activation of the hippocampus would also result in NMDA receptor hypo-function, again leading to poor memory retrieval.

A possible criticism of this study is that patients seem to usually display memory encoding deficits rather than memory retrieval, and that a memory retrieval deficit may be linked to the cortex rather than from the hippocampus depending on the type of memory involved (short term vs. long term memory). Particularly, it is important to note that the different models presented here investigated different types of memory (Baddeley, 1987). Siekmeier (2009) was modelling deficits in verbal short-term memory, while Murray et al. (2012) and Wang (2006) were investigating spatial working-memory networks. These two types of memory are known to involve different cortical processes and memory systems.

4.2.2. Realistic biophysical models – cortical stability and signal-to-noise ratio

Earlier attractor network models were the precursors of the latest biophysical models, which use AMPA, NMDA and GABA receptors to model working-memory (Loh et al., 2007; Rolls et al., 2008). In these models, the balance between inhibitory (GABA) and excitatory signals (AMPA/NMDA) is critical. First, in a combined experimental and computational setting, Wolf et al. (2005) studied the bistability (i.e. the switching between an up or down state) of medium spiny neurons in the Nucleus Accumbens (NAcc), which has been proposed to serve for gating purposes in working-memory (Gruber et al., 2006). Their model predicted that the medium spiny neurons (MSN) would require sustained excitatory inputs (from about 1000 afferent) in order to maintain a stable depolarized (up) state. In this model, NMDA receptor hypofunction is predicted to lead to an inability for MSN to express bistable activity and to impede gating or integration of information. Another study by Loh et al. (2007) addressed the interactions between inhibitory (GABA) and excitatory (NMDA/AMPA) activity on the dynamics of a working-memory attractor network. The authors found that an imbalance in excitation or inhibition led to the instability of the whole system, resulting in unstable working-memory. Such instability resulted in changes in the attractor landscape. Decreased excitatory activity led to jumps from one attractor to another due to an increased stochastic firing of the neurons in combination with shallower attractor states. As a result, memories were unstable. In this model, a decrease in both excitation (NMDA) and inhibition (GABA) results in a flat attractor landscape. The authors argue that a flat attractor landscape in temporal areas would lead to jumps between trains of thoughts. This prediction is also in line with previous experiments showing excessive amounts of noise in the temporal (auditory) cortices of patients, especially during auditory hallucinations. Interestingly, although supporting the glutamate hypothesis, the authors managed to adapt the model using the work from Durstewitz and Seamans (2008), so as to also account for the role of DA in modulating network activity. The authors found that intermediate levels of DA modulate the SNR in frontal areas. That is, D1 receptor activation enhance both excitatory (NMDA) and inhibitory (GABA) activity resulting in an increased stability of the network (increase in SNR), while D2 receptor activation has the opposite effect and reduce the signal-to-noise ratio. The authors argue that this mechanism could potentially explain the effects of anti-psychotic medications by

stabilizing deficient attractor networks through a decrease in D2 receptors activity.

It is worth noting however, that the majority of dopaminergic receptors in the frontal cortex seem to be of the D1r subtype (Howes and Kapur, 2009) and that D2r activation have previously been found to have no effect on WM networks (Wang et al., 2004). Again, while these studies provide interesting insights on the possible link between DA, SNR and attractor dynamics in schizophrenia, the model predictions are difficult to relate to experimental data, which mostly consists of behavioral and/or imaging data.

5. The GABA hypothesis

5.1. Experimental evidence

Lewis and Hashimoto (2005) observed anomalies in GABAergic interneurons of schizophrenia patients. Namely, they found that GABA synthesis and re-uptake was altered and diminished in the dlPFC leading to disrupted gamma oscillations and de-synchronization. Also, DA neurons appear to provide direct synaptic input to parvalbumin-expressing GABA interneurons in the dlPFC of primates, suggesting a possible modulation of GABAergic inhibition through DA activation (Lewis and Hashimoto, 2005).

5.2. Models and support

Very few computational models support the GABAergic hypothesis alone, but rather integrate GABAergic inhibition with NMDA hypofunction to model biologically realistic simulations of cortical function, stability and synchrony (Loh et al., 2007; Murray et al., 2012; Rolls et al., 2008).

5.2.1. Cortical stability

Each of the models that investigated the GABA hypothesis explored different aspects of inhibitory dysfunction in schizophrenia (Spencer, 2009; Tanaka, 2008). In Tanaka (2008), the effects of GABAergic activation through dopamine D1r modulation were investigated using a pure mathematical model of balanced inhibitory and excitatory activity. The author established through parameter exploration that for intermediate levels of D1r activation, GABA inhibits noise in the dlPFC circuitry of working-memory (increased SNR). The model of Spencer (2009) investigated the link between the GABA and the dysconnection hypothesis (which we detail below). Their integrate-and-fire model suggests that a deficient inhibition would lead to disrupted γ-rhythms as observed in schizophrenia. γ-rhythms, if disrupted, would disturb cortico-cortical synchrony and eventually result in a functional dysconnection syndrome.

6. The dysconnection hypothesis

6.1. Experimental evidence

The dysconnection hypothesis states that schizophrenia is associated with reduced synaptic connectivity (in which case it is sometimes spelt 'disconnection' – lack of connectivity) or dysfunctional connectivity (i.e. dysconnection – abnormal functional connectivity) primarily in the mesocortical pathway (i.e. midbrain dopamine and serotonin afferent to the PFC) and between cortical areas such as the frontal cortex and the temporal lobes (Friston, 1996; Lawrie et al., 2002; Pettersson-Yeo et al., 2011; Stephan et al., 2009). This theory is supported by several post-mortem or neuroimaging studies revealing anatomical (Kubicki et al., 2005, 2007; Samartzis et al., 2014) and functional dysconnection (Dauvermann et al., 2013; Dima et al., 2010; Lawrie et al., 2002) in patients (for a review see: Friston et al., 2016). It is known that the normal developmental course of the mammalian brain begins with an over-elaboration of neuritic processes, which is

then followed by a gradual reduction of synaptic density during adolescence, reaching about 60% of maximum levels in early adulthood (McGlashan and Hoffman, 2000). Interestingly, the end of this developmental timeline coincides with the age of onset of psychotic symptoms (first episode), suggesting a late neurodevelopmental dysfunction during adolescence. Several post-mortem examinations later found reduced spine densities and smaller dendritic arbors on prefrontal pyramidal cells of schizophrenia patients (Stephan et al., 2009). Additionally, decreased synaptic protein messengers and synaptophysin were found in the dlPFC of patients. Together these findings provide a possible explanation for the observed decreased neuropil without neural loss found previously in schizophrenia (McGlashan and Hoffman, 2000). It is worth noting, however, that decreased neuropil appears in other mental disorders and is not specific to psychotic illness.

6.2. Models and support

6.2.1. Cortical stability

Computational models of the dysconnection hypothesis can be classified into three subcategories. First, simple Hopfield networks were used to study positive symptoms (David, 1994; Hoffman, 1987; Hoffman and Dobscha, 1989; Seeman, 1994). In these models, dysconnection is usually implemented by 'pruning' the synaptic connections between the units of the networks after training. The pruning strategy adopted is a Darwinian 'evolutionary' process, which eliminates weak and spatially distant connections by setting their weights to zero. This eventually results in an inability for the network to flow into previously learnt patterns of activation and recover memories. When excessive pruning is performed two types of behaviors emerge. First, the network produces generalizations or 'loose associations', by merging parts of distinct memory patterns into a single one, which was interpreted as a potential explanation for bizarre trains of thoughts (thought disorder). Secondly, the network could elicit spontaneous patterns of activations, i.e. relentlessly recovering the same memory output irrespective of the input presented or recovering new memory patterns unknown to the model. The authors argued that the spontaneous emergence of new memories was homologous to hallucinations. was the main instigator of this hypothesis in the field of schizophrenia. His early models qualitatively supported the hypothesis of an excessive pruning or memory overload in the disorder (Hoffman, 1987; Hoffman and Dobscha, 1989). However, simulations of these effects were quantitatively unrealistic as up to 80% of 'evolutionary' pruning was required for hallucinations to emerge. This specific hypothesis was later disputed by David (1994), who on the contrary proposed that positive symptoms may emerge from an hyper-connectivity due to a deficit of the neurodevelopmental pruning process. However, these conclusions appear to be contradicted by experimental findings of reduced grey matter and connectivity in schizophrenia. Following suggestions, other researchers sought to expand the model to account for the effects of environmental stress and dopamine modulation (Chen, 1994, 1995; Seeman, 1994). These effects were expressed in terms of memory overload and increased network temperature. The models provided a possible link between the positive symptoms and cognitive deficits by incorporating dopamine as a signal-to-noise enhancer, where too little or too much dopamine was detrimental to the signal. In the mid 90s, computational studies started investigating the dysconnectivity hypothesis from a different perspective, suggesting that positive symptoms might stem from secondary self-repairing properties of the brain following cortico-cortical synaptic pruning. When the cortico-cortical inputs connecting to the working-memory units of the network were degraded, the PFC tend to compensate by updating its local weights in order to recover memory patterns. This, in turn, would lead to increased WM noise resulting in the spontaneous retrieval of memories in the absence of external inputs (Horn and Ruppin, 1995; Ruppin et al., 1996).

The second group of dysconnection models used three-layer perceptrons to study hallucinated voices in patients with schizophrenia (de

la Fuente-Sandoval et al., 2005; Hoffman, 1997; Hoffman and McGlashan, 1993a,b, 1997, 1999, 2001, 2006; McGlashan and Hoffman, 2000). These networks were trained in an ad-hoc manner to associate inputs (phonemes) and outputs (words) using back-propagation. The network then relied on an intermediate layer representing verbal working-memory to disambiguate current phonemes. In the first implementations of this model, the working-memory module was only a delayed copy of the hidden layer (i.e. temporary buffer) used to bias and compute temporally successive inputs (de la Fuente-Sandoval et al., 2005; Hoffman, 1997; Hoffman and McGlashan, 1993a,b, 1997, 1999, 2001; McGlashan and Hoffman, 2000). Later models modified the working-memory so as to use a Hopfield network within the hidden layer (Hoffman and McGlashan, 2006). To account for the dysconnection syndrome, synaptic connections were removed following an *'evolutionary'* approach as described previously. These models generated particularly interesting predictions, whereby synaptic pruning improved the performance of word recognition by 50% when pruning up to 64% of the connections. However, above 77% of pruning, hallucinated words – i.e. words detected without input – started to occur and performance decreased drastically. The authors suggested that synaptic pruning during the neurodevelopmental stage of late adolescence might actually be beneficial as it would improve recall performance while reducing energetic costs. The model also suggests that a failure to stop normal synaptic pruning in early adulthood could account for the onset of the disorder. Neuronal loss (Hoffman, 1997) and deregulated hypo-dopaminergic modulation (Hoffman and McGlashan, 2006; McGlashan and Hoffman, 2000) were also addressed in this framework. However, both failed to initiate "hallucinations". Hypo-dopaminergic modulation was implemented in this model as a shift of the bias to each neural unit in WM (i.e. hidden layer), which protected over-pruned networks against hallucinations (Hoffman and McGlashan, 2006). Interestingly, under the assumption that dopaminergic modulation does alter the bias of WM units, the model could successfully account for the effects of anti-psychotics against positive symptoms.

