Predictive Processing and Source Monitoring in the Psychosis Continuum

Clara Sue Humpston

This Thesis is submitted for the degree of Doctor of Philosophy at Cardiff University

2017

Summary

Schizophrenia is a serious and debilitating mental illness, and sufferers frequently experience a multitude of symptoms. Of particular interest to the current Thesis are psychotic symptoms including delusions, hallucinations and associated self-disturbances such as interference in the agency and ownership of thoughts and actions. Since the disorder was first described over a century ago, research into the pathogenesis of schizophrenia has advanced greatly. However, there are still large gaps in the current knowledge and understanding of the neuropsychological bases of this devastating illness. The current Thesis adopts a cognitive neuropsychiatric approach and applies a continuum model to the construct of psychosis. The aim of the current Thesis was to incorporate theories such as the source monitoring and the predictive processing frameworks across a range of behavioural tasks, in order to investigate some of the neuropsychological deficits in schizotypy and early psychotic symptoms.

Healthy individuals with schizotypal traits and patients with early psychosis who did not yet meet a full diagnosis of schizophrenia underwent a battery of behavioural paradigms, with each task aimed at a different aspect of predictive processing and source monitoring. In healthy individuals, nonclinical psychosis-like experiences measured with schizotypy scales were significantly associated with difficulties in the source monitoring of actions, in particular deficits in reality monitoring and internal source monitoring. However, no significant relationships were found for the predictive processing tasks, which focused on the perceptual (force-matching), associative (Kamin blocking) and motivational (reversal learning) domains. In the patients with first episode psychosis, positive psychotic symptoms were not significantly correlated with specific deficits in either category of tasks, although this study was under-powered and strong conclusion could not be drawn. Nevertheless, these findings have provided support for partial dimensionality in psychosis vulnerability and will serve as foundations for future research on a larger scale.

Acknowledgements

This Thesis would not have been made possible with the help and support from many individuals both within and outside Cardiff University. Firstly, I would like to thank my supervisors Professor David Linden and Dr Lisa Evans for giving me the opportunity to carry out my own project, and their continued guidance and support throughout the last three years. I really could not have asked for more.

Secondly, I would like to thank Dr Christoph Teufel for offering his expertise in psychophysics and kindly letting me borrow the force-matching equipment essential for my research. I would like to thank Dr Niklas Ihssen for his knowledge in reversal learning, Dr James Walters, Dr Stefanie Linden and Dr Gemma Williams for their assistance with patient recruitment. My thanks also go to colleagues at Research Governance, especially Helen Falconer, for her support with the process of obtaining ethical and governance approvals for the patient study.

Thirdly, I must express my gratitude towards my collaborators in the NHS. I would like to thank Dr Bala Oruganti, Norman Young and Dr Ann-Marie Dunsby from Cardiff and Vale University Health Board, Chris Kelly and his Early Intervention for Psychosis team from Aneurin Bevan University Health Board, Mark Bartholomew, Bradley Jones, Abigail Jones, and Laura Armstrong-James from Avon and Wiltshire Mental Health Partnership Trust. Patient recruitment is always an arduous process and these individuals have dedicated their time and effort to help wherever possible.

Fourthly, I thank all the participants, regardless of their clinical status, who volunteered to take part in my studies. Thank you for offering your time and your openness to share your stories. Most importantly, thank you for your trust in my research and the contributions you have made not only to the completion of my PhD, but also the potential to help many others. I truly value all your input and willingness to recall your experiences, even if they are sometimes distressing, personal and poignant.

Lastly, there is one special person I must thank. He is the one who understands me the most and who has always been by my side. Thank you, my beloved husband John, for accompanying me through all my endeavours and pursuits, through disappointments as well as rewards. I am eternally grateful.

Table of Contents

List of Tables

List of Figures

Chapter One: An Overview of Psychosis

1.1 Introduction

The purpose of this Chapter is to introduce the concept of psychosis and its constituent symptoms such as delusions and hallucinations on multiple levels of explanation, from neurobiology to phenomenology. Drawing from current literature, this Chapter will also discuss the heterogeneity of psychotic symptoms, different approaches to the study of psychosis and implications for research.

1.1.1 What is 'Psychosis'?

The etymology of the word 'psychosis' derives from Greek, *psyche* (mind, soul) and *osis* (ailment, condition): in other words, psychosis is an illness of the mind or soul. However, the clinical picture is much more complex. Often synonymous with 'reality distortions' or 'a loss of touch with reality', psychosis disrupts if not destroys the sufferer's sense of reality and therefore has significant implications not only on a personal or medical level, but also on a philosophical level as to how a person's selfhood is embedded in the world. Nevertheless, psychosis itself is not viewed as a disease entity but a group of symptoms within other psychiatric syndromes such as schizophrenia, schizoaffective disorder, some personality disorders or even organic brain diseases such as dementia. Psychosis is of particular importance to the diagnosis of schizophrenia: in fact, it has been referred to as 'the clinical hallmark of schizophrenia' (Kapur, 2003, p.13). The terms 'psychosis' and 'schizophrenia' are sometimes (incorrectly) used interchangeably due to the close relationship between the two concepts. Even psychosis itself is not necessarily a singular concept (although there is the notion of 'unitary psychosis'; see Berrios and Beer, 1994) but consists of three major domains, namely delusions, hallucinations and thought disorder which are also called the 'positive' (i.e. *added* to the subject's experience) symptoms of schizophrenia.

Even the very definitions of delusions and hallucinations have not yet reached absolute consensus. However, the 'standard' definitions may seem straightforward at first glance: delusions are 'false beliefs about which a person is firmly convinced and

is impervious to outside contradictory evidence'. Hallucinations, on the other hand, are defined as 'sensory perceptions in the absence of any externally generated stimulus' (Lindenmayer and Khan, 2006). This Thesis will focus exclusively on delusions and hallucinations in the context of schizophrenia-spectrum psychoses.

1.1.2 Classification and Diagnosis

The earliest attempt to classify schizophrenia and related psychotic disorders can be traced back to Karl Kahlbaum's descriptions of catatonia in 1868 (Barnes, Saunders, Walls, Saunders, Kirk, 1986). This influenced Emil Kraepelin's dichotomous notions of 'dementia praecox' (*lit*. early dementia; see Bleuler, 1950) as the precedent to schizophrenia and 'manic-depressive illness' which evolved to what clinicians call bipolar affective disorder today. To Kraepelin, dementia praecox followed a chronic and deteriorating course with poor outcomes whereas manic depression often had an episodic course and more favourable outcomes. Soon after Kraepelin's formulation, Eugen Bleuler first coined the term 'schizophrenia' or 'the group of schizophrenias' which focused more on signs and symptoms rather than course and outcome (Andreasen and Carpenter, 1993). Bleuler suggested that the term 'schizophrenia' (*lit*. split mind) should replace 'dementia praecox' because the consequence of the latter was not always as severe and debilitating as Kraepelin thought. Bleuler thought the key feature of schizophrenia was in fact the dissociation and fragmentation of normally integrated mental processes such as thought and affect, a feature which was different from and more persistent than delusions and hallucinations.

However, perhaps the most influential contribution to the nosology of schizophrenia-spectrum psychoses which is reflected in the diagnostic systems today is that by Kurt Schneider. Schneider brought the focus back on specific types of delusions and hallucinations which were thought to be pathognomonic (i.e. defining features) of a schizophrenic psychosis. These 'Schneiderian First-rank Symptoms' include bizarre and implausible forms of delusion (e.g. thought interference symptoms such as thought insertion, broadcast and withdrawal, as well as passivity/delusions of control) partly based on Karl Jaspers' concept of incomprehensibility, and third-person auditory-verbal hallucinations which provide a running commentary on the patient's behaviour or voices arguing with one another. Interestingly, all first-rank symptoms constitute a breach in the patient's ego-boundary (i.e. breaking down of the

demarcation between self and other), especially thought interference (for example, in thought insertion the patient reports having thoughts that they did not author and did not belong to them, sometimes 'transmitted' from a specific external agent or entity), thought echo (patient hearing one's own voice spoken aloud) and delusions of control. These seemingly unintelligible and outright impossible phenomena which have no basis in external or consensual reality have attracted great interest, both in terms of clinical presentation and philosophical investigation, to the degree that some theorists propose that (schizophrenic) psychosis is essentially an extreme form of self-disturbance (Sass and Parnas, 2003).

Two of the most widely implemented diagnostic systems, the International Classification of Diseases (ICD) by the World Health Organisation and the American Psychiatric Association's Diagnostic and Statistical Manual of Mental Disorders (DSM), subdivide the psychosis spectrum into individual disorders from full-blown schizophrenia (which comprises of multiple subtypes including paranoid, disorganised, catatonic, etc.) to schizotypal disorder (schizotypal *personality* disorder in the DSM), affective psychoses (e.g. schizoaffective disorder and psychotic depression) and to brief and reactive psychotic disorder. Such distinctions are not always made on the basis of phenomenology but on duration and severity. Whereas the ICD is particularly influenced by Schneiderian first-rank symptoms in the diagnosis of schizophrenia (the presentation of only one of such symptoms is sufficient provided the duration is a month or longer), the DSM in general has less of an emphasis on the actual content of the delusions and hallucinations and requires a duration of at least six months (Table 1.1). Subtypes of schizophrenia have also been removed in DSM-5, as well as the 'bizarre delusion' criterion.