Other models studied cognitive impairments in schizophrenia at specific tasks such as facial affect recognition using a three-layer perceptron (Johnston et al., 2001), episodic memory deficits using a connectionist framework (Meeter et al., 2002), and semantic priming using interconnected Hopfield networks (Siekmeier and Hoffman, 2002). All these studies converged to similar conclusions, namely that synaptic pruning was found to degrade the performance of the network. However, the causes of an excessive pruning mechanism remains unknown and was largely left untouched in these studies. While genetic factors could be at play, no experimental study has found a common genetic component that would be responsible for this developmental deficit.

It is important to mention that while Hoffman was the most prominent scientist defending the dysconnection hypothesis through models of excessive pruning processes (10 out of the 15 published modelling studies on dysconnection), in later studies the author tested other competing hypotheses of schizophrenia in a story-recall task (Grasemann et al., 2009; Hoffman et al., 2011). In that study, the authors found that only the dysconnection and aberrant salience hypotheses could account for positive symptoms, but that the aberrant salience hypothesis accounted best for the performance at story learning and recall in the subgroup of schizophrenia patients exhibiting delusions. Finally, recent work from Whitford et al. (2012) hypothesized that frontal myelin damage in schizophrenia would lead to delays in the transmission of efference copies and corollary discharge (copies of motor commands and predicted sensory feedback). This delay would result in an asynchrony between proprioception (sensory feedback) and corollary discharge leading to sensory discrepancies. In such cases, a subject would perceive these sensory discrepancies as-if their own actions were not self-generated. This would result in delusions of control, that is, the false belief that an external force controls one's thoughts and behavior.

A recent review of the advances and accumulated empirical

evidence regarding the dysconnection hypothesis (Friston et al., 2016) suggest that the dysconnection syndrome may result from abnormal neuromodulation or synaptic gain in neuronal microcircuits, where microcircuits implement hierarchical inference and abstraction over multiple layers of the hierarchy using a predictive coding framework. They propose that the dysconnection hypothesis can be understood within the framework of Bayesian inference (see below). In this framework, abnormal synaptic gain would disrupt the inference mechanism, leading to incorrect beliefs (false inference) about the causes of events and sensory percepts, leading to psychotic symptoms.

7. The Bayesian inference hypothesis

Recently the brain has been viewed as a complex processing machine used to interpret sensory inputs in order to make sense of the environment (Franklin and Wolpert, 2011; Friston, 2005, 2010; Wolpert et al., 2011). According to this theory, the brain evolved to interpret and infer the cause and consequences from the environment in order to predict future outcomes in the environment and minimize surprises. This framework assumes that cognition can be described in terms of Bayesian inference, where subjects combine optimally sensory evidence (the likelihood, for e.g. sensory inputs about a visual scene) and prior knowledge or expectations (the prior, for e.g. knowledge about the frequency of certain objects in the environment) so as to form probability distributions relevant to the task at hand (e.g. how likely is it to see a particular object in this environment?). It is argued that using this framework, if perceptual biases or illusions occur, they would lead to an effect of surprise. This surprise would then require to logically explain these abnormal percepts by updating the internal model of the environment, resulting in false beliefs akin to delusions (Corlett and Fletcher, 2015; Corlett et al., 2007a,b, 2009a,c, 2016; Fletcher and Frith, 2009). Delusional content, would in turn bias expectations of future outcomes in the environment resulting in stronger perceptual biases (i.e. illusions or hallucinations). This spiralling effect would gradually result in stronger, more salient illusions and false-beliefs, eventually leading to full-blown complex hallucinations and deeply anchored delusions.

7.1. Experimental evidence

The Bayesian brain hypothesis of schizophrenia has received relatively little empirical testing and validation, with the exception of studies investigating illusions (Crawford et al., 2010; Dima et al., 2009, 2010; Horton and Silverstein, 2011; Keane et al., 2013; Silverstein and Keane, 2011a,b; Tschacher et al., 2006; Williams et al., 2010) or explicit statistical learning (Averbeck et al., 2011; Evans et al., 2012; Freeman et al., 2008, 2014; Garety et al., 2013; Garety and Freeman, 2013; Huq et al., 1988; Joyce et al., 2013; Speechley et al., 2010). Such studies however, tend to investigate either illusions or learning in isolation (i.e. not first the acquisition of expectations through learning and then the influences of these expectations on perception).

For illusory perception, patients with schizophrenia have been found to be less susceptible than healthy controls at the hollow mask illusion (Dima et al., 2009, 2010; Keane et al., 2013), motion-induced blindness (Tschacher et al., 2006), illusory motion (Crawford et al., 2010), the size-weight illusion (Williams et al., 2010) and the Ebbinghaus illusion (Horton and Silverstein, 2011) (for reviews of perception in schizophrenia see Notredame et al., 2014; Silverstein and Keane, 2011a,b). In healthy controls, Schmack et al. (2013) recently demonstrated that the magnitude of expectation-driven illusions correlated with delusional ideation. That is, in line with previous studies on perceptual illusions in schizophrenia (Crawford et al., 2010; Dima et al., 2009, 2010; Keane et al., 2013; Tschacher et al., 2006; Williams et al., 2010), the authors found that the stronger the delusions of healthy controls, the less likely these were to have their percepts affected by expectations (Schmack et al., 2013). Recently, Teufel et al. (2015)

described on the contrary how healthy subjects who demonstrated schizotypal features, as well as patients with subclinical levels of psychotic symptoms, relied significantly more than controls on their priors to disambiguate noisy sensory information in visual perceptual tasks. This suggests that subjects displaying mild to moderate levels of psychosis (akin to levels observed in the prodromal phase of schizophrenia) may show a stronger reliance on their prior relative to sensory evidence (or an equivalent down-weighting of sensory evidence relative to prior information). This is consistent with the idea that patients with schizophrenia (or controls with mild forms of psychotic symptoms) might have a deficit of perceptual inference or in the acquisition of expectations. These findings have been sometimes reconciled by proposing that whether prior knowledge had stronger or weaker impact in schizophrenia compared to controls depended on the level of the predictions: low-level sensory predictions would have weaker impact in schizophrenia whereas higher-level (more cognitive) predictions would have a stronger impact.

7.2. Models and support

While a number of authors have proposed theories suggesting that schizophrenia could be linked with Bayesian inference deficits (Corlett et al., 2009a, 2016; Fletcher and Frith, 2009; Frith and Friston, 2012; Jardri and Cachia, 2013), relatively few computational models have been implemented to provide a quantitative and mechanistic account of delusions and hallucinations using this framework (Adams et al., 2013; Deneve and Jardri, 2016; Jardri and Deneve, 2013; Jardri et al., 2016; Notredame et al., 2014).

In Adams et al. (2013), the authors argue that psychosis may stem from an abnormal encoding of the metacognitive level of precision, i.e. the confidence about the precision (i.e. inverse of standard deviation) of incoming information (the sensory likelihood) or confidence about the precision of beliefs about the world (the prior). Particularly, the authors argue that the expectations (prior beliefs) of patients are weaker than they ought to be, resulting in too much emphasis on sensory evidence and a high state of 'surprise' since sensory observations are unexpected (weak prior expectations). So as to minimize surprise, compensatory changes in the precision of prior information or sensory likelihood ensue, leading to cognitive, negative or positive symptoms. Using a predictive coding model of hierarchical inference, the authors then demonstrate how weaker priors could explain deficits in tasks such as in the oddball stimuli (mismatch negativity), smooth eye-pursuit and the force-matching task. In this framework, trait symptoms such as cognitive deficits and negative symptoms may be explained by weaker priors relative to sensory evidence. Conversely, the authors argue that state symptoms such as hallucinations and delusions may be better explained in terms of an increase in prior precision relative to sensory evidence.

Interestingly, this framework could potentially account for the occurrence of positive symptoms alone in first episode psychosis (and prodromal phase) due to stronger priors, and account for trait symptoms (e.g. cognitive deficits and negative symptoms) due to weaker priors in chronic schizophrenia. However, it is unclear how both a decrease in prior precision for trait symptoms and an increase in prior precision for transient state symptoms might cohabit (e.g. in relapsing-remitting patients or when chronic patients suspend pharmacological treatment). This also appears to be at odds with the co-occurrence of decreased sensitivity to perceptual illusions (suggesting a weaker prior) and delusional ideation (suggesting stronger priors), both in healthy (Schmack et al., 2013) and patient populations (Silverstein and Keane, 2011a,b).

Similarly to Adams et al. (2013), Jardri and Deneve (2013) proposed a hierarchical Bayesian inference model using an algorithm known as 'belief propagation', where each level of the hierarchy produces inference and abstraction over lower levels. In this model, the authors argue that bottom-up sensory evidence and top-down predictions could be reverberated throughout the hierarchy due to poor

GABAergic inhibition. Particularly, the authors make the predictions that an impairment in bottom-up inhibition would result in sensory evidence being reverberated throughout the hierarchy and wrongly interpreted as if it were prior beliefs. Conversely impaired top-down inhibition would result in prior knowledge being reverberated as sensory evidence. That is, impaired inhibition would result in an over-estimation (overconfidence) in sensory evidence and/or prior knowledge depending on the inhibition circuits affected (bottom-up vs. top-down respectively). Interestingly, a symmetric impairment of both bottom-up and top-down inhibition resulted in no deficit in belief inference itself, but lead to an abnormally high confidence in that belief. This phenomenon is argued to be reminiscent to the jumping-to-conclusion bias and the relative imperviousness of delusional beliefs to contradictory evidence. While hallucinations and delusions may more readily be explained by a deficit in top-down inhibition resulting in stronger priors, the authors argue that converging lines of evidence such as the reduced susceptibility to illusions and the jumping-to-conclusions bias suggest that bottom-up inhibition may be selectively impaired in schizophrenia. That is, impaired bottom-up inhibition would result in sensory evidence reverberating back through the hierarchy as if they were priors, resulting in increased precision of sensory evidence (or an equivalent under-weighing of the prior relative to sensory evidence). The authors argue that hallucinations could still occur in this context when sensory information is particularly noisy. As a result, Jardri and Deneve (2013) conclude that an impairment of bottom-up inhibition could account for most of the deficits of inference observed in schizophrenia including hallucinations and delusions.

It is interesting to note however that the conclusions of Jardri and Deneve (2013) and Jardri et al. (2016) appear to be in contrast with those of Adams et al. (2013) for state symptoms where Adams et al. (2013) argue that delusions and hallucinations stem from an increased precision of the prior (stronger prior relative to sensory evidence). The conclusions of Jardri and colleagues are however in line with Adams et al. (2013) model of trait symptoms (i.e. cognitive deficits and negative symptoms result from an under-weighting of the prior – decreased 'precision' of the prior – and an over-estimation of the strength of sensory evidence – increased 'precision' of sensory evidence).

A couple of points can be discussed in relation to these studies: First, it is important to mention that although behavioral performance (e.g. perception in psychophysics tasks) may appear to be approximately Bayesian optimal, a multitude of implementations could give rise to similar behaviors and that Bayes optimal behavior (i.e. 1st level of analysis) does not necessarily require the brain to perform Bayesian inference at the neural level (i.e. 3rd level of analysis; see: Jacobs and Kruschke, 2011 and Sanborn and Chater, 2016). While both studies used predictive coding models as methods for approximate Bayesian inference, these are just one of many possible implementations of how inference may be implemented at the neural substrate, and that others mechanisms have also been suggested (e.g. including but not limited to: neural sampling Berkes et al., 2011, probabilistic population codes Ma et al., 2006, etc. Gershman and Beck, 2016; Lochmann and Deneve, 2011). Particularly, Deneve and Jardri (2016) called for neurophysiological validation to be carried out, as a specific neural implementation of belief propagation (i.e. inference) has yet to be identified in the brain (Deneve and Jardri, 2016). While the models presented above provided an elegant conceptual illustration of how changes in precision may lead to psychotic symptoms using synthetic data, it would be extremely valuable to see whether these models yield similar conclusions once fitted to experimental data in schizophrenia. For example, Adams et al. (2016) recently demonstrated how such computational models can be used to successfully infer the precision of beliefs at an eye-pursuit task in healthy controls. Similarly, Jardri et al. (2017) used a modification of the beads task in healthy controls and chronic medicated patients with schizophrenia. In this task, both the prior (i.e. prior probability of each urn being selected) and likelihoods (i.e. how likely is it to find a red/blue bead in each urn) were explicitly given on every trial. The

participants then used a visual analog scale to report their confidence as to where a bead originated from. Responses were then used to identify whether participants could optimally combine prior information and likelihood. The authors found that their circular inference model (Jardri and Deneve, 2013) best accounted for the patients' behavior. Surprisingly however, they found that the integration of prior and likelihoods was suboptimal in all participants, but more so in patients than controls. In line with their earlier theoretical predictions, patients' behavior was best explained by a model that over-counted sensory evidence, suggesting that patients were unable to incorporate prior information optimally. The degree to which patients over-counted sensory information correlated with the severity of their positive symptoms (nonclinical delusion severity as measured by the PDI). Surprisingly, however, considering that the PDI was originally designed to measure delusional ideation in the general population, this correlation did not hold in the healthy control sample. Interestingly, the experiment required participants to memorize and maintain the prior information in working memory in order to combine the two sources of information optimally. The two groups were not matched for working memory deficits (patients had significantly worse WM than controls). As a result, the differences in model parameters could largely be accounted for by WM deficits in the schizophrenia sample, which could also explain why the correlation between model parameters and positive or negative symptoms only occurred in the patient sample. Further experimental investigation need to be carried out in patients, to quantitatively test how patients weigh their likelihood and priors during decisions tasks (Adams et al., 2016; Jardri et al., 2016). It will be essential that future studies attempt as best as possible to remove potential confounds related to patients' potential deficits in dimensions not directly assessed by the task (e.g. IQ, WM).

Finally, it is unclear how the proposed models would scale for realistic problems. For example, the conceptual illustration provided in Jardri and Deneve (2013) (i.e. "Forest → Tree → Leaves → Green") was a singly-connected graphical model. For this particular graphical model (or more generally for hierarchies involving trees and poly-trees which doesn't include any loops), belief propagation (BP) is known to perform exact inference. However, BP is not guaranteed to converge to a stable equilibrium when used with loopy graphical models (Pearl, 1982). It is likely that graphical models representing complex world statistical relationships may not be reducible to a tree-like structure. Loopy BP could then lead to widely oscillating beliefs that may never converge, beliefs that converge to poor approximations or incorrect equilibriums which lead to beliefs that are not globally consistent (i.e. beliefs that are not coherent with respect to all the nodes of the network). This may provide a different explanation for the false inference mechanisms than impaired top-down or bottom-up inhibition. Secondly, the examplar graphical model used in the study was correctly depicting realistic statistical relationships between elements of the world's hierarchy (e.g. "Forest → Tree → Leaves → Green"). It is unknown however, whether patients with schizophrenia have (or are able to learn) the correct internal graphical model of their environment. That is – in patients – if the internal model of the world is not representative of the true structure of the environment, inference processes could lead to beliefs that are subjectively consistent with the patients' incorrect internal model of the world, but incoherent with respect to the true model of the environment.

Such a framework could explain delusional beliefs, and in particular the cases where patients tend to assume causality between unrelated events and appear to be impervious to evidence contrary to their delusional beliefs. Specifically, patients rather appear to integrate contrary evidence as supporting their delusional beliefs, suggesting that the contrary evidence provided may be consistent with their internal model of the world.

The false Bayesian inference hypothesis of schizophrenia holds much hope for a unifying framework explaining the variety of symptoms expressed in patients. However, sensitive behavioral tasks are now

needed to validate the models' assumptions at the behavioral level before attempting to make predictions at the neural level. Particularly, it appears essential to test empirically, whether:

1. Patients behave as Bayesian observers in simple tasks, correctly mixing prior information and sensory evidence (likelihood) when these are explicitly provided.
2. Patients can acquire relevant statistical information about the world in the form of prior distributions.
3. Patients learn a coherent graphical model (causal structure) of the world.

8. Discussion

In this review, we described promising models, which support various hypotheses of schizophrenia's pathophysiology. While, none of these computational studies could account for the variety and complexity of symptoms found in the disorder, most studies focusing on cognitive deficits appear to support the dopaminergic hypothesis. Computationally, cognitive deficits appear to stem from:

1. A weak frontal dopaminergic D1r activation (dlPFC), resulting in a decreased frontal SNR and deficient working memory.
2. Excessive striatal D2r activation, leading to impairments in prediction-error signalling, essential for associative learning and goal-directed behavior.

Interestingly, a recent study from Collins et al. (2014) managed to disentangle the influence of working-memory from that of reinforcement-learning using a novel behavioral task and computational modelling in chronic schizophrenia. The study suggested that the generalized cognitive deficits observed in schizophrenia appeared to result from reduced working memory capacity and reliability rather than from deficient reinforcement learning (Collins et al., 2014). Particularly it appears that in the cognitive and negative symptom domains, schizophrenia is associated with reduced adaptive phasic DA release to relevant stimuli (Gold et al., 2015; Maia and Frank, 2016), while it is associated with increases in spontaneous phasic DA release to innocuous stimuli in the positive domain (Maia and Frank, 2016).

However, schizophrenia is an heterogeneous disorder, expressing itself through unique combinations of symptoms in every patient. Therefore, several of the current hypotheses discussed in this review might jointly be responsible for the wide variety of behaviors and symptoms observed in the disorder.

For example, biophysically realistic models of working-memory elegantly demonstrated how the balance between inhibitory GABAergic and excitatory glutamatergic activity is crucial to the proper functioning of realistic attractor networks (cortical stability). In these models, a deficit in either excitatory glutamate or inhibitory GABAergic activity led to impaired working-memory dynamics, cognitive impairments and arguably to some form of positive symptoms.

Alternatively, models supporting a dysconnection syndrome were able to successfully demonstrate how excessive pruning in cortical networks could lead to positive symptoms (spurious attractors), as well as predicting the neurodevelopmental timeline of schizophrenia, providing for the first time an explanation for the late adolescence onset of the disorder (Hoffman and McGlashan, 2006; McGlashan and Hoffman, 2000). Less experimental evidence was found to support an association between neurotransmitter dysfunctions and cortico-cortical dysconnection. However, we would argue that both the glutamate and GABAergic hypothesis could lead to a dysconnection syndrome. Specifically, weaker synapses could emerge following a prefrontal NMDA receptor hypo-function. Such synapses could then be pruned away during the Darwinian *'evolutionary'* neurodevelopmental process proposed by (Hoffman and McGlashan, 2006; McGlashan and Hoffman, 2000). That is, weaker synapses would lead to an over-pruning of

frontal cortices, resulting in a *'physical dysconnection syndrome'*, as presented by the synaptic runaway model of Greenstein-Messica and Ruppin (1998). Alternatively, GABAergic inhibition appears to be essential to the generation of γ-band rhythms. Aberrant γ-oscillations, is argued to result in an asynchrony between cortical regions (Spencer, 2009), leading to a reduced ability to transmit information between cortical regions, (*'functional dysconnection'*). As a result, we argue that a dysconnection syndrome might be secondary to an incorrect balance between excitatory and inhibitory activity in cortical regions, leading to either excessive synaptic pruning during adolescence (*'physical dysconnection'*) or an impossibility to synchronize information across cortical regions (*'functional dysconnection'*). Finally, it is worth mentioning that NMDA receptor blockade has been found to result in strong changes of dopaminergic midbrain neurons (Jentsch and Roth, 1999), while chronic ketamine abuse has been found to upregulate dopamine D1r in the PFC (Narendran et al., 2005). It is therefore possible that the dopamine dysfunction observed in schizophrenia could be secondary to a generalized NMDA receptor hypo-function (Lewis and Gonzalez-Burgos, 2006).

In a recent review of computational studies in schizophrenia research, Rolls and Deco (2011) called for further investigation using bottom-up modelling approaches. The authors argued that using realistic biophysical models of attractor network and cortical dynamics, one could explore in much detail the interactions between neurotransmitter functions and produce precise predictions about the states of the neural networks in schizophrenia. We agree with the authors that the abstract modelling of decreased signal-to-noise ratio in schizophrenia can successfully give place to more refined biophysical models in order to account for our current knowledge of network dynamics and neurotransmitter function. Such models' plausibility and realism might however often appear limited, as they tend to focus on simulating one single (or sometimes a few) cortical area(s) in isolation, with little diversity in terms of cell and receptors types or structure. This is in part due to the inherent mathematical intractability and parameter identifiability constraints of models with complex network interactions and dynamics, resulting in an over simplification of the models' components (that is, only the units that are believed to substantially affect the model's behavior or are relevant to the hypothesis being tested are implemented).