ICD-10	DSM-5
At least one of: • Thought echo, thought insertion/withdrawal/broadcast • Passivity, delusional perception • Third person auditory hallucination, running commentary • Persistent bizarre delusions	*Two or more of:* • Hallucinations • Bizarre or nonbizarre delusions • Disorganised speech • Grossly disorganised or catatonic behaviour • Negative symptoms
Or two or more of: • Persistent hallucinations • Thought disorder • Catatonic behaviour • Negative symptoms • Significant behaviour change	*At least one symptom must be:* • Hallucinations • Bizarre or nonbizarre delusions • Disorganised speech
Duration: • At least 1 month	*Duration:* • Continuous disturbance for 6 months (attenuated and residual symptoms) • Acute symptoms for at least 1 month (less if successfully treated)
Exclusion criteria: • Mood disorders, schizoaffective disorder • Overt brain disease • Drug intoxication or withdrawal	*Exclusion criteria:* • Mood disorders, schizoaffective disorder • Overt brain disease • Drug intoxication or withdrawal

Table 1.1. Comparison between ICD-10 and DSM-5 criteria for schizophrenia. ICD, International Classification of Diseases; DSM, Diagnostic and Statistical Manual of Mental Disorders.

Schizophrenia is perhaps the prototype of what clinicians call 'non-affective functional (as opposed to organic) psychosis'; despite the frequent occurrence of depressive symptoms *alongside* a schizophrenic illness (Mulholland and Cooper, 2000), symptoms of psychosis themselves are often viewed as non-affective. For example, although a patient with schizophrenia can exhibit significant depressed mood, it may be a secondary response to the frightening delusions and hallucinations they experience and not primary symptoms of the psychosis itself. By contrast, a diagnosis of affective psychoses in the case of schizoaffective disorder requires an almost equally prominent mixture of schizophrenic and affective symptoms (manic, depressive or mixed) which need to have been present simultaneously or within a few

days of each other, but psychotic symptoms must have also been present *without* prominent mood symptoms for at least two weeks. This differentiates schizoaffective disorder from bipolar or unipolar depressive disorder with psychotic features as in the latter situations mood disturbances overshadow psychotic symptoms.

Nevertheless, despite similarities in modern diagnostic systems, the definition and incidence of schizoaffective disorder have faced inconsistencies and controversies. In fact, the modern definition of schizoaffective disorder departs quite substantially from its original description in the 1930's, where the onset of emotional disturbances mixed with psychosis was sudden and patients often recovered fairly soon (Kasanin's definition; see Brockington and Leff, 1979). It is generally agreed that the prevalence of schizoaffective disorder is less than that of schizophrenia, often with better prognosis and outcome, but poorer than that of mood disorders with psychotic features (Harrow, Grossman, Herbener and Davies, 2000). In particular, the presence of mood-incongruent psychotic symptoms (e.g. delusional content that is not relevant to the mood of the patient) in affective psychoses seem to worsen the outcome compared with those with mood-congruent (e.g. a severely depressed patient with delusions of guilt) psychotic symptoms. Such observations place schizoaffective disorder in the middle of the spectrum between prominent mood disorders and 'pure' schizophrenia without an affective component (Cheniaux et al., 2008); indeed, first-degree relatives with schizoaffective disorder have been shown to be at higher risk for both mood disorders and schizophrenia (Tsuang, 1991).

There have been no definitive external validity criteria or reliable biomarkers for classifying and diagnosing psychotic disorders (Peralta and Cuesta, 2003, 2005); as a result, diagnosis is still determined by and dependent on descriptive psychopathology. More recently however, new approaches to the diagnosis and classification of psychotic disorders and psychiatric illnesses as a whole have emerged, largely due to the latest advancements in aetiology, neuroimaging and genetics research (Owen, 2014). The most prominent example is probably the Research Domain Criteria (RDoC) project by the National Institute of Mental Health (NIMH) in the US, as the former director of which has recently announced that research projects should distance themselves from the 'traditional' categorical approach of the DSM and implement dimensional approaches such as the RDoC. RDoC has four dimensions which cover different levels of explanations for disease mechanisms including domains of

functioning (e.g. positive and negative valence systems), units of analysis (from genes to behaviour), developmental aspects and environmental aspects (Cuthbert, 2014). In other words, the RDoC approach integrates the neural with the behavioural, the biological with the environmental and is conceived as a work in progress depending on the evidence available. It has already been applied to the study of hallucinations amongst other psychiatric symptoms (e.g. Ford et al., 2014). There has been concerns that an overhaul of the current DSM/ICD systems will be instigated by the RDoC (Owen, 2014) and the latter has attracted a number of criticisms such as ignoring the first-person experience and being too idealistic due to uncertainties in our current knowledge of the brain-behaviour relationship. Nevertheless, although it is not likely that the RDoC will (fully) replace the DSM/ICD systems in the foreseeable future, it does encourage new ways of thinking and can act as a valuable tool alongside the current systems. As the dimensional approach to psychosis as a whole (Section 1.5) gains acceptance more and popularity, the RDoC framework may well become increasingly influential.

1.1.3 Symptomatology and Treatment

As discussed above, the symptomatology of psychosis is multifaceted and heterogeneous. Consequently there is no consensus as to what is truly pathognomonic of a psychotic disorder such as schizophrenia. Delusions, hallucinations and formal thought disorder constitute the (largely descriptive) psychopathology of psychosis but are by no means exclusive to any single disorder. In addition, psychotic symptoms are often extremely frightening experiences (especially at first onset) to the suffering individual and can undoubtedly lead to severe distress and sometimes debilitating impairments in one's functioning. Furthermore, the estimated lifetime prevalence of completed suicide in patients of schizophrenia is around 5% (Palmer, Pankratz and Bostwick, 2005) and the risk is particularly high after the first episode (in the ICD there is a diagnosis of post-psychotic depression) which is frequently when psychotic symptoms remit and depressive symptoms begin as the individual realises the devastating effects schizophrenia had on their life and prospects for their future (although evidence for this remains inconclusive; see Hawton, Sutton, Haw, Sinclair, and Deeks, 2005). Apart from depressive symptoms after the first episode, the presence of delusions and hallucinations during the acute phase can (understandably)

render a person extremely anxious and fearful, especially when the symptoms involve themes of persecution or alien control for example. Therefore, early intervention and effective treatment can be literally life-saving.

The first line treatment strategy of psychotic disorders remains the use of antipsychotic medications. Since the introduction of chlorpromazine and haloperidol (also called first-generation antipsychotics or FGA) in the 1950s and the 1960s, their efficacy and effectiveness in treating positive symptoms (although at the same time they can exacerbate negative symptoms such as lack of motivation, poverty of speech and catatonia) have played a major role in the de-institutionalisation of chronically ill patients from psychiatric hospitals back into the community. Interestingly the discovery of FGA was almost entirely serendipitous and their mechanism of action (dopamine D_2 receptor antagonism) was only determined after their widespread use in the clinic. Nowadays there are more options in the treatment of psychosis, including second-generation antipsychotics (SGA; e.g. olanzapine, risperidone) which are antagonists at both dopamine D_2 and serotonin 5-HT_{2A} receptors. However, despite having a supposedly more favourable side effect profile compared to FGA (e.g. reduced extrapyramidal symptoms), SGA are not generally deemed as *more* efficacious, with the exception of clozapine, and can contribute to other serious side effects such as metabolic syndromes (Stroup et al., 2003; Lieberman, 2007; McEvoy et al., 2005). Although dopamine antagonism has been the 'gold standard' of antipsychotic action, in 2002 a new drug, aripiprazole, broke this tradition by acting as a dopamine partial agonist (Shapiro et al., 2003). Marketed as a 'dopamine-system stabiliser', aripiprazole regulates dopamine concentrations in different areas of the brain depending on endogenous levels of the neurotransmitter (i.e. acting as an agonist in hypo-dopaminergic states and as an antagonist in hyper-dopaminergic states). Curiously however, aripiprazole has not proven to be more efficacious than other antipsychotics, either, albeit possessing minimal extrapyramidal or metabolic side effects. Agonists at metabotropic glutamatergic receptors are being trialled as the latest class of antipsychotics (e.g. Patil et al., 2007) alongside others such as those acting on GABA and glycine.

On the other hand, psychological therapies for psychosis have been increasingly recognised as effective and (seemingly at least) side effect-free treatments for psychosis, which has been considered by some as a purely biological disorder and

received less input from psychotherapeutic endeavours, despite earlier efforts in psychoanalysis (Eissler, 1951). Currently the most widely accepted and implemented psychotherapy for psychosis is cognitive-behavioural therapy (CBTp). Some recent evidence even suggests that CBTp can be effective in alleviating psychotic symptoms *in place of* antipsychotic drugs (Morrison et al., 2014), although such findings have sparked considerable controversy and a meta-analysis in the same year suggest that CBT for psychosis only has therapeutic effectiveness in the small range, at least when measuring statistical effect sizes (Jauhar, McKenna, Radua, Fung, Salvador, and Laws, 2014).

The 'safety' of CBTp has also been called into question: one school of thought by phenomenologically-informed theorists and clinicians is the hyperreflexivity model of psychosis (e.g. Sass and Parnas, 2003), where the patient experiences a pervasive, immense and extreme self-absorption to the extent that self-awareness becomes an external focus of observation. CBTp can therefore be counter-productive as it may exacerbate such self-disturbances by stimulating hyperreflexive processes (Pérez Álvarez, García Montes, Vallina Fernández, Perona Garcelán, and Cuevas Yust, 2011; Nelson and Sass, 2009) and hence not completely 'side effect-free' as some believe. Other psychological and psychosocial therapies such as family intervention (as opposed to individual-based therapies) can also play a key role in the management of early psychosis, and such therapies work most effectively when combined with an accurate evaluation of the patient's individual needs (Penn, Waldheter, Perkins, Mueser, and Lieberman, 2005; Bird, Premkumar, Kendall, Whittington, Mitchell, and Kuipers, 2010). More research evidence is clearly needed in the fields of both pharmacology and psychotherapy; perhaps a personalised approach is the most beneficial given the sheer heterogeneity of psychosis, although the utilisation of 'tailor-made' treatment regimens might still be unrealistic due to logistic (cost, resource, professional education etc.) constraints.