However, Rolls and Deco (2011) argue that high-level, abstract, behavioral or descriptive models (i.e. phenomenological models) have no construct validity since these do not map to realistic brain function and as a result fail to produce testable predictions. In this review, we take a different standpoint. First, we would first argue that top-down and bottom-up modelling approaches are not incompatible and that the choice of modelling strategy depends on the scientific question and disorder being investigated. Particularly, biophysical models have a potential to lead to strong and testable predictions about the neural substrate, where electrophysiological and optogenetic tools are available to manipulate, test, validate, – and if necessary, revise – the models and their predictions. While predictions at that level can be tested in animal models of particular psychiatric disorders (e.g. drug abuse, addiction, gambling, compulsion, anxiety, etc.), they are more limited in disorders that appear to be unique to humans (e.g. psychosis – where electrophysiological experimentation is impossible, or limited to exceptional case of patients undergoing brain surgery –, although some animal models do exist, see Forrest et al., 2014). Particularly, since cognitive, positive and negative symptoms are the most stable, readily available, and quantitatively measureable effects of the disorder across patients, high-level descriptive models can yield strong testable predictions at the behavioral level and serve as a starting point for addressing the underlying neural processes.

Overall, the Bayesian brain hypothesis of psychosis seems very promising but has received relatively little theoretical and empirical investigation in comparison to the other hypotheses. In fact, in comparison to the models supporting the GABAergic, glutamate or dopamine hypothesis, the Bayesian brain hypothesis provides a high-level construct that makes strong testable predictions that can be validated experimentally at the behavioral level. That is, one could potentially test several the assumptions of the framework, and if proven successful, start to investigate the underlying neural processes that may have gone awry.

If we attempt to synthesize our findings across models and symptoms, we find strong experimental and modelling evidence suggesting that dopamine is the final common pathway to psychotic symptoms (Abi-Dargham, 2004; Howes et al., 2016; Howes and Murray, 2014; Maia and Frank, 2016). However dopaminergic dysfunction may be secondary to other upstream neurotransmitter dysfunctions. In fact, a growing body of evidence suggests that aberrant dopaminergic neurotransmission in the striatum may result from upstream glutamatergic dysfunction (e.g. NMDA hypofunction – Abi-Dargham, 2004; Grace, 2016; Howes et al., 2016), possibly combined with deficient GABAergic inhibition (Abi-Dargham, 2004; Grace, 2016). Aside from the obvious contributions to positive symptoms, dopamine may also be of particular interest when considering both cognitive and negative symptoms (Maia and Frank, 2016). For example, it has long been assumed that cognitive deficits in schizophrenia were the result of poor dopaminergic transmission, leading to either poor working memory gating and maintenance (a component essential to all cognition) and/or to the abnormal acquisition of reward and punishment contingencies. This resonates with earlier models of dopamine function which suggest that in large parts, the generalized cognitive deficits in schizophrenia may be due to deficient working memory potentially resulting from poor DA transmission (e.g. Cohen et al., 1996) or/and alternatively to Glutamatergic dysfunction (Murray et al., 2012; Wang, 2006). Recently however, using finely tuned experimental designs, it has become possible to dissociate the contributions of WM deficits to that of reward learning (Collins et al., 2014). Using these approaches with patients at different time points across the time course of the disorder may be particularly useful. This would help elucidating which of these systems is impaired in schizophrenia, and whether any of these deficits is secondary to antipsychotic medication (Artaloytia et al., 2006; Maia and Frank, 2016). Finally, we know that dopaminergic signalling is critical for the subjective valuation of actions (Schultz, 2016), that is, in deriving a subjective value from potential reward benefits given their effort costs. This could provide partial explanation for some negative symptoms. If the valuation system is impaired due to inappropriate dopaminergic signalling, inaccurate sensitivities to rewards and efforts costs may negatively affect the patients' willingness to engage in activities leading to poor motivation, and a gradual worsening of social-economic functioning, and eventually severe isolation.

Acknowledgments

References

Aakerlund, L., Hemmingsen, R., 1998. Neural networks as models of psychopathology. Biol. Psychiatry 43 (7), 471–482.

Abboud, R., Roiser, J.P., Khalifeh, H., Ali, S., Harrison, I., Killaspy, H.T., Joyce, E.M., 2016. Are persistent delusions in schizophrenia associated with aberrant salience? Schizophr. Res. Cogn. 4, 32–38.

Abi-Dargham, A., 2004. Do we still believe in the dopamine hypothesis? New data bring new evidence. Int. J. Neuropsychopharmacol. 7 (Suppl. 1), S1–5.

Adams, R.A., Bauer, M., Pinotsis, D., Friston, K.J., 2016. Dynamic causal modelling of eye movements during pursuit: confirming precision-encoding in V1 using MEG. NeuroImage 15 (132), 175–189.

Adams, R.A., Stephan, K.E., Brown, H.R., Frith, C.D., Friston, K.J., 2013. The computational anatomy of psychosis. Front. Psychiatry 4, 47.

Ahn, W.-Y., Busemeyer, J.R., 2016. Challenges and promises for translating computational tools into clinical practice. Curr. Opin. Behav. Sci. 11, 1–7.

Akbarian, S., Sucher, N.J., Bradley, D., Tafazzoli, A., Trinh, D., Hetrick, W.P., Potkin, S.G., Sandman, C.A., Bunney, W.E., Jones, E.G., 1996. Selective alterations in gene expression for NMDA receptor subunits in prefrontal cortex of schizophrenics. J. Neurosci. 16 (1), 19–30.

Allardyce, J., Suppes, T., van Os, J., 2007. Dimensions and the psychosis phenotype. Int. J. Methods Psychiatr. Res. 16 (Suppl. 1), S34–S40.

Amos, A., 2000. A computational model of information processing in the frontal cortex and basal ganglia. J. Cogn. Neurosci. 12 (3), 505–519.

Andrew, A., Knapp, M., McCrone, P.R., Parsonage, M., Trachtenberg, M., 2012. Effective Interventions in Schizophrenia: The Economic Case: A Report Prepared for the Schizophrenia Commission. Rethink Mental Illness, London.

Anticevic, A., Corlett, P.R., 2012. Cognition-emotion dysinteraction in schizophrenia. Front. Psychol. 3, 392.

Anticevic, A., Gancsos, M., Murray, J.D., Repovs, G., Driesen, N.R., Ennis, D.J., Niciu, M.J., Morgan, P.T., Surti, T.S., Bloch, M.H., Ramani, R., Smith, M.A., Wang, X.-J., Krystal, J.H., Corlett, P.R., 2012. NMDA receptor function in large-scale anticorrelated neural systems with implications for cognition and schizophrenia. Proc. Natl. Acad. Sci. U.S.A. 109 (41), 16720–16725.

Anticevic, A., Repovs, G., Corlett, P.R., Barch, D.M., 2011. Negative and nonemotional interference with visual working memory in schizophrenia. Biol. Psychiatry 70 (12), 1159–1168.

Artaloytia, J.F., Arango, C., Lahti, A., Sanz, J., Pascual, A., Cubero, P., Prieto, D., Palomo, T., 2006. Negative signs and symptoms secondary to antipsychotics: a double-blind, randomized trial of a single dose of placebo, haloperidol, and risperidone in healthy volunteers. Am. J. Psychiatry 163 (3), 488–493.

Austin, M.P., Mitchell, P., Goodwin, G.M., 2001. Cognitive deficits in depression. Br. J. Psychiatry 178 (3), 200–2006.

Averbeck, B.B., Evans, S., Chouhan, V., Bristow, E., Shergill, S.S., 2011. Probabilistic learning and inference in schizophrenia. Schizophr. Res. 127 (1–3), 115–122.

Baddeley, A.D., 1987. Working Memory. Oxford University Press.

Berkes, P., Orbán, G., Lengyel, M., Fiser, J., 2011. Spontaneous cortical activity reveals hallmarks of an optimal internal model of the environment. Science (New York, N.Y.) 331 (6013), 83–87.

Berridge, K., 1998. What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? Brain Res. Brain Res. Rev.

Berry, N., Jobanputra, V., Pal, H., 2003. Molecular genetics of schizophrenia: a critical review. J. Psychiatry Neurosci. 28 (6), 415–429.

Bertolino, A., Blasi, G., 2009. The genetics of schizophrenia. Neuroscience 164 (1), 288–299.

Bhugra, D., 2005. The global prevalence of schizophrenia. PLoS Med. 2 (5) e151–quiz e175.

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 (3), 312–328.

Braver, T.S., Cohen, J.D., 1999. Dopamine, cognitive control, and schizophrenia: the gating model. Prog. Brain Res. 121, 327–350.

Brennand, K.J., Simone, A., Jou, J., Gelboin-Burkhart, C., Tran, N., Sangar, S., Li, Y., Mu, Y., Chen, G., Yu, D., McCarthy, S., Sebat, J., Gage, F.H., 2011. Modelling schizophrenia using human induced pluripotent stem cells. Nature 473 (7346), 221–225.

Bromberg-Martin, E.S., Matsumoto, M., Hikosaka, O., 2010. Dopamine in motivational control: rewarding, aversive, and alerting. Neuron 68 (5), 815–834.

Caldwell, C.B., Gottesman, I.I., 1990. Schizophrenics kill themselves too: a review of risk factors for suicide. Schizophr. Bull. 16 (4), 571–589.

Callaway, E., Naghdi, S., 1982. An information processing model for schizophrenia. Arch. Gen. Psychiatry 39 (3), 339–347.

Carr, V.J., Wale, J., 1986. Schizophrenia: an information processing model. Aust. N. Z. J. Psychiatry 20 (2), 136–155.

Carter, J.R., Neufeld, R.W.J., 2007. Cognitive processing of facial affect: connectionist model of deviations in schizophrenia. J. Abnorm. Psychol. 116 (2), 290–305.

Chekroud, A.M., Zotti, R.J., Shehzad, Z., Gueorguieva, R., Johnson, M.K., Trivedi, M.H., Cannon, T.D., Krystal, J.H., Corlett, P.R., 2016. Cross-trial prediction of treatment outcome in depression: a machine learning approach. Lancet Psychiatry 3 (3), 243–250.

Chen, E.Y., 1994. A neural network model of cortical information processing in schizophrenia. I. Interaction between biological and social factors in symptom formation. Can. J. Psychiatry [Revue canadienne de psychiatrie] 39 (8), 362–367.

Chen, E.Y., 1995. A neural network model of cortical information processing in schizophrenia. II. Role of hippocampal-cortical interaction: a review and a model. Can. J. Psychiatry [Revue canadienne de psychiatrie] 40 (1), 21–26.

Chubb, J.E., Bradshaw, N.J., Soares, D.C., Porteous, D.J., Millar, J.K., 2008. The DISC locus in psychiatric illness. Mol. Psychiatry 13 (1), 36–64.

Cohen, J.D., Braver, T.S., O'Reilly, R.C., 1996. A computational approach to prefrontal cortex, cognitive control and schizophrenia: recent developments and current challenges. Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. 351 (1346), 1515–1527.

Cohen, J.D., Servan-Schreiber, D., 1992. Context, cortex, and dopamine: a connectionist approach to behavior and biology in schizophrenia. Psychol. Rev. 99 (1), 45–77.

Cohen, J.D., Servan-Schreiber, D., 1993. A theory of dopamine function and its role in cognitive deficits in schizophrenia. Schizophr. Bull. 19 (1), 85–104.

Collins, A.G.E., Brown, J.K., Gold, J.M., Waltz, J.A., Frank, M.J., 2014. Working memory contributions to reinforcement learning impairments in schizophrenia. J. Neurosci. 34 (41), 13747–13756.

Colombo, M., Series, P., 2012. Bayes in the brain – on Bayesian modelling in neuroscience. Br. J. Philos. Sci. 63, 697–723.

Cools, R., D'Esposito, M., 2011. Inverted-U-shaped dopamine actions on human working memory and cognitive control. Biol. Psychiatry 69 (12), e113–25.

Corlett, P.R., Cambridge, V.C., Gardner, J.M., Pigott, J.S., Turner, D.C., Everitt, J.C., Arana, F.S., Morgan, H.L., Milton, A.L., Lee, J.L., Aitken, M.R.F., Dickinson, A., Everitt, B.J., Absalom, A.R., Adapa, R., Subramanian, N., Taylor, J.R., Krystal, J.H., Fletcher, P.C., 2013. Ketamine effects on memory reconsolidation favor a learning model of delusions. PLOS ONE 8 (6), e65088.

Corlett, P.R., Fletcher, P.C., 2012. The neurobiology of schizotypy: fronto-striatal prediction error signal correlates with delusion-like beliefs in healthy people. Neuropsychologia 50 (14), 3612–3620.

Corlett, P.R., Fletcher, P.C., 2015. Delusions and prediction error: clarifying the roles of behavioural and brain responses. Cogn. Neuropsychiatry 20 (2), 95–105.

Corlett, P.R., Frith, C.D., Fletcher, P.C., 2009a. From drugs to deprivation: a Bayesian framework for understanding models of psychosis. Psychopharmacology 206 (4), 515–530.

Corlett, P.R., Honey, G.D., Fletcher, P.C., 2007a. From prediction error to psychosis: ketamine as a pharmacological model of delusions. J. Psychopharmacol. (Oxford, England) 21 (3), 238–252.

Corlett, P.R., Honey, G.D., Fletcher, P.C., 2016. Prediction error, ketamine and psychosis: an updated model. J. Psychopharmacol. 30 (11), 1145–1155.

Corlett, P.R., Honey, G.D., Krystal, J.H., Fletcher, P.C., 2011. Glutamatergic model psychoses: prediction error, learning, and inference. Neuropsychopharmacology 36 (1), 294–315.

Corlett, P.R., Krystal, J.H., Taylor, J.R., Fletcher, P.C., 2009b. Why do delusions persist? Front. Human Neurosci. 3, 12.

Corlett, P.R., Murray, G.K., Honey, G.D., Aitken, M.R.F., Shanks, D.R., Robbins, T.W., Bullmore, E.T., Dickinson, A., Fletcher, P.C., 2007b. Disrupted prediction-error signal in psychosis: evidence for an associative account of delusions. Brain 130 (Pt 9), 2387–2400.

Corlett, P.R., Simons, J.S., Pigott, J.S., Gardner, J.M., Murray, G.K., Krystal, J.H., Fletcher, P.C., 2009c. Illusions and delusions: relating experimentally-induced false memories to anomalous experiences and ideas. Front. Behav. Neurosci. 3, 53.

Crawford, T.J., Hamm, J.P., Kean, M., Schmechtig, A., Kumari, V., Anilkumar, A.P., Ettinger, U., 2010. The perception of real and illusory motion in schizophrenia. Neuropsychologia 48 (10), 3121–3127.

Cuthbert, B.N., Insel, T.R., 2013. Toward the future of psychiatric diagnosis: the seven pillars of RDoC. BMC Med. 11, 126.

Dauvermann, M.R., Whalley, H.C., Romaniuk, L., Valton, V., Owens, D.G.C., Johnstone, E.C., Lawrie, S.M., Moorhead, T.W.J., 2013. The application of nonlinear dynamic causal modelling for fMRI in subjects at high genetic risk of schizophrenia. NeuroImage 73, 16–29.

David, A.S., 1994. Dysmodularity: a neurocognitive model for schizophrenia. Schizophr. Bull. 20 (2), 249–255.

David, A.S., 2010. Why we need more debate on whether psychotic symptoms lie on a continuum with normality. Psychol. Med. 40 (12), 1935–1942.

Dayan, P., Abbott, L.F., 2005. Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems. MIT Press.

Dayan, P., Durstewitz, D., 2016. Computational modelling: after the middle ages of computational neuroscience. Curr. Opin. Behav. Sci. 11, 5–6.

de la Fuente-Sandoval, C., Portillo, V., Fresán, A., Apiquian, R., 2005. Replication of a computer model of auditory hallucinations in schizophrenia. Actas españolas de psiquiatría 33 (3), 141–146.

Deneve, S., Jardri, R., 2016. Circular inference: mistaken belief, misplaced trust. Curr. Opin. Behav. Sci. 11, 40–48.

Dima, D., Dietrich, D.E., Dillo, W., Emrich, H.M., 2010. Impaired top-down processes in schizophrenia: a DCM study of ERPs. NeuroImage 52 (3), 824–832.

Dima, D., Roiser, J.P., Dietrich, D.E., Bonnemann, C., Lanfermann, H., Emrich, H.M., Dillo, W., 2009. Understanding why patients with schizophrenia do not perceive the hollow-mask illusion using dynamic causal modelling. NeuroImage 46 (4), 1180–1186.

Diwadkar, V.A., Flaugher, B., Jones, T., Zalányi, L., Ujfalussy, B., Keshavan, M.S., Erdi, P., 2008. Impaired associative learning in schizophrenia: behavioral and computational studies. Cogn. Neurodyn. 2 (3), 207–219.

DSM-5, 2013. Diagnostic and Statistical Manual of Mental Disorders, 5th ed. DSM-5®. American Psychiatric Association.

DSM-IV-TR, 2000. Diagnostic and Statistical Manual of Mental Disorders, 4th ed. DSM-IV-TR®. American Psychiatric Association.

Durstewitz, D., Seamans, J.K., 2008. The dual-state theory of prefrontal cortex dopamine function with relevance to catechol-o-methyltransferase genotypes and schizophrenia. Biol. Psychiatry 64 (9), 739–749.

Evans, S., Almahdi, B., Sultan, P., Sohanpal, I., Brandner, B., Collier, T., Shergill, S.S., Cregg, R., Averbeck, B.B., 2012. Performance on a probabilistic inference task in healthy subjects receiving ketamine compared with patients with schizophrenia. J. Psychopharmacol. 26 (9), 1211–1217.

Fletcher, P.C., Frith, C.D., 2009. Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia. Nat. Rev. Neurosci. 10 (1), 48–58.

Forrest, A.D., Coto, C.A., Siegel, S.J., 2014. Animal models of psychosis: current state and future directions. Curr. Behav. Neurosci. Rep. 1 (2), 100–116.

Frank, M.J., 2008. Schizophrenia: a computational reinforcement learning perspective. Schizophr. Bull. 34 (6), 1008–1011.

Frank, M.J., Claus, E.D., 2006. Anatomy of a decision: striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal. Psychol. Rev. 113 (2),

300–326.

Frank, M.J., Seeberger, L.C., O'Reilly, R.C., 2004. By carrot or by stick: cognitive re-inforcement learning in Parkinsonism. Science (New York, N.Y.) 306 (5703), 1940–1943.

Franklin, D.W., Wolpert, D.M., 2011. Computational mechanisms of sensorimotor control. Neuron 72 (3), 425–442.

Freeman, D., Pugh, K., Garety, P., 2008. Jumping to conclusions and paranoid ideation in the general population. Schizophr. Res. 102 (1–3), 254–260.

Freeman, D., Startup, H., Dunn, G., Černis, E., Wingham, G., Pugh, K., Cordwell, J., Mander, H., Kingdon, D., 2014. Understanding jumping to conclusions in patients with persecutory delusions: working memory and intolerance of uncertainty. Psychol. Med. 44 (14), 3017–3024.

Friston, K., Brown, H.R., Siemerkus, J., Stephan, K.E., 2016. The dysconnection hy-pothesis. Schizophr. Res. 176 (2–3), 83–94.

Friston, K.J., 1996. Theoretical neurobiology and schizophrenia. Br. Med. Bull. 52 (3), 644–655.

Friston, K.J., 2005. A theory of cortical responses. Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. 360 (1456), 815–836.

Friston, K.J., 2010. The free-energy principle: a unified brain theory? Nat. Rev. Neurosci. 11 (2), 127–138.

Frith, C.D., Friston, K.J., 2012. False perceptions and false beliefs: understanding schi-zophrenia. Working Group on Neurosciences and the Human Person: New Perspectives on Human Activities. Pontif. Acad. Sci. 121 (8–10), 320.

Frith, C.D., Leary, J., Cahill, C., Johnstone, E.C., 1991. Performance on psychological tests. Demographic and clinical correlates of the results of these tests. Br. J. Psychiatry (Suppl. 13), 26–29 44–46.

Fusar-Poli, P., Meyer-Lindenberg, A., 2013. Striatal presynaptic dopamine in schizo-phrenia. Part II: meta-analysis of [18F/11C]-DOPA PET studies. Schizophr. Bull. 39 (1), 33–42.

Garety, P., Joyce, E.M., Jolley, S., Emsley, R., Waller, H., Kuipers, E., Bebbington, P., Fowler, D., Dunn, G., Freeman, D., 2013. Neuropsychological functioning and jumping to conclusions in delusions. Schizophr. Res. 150 (2–3), 570–574.

Garety, P.A., Freeman, D., 2013. The past and future of delusions research: from the inexplicable to the treatable. Br. J. Psychiatry 203 (5), 327–333.

Gershman, S.J., Beck, J.M., 2016. Complex probabilistic inference: from cognition to neural computation. Comput. Models Brain Behav.

Gilmour, G., Dix, S., Fellini, L., Gastambide, F., Plath, N., Steckler, T., Talpos, J., Tricklebank, M., 2012. NMDA receptors, cognition and schizophrenia – testing the validity of the NMDA receptor hypofunction hypothesis. NeuroImage 62 (3), 1401–1412.

Gold, J.M., Waltz, J.A., Frank, M.J., 2015. Effort cost computation in schizophrenia: a commentary on the recent literature. Biol. Psychiatry 78, 747–753.

Grace, A.A., 1991. Phasic versus tonic dopamine release and the modulation of dopamine system responsivity: a hypothesis for the etiology of schizophrenia. Neuroscience 41 (1), 1–24.