1.2 Delusions

1.2.1 Types of Delusions

This section of the Chapter focuses on the symptom domain of delusion, which is by no means a homogeneous construct. Delusions may be monothematic (focusing on one theme) or polythematic (possessing multiple themes), functional (without known or detectable brain lesion or damage) or organic (occurring after brain injury), primary (occurring not as a consequence of other mental events, e.g. a hallucination) or secondary (occurring as a direct result or elaboration of other psychopathology), circumscribed (limited to the content of the present delusion only) or systemised/elaborated (leading to the formation and integration of other delusions) depending on the level of explanation. The most common themes of delusional content are summarised in Table 1.2:

Type of delusion	Theme
Delusions of control and thought interference	The control over one's thought processes, feelings and actions are no longer one's own, e.g. thought insertion, withdrawal, broadcasting, passivity phenomena of volition/affect/actions and somatic passivity
Delusions of persecution	One is being persecuted by an often omnipotent and malevolent organisation or a group of individuals
Delusions of reference	Other people, objects or entirely coincidental events carry a special message or meaning specifically related to the individual
Delusions of grandeur	One is special or powerful in some way, e.g. a world-famous billionaire
Delusions of misidentification	A familiar individual has been replaced by an imposter (Capgras delusion) or disguising as a stranger (Fregoli delusion)
Delusions of guilt	One is responsible for disastrous event or has committed a despicable crime and deserves punishment e.g. being personally responsible for a natural disaster
Delusions of jealousy (Othello Syndrome)	One's romantic partner or spouse is unfaithful e.g. having an affair
Delusions of love/erotomania (de Clérambault Syndrome)	One is loved by someone who is usually of higher social status and with whom one has little or no contact
Somatic/hypochondriacal delusions	Themed around the body or physical illness e.g. one has a terminal illness or severe disfigurement
Nihilistic delusions	One has ceased to exist or is dead, one's internal organs are decaying, the world is about to end (e.g. Cotard delusion)
Religious delusions	One has a special relationship with God/prominent religious figures, has supernatural powers (cf. delusions of grandeur) or is persecuted/possessed by the devil (cf. delusions of persecution/control)

Table 1.2. Common delusional themes. Adapted from Bajorek and Stockmann (2012).

It is worth noting that although in the table above thought interference is categorised as a delusional belief (see Section 1.2.5), some more philosophically-

informed theorists have argued that forms of thought interference (especially thought insertion) and passivity are 'duplex phenomena' and the delusional elaboration that commonly follows is not essential for the experience of thought interference (e.g. Humpston and Broome, 2016). In other words, the formation of a delusion *per se* neither completes nor constitutes the totality of the anomalous experience. Furthermore, the types and contents of delusions present are not always indicative of a particular diagnosis (e.g. persistent delusional disorder versus paranoid schizophrenia). This challenges the notion of first-rank symptoms and add to the difficulties in differential diagnosis, although the prevalence of certain delusions does vary across disorders (Kendler, 1980; Appelbaum et al., 1999).

Concepts related to delusions include overvalued ideas (an unusual/atypical but socially acceptable belief which is firmly held but not fixed with delusional intensity or conviction), delusional mood/atmosphere (as a process in the psychosis prodrome; see Section 1.4) and delusional perception (the attachment of a delusional meaning to a normal perception, usually signifies the genesis of a primary delusion).

1.2.2 Formation of Delusional Thought

How is a delusion formed? It is a question that has attracted much debate, interest and attention yet very little consensus. Theories for delusion formation have varied and evolved over time: the one-factor account of delusion by Maher (1974, 1999, 2006) proposes that delusions form as a perfectly normal and logical response to, and a consequence of, abnormal perceptual experiences. For example, the experience of hearing abusive voices when nobody is present is sufficient to lead to the delusional elaboration that one is persecuted by spirits. However, this does not explain why not everyone who hears voices will develop delusions, let alone the same type of delusions. On the other hand, the two-factor account advocated by Coltheart and colleagues (Davies, Coltheart, Langdon, and Breen, 2001; Coltheart, Langdon, and McKay, 2007; Coltheart, 2007) maintains that there is a second factor dissociable from anomalous perceptual experience. Whereas the first factor (abnormal perception) explains the content of the delusion, the second factor (deficit in belief evaluation) explains how the delusion is adopted and maintained. Nevertheless, the two-factor account has its own problems. This theory has, first of all, limited scope; it tends to be only applicable to monothematic delusions (e.g. Capgras and other delusions of misidentification).

Secondly, although Coltheart and colleagues suggest the second factor is an abnormality in belief evaluation systems, there is no agreement over a sufficiently specific and satisfactory definition of the second factor e.g. could it be a reasoning bias?

Most recently, a new approach has emerged which is in fact the focus of this Thesis: that is, the role of prediction error signalling in the formation of delusions (and also hallucinations). Predictive processing centres on the idea that perceptions, cognitions and motivational pursuits are the combined results of the agent's actively generating models about the external world, responding to mismatches and discrepancies between expectation and actual outcome (i.e. minimising prediction errors) in order to draw inferences and update the model to an optimal state. The concept of predictive coding is not a new one, however, Corlett and colleagues are amongst the first to apply it to psychopathology research (Corlett et al., 2007, 2009, 2010). The significance and application of this approach will be discussed throughout the current Thesis, hence only a brief overview is offered here in this Section. According to Corlett and colleagues, Kapur's theory (Kapur, 2003) of aberrant salience in psychosis is a result of abnormal prediction error signalling (i.e. unnecessary salience is allocated to irrelevant stimuli due to faulty and excessive error signals with a high level of noise which give the 'wrong impression' that there is a mismatch between expectation and actual outcome).

Also, this approach blurs the boundary between belief and perception which can be incorporated into a single principle of prediction error minimisation and which follows Bayesian inference (e.g. Fletcher and Frith, 2009). For example, in delusions of control, it may be that prediction errors arising from a mismatch between the expected and actual sensory stimuli (see forward model, below) of a self-initiated movement are not cancelled out or minimised. This would not only contribute to the perception of heightened sensation but may also lead to the belief that the movement was not initiated by oneself. As a result, this account is often viewed as another type of the one-factor approach, although some more recent theorists have also argued for compatibility within the two-factor approach (Miyazono, Bortolotti, and Broome, 2015).

1.2.3 Neurobiological and Cognitive-Psychological Theories

It is extremely unlikely that any single theory is sufficient to encapsulate every aspect of delusion; therefore, it should not be surprising that multiple levels of explanation are necessary if not essential for the study of such a complex phenomenon. This kind of complexity starts from the molecular (e.g. dopamine receptor gene expression, function and structure) to the cognitive (e.g. reasoning bias and attributional style) and continues right up to the level of phenomenology and subjective experience of psychosis.

The abovementioned mechanisms of antipsychotic drugs (note they are anti*psychotic* and not necessarily 'anti-*schizophrenic'*: see Kapur, 2004) might provide some insight into the neurochemistry of psychosis and point towards the pivotal role of dopamine. Indeed, the dopamine hypothesis (after much revision) is still the dominant school of thought in the biomedical model of psychotic symptoms. In its latest revision by Howes and Kapur (2009), presynaptic striatal hyperdopaminergia is deemed as the 'final common pathway' to psychosis which is the end product of multiple genetic, biochemical and environmental risk factors. Further, positron emission tomography (PET) imaging studies have shown that in patients resistant to antipsychotic treatment, dopamine function remains unaffected (in fact the results suggest that glutamate, and not dopamine, concentration is elevated in this group of patients) which potentially explains their resistance to dopaminergic drug effects (Demjaha et al., 2012, 2014). This has led to a very recent theory, that there may be a hyperdopaminergic and *normo*dopaminergic division in the classification of schizophrenia (Howes and Kapur, 2014). It must be borne in mind that such neurobiological abnormalities underlie psychosis as a whole and not just delusions; to date, the prediction error theory of delusion formation (see above) is probably the best attempt to apply dopaminergic dysregulation specifically to delusions (Corlett et al., 2010) although in this model there is no absolute difference between a delusional thought and a hallucinatory percept. Therefore, mechanisms differentiating various psychotic symptoms must occur at a higher (i.e. cognitive and experiential) level of explanation. Indeed, Corlett and Fletcher (2012) found that although the neurobiological bases of clinical and nonclinical delusions were similar in terms of striatal prediction error signalling, at a higher frontal cortical level the error responses were significantly associated with intrusion and distress, which were considered more

akin to the qualities consistent with clinical delusions – potentially separating healthy schizotypy with the manifestation of a psychotic disorder on a neural level.

Delusions as beliefs are difficult to define (in fact some theorists argue against the notion that they are beliefs *at all*: see Section 1.2.5) and their cognitive theories range from biases in probabilistic reasoning (e.g. jumping to conclusions based on little or insufficient evidence), attributional style (e.g. externalising bias), attention allocation (cf. aberrant salience) to metacognitive processes such as theory of mind and threat perception (Gilleen and David, 2005; Bell, Halligan, and Ellis, 2006a; Garety, Bebbington, Fowler, Freeman, and Kuipers, 2007). To present a comprehensive review of all the (neuro)psychological theories involved in delusion is beyond the scope of the current Thesis; with the increasing research into neuropsychological models of delusions and their emerging therapeutic application, there is clearly more focus on the psychology and not just the biology of delusional thinking. Indeed, critics of the biomedical approach often claim that biology is far too reductionist to explain psychological processes and undermines the value of the first-person perspective. Nevertheless, although there is much controversy as to whether the mind is simply an epiphenomenon of the brain (still, it should require little doubt that mental processes cannot take place without an organic 'container'- that is, the brain), psychological theories can be equally reductionist if viewed in isolation from biology, and neither approach can truly advance further until they are working in unity. Perhaps this is one of the reasons why the new field of cognitive neuropsychiatry (which is the approach adopted in the current Thesis) is so attractive as it is devoted to the integration between clinical presentation and cognitive mechanisms, which are essentially products of neural activities.