Grace, A.A., 2016. Dysregulation of the dopamine system in the pathophysiology of schizophrenia and depression. Nat. Rev. Neurosci. 17 (8), 524–532.

Gradin, V.B., Kumar, P., Waiter, G., Ahearn, T., Stickle, C., Milders, M., Reid, I., Hall, J., Steele, J.D., 2011. Expected value and prediction error abnormalities in depression and schizophrenia. Brain 134 (Pt 6), 1751–1764.

Grasemann, U., Miikkulainen, R., Hoffman, R.E., 2009. Hyperlearning: a connectionist model of psychosis in schizophrenia. In: Taatgen, N.A., van Rijn, H. (Eds.), Proceedings of the 31st Annual Meeting of the Cognitive Science Society. Amsterdam, The Netherlands.

Gray, J.A., Feldon, J., Rawlins, J.N.P., Hemsley, D.R., Smith, A.D., 1991. The neu-ropsychology of schizophrenia. Behav. Brain Sci. 14 (20), 1–20.

Greenstein-Messica, A., Ruppin, E., 1998. Synaptic runaway in associative networks and the pathogenesis of schizophrenia. Neural Comput. 10 (2), 451–465.

Grossberg, S., 1999. Neural models of normal and abnormal behavior: what do schizo-phrenia, Parkinsonism, attention deficit disorder, and depression have in common? Prog. Brain Res. 121, 375–406.

Grossberg, S., 2000. The imbalanced brain: from normal behavior to schizophrenia. Biol. Psychiatry 48 (2), 81–98.

Gruber, A.J., Dayan, P., Gutkin, B.S., Solla, S.A., 2006. Dopamine modulation in the basal ganglia locks the gate to working memory. J. Comput. Neurosci. 20 (2), 153–166.

Hall, J., Romaniuk, L., McIntosh, A.M., Steele, J.D., Johnstone, E.C., Lawrie, S.M., 2009. Associative learning and the genetics of schizophrenia. Trends Neurosci. 32 (6), 359–365.

Harrison, P.J., 1999. The neuropathology of schizophrenia: a critical review of the data and their interpretation. Brain 122 (4), 593–624.

Hoffman, R.E., 1987. Computer simulations of neural information processing and the schizophrenia-mania dichotomy. Arch. Gen. Psychiatry 44 (2), 178–188.

Hoffman, R.E., 1997. Neural network simulations, cortical connectivity, and schizo-phrenic psychosis. MD Comput.: Comput. Med. Pract. 14 (3), 200–208.

Hoffman, R.E., Dobscha, S.K., 1989. Cortical pruning and the development of schizo-phrenia: a computer model. Schizophr. Bull. 15 (3), 477–490.

Hoffman, R.E., Grasemann, U., Gueorguieva, R., Quinlan, D.M., Lane, D., Miikkulainen, R., 2011. Using computational patients to evaluate illness mechanisms in schizo-phrenia. Biol. Psychiatry 69 (10), 997–1005.

Hoffman, R.E., McGlashan, T.H., 1993a. Neurodynamics and schizophrenia research: editors' introduction. Schizophr. Bull. 19 (1), 15–19.

Hoffman, R.E., McGlashan, T.H., 1993b. Parallel distributed processing and the emer-gence of schizophrenic symptoms. Schizophr. Bull. 19 (1), 119–140.

Hoffman, R.E., McGlashan, T.H., 1997. Synaptic elimination, neurodevelopment, and the mechanism of hallucinated "voices" in schizophrenia. Am. J. Psychiatry 154 (12), 1683–1689.

Hoffman, R.E., McGlashan, T.H., 1999. Using a speech perception neural network si-mulation to explore normal neurodevelopment and hallucinated 'voices' in schizo-phrenia. Prog. Brain Res. 121, 311–325.

Hoffman, R.E., McGlashan, T.H., 2001. Neural network models of schizophrenia. Neuroscientist 7 (5), 441–454.

Hoffman, R.E., McGlashan, T.H., 2006. Using a speech perception neural network com-puter simulation to contrast neuroanatomic versus neuromodulatory models of au-ditory hallucinations. Pharmacopsychiatry 39 (Suppl. 1), S54L 64.

Hoffmann, H., Kupper, Z., 1997. Relationships between social competence, psycho-pathology and work performance and their predictive value for vocational re-habilitation of schizophrenic outpatients. Schizophr. Res. 23 (1), 69–79.

Honey, G.D., O'loughlin, C., Turner, D.C., Pomarol-Clotet, E., Corlett, P.R., Fletcher, P.C., 2006. 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 31 (2), 413–423.

Hopfield, J.J., 1982. Neural networks and physical systems with emergent collective computational abilities. Proc. Natl. Acad. Sci. U.S.A 79, 2554–2558.

Hopfield, J.J., 1984. Neurons with graded response have collective computational properties like those of two-state neurons. Proc. Natl. Acad. Sci. U.S.A. 81, 3088–3092.

Horn, D., Ruppin, E., 1995. Compensatory mechanisms in an attractor neural network model of schizophrenia. Neural Comput. 7 (1), 182–205.

Horton, H.K., Silverstein, S.M., 2011. Visual context processing deficits in schizophrenia: effects of deafness and disorganization. Schizophr. Bull. 37 (4), 716–726.

Howes, O.D., Kapur, S., 2009. The Dopamine hypothesis of schizophrenia: version III – the final common pathway. Schizophr. Bull. 35 (3), 549–562.

Howes, O.D., McCutcheon, R., Owen, M.J., Murray, R., 2016. The role of genes, stress and dopamine in the development of schizophrenia. Biol. Psychiatry.

Howes, O.D., Murray, R.M., 2014. Schizophrenia: an integrated sociodevelopmental-cognitive model. Lancet 383 (9929), 1677–1687.

Hsu, P.-C., Yang, U.-C., Shih, K.-H., Liu, C.-M., Liu, Y.-L., Hwu, H.-G., 2008. A protein interaction based model for schizophrenia study. BMC Bioinformatics 9 (Suppl 12), S23.

Huq, S.F., Garety, P.A., Hemsley, D.R., 1988. Probabilistic judgements in deluded and non-deluded subjects. Q. J. Exp. Psychol. A Hum. Exp. Psychol. 40 (4), 801–812.

Huth, A.G., de Heer, W.A., Griffiths, T.L., Theunissen, F.E., Gallant, J.L., 2016. Natural speech reveals the semantic maps that tile human cerebral cortex. Nature 532 (7600), 453–458.

Huys, Q.J., 2013. Computational Psychiatry. Encyclopaedia of Computational Neuroscience.

Huys, Q.J.M., Maia, T.V., Frank, M.J., 2016. Computational psychiatry as a bridge from neuroscience to clinical applications. Nat. Neurosci. 19, 404–413.

Huys, Q.J.M., Moutoussis, M., Williams, J., 2011. Are computational models of any use to psychiatry? Neural Netw. 24 (6), 544–551.

Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D.S., Quinn, K., Sanislow, C., Wang, P., 2010. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. Am. J. Psychiatry 167 (7), 748–751.

Insel, T.R., 2010. Rethinking schizophrenia. Nature 468 (7321), 187–193.

Jacobs, R.A., Kruschke, J.K., 2011. Bayesian learning theory applied to human cognition. Wiley Interdisc. Rev. Cogn. Sci. 2 (1), 8–21.

Jardri, R., Cachia, A., 2013. The Neuroscience of Hallucinations.

Jardri, R., Deneve, S., 2013. Circular inferences in schizophrenia. Brain 136 (11), 3227–3241.

Jardri, R., Duverne, S., Litvinova, A.S., Deneve, S., 2017. Experimental evidence for cir-cular inference in schizophrenia. Nat. Commun. 8, 14218.

Jardri, R., Hugdahl, K., Hughes, M., Brunelin, J., Waters, F., Alderson-Day, B., Smailes, D., Sterzer, P., Corlett, P.R., Leptourgos, P., Debbané, M., Cachia, A., Deneve, S., 2016. Are hallucinations due to an imbalance between excitatory and inhibitory influences on the brain? Schizophr. Bull. 42 (5), 1124–1134.

Javanbakht, A., 2005. The theory of bowl and bugs: a model for the explanation of the coexistence of psychological and biological etiologies in the psychosis. J. Am. Acad. Psychoanal. Dyn. Psychiatry 33 (2), 363–375.

Javanbakht, A., 2006. Sensory gating deficits, pattern completion, and disturbed fronto-limbic balance, a model for description of hallucinations and delusions in schizo-phrenia. Med. Hypo. 67 (5), 1173–1184.

Javitt, D.C., 1987. Negative schizophrenic symptomatology and the PCP (phencyclidine) model of schizophrenia. Hillside J. Clin. Psychiatry 9 (1), 12–35.

Javitt, D.C., Zukin, S.R., 1991. Recent advances in the phencyclidine model of schizo-phrenia. Am. J. Psychiatry 148 (10), 1301–1308.

Jentsch, J., Roth, R., 1999. The neuropsychopharmacology of phencyclidine: from NMDA receptor hypofunction to the dopamine hypothesis of schizophrenia. Neuropsychopharmacology 20 (3), 201–225.

Jobe, T.H., Harrow, M., Martin, E.M., Whitfield, H.J., Sands, J.R., 1994. Schizophrenic deficits: neuroleptics and the prefrontal cortex. Schizophr. Bull. 20 (3), 413–416 discussion 417–421.

Johns, L.C., van Os, J., 2001. The continuity of psychotic experiences in the general population. Clin. Psychol. Rev. 21 (8), 1125–1141.

Johnston, P.J., Katsikitis, M., Carr, V.J., 2001. A generalised deficit can account for problems in facial emotion recognition in schizophrenia. Biol. Psychol. 58 (3), 203–227.

Johnstone, E.C., Leary, J., Frith, C.D., Owens, D.G.C., 1991. Disabilities and circum-stances of schizophrenic patients – a follow-up study. Police contact. Br. J. Psychiatry (Suppl. 13), 37–39 44–46.

Jones, P.B., Rodgers, B., Murray, R.M., Marmot, M., 1994. Child development risk factors for adult schizophrenia in the British 1946 birth cohort. Lancet 344 (8934), 1398–1402.

Joyce, D.W., Averbeck, B.B., Frith, C.D., Shergill, S.S., 2013. Examining belief and confidence in schizophrenia. Psychol. Med. 43, 2327–2338.

Kandel, E.R., Schwartz, J., Jessell, T.M., Siegelbaum, S., Hudspeth, A.J., 2013. Principles of Neural Science, 5th ed. McGraw Hill Professional.

Kapur, S., 2003. Psychosis as a state of aberrant salience: a framework linking biology, phenomenology, and pharmacology in schizophrenia. Am. J. Psychiatry.