1.2.4 Measuring Delusions

Delusions have long been considered as a solely clinical concept (i.e. as a symptom or symptom cluster of an identifiable mental illness); as such, their measurement and evaluation are often limited to methods used in clinical interviews. Interestingly, instruments dedicated to the measurement of delusions alone in a clinical setting are rare and are often incorporated with other positive symptoms such as hallucinations. Examples of this kind include the Positive and Negative Syndrome Scale (PANSS, Kay, Fiszbein, and Opfer, 1987), Scale for the Assessment of Positive Symptoms

(SAPS, Andreasen, 1984) and the PSYchotic symptom RATing Scales (PSYRATS, Haddock, McCarron, Tarrier, and Faragher, 1999) which consists of subscales for delusions and hallucinations, respectively. As Bell et al. (2006b) suggest, although the diagnosis of delusions in general is rather reliable it does not seem to extend to the concept of 'bizarre delusions'. Bizarre delusions are those that are entirely 'nonsensical' and impossible to happen in real life (such as those that defy the laws of physics) and until recently, their manifestation alone could grant a diagnosis of schizophrenia (in the latest instalment of the DSM, DSM-5, bizarre delusions can also occur in delusional disorder). The lack of reliability in measuring the 'bizarreness' in these types of delusions could be (at least partly) due to the observation that the expression of delusional thought is beyond verbal description and should be viewed not by its 'literal' meaning but in the context of the person's general experiential dimensions (Cermolacce, Sass, and Parnas, 2010).

The contemporary notion that subclinical or even nonclinical psychotic symptoms (Verdoux and van Os, 2002) are present and more prevalent than previously speculated in the general population who otherwise have no need for care called for the continuum approach to psychosis (Section 1.5 of this Chapter). This gave rise to the development of new instruments and scales. One notable example and one that is used in the current Thesis is the Peters et al. Delusions Inventory (PDI; Peters et al., 1999, 2004) which, despite using clinical methods (the Present State Examination) as a template, is designed to measure delusional ideation in (subjectively) healthy populations (although it can also be used in psychiatric populations with active delusions) and adopts a multidimensional approach in its three subscales (distress, preoccupation and conviction). The PDI is likely the most widely known and implemented psychometric scale for measuring nonclinical delusions and has been translated into several other languages. Its psychometric properties, reliability and validity will be detailed in Chapter 4. Other nonclinical measurements, focusing more specifically on paranoia and persecutory delusions, include the Paranoia Scale (Fenigstein and Vanable, 1992), the Green et al. Paranoid Thought Scales (GPTS, which can also apply to psychiatric populations; see Green et al., 2008), the Paranoia Checklist (Freeman et al., 2005; again, it can be used in clinical settings as well) the Beliefs About Paranoia Scale (BAPS, Gumley, Gillan, Morrison, and Schwannauer,

2011) and the Persecutory Ideation Questionnaire (PIQ, McKay, Langdon, and Coltheart, 2006).

The 21-item version of the PDI (PDI-21) is chosen for this Thesis because it does not focus exclusively on one type of delusional ideation (which is an advantage because of the lack of symptom specificity in psychosis-like experiences) and is comparatively more in-depth (having three subscales). In a validation study for the psychometric properties of PDI-21 (compared with the original PDI-40), the authors (Peters et al., 2004) found an average yes/no endorsement rate of 29.8% in healthy individuals which was only slightly higher than that in the 40-item version (25.2%). In a sample with clinical delusions, the yes/no endorsement rate was 53.2% for the 21-item version (51.6% in the 40-item version). The most striking difference that set apart healthy versus deluded samples, however, was their conviction, distress, and preoccupation sub-scores. As the authors stated: 'It is not *what* you think, it is *how* you think about it' (p. 1013, original italics).

1.2.5 Are Delusions Beliefs?

Both the one- and two-factor accounts are *doxastic* (*lit*. belief-related) approaches to the understanding of delusion which intrinsically assume that (just as defined in the DSM and the ICD) delusions are fixed false *beliefs*. The widespread acceptance of this approach, which has been consistently defended by doxastic theorists (e.g. Bayne and Pacherie, 2005), is reflected in the current measurement, diagnosis and treatment methods of delusional symptoms. However, there are also theorists who argue against the doxastic account. Indeed, many of the key features of delusional beliefs, such as absolute conviction and incomprehensibility (in the case of bizarre delusions) may not be accurate reflections of the actual phenomenology of delusional *experiences*. Alternative features based on first-person accounts, such as subjectivity, ambivalence and double-bookkeeping (the observation that delusional subjects are often 'in two minds' about the veridicality of their ideas which is a direct contrast to the notion of conviction and incorrigibility; see Sass, 2004) have emerged and are mainly supported by *non-doxastic* theorists. Nevertheless, the arguments put forward by non-doxastic theorists are not necessarily simply limited to personal-level explanations.

Gerrans (2001, 2013, 2014) is one such theorist who eloquently defends the non-doxastic approach with evidence not only from phenomenology but also from basic

neurobiology, which is compatible with the prediction error theory. He proposes that delusional *thoughts* are the result of excessive activities in the default mode network (e.g. ventromedial prefrontal cortex) overwhelmed with hypersalient information but is unsupervised by de contextualised processing (governed by the more dorsolateral prefrontal regions of the brain, such as parts of the central executive network) and generally refers to delusions as 'default thoughts'. Their relevance to prediction error, as Gerrans argues, is that malfunctions in the sense of agency or ego-boundary (e.g. in thought insertion and many other first-rank symptoms) occur when error signals are not minimised or corrected and instead are propagated up towards the *highest level* in the cognitive hierarchy where delusions are generated. His approach is attractive because it incorporates evidence from multiple levels of investigation (biological, psychological and phenomenological) and is not confined to a single theoretical framework. If faulty prediction error signalling really is the definitive underlying mechanism behind delusions (and hallucinations) from this point of view perhaps the conceptualisation of delusion needs to reorient itself away from the doxastic approach, although in the meantime the debate will certainly continue between one- and two-factor, doxastic and non-doxastic theorists.

1.3 Hallucinations

1.3.1 Types of Hallucinations

Just like the notion that delusions are 'false beliefs' (at least according to the doxastic account), hallucinations are considered 'false perceptions' (perceptions without a corresponding sensory stimulus, unlike *illusions* which are misinterpretations of real stimuli). A hallucination can occur in any of the five sensory modalities (auditory, visual, olfactory, gustatory and somatic) but the auditory type is particularly linked to psychosis and specific forms of auditory-verbal hallucinations (AVHs) constitute some of the Schneiderian first-rank symptoms (e.g. voices providing a running commentary on the subject's activities or voices heard arguing with one another). Apart from being classified by the sensory modality in which a hallucination manifests, the actual form of the hallucination also grants further classification which is summarised in Table 1.3.

Type of hallucination	Examples
Simple or elementary hallucinations	Simple unstructured sounds, single words lacking grammatical complexity or simple flashes of light
Complex hallucinations	Complex sounds such as voices in the form of structured sentences, music, or complex images of people or objects
Functional hallucinations	Hallucinations triggered by a simultaneous external stimulus of the same modality e.g. hearing voices in traffic noise
Reflex hallucinations	Hallucinations triggered by a simultaneous external stimulus of a different modality e.g. seeing a human face when a door slams shut
Hypnogogic and hypnopompic hallucinations	Hallucinations that occur when falling asleep and waking up, respectively
Extracampine hallucinations	Hallucinations perceived as outside the limits of the sensory field e.g. hearing astronauts speaking in space

Table 1.3. Examples of hallucination types. Adapted from Bajorek and Stockmann (2012).

Due to its clinical significance to the differential diagnosis of psychosis, this Section of the Chapter discusses exclusively the auditory-verbal modality of hallucinations in schizophrenia-spectrum psychoses. However, AVHs are not a pathognomonic marker of either psychosis or schizophrenia and can occur, like delusions, in a range of other neuropsychiatric disorders such as Alzheimer's disease, temporal lobe epilepsy, dissociative disorders and some mood and personality disorders (e.g. Borderline Personality Disorder). Although third-person AVHs (i.e. those heard referring to the patient as he/she) are historically thought to be more typical of a schizophrenic pathology and second-person AVHs (i.e. voices addressing the patient directly as 'you') more relevant to psychotic depression, phenomenological surveys of AVH experiences have shown no conclusive evidence supporting this claim (Nayani and David, 1996; McCarthy-Jones, Trauer, Mackinnon, Sims, Thomas, and Copolov, 2014) and neither is there any evidence on the significance of internal versus external hallucinations (i.e. voices heard inside or outside the head) in psychosis. Nevertheless, some types of AVHs may require more clinical intervention (i.e. abusive, persecutory or command hallucinations) than for example positive, comforting or reassuring voices experienced by individuals without a need for care (see Johns et al., 2014).

The observation that there are individuals who experience persistent AVHs yet never come to the attention of mental health professionals has contributed to the debate about the dimensional nature of nonclinical and subclinical psychotic symptoms (Section 1.5.2) and to many alternative (i.e. not conforming to the 'mainstream' biomedical model) views about AVHs. It must be noted that although such views are gaining popularity and acceptance in patient groups, they have not yet been subjected to the same level of scientific inquiry or rigour as the biomedical model and hence must be interpreted with much caution.

1.3.2 Neurobiology of Auditory-Verbal Hallucinations

Despite the uncertainty and lack of consensus for a unifying model of AVHs, there is strong evidence for their biological basis. Once again, on a neurochemical level it is the action of antipsychotic drugs in reducing psychotic symptoms and the propensity for certain psychoactive substances (e.g. amphetamine, ketamine, LSD) to induce psychotic-like hallucinations in their users that link dysfunctional neurotransmitter systems (e.g. dopamine, glutamate, serotonin) with the formation of hallucinations. However, such dysfunction is not necessarily specific to the pathogenesis of hallucinations but to that of psychosis as a whole as discussed above. For example, ketamine is used to model psychosis with all its complexity and not just hallucinations or delusions (Lahti, Koffel, LaPorte, and Tamminga, 1995; Corlett et al., 2007) and abnormal dopamine synthesis capacities have been shown in both nonclinical populations and treatment-resistant schizophrenia patients who experience both delusions and hallucinations (Howes et al., 2012, 2013).

The advent of structural and functional neuroimaging techniques and subsequent investigations in the past twenty years have enabled researchers to find a substantial amount of evidence for the neural bases of AVHs, even though such evidence remains inconclusive and insufficient for a complete understanding of AVHs. Earlier studies suggest that activation in an extensive network of cortical and subcortical areas may be involved in the pathogenesis of AVHs: these areas include the inferior frontal, insular, anterior cingulate and temporal cortices bilaterally, with a greater lateralisation to the *right* hemisphere in contrast to the usual left lateralisation of language areas (Shergill et al., 2000a; see also Hugdahl et al., 2008; Sommer et al., 2008).

The same group (Shergill et al., 2000b) also studied auditory imagery and verbal self-monitoring in schizophrenia patients who had a history of prominent AVHs and found relatively attenuated activation patterns in inner-speech processing areas but not in inner-speech generation areas compared to those in the control group. The self-monitoring model of AVHs is discussed in more details below. Structurally, on the other hand, decreases in local volumes of the right middle and inferior frontal gyri (containing a right hemisphere analogue of Broca's area), the left transverse temporal gyrus (including the primary auditory cortex; e.g. Dierks et al., 1999) and the inferior part of the left supramarginal gyrus have been correlated with the severity of AVHs (Gaser, Nenadic, Volz, Büchel, and Sauer, 2004). In addition, functional connectivity studies have demonstrated that schizophrenia patients have altered connectivity between areas responsible for self-referential processing (e.g. the medial prefrontal cortex, part of the cortical midline structures) and the superior temporal gyrus, for example (Allen et al., 2012).

1.3.3 Cognitive-Psychological Theories

The abovementioned model, in which inner-speech is misattributed to an external source involving faulty self-monitoring, is perhaps one of the most influential cognitive theories of AVHs. The forward or 'comparator' model was originally proposed by Frith, Blakemore, and Wolpert (2000) in an attempt to explain passivity phenomena and related symptoms involving the misattribution of agency such as delusions of control; more recently it has also been applied to AVHs as the view of thought/inner speech as action is increasingly accepted by researchers. In the forward model, the brain issues an *efference copy* of the motor command required to achieve a particular goal based on the estimated current state and the desired end-state in order to *predict* the consequence of executing the motor command. If the actual sensory feedback matches the predicted state, awareness of performing the motor action will remain based on the predicted state which occurred before the actual execution of the motor command so that the subject has the experience of agency (i.e. awareness that the action is initiated by oneself) and then the predicted sensory feedback is cancelled out by the actual feedback, leading to sensory attenuation of the motor act. However, if there is a mismatch between the predicted and the actual sensory feedback the actual feedback will not be cancelled out, which leads to the feeling of external control (due

to increased activities in the parietal cortex) and a lack of wilfulness for that action. A diagrammatical representation of the forward model is shown in Figure 1.1 below (adapted from Frith et al., 2000, as cited by Jones and Fernyhough, 2007). It is also worth noting that the sense of agency could be viewed as two-fold (e.g. Synofzik, Vosgerau, and Newen, 2008) where there is a clear difference between the pre-reflective *feelings* of agency and the reflective or even meta-cognitive *judgements* of agency. It is thought that the latter form of agency is more closely related to the formation of delusions of control and the lack of agential feelings alone is insufficient for a delusion to form.

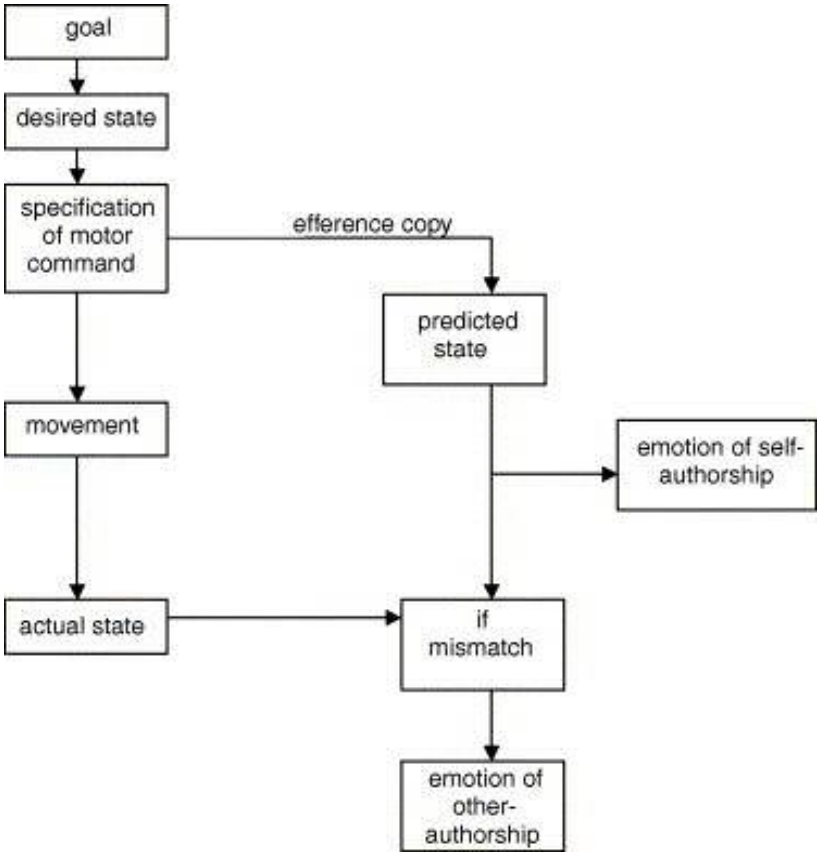

Figure 1.1. Forward model of motor control (adapted from Frith et al., 2000, as cited by Jones and Fernyhough, 2007).

The hallucinating subject often reports AVHs as unintended, involuntary and intrusive, which has led to the hypothesis of inner speech or 'verbal thought' as being wrongly attributed to an external force so that the sense of agency is lost in AVHs. Under the assumption that inner speech is indeed reliant on motor commands, it would seem that such non-vocal or sub-vocal thoughts are the 'raw materials' for AVHs and the forward model can therefore apply. Nevertheless, according to Jones and

Fernyhough (2007)'s account, which is a revision of Seal, Aleman, and McGuire (2004)'s first application, the efference copy is not sent and hence there is no predicted state of the motor command for inner speech, leading to a mismatch between the actual sensory feedback and the (lack of) predicted feedback. Their adaptation is shown in Figure 1.2 below. It is not difficult to notice the significance of the forward model to the concept of self-monitoring and how it relates to prediction errors which are targets of investigation in the current Thesis: a prediction error would be generated from the mismatch between predicted and actual feedback which in itself reflects deficits in self- or reality-monitoring.

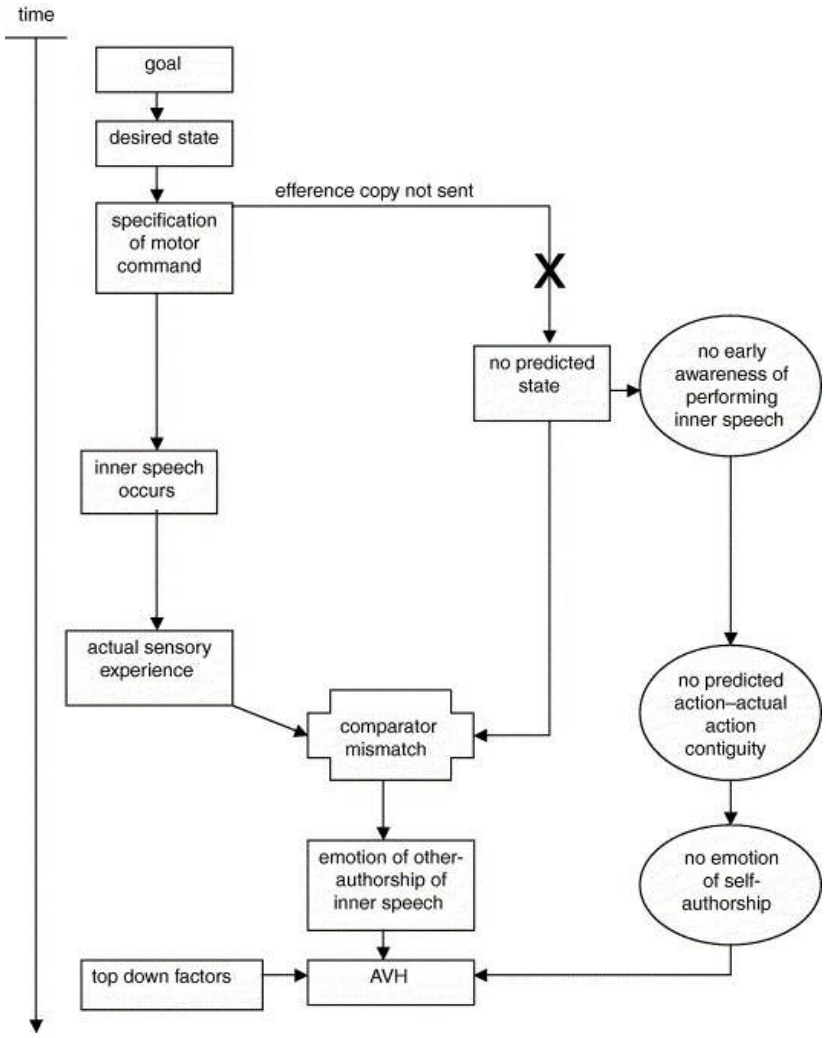

Figure 1.2. Forward model of motor control as applied to AVHs (adapted from Jones and Fernyhough, 2007).