Keane, B.P., Silverstein, S.M., Wang, Y., Papathomas, T.V., 2013. Reduced depth inversion illusions in schizophrenia are state-specific and occur for multiple object types and viewing conditions. J. Abnorm. Psychol. 122 (2), 506–512.

Knapp, M., Mangalore, R., Simon, J., 2004. The global costs of schizophrenia. Schizophr. Bull. 30 (2), 279–293.

Kreczmanski, P., Heinsen, H., Mantua, V., Woltersdorf, F., Masson, T., Ulfig, N., Schmidt-Kastner, R., Korr, H., Steinbusch, H.W.M., Hof, P.R., Schmitz, C., 2007. Volume, neuron density and total neuron number in five subcortical regions in schizophrenia. Brain 130 (Pt 3), 678–692.

Kubicki, M., McCarley, R., Westin, C.-F., Park, H.-J., Maier, S., Kikinis, R., Jolesz, F.A., Shenton, M.E., 2007. A review of diffusion tensor imaging studies in schizophrenia. J. Psychiatr. Res. 41 (1-2), 15–30.

Kubicki, M., Westin, C.-F., McCarley, R.W., Shenton, M.E., 2005. The application of DTI to investigate white matter abnormalities in schizophrenia. Ann. N. Y. Acad. Sci. 1064, 134–148.

Lawrie, S.M., Buechel, C., Whalley, H.C., Frith, C.D., Friston, K.J., Johnstone, E.C., 2002. Reduced frontotemporal functional connectivity in schizophrenia associated with auditory hallucinations. Biol. Psychiatry 51 (12), 1008–1011.

Lawrie, S.M., Hall, J., McIntosh, A.M., Owens, D.G.C., Johnstone, E.C., 2010. The 'continuum of psychosis': scientifically unproven and clinically impractical. Br. J. Psychiatry 197 (6), 423–425.

Lawrie, S.M., McIntosh, A.M., Hall, J., Owens, D.G.C., Johnstone, E.C., 2008. Brain structure and function changes during the development of schizophrenia: the evidence from studies of subjects at increased genetic risk. Schizophr. Bull. 34 (2), 330–340.

Lewis, D., Hashimoto, T., 2005. Cortical inhibitory neurons and schizophrenia. Nat. Rev. 6, 312–324.

Lewis, D.A., Gonzalez-Burgos, G., 2006. Pathophysiologically based treatment interventions in schizophrenia. Nat. Med. 12 (9), 1016–1022.

Linscott, R.J., van Os, J., 2010. Systematic reviews of categorical versus continuum models in psychosis: evidence for discontinuous subpopulations underlying a psychometric continuum. Implications for DSM-V, DSM-VI, and DSM-VII. Annu. Rev. Clin. Psychol. 6, 391–419.

Lochmann, T., Deneve, S., 2011. Neural processing as causal inference. Curr. Opin. Neurobiol. 21 (5), 774–781.

Loh, M., Rolls, E.T., Deco, G., 2007. A dynamical systems hypothesis of schizophrenia. PLoS Comput. Biol. 3 (11), e228.

Lysaker, P., Bell, M., 1995. Negative symptoms and vocational impairment in schizophrenia: repeated measurements of work performance over six months. Acta Psychiatr. Scand. 91 (3), 205–208.

Ma, W.J., Beck, J.M., Latham, P.E., Pouget, A., 2006. Bayesian inference with probabilistic population codes. Nat. Neurosci. 9 (11), 1432–1438.

Maia, T.V., Frank, M.J., 2011. From reinforcement learning models to psychiatric and neurological disorders. Nat. Neurosci. 14 (2), 154–162.

Maia, T.V., Frank, M.J., 2016. An integrative perspective on the role of dopamine in schizophrenia. Biol. Psychiatry S0006-3223 (16), 32425–32428.

Mangalore, R., Knapp, M., 2007. Cost of schizophrenia in England. J. Mental Health Policy Econ. 10 (1), 23–41.

Manoach, D.S., Press, D.Z., Thangaraj, V., Searl, M.M., Goff, D.C., Halpern, E., Saper, C.B., Warach, S., 1999. Schizophrenic subjects activate dorsolateral prefrontal cortex during a working memory task, as measured by fMRI. Biol. Psychiatry 45 (9), 1128–1137.

Marr, D., 1982. Vision. A Computational Investigation Into the Human Representation and Processing of Visual Information. Freeman, New York, NY, USA.

McDonald, C., Murray, R.M., 2000. Early and late environmental risk factors for schizophrenia. Brain Res. Brain Res. Rev. 31 (2–3), 130–137.

McGlashan, T.H., Hoffman, R.E., 2000. Schizophrenia as a disorder of developmentally reduced synaptic connectivity. Arch. Gen. Psychiatry 57 (7), 637–648.

McIntosh, A.M., Harrison, L.K., Forrester, K., Lawrie, S.M., Johnstone, E.C., 2005. Neuropsychological impairments in people with schizophrenia or bipolar disorder and their unaffected relatives. Br. J. Psychiatry 186, 378–385.

Meeter, M., Murre, J.M.J., Talamini, L.M., 2002. A computational approach to memory deficits in schizophrenia. Neurocomputing 44, 929–936.

Meyer-Lindenberg, A., Kohn, P., Kolachana, B., 2005. Midbrain dopamine and prefrontal function in humans: interaction and modulation. Nat. Neurosci. 8, 594–596.

Monchi, O., Taylor, J.G., Dagher, A., 2000. A neural model of working memory processes in normal subjects, Parkinson's disease and schizophrenia for fMRI design and predictions. Neural Netw. 13 (8–9), 953–973.

Montague, P.R., Dolan, R.J., Friston, K.J., Dayan, P., 2012. Computational psychiatry. Trends Cogn. Sci. 16 (1), 72–80.

Montague, P.R., Hyman, S.E., Cohen, J.D., 2004. Computational roles for dopamine in behavioural control. Nature 431 (7010), 760–767.

Moore, J.W., Turner, D.C., Corlett, P.R., Arana, F.S., Morgan, H.L., Absalom, A.R., Adapa, R., de Wit, S., Everitt, J.C., Gardner, J.M., Pigott, J.S., Haggard, P., Fletcher, P.C., 2011. Ketamine administration in healthy volunteers reproduces aberrant agency experiences associated with schizophrenia. Cogn. Neuropsychiatry 1–18.

Mortensen, P.B., Pedersen, C.B., Westergaard, T., Wohlfahrt, J., Ewald, H., Mors, O., Andersen, P.K., Melbye, M., 1999. Effects of family history and place and season of birth on the risk of schizophrenia. N. Engl. J. Med. 340 (8), 603–608.

Moustafa, A.A., Cohen, M.X., Sherman, S.J., Frank, M.J., 2008a. A role for dopamine in temporal decision making and reward maximization in parkinsonism. J. Neurosci. 28 (47), 12294–12304.

Moustafa, A.A., Sherman, S.J., Frank, M.J., 2008b. A dopaminergic basis for working memory, learning and attentional shifting in Parkinsonism. Neuropsychologia 46 (13), 3144–3156.

Murray, G.K., Corlett, P.R., Clark, L., Pessiglione, M., Blackwell, A.D., Honey, G.D., Jones, P.B., Bullmore, E.T., Robbins, T.W., Fletcher, P.C., 2008. Substantia nigra/ventral tegmental reward prediction error disruption in psychosis. Mol. Psychiatry 13 (3), 267–276.

Murray, J.D., Anticevic, A., Gancsos, M., Ichinose, M., Corlett, P.R., Krystal, J.H., Wang, X.-J., 2012. Linking microcircuit dysfunction to cognitive impairment: effects of disinhibition associated with schizophrenia in a cortical working memory model. Cereb. Cortex (New York, N.Y., 1991) 24 (4), 859–872.

Nakazawa, K., Zsiros, V., Jiang, Z., Nakao, K., Kolata, S., Zhang, S., Belforte, J.E., 2012. GABAergic interneuron origin of schizophrenia pathophysiology. Neuropharmacology 62 (3), 1574–1583.

Narendran, R., Frankle, W.G., Keefe, R., Gil, R., Martinez, D., Slifstein, M., Kegeles, L.S., Talbot, P.S., Huang, Y., Hwang, D.-R., Khenissi, L., Cooper, T.B., Laruelle, M., Abi-Dargham, A., 2005. Altered prefrontal dopaminergic function in chronic recreational ketamine users. Am. J. Psychiatry 162 (12), 2352–2359.

Notredame, C.-E., Pins, D., Deneve, S., Jardri, R., 2014. What visual illusions teach us about schizophrenia. Front. Integr. Neurosci. 8.

Olney, J.W., Newcomer, J.W., Farber, N.B., 1999. NMDA receptor hypofunction model of schizophrenia. NeuroImage 33 (6), 523–533.

O'Reilly, R.C., Frank, M.J., 2006. Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia. Neural Comput. 18 (2), 283–328.

Papanastasiou, E., Stone, J.M., Shergill, S., 2013. When the drugs don't work: the potential of glutamatergic antipsychotics in schizophrenia. Br. J. Psychiatry 202 (2), 91–93.

Pearl, J., 1982. Reverend Bayes on Inference Engines.

Peled, A., Geva, A.B., 2000. The perception of Rorschach inkblots in schizophrenia: a neural network model. Int. J. Neurosci. 104 (1–4), 49–61.

Peters, E., Joseph, S., Day, S., Garety, P., 2004. Measuring delusional ideation: the 21-item Peters et al. delusions inventory (PDI). Schizophr. Bull. 30 (4), 1005–1022.

Peters, E.R., Joseph, S.A., Garety, P.A., 1999. Measurement of delusional ideation in the normal population: introducing the PDI (Peters et al. delusions inventory). Schizophr. Bull. 25 (3), 553–576.

Pettersson-Yeo, W., Allen, P., Benetti, S., McGuire, P., Mechelli, A., 2011. Dysconnectivity in schizophrenia: where are we now? Neurosci. Biobehav. Rev. 35 (5), 1110–1124.

Powell, S.B., Zhou, X., Geyer, M.A., 2009. Prepulse inhibition and genetic mouse models of schizophrenia. Behav. Brain Res. 204 (2), 282–294.

Redish, A.D., Jensen, S., Johnson, A., 2008. A unified framework for addiction: vulnerabilities in the decision process. Behav. Brain Sci. 31 (4), 415–437 discussion 437–487.

Roiser, J.P., Howes, O.D., Chaddock, C.A., Joyce, E.M., McGuire, P., 2013. Neural and behavioral correlates of aberrant salience in individuals at risk for psychosis. Schizophr. Bull. 39 (6), 1328–1336.

Roiser, J.P., Stephan, K.E., den Ouden, H.E.M., Barnes, T.R.E., Friston, K.J., Joyce, E.M., 2009. Do patients with schizophrenia exhibit aberrant salience? Psychol. Med. 39 (2), 199–209.