However, the inner speech framework is not the only cognitive theory applicable to the understanding of AVHs. Jones (2010) argues that there are sub-categories of AVHs which may be more memory-like than inner speech-like; this argument is

supported by a recent phenomenological survey (McCarthy-Jones et al., 2014). In this survey of 199 psychiatric patients, the authors were able to distinguish four subtypes of AVHs using a semi-structured interview. The most prominent subtype was 'Constantly Commanding and Commenting AVHs' (widely associated with schizophrenia), with the other three being 'Replay AVHs', 'Own Thought AVHs' and 'Nonverbal Auditory Hallucinations'. The phenomenological nuances of AVHs are discussed in Section 1.3.5 below. In addition, top-down mechanisms involving the meaning and beliefs about hallucinations are also heavily implicated (what are voices without a listener?), such as in delusions of persecution arising secondary from hearing abusive voices. The beliefs about, and distress caused by, the experience of AVHs are key factors differentiating clinical (e.g. in schizophrenia) and nonclinical populations (Waters et al., 2012; Johns et al., 2014).

1.3.4 Measuring Hallucinations

The multifaceted nature of hallucinations means that their measurement can be as challenging as the measurement of delusions discussed above (Section 1.2.4.). Despite the potential difficulty in capturing their full features, there are various clinical and nonclinical psychometric scales designed for the assessment of AVHs (and also hallucinations in other sensory modalities). The abovementioned Positive and Negative Syndrome Scale (PANSS), Scale for the Assessment of Positive Symptoms (SAPS) and the PSYchotic symptom RATing Scales (PSYRATS, hallucinations subscale), with the addition of the Auditory Hallucinations Rating Scale (AHRS, Hoffman et al., 2003) and also other scales aiming to measure beliefs and actions associated with AVHs (e.g. Voice and You, Hayward, Denney, Vaughan, and Fowler, 2008; Voices Acceptance and Action Scale, Shawyer, Ratcliff, Mackinnon, Hayes, and Copolov, 2007; Beliefs about Voices Questionnaire, Chadwick, Lees, and Birchwood, 2000) all fall into the clinical category, whereas the measurement of subclinical and nonclinical hallucinations is somewhat more varied and has relatively less focus on AVHs alone.

The scale adopted in the studies in this Thesis is the Cardiff Anomalous Perceptions Scale (CAPS; Bell, Halligan, and Ellis, 2006c) which offers a comprehensive assessment of a range of perceptual aberrations across sensory modalities in the general population, unlike the focus on the auditory-verbal domain

in clinical populations. Like the PDI, it also consists of three subscales (distress, intrusiveness and frequency) and taps into both the presence of anomalous experiences and changes in clarity and intensity of perceptions. Its psychometric properties are discussed in Chapter 4; the CAPS is chosen in preference to other more AVH-specific scales because of the lack of specificity in psychosis-like experiences and at-risk mental states (see Section 1.4.1) in nonclinical populations. Other prominent examples include the widely-used Launay-Slade Hallucination Scale (LSHS) which is also designed for use in the general population and measures unusual perceptual experiences across modalities (depending on the version, the focus can vary from vivid or intrusive thoughts to auditory/visual hallucinations), and the Perceptual Aberration Scale (PAS). However, once again the CAPS is chosen for this Thesis because it is comparatively more in-depth (e.g. having three subscales) and covers a wider range of experiences.

1.3.5 Soundless Voices and Audible Thoughts

The phenomenology of some first-rank symptoms, especially those of thought interference (e.g. thought insertion), challenges the traditional definitions of both delusions *and* hallucinations and is in fact thought to be somewhere in-between. Combined with the more recent concept that belief (at least from a doxastic point of view) and perception are interlinked and arise from the same cognitive processing hierarchy, it is perhaps not surprising that some theorists have proposed the idea of 'soundless voices and audible thoughts' (e.g. Humpston and Broome, 2016; Jones and Luhrmann, 2016) to capture the phenomenology of AVHs and related symptoms which, according to these authors, lies on a spectrum (or even spectra) of varying degrees of audibility and externality. Indeed, first-person reports from patients often suggest that unlike the auditory hallucinations in alcohol hallucinosis or even in deaf individuals, AVHs or 'voices' heard by patients with psychosis frequently *lack* the sensory and acoustic features associated with external sounds and are more akin to 'a sense of being spoken to' or 'inaudible voices' (Moritz and Larøi, 2008; Waters and Jardri, 2014).

Furthermore, the reality of voices perceived by patients is often more reliant on the lack of control rather than clarity and loudness when asked to tell apart ordinary verbal thought from AVHs (Hoffman, Varanko, Gilmore, and Mishara, 2008).

Therefore, it can be difficult for patients to subjectively differentiate AVHs from other intrusions such as thought insertion. This is not to say that truly external and audible hallucinations cannot occur in schizophrenia but is simply to challenge the definition that *all* AVHs have to possess clear sensory qualities.

In fact, it is probably the second-order *appraisal* of mental events that attributes acoustic features to the experience of 'hearing' voices: in other words, it is possible that even when a patient reports voice-hearing through his ears, it is still more of a simulation than an actual perception. It must be pointed out that the patient is *not* lying when they say they are hearing voices even though there is no real sound being perceived; instead, it simply means they have no better way to label the experience otherwise. Certainly, this directly contradicts the orthodox conceptualisation (false sensory perception without external corresponding stimuli) of AVHs or any hallucinatory experience by limiting the sensory element, but it has been shown to be more consistent with patient reports at least in psychotic AVHs (i.e. the phenomenology of AVHs in organic disorders may well be drastically different).

Another related concept, thought echo, is not classed as a 'true' hallucination due to its 'own-thought' nature (i.e. one's own thoughts diffusing outwards) but may actually be *perceived* as more audible than internal hallucinations, for example. Once again, this is directly dependent on how the subject appraises the experience: an internal hallucinator might be confident that his voices cannot be heard by others because they are solely within the confines of their head whereas another person experiencing thought echo might be extremely worried that their thoughts are heard by everyone else (which can lead to a delusional elaboration) because they *themselves* can 'hear' them spoken aloud. In this sense, once thought echo loses the 'own-thought' nature (i.e. authorship and subjectivity) it will become 'true' external AVHs, indicating that psychotic symptoms or even psychological phenomena in general are not static and can often morph into one another and therefore cannot be understood as isolated mental events.

1.4 The Psychosis Prodrome

The term 'prodrome' denotes the period before frank illness onset where attenuated and nonspecific (yet still pathological) symptoms may be present alongside a certain

degree of functional decline or impairment. It is somewhat debatable because 'prodrome' seems to suggest inevitable transition which is simply not the case. In fact, some have argued that it should be reserved in a solely retrospective sense (Yung and McGorry, 1996) as the presence of prodromal symptoms only predicts transition in around 30% of high-risk individuals (Fusar-Poli et al., 2012). Nevertheless, it is still a useful concept for studying the aetiology of psychosis and has important implications for treatment provision (e.g. early intervention services). In fact, the transition rate has been steadily decreasing over the past few years, perhaps at least partly due to the improvements in service provision and earlier access to treatment (Wiltink, Velthrost, Nelson, McGorry, and Young, 2015; Nelson et al., 2016).

The early stages of psychosis, including the ultra-high risk state, are defined as 'a period of escalating severity of symptoms or functional decline that lies between the end of the relatively asymptomatic premorbid phase and the beginning of the frankly psychotic phase of schizophrenic psychosis' (Miller et al., 2003). Rather than a stable and enduring trait as it is often the case in schizotypy, the transitioning period from the prodrome to an episode of florid psychosis is characterised by escalating severity in a relatively short period of time. The duration of the prodrome can range between a few months to two years, sometimes longer (Yung and McGorry, 1996); however, as the prodrome is a retrospective concept, it can be difficult to differentiate between 'true' prodrome and untreated psychosis.

1.4.1 Assessments of Prodromal Psychosis

The psychometric instrument for early psychosis chosen in the current project is the Structured Interview for Prodromal Syndromes (SIPS) and the Scale of Prodromal Symptoms (SOPS; see Miller et al., 2003) for its relatively straightforward rating scales and fully structured nature, in addition to the fact that training opportunities and expertise in using the interview are already available within Cardiff University. Similar to the Comprehensive Assessment of the At-Risk Mental State (CAARMS) instrument mentioned below, the SIPS has three prodromal criteria (APS, BIPS and GRD) but with different duration requirements from the CAARMS:

1) Attenuated positive symptom syndrome (APS): One or more of the 5 positive items scoring in the prodromal range of 1-5, AND symptoms beginning in the

past year or increasing 1 or more points within the past year, AND symptoms occurring at least once per week for the last month.

2) Brief intermittent psychotic symptom syndrome (BIPS): One or more of the 5 positive items scoring in the psychotic range of 6, AND symptoms beginning in the past 3 months, AND symptoms occurring currently at least several minutes a day, once per month.
3) Genetic risk and deterioration syndrome (GRD): First degree relative with history of any psychotic disorder, OR criteria for schizotypal personality disorder met, AND a drop of at least 30% in the last year in the Global Assessment of Functioning (GAF).

The two related concepts, at-risk mental state (ARMS) and clinical ultra-high risk state (UHR) are often used interchangeably with the psychosis prodrome concept. Unlike 'prodrome', however, ARMS and UHR are not necessarily retrospective in nature and thus do not always indicate transition to a clinical status. The most widely accepted criteria for UHR status are those used by Yung and colleagues from the Personal Assessment and Crisis Evaluation (PACE) Clinic in Australia who also devised the CAARMS. The ARMS criteria are usually regarded as the presence of one or more of the following (Yung et al., 2005, 2008):

1) Attenuated (subthreshold) psychotic symptoms (APS) within the previous 12 months.
2) Brief limited intermittent psychotic symptoms (BLIPS): history of brief self-limited psychotic symptoms which spontaneously resolve within the previous 12 months.
3) Trait group: presumed genetic vulnerability to psychotic disorder (either schizotypal personality disorder or family history of psychotic disorder in a first degree relative) plus persistent low functioning for at least 1 month within the previous 12 months.

As mentioned above, only around 30% of UHR or ARMS individuals will actually transition to frank psychosis. Therefore it is important for clinicians to be aware what the risk factors are and which ones might have the highest predictive value for later transition. Previous studies (e.g. Mason, Startup, Halpin, Schall, Conrad, and Carr, 2004) have found that when measured using psychometric scales, the most

reliable predictor is the presence of high schizotypal personality traits (see Section 1.5.3 for details on schizotypy and risk for psychosis). Other risk factors with high predictive values include auditory hallucinations, odd beliefs and magical thinking, poor premorbid social adjustment and functioning, blunted/flattened affect and anhedonia/social withdrawal.

However, just like the observation that there is no single symptom that is pathognomonic of psychosis, none of these risk factors is able to fully predict or 'guarantee' a later transition. More recently, presence of basic self-disturbances (e.g. altered sense of ego-boundary and depersonalisation of thought agency/ownership) has also been shown to strongly predict eventual psychosis onset in UHR individuals (Nelson, Thompson, and Yung, 2012).

From a neurobiological-neuropsychological point of view, UHR individuals already exhibit deficits in working memory, attention, executive control and general intelligence (Pflueger, Gschwandtner, Stieglitz, and Riecher-Rössler, 2007). Neural activation pattern abnormalities in this group are thought to be on an intermediate level (i.e. more severe than controls but milder than patients with established psychosis) which furthers the notion of a continuum approach to psychotic symptoms and levels of dysfunction (Broome et al., 2009). Impairments in functional connectivity between posterior, but not anterior, hippocampal regions and the prefrontal cortex have also been found to be disrupted in both unmedicated UHR individuals and patients who have just received a diagnosis of (first episode) schizophrenia (Benetti, Mechelli, Picchioni, Broome, Williams, and McGuire, 2009) indicating that these brain abnormalities may be present from the very early stages (sometimes even when the individual is still asymptomatic). However, as an individual responds to timely and effective treatment, the potential progression to chronic schizophrenia and further decline in functioning can be prevented (Wood, Yung, McGorry, and Pantelis, 2011).

1.4.2 Stages of the Psychosis Prodrome

Two main approaches to the staging of psychosis have been proposed, one based on (psycho)pathological measures/clinical presentation (Wood et al., 2011) and the other on phenomenology (Conrad's 'Beginning Schizophrenia' model, see Mishara, 2009). The former approach focuses on disease progression and levels of abnormality (see Table 1.4) although the transition between stages is considered bidirectional. In other

words, even physiological changes may be reversed with treatment. Wood and colleagues also argue that such treatment, whether pharmacological or psychological, should be most effective in the early stages and 'milder' in terms of side effect profiles (choosing medications other than clozapine for example, which is an antipsychotic with severe side effects but can be highly effective in treatment-resistant schizophrenia).

Clinical Stage	Definition	Example Populations
0	Increased risk of psychotic or severe mood disorder. No symptoms currently.	First-degree teenage relatives of probands.
Ia	Mild or nonspecific symptoms, including neurocognitive deficits, subthreshold psychotic-like experiences associated with only mild functional decline, mood or anxiety symptoms that are distressing and/or have resulting in help-seeking with or without comorbid substance abuse. Mild functional change or decline.	Individuals referred to youth mental health service by primary care physicians or school counselors.
Ib	Ultra high risk: Moderate but subthreshold symptoms, with/without mild to moderate neurocognitive changes, comorbid substance abuse, and with functional decline to caseness.	Individuals who meet CAARMS criteria.
II	Full threshold disorder with moderate-severe symptoms, neurocognitive deficits, and functional decline.	Individuals who meet psychosis criteria as defined by the CAARMS.
III	Incomplete remission or recurrence/relapse.	Individuals with a recurrence of stage II psychosis.
IV	Severe, persistent, and unremitting illness as judged on symptoms and disability criteria.	Patients with established schizophrenia with functional impairment.

Table 1.4. Clinical staging model of psychosis (adapted from Wood et al., 2011). CAARMS, Comprehensive Assessment of the At-Risk Mental State.

Conrad's model of beginning schizophrenia was proposed long before the advent of sophisticated neuroimaging techniques but is still influential due to its particular significance for the phenomenological investigation of early psychosis. Table 1.5 outlines the earlier stages in Conrad's model (he also had his own terminology for the later stages, e.g. apocalyptic-catatonic, consolidation/partial remission and residual defect).

Stage	Term
I	Trema (derived from Greek, colloquial for stage fright) Meaning Delusional mood (or atmosphere) Characteristics Undefinable, but increasingly upsetting quality spreads from salient aspects to entire perceptual field. Patient feels anticipatory excitement, suspiciousness, alienation, fear, guilt, depression, or combination of these. Patient may perform abrupt, seemingly meaningless actions
II	Apophany (Greek apo [away from] + phaenein to show → revelation) Meaning Delusion as revelation (Aha-Erlebnis) Characteristics Perceptual Gestalt experienced incompletely in terms of its expressive rather than its objective material holistic qualities. Inability to transcend current perspective or to shift frame of references. Abnormal connectedness between seemingly unrelated meanings. Delusional perception, misidentification. Relentless ("monotonous, repetitive") spreading of the delusion as both "elastic" and fixed to new gestalts.[7,18,26] Progression of delusions from external to inner "space", delusional body sensations. Patient uncritically receptive and unable to detach, as if trapped between sleeping and waking. Thought insertion, thought broadcasting, hallucinations
III	Anastrophe (Greek, ana- (back) + strephein (to turn) → turning back) Meaning Patient feels self to be passive middle point (subject-directed complement to world-directed apophany) Characteristics Delusions of reference. Events and perceptions are related to self

Table 1.5. Conrad's stage model of Beginning Schizophrenia (adapted from Mishara, 2010).

It is apparent that the three stages (trema, apophany and anastrophe) in this model can 'match' onto the clinical staging model (clinical stages Ia, Ib and II respectively)

by Wood and colleagues. Especially relevant to the psychosis prodrome is the presence of delusional mood/atmosphere as an 'ideal' environment for the subsequent formation of delusions and/or hallucinations and the subtle changes in thinking, perception in behaviour may reflect such 'primordial' forms of emerging psychotic symptoms. Once again, the sense of perplexity and confusion experienced by the at-risk individual during this period (long before any indication of an impending psychotic episode) is very much nonspecific despite having (retrospective) clinical significance, and therefore not considered a screening tool for psychosis.

However, with the support from more recent studies where early shifts in one's self-world relationship and existential orientation (i.e. self-disturbances; Nelson et al., 2012) could in fact be trait markers for the clinical manifestation of illness that can eventually be used to reduce false-positive rates (Nelson, Yung, Bechdolf, and McGorry, 2008). Such emphasis on self-disorders is not without biological basis (e.g. Postmes, Sno, Goedhart, van der Stel, Heering, and de Haan, 2014): for instance, the roles of the default mode network (DMN; see also Section 1.2.5) and cortical midline structures (CMS) in self-referential processing (for a comprehensive review, see Nelson et al., 2009).

1.5 The Continuum Approach

The dichotomous classification of psychosis has been challenged on numerous occasions; in particular, the dimensional approach of a 'Psychosis Continuum' (or even continua) that ranges from transient suspiciousness in the general population without any need for care and extends into clinical patients with severe paranoia has proven to be a highly influential alternative (van Os et al., 2000, 2009; Kaymaz and van Os, 2010).

According to this view, many of the symptoms seen in schizophrenia, such as paranoid ideation and hearing voices, can also be found in the general population, albeit to a milder or attenuated degree which would normally cause much less distress to the experiencing individual (Freeman et al., 2008; Johns et al., 2014). In this sense, the psychosis prodrome is but a component of the continuum linking nonclinical psychotic experiences and clinical disorder. Symptoms of schizophrenia themselves, on the other hand, are not necessarily pathognomonic of the illness (Verdoux and van

Os, 2002) Interestingly, persistence of these unusual experiences is not always an indicator of impairment, as many individuals can, for example, hear voices most of the day but do not cross the clinical threshold for a diagnosable psychotic disorder (Powers, Kelley, and Corlett, 2017), often due to the lack of distress.

Individuals who report such nonclinical psychosis-like experiences are often viewed from two angles, one from a latent psychosis risk perspective (Debbané et al., 2015) and the other from a normally distributed schizotypal personality trait perspective (Claridge, 1997). These two schools of thought have not yet reached agreement. Additionally, the continuum approach is supported by a relatively recent factor analysis of the structures of different diagnostic systems used for schizophrenia (Peralta and Cuesta, 2005).

1.5.1 Schizotypy

Schizotypy in the current Thesis is defined as nonclinical and subclinical levels of subjective personality traits and unusual experiences akin to those found in schizophrenia-spectrum psychoses, but do not lead to functional impairment or a need for care. There are two main schools of thought regarding schizotypy: one from the study of extended phenotypes of schizophrenia (e.g. milder forms of perceptual and behavioural oddities in unaffected first-degree relatives of schizophrenia patients), and the other from that of individual differences/personality approach. This however does not mean there is no overlap between the two approaches; indeed, some argue (Hewitt and Claridge, 1989, Johns and van Os, 2001, Kendler et al., 2008) that schizotypal traits result from a combination of genetic, environmental and personality factors which, like their counterparts in other fields of medicine, are normally distributed in the general population (i.e. *everyone* has a certain level of expression of these traits). Nevertheless, the division between these viewpoints may be seen as a reflection in two of the major diagnostic systems: in the DSM, schizotypal *personality* disorder is classed (as its name suggests) as a cluster A personality disorder whereas in the ICD, schizotypal disorder is grouped with other schizophrenia-spectrum psychoses.