Rolls, E.T., Deco, G., 2011. A computational neuroscience approach to schizophrenia and its onset. Neurosci. Biobehav. Rev. 35 (8), 1644–1653.

Rolls, E.T., Loh, M., Deco, G., Winterer, G., 2008. Computational models of schizophrenia and dopamine modulation in the prefrontal cortex. Nat. Rev. Neurosci. 9 (9), 696–709.

Romaniuk, L., Honey, G.D., King, J.R.L., Whalley, H.C., McIntosh, A.M., Levita, L., Hughes, M., Johnstone, E.C., Day, M., Lawrie, S.M., Hall, J., 2010. Midbrain activation during Pavlovian conditioning and delusional symptoms in schizophrenia. Arch. Gen. Psychiatry 67 (12), 1246–1254.

Ruppin, E., 1995. Neural modelling of psychiatric disorders. Network Comput. Neural Syst. 6 (4), 635–656.

Ruppin, E., Reggia, J.A., Horn, D., 1996. Pathogenesis of schizophrenic delusions and hallucinations: a neural model. Schizophr. Bull. 22 (1), 105–123.

Samartzis, L., Dima, D., Poli, P.F., Kyriakopoulos, M., 2014. White matter alterations in early stages of schizophrenia. A systematic review of diffusion tensor imaging studies. J. Neuroimaging 24, 101–110.

Sanborn, A.N., Chater, N., 2016. Bayesian brains without probabilities. Trends Cogn. Sci. 20 (12), 883–893.

Schmack, K., Gòmez-Carrillo de Castro, A., Rothkirch, M., Sekutowicz, M., Rössler, H., Haynes, J.-D., Heinz, A., Petrovic, P., Sterzer, P., 2013. Delusions and the role of beliefs in perceptual inference. J. Neurosci. 33 (34), 13701–13712.

Schmajuk, N., 2005. Brain-behaviour relationships in latent inhibition: a computational model. Neurosci. Biobehav. Rev. 29 (6), 1001–1020.

Schultz, W., 2016. Dopamine reward prediction-error signalling: a two-component response. Nat. Rev. Neurosci. 17 (3), 183–195.

Schultz, W., Dayan, P., Montague, P.R., 1997. A neural substrate of prediction and reward. Science (New York, N.Y.) 275 (5306), 1593–1599.

Seeman, M.V., 1994. Neural networks and schizophrenia. Can. J. Psychiatry [Revue canadienne de psychiatrie] 39 (8), 353.

Serretti, A., Mandelli, L., Bajo, E., Cevenini, N., Papili, P., Mori, E., Bigelli, M., Berardi, D., 2009. The socio-economical burden of schizophrenia: a simulation of cost-offset of early intervention program in Italy. Eur. Psychiatry 24 (1), 11–16.

Sherman, A.D., Davidson, A.T., Baruah, S., Hegwood, T.S., Waziri, R., 1991. Evidence of glutamatergic deficiency in schizophrenia. Neurosci. Lett. 121 (1-2), 77–80.

Siekmeier, P.J., 2009. Evidence of multistability in a realistic computer simulation of

hippocampus subfield CA1. Behav. Brain Res. 200 (1), 220–231.

Siekmeier, P.J., Hasselmo, M.E., Howard, M.W., Coyle, J., 2007. Modeling of context-dependent retrieval in hippocampal region CA1: implications for cognitive function in schizophrenia. Schizophr. Res. 89 (1-3), 177–190.

Siekmeier, P.J., Hoffman, R.E., 2002. Enhanced semantic priming in schizophrenia: a computer model based on excessive pruning of local connections in association cortex. Br. J. Psychiatry 180, 345–350.

Silverstein, S.M., Keane, B.P., 2011a. Perceptual organization impairment in schizophrenia and associated brain mechanisms: review of research from 2005 to 2010. Schizophr. Bull. 37 (4), 690–699.

Silverstein, S.M., Keane, B.P., 2011b. Vision science and schizophrenia research: toward a re-view of the disorder. Editors' introduction to special section. Schizophr. Bull. 37 (4), 681–689.

Smith, A.J., Becker, S., Kapur, S., 2003. From dopamine to psychosis: a computational approach. Knowledge-Based Intellignet Information and Engineering Systems, Pt 2, Proceedings 2774 (Chapter 152). pp. 1115–1121.

Smith, A.J., Becker, S., Kapur, S., 2005. A computational model of the functional role of the ventral-striatal D2 receptor in the expression of previously acquired behaviors. Neural Comput. 17 (2), 361–395.

Smith, A.J., Li, M., Becker, S., Kapur, S., 2004. A model of antipsychotic action in conditioned avoidance: a computational approach. Neuropsychopharmacology 29 (6), 1040–1049.

Smith, A.J., Li, M., Becker, S., Kapur, S., 2007. Linking animal models of psychosis to computational models of dopamine function. Neuropsychopharmacology 32 (1), 54–66.

Sokolov, E.N., 1960. In: Brazier, M.A. (Ed.), Neuronal Models and The Orienting Reflex.

Solinski, S., Jackson, H.J., Bell, R.C., 1992. Prediction of employability in schizophrenic patients. Schizophr. Res. 7 (2), 141–148.

Speechley, W.J., Whitman, J.C., Woodward, T.S., 2010. The contribution of hypersalience to the "jumping to conclusions" bias associated with delusions in schizophrenia. J. Psychiatry Neurosci. 35 (1), 7–17.

Spencer, K.M., 2009. The functional consequences of cortical circuit abnormalities on gamma oscillations in schizophrenia: insights from computational modeling. Front. Hum. Neurosci. 3, 33.

Spitzer, M., 1995. A neurocomputational approach to delusions. Comprehens. Psychiatry 36 (2), 83–105.

Stephan, K.E., Friston, K.J., Frith, C.D., 2009. Dysconnection in schizophrenia: from abnormal synaptic plasticity to failures of self-monitoring. Schizophr. Bull. 35 (3), 509–527.

Stephan, K.E., Mathys, C., 2014. Computational approaches to psychiatry. Curr. Opin. Neurobiol. 25, 85–92.

Strauss, J.S., 1969. Hallucinations and delusions as points on continua function. Arch. Gen. Psychiatry 21 (5), 581–586.

Suslow, T., Schonauer, K., Ohrmann, P., 2000. Prediction of work performance by clinical symptoms and cognitive skills in schizophrenic outpatients. J. Nervous Mental Dis.

Tanaka, S., 2008. Dysfunctional GABAergic inhibition in the prefrontal cortex leading to "psychotic" hyperactivation. BMC Neurosci. 9, 41.

Teufel, C., Subramanian, N., Dobler, V., Perez, J., Finnemann, J., Mehta, P.R., Goodyer, I.M., Fletcher, P.C., 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 (43), 13401–13406.

Torregrossa, M.M., Corlett, P.R., Taylor, J.R., 2011. Aberrant learning and memory in addiction. Neurobiol. Learn. Memory 96 (4), 609–623.

Tretter, F., Albus, M., 2007. "Computational Neuropsychiatry" of working memory disorders in schizophrenia: the network connectivity in prefrontal cortex – Data and models. Pharmacopsychiatry 40, S2–S16.

Tschacher, W., Schuler, D., Junghan, U., 2006. Reduced perception of the motion-induced blindness illusion in schizophrenia. Schizophr. Res. 81 (2–3), 261–267.

van Os, J., Hanssen, M., Bijl, R.V., Ravelli, A., 2000. Strauss (1969) revisited: a psychosis continuum in the general population? Schizophr. Res. 45 (1–2), 11–20.

van Os, J., Kapur, S., 2009. Schizophrenia. Lancet 374 (9690), 635–645.

Vijayraghavan, S., Wang, M., Birnbaum, S.G., Williams, G.V., Arnsten, A.F.T., 2007. Inverted-U dopamine D1 receptor actions on prefrontal neurons engaged in working memory. Nat. Neurosci. 10 (3), 376–384.

Waltz, J.A., Frank, M.J., Robinson, B.M., Gold, J.M., 2007. Selective reinforcement learning deficits in schizophrenia support predictions from computational models of striatal-cortical dysfunction. Biol. Psychiatry 62 (7), 756–764.

Waltz, J.A., Schweitzer, J.B., Gold, J.M., Kurup, P.K., Ross, T.J., Salmeron, B.J., Rose, E.J., McClure, S.M., Stein, E.A., 2009. Patients with schizophrenia have a reduced neural response to both unpredictable and predictable primary reinforcers. Neuropsychopharmacology 34 (6), 1567–1577.

Wang, M., Vijayraghavan, S., Goldman-Rakic, P.S., 2004. Selective D2 receptor actions on the functional circuitry of working memory. Science (New York, N.Y.) 303 (5659), 853–856.

Wang, X.-J., 2006. Toward a prefrontal microcircuit model for cognitive deficits in schizophrenia. Pharmacopsychiatry 39 (Suppl. 1), S80L 7.

Weinberg, D., Shahar, G., Davidson, L., McGlashan, T.H., Fennig, S., 2009. Longitudinal associations between negative symptoms and social functioning in schizophrenia: the moderating role of employment status and setting. Psychiatry 72 (4), 370–381.

Whitford, T.J., Ford, J.M., Mathalon, D.H., Kubicki, M., Shenton, M.E., 2012. Schizophrenia, myelination, and delayed corollary discharges: a hypothesis. Schizophr. Bull. 38 (3), 486–494.

Williams, L.E., Ramachandran, V.S., Hubbard, E.M., Braff, D.L., Light, G.A., 2010. Superior size-weight illusion performance in patients with schizophrenia: evidence for deficits in forward models. Schizophr. Res. 121 (1–3), 101–106.

Winterer, G., Weinberger, D.R., 2004. Genes, dopamine and cortical signal-to-noise ratio in schizophrenia. Trends Neurosci. 27 (11), 683–690.

Wise, R.A., 1978. Neuroleptics and operant behavior: the anhedonia hypothesis. Behav. Brain Sci. 5 (1), 39–53.

Wolf, J.A., Moyer, J.T., Lazarewicz, M.T., Contreras, D., Benoit-Marand, M., O'Donnell, P., Finkel, L.H., 2005. NMDA/AMPA ratio impacts state transitions and entrainment to oscillations in a computational model of the nucleus accumbens medium spiny projection neuron. J. Neurosci. 25 (40), 9080–9095.

Wolpert, D.M., Diedrichsen, J., Flanagan, J.R., 2011. Principles of sensorimotor learning. Nat. Rev. Neurosci. 12, 739–751.

World Health Organisation, 1996. Schizophrenia and Public Health. http://www.who.int/mental_health/media/en/55.pdf.

World Health Organization, 2009. The International Statistical Classification of Diseases and Related Health Problems 2008.