It has also been proposed that schizotypy itself is multidimensional (dimensions within the psychosis continuum), representing underlying vulnerabilities across personality, subclinical and clinical phenomenology. In other words, schizophrenia is not a separate entity from schizotypy and neither is schizotypy simply an analogue of

schizophrenia; instead, schizophrenia is the most extreme *expression* of schizotypy (Kwapil and Barrantes-Vidal, 2014). Also, the presence of schizotypal *traits* cannot be equated to the psychosis prodrome (at best, it conveys psychosis-proneness or risk *states*) let alone an 'inevitable' transition to clinical status due to protective factors and potential reversal against the stages of the 'psychosis trajectory'.

As a result, the current Thesis prefers the individual differences approach and measures overall schizotypy (in addition to more specific psychosis-like experiences such as those measured by the PDI-21 and the CAPS) using the Oxford-Liverpool Inventory of Feelings and Experiences (O-LIFE; Mason et al., 1995; Mason and Claridge, 2006). The O-LIFE consists of four subscales (unusual experiences, cognitive disorganisation, introvertive anhedonia and impulsive nonconformity) which aim to reflect the positive, negative and disorganised dimensions of schizotypy in the general population only. It is chosen over the Schizotypal Personality Questionnaire (SPQ; Raine, 1991; Vollema, Sitskoorn, Appels, and Kahn, 2002) because it is less driven by a disease model. The O-LIFE is derived from the personality framework and it has been designed to give a normal distribution of responses. Consequently, it is made very clear that the psychometric scales used in the current project are nonclinical, hence they do not measure schizophrenia and cannot offer information on the respondent's mental health in general (see Chapters 4 and 5).

1.5.2 Nonclinical and Subclinical Psychotic Symptoms

It is difficult to define what is 'purely' nonclinical and what is sub-threshold given the continuum approach; they could simply be different levels of expression of the same schizotypal traits (with nonclinical considered even milder than subclinical psychotic symptoms). Indeed, phenomenologically it is probably a case of quantitative rather than just qualitative difference in experience between nonclinical and clinical populations (i.e. intensity and duration and not just the content of the psychotic experience). Empirical studies demonstrating the presence of the psychosis continuum can be traced back to at over a decade ago at least (if not further back – Sidgwick and colleagues' report on the Consensus of Hallucinations surveyed approximately 17,000 adults – see Beavan, Read, and Cartwright, 2011) when Stefanis and colleagues measured positive, negative and depressive symptoms as related to psychosis in a large general population sample (Stefanis et al., 2002). The authors found that the three

symptom dimensions were intercorrelated and had a distribution in the general population. Depressive symptoms were prominent in members of the general population experiencing psychosis-like symptoms. Prior to this Ohayon (2000) used general population samples from three European countries (with a total number of over 13,000 participants) to investigate the prevalence of hallucinations in all sensory modalities. Ohayon found that around 38% of all respondents reported hallucinatory experience, with hypnogogic and hypnopompic hallucinations being the most common. Nevertheless, the samples were not purely nonclinical as although they were drawn from the general populations they also included psychiatric patients receiving care in the community.

Perhaps a more suitable example comes from Hanssen, Bak, Bijl, Vollebergh, and van Os (2005) who followed up over 7,000 individuals for two years, again from the general population, for the presence, (dis)continuity and outcome of positive psychotic experiences. The authors found that the incidence of such experiences was around 100 times greater than what had been originally estimated as the incidence of clinical psychotic disorders, and the most likely outcome for these experiences was spontaneous resolution/discontinuity. In those (a much smaller number of individuals) who had more persistent experiences, there was an equally large likelihood of subclinical and clinical 2-year outcomes. They conclude that 'emotional appraisal and degree of intrusiveness of psychotic experiences are important modifiers not for continuity per se but for clinical outcome specifically' (p. 181, a finding that is supported by Johns et al. (2014)'s analysis.

A more recent study by Freeman, McManus, Brugha, Meltzer, Jenkins, and Bebbington. (2011) is also worth mentioning: the authors specifically measured the prevalence of paranoia and paranoid thinking in the general population with different levels of severity and found that around 20% of respondents (obtained from the Adult Psychiatric Morbidity Survey in England) reported feelings or ideas that others are against them, whereas a much lower proportion (less than 2%) reported that others are plotting against them to cause serious harm. Their findings again demonstrate a multidimensional nature of psychosis-like experiences; interestingly, the prevalence of those reporting severe paranoia is still lower than the prevalence of those diagnosed with schizophrenia (estimated to be around 1%), suggesting it is very likely that other

factors (e.g. coping methods, appraisal, social and cultural influences) act as modulators in the development of clinical disorders.

1.5.3 Schizotypy and Risk for Psychosis

Given that the concept of schizotypy in the current Thesis adopts an individual difference/personality perspective rather than necessarily a risk marker one, and the fact that there is still insufficient evidence supporting a solid association between schizotypy and later development of psychosis (although *retrospectively* schizotypy can be found in the early stages of many psychosis patients; Morrison et al., 2002; Perkins, Leserman, Jarskog, Graham, Kazmer, and Lieberman, 2000), predicting the onset of schizophrenia-spectrum disorders from schizotypal traits alone would perhaps prove overly ambitious. Previous research such as the Edinburgh High-Risk Study (e.g. Miller, Byrne, Hodges, Lawrie, Owens, and Johnstone, 2002; Johnstone, Ebmeier, Miller, Owens, and Lawrie, 2005) has identified multiple factors underlying the heterogeneity in schizotypal traits. They found that isolated schizotypal signs seldom lead to an impending onset of schizophrenia and hence are not reliable precursors. The risk therefore is more likely to be highest when a combination of factors (in the Edinburgh study, the factors are social withdrawal, psychotic symptoms, socio-emotional dysfunction and odd behaviour) are present; it is also apparent that psychotic or positive schizotypal traits do not necessarily lead to psychotic outcomes. It is also possible that amongst all the potential predictors, there is an interaction effect between schizotypy measures and other variables.

Another important difference lies between *state* and *trait* markers: as discussed above, clinical high risk is a state whereas schizotypal personality is a trait. As Debbané, Eliez, Badoud, Conus, Flückiger, and Schultze-Lutter (2014) argue, schizotypy is probably more useful when acting as a 'distal risk marker' and may not be as predictive as other indicators in the UHR state criteria in identifying those at *imminent* risk. However, the predictive value of schizotypy seems to originate from its ability to detect psychosis-prone subjects in the community who are otherwise healthy. Therefore, prospective research in large general population samples which combines state and trait markers is needed in order to devise potentially more sensitive *clinical* measures for predicting psychosis onset. For the purposes of the current

research, all measures used are nonclinical and as such, they are neither able nor intended to offer any clinical information.

1.6 Conclusion

This Chapter has offered an overview of the diagnosis, symptomatology, assessment and treatment of schizophrenia-spectrum psychoses with a special emphasis on positive symptoms (delusions and hallucinations), integrating multiple levels of explanation from the molecular to the experiential. It has also discussed the various aspects of psychosis risk states and the concept of schizotypy which are central to the present Thesis. Debatable topics such as whether all delusions are beliefs and all hallucinations are sensory are also included as part of the phenomenological argument based on the latest theoretical and empirical research. The next Chapter focuses on the role of prediction error in the pathogenesis of positive psychotic symptoms and provides a review and synthesis of current literature.

Chapter Two: Predictive Processing and the Pathogenesis of Psychosis

2.1 Introduction

One of the most widely accepted theories of first-rank psychotic symptoms such as hallucinations and passivity phenomena focus on source monitoring deficits; that is, 'a difficulty distinguishing between the origins of endogenous and exogenous stimuli' (for a detailed review, see Nelson, Whitford, Lavoie, Sass, 2014a) as a result of abnormal neural mechanisms involved in the differentiation of internal/external boundaries of stimuli. Another emerging approach based on the concepts of predictive coding and prediction error is often applied to the processes of belief and delusion formation (Corlett, Taylor, Wang, Fletcher, and Krystal, 2010); however, relatively little research has been done in an attempt to link the two theories in terms of a common pathway. Considering delusions and hallucinations are intricately related phenomena phenomenologically (Maher, 2006; Humpston and Broome, 2016), it should not be a surprise that in cognitive terms, source monitoring and predictive coding are interrelated or even interdependent processes. In the following two Sections the aim is to review the relevant literature for both accounts and link the main cognitive neuropsychiatric approaches to the pathogenesis of delusions and hallucinations, with predictive processing as the centre stage.

A prediction error (PE) occurs when there is a mismatch or discrepancy between the expectation of an experience and the actual experience itself. It has been suggested that PEs are 'a general neural coding strategy' present in the whole brain which are involved in perception, cognition and motivational control (den Ouden, Kok, and de Lange, 2012). Consequently, PEs are divided into three main classes: perceptual, cognitive and motivational PEs. Sometimes the first two classes are further grouped together because unlike motivational PEs, perceptual and cognitive PEs only report the degree of surprise or deviation from prior experiences whereas motivational PEs carry a 'sign' or valance (better or worse) of the outcome experienced which can drive behaviour (e.g. reward-seeking).

2.2 Types of Predictive Processing

Perceptual PEs are most frequently investigated by electrophysiological studies, in particular those using electroencephalography (EEG). Such studies often focus on the neural responses to surprising and unexpected stimuli with the oddball paradigm, in which an 'oddball' stimulus that stands out from a series of repetitive and standard stimuli leading to a phenomenon called mismatch negativity (MMN). MMN has been reported in both auditory and visual sensory modalities (Näätänen, Paavilainen, Rinne, and Alho, 2007; Stagg, Hindley, Tales, and Butler, 2004; Czigler, 2007). Interestingly, it has also been found that multisensory regions such as the superior temporal sulcus (STS) are functionally connected to unimodal regions (e.g. auditory and visual cortices). The strength of the violation of prior expectations positively correlated with the STS oscillatory response in terms of both frequency (high or slow) and spatial distribution of cortical activity (Arnal, Morillon, Kell, and Giraud, 2009; Arnal, Wyart, and Giraud, 2011), suggesting a predictive hierarchy of message transmission across modalities (den Ouden et al., 2012).

On the other hand, cognitive PEs arise from cortical areas responsible for higher-order representations which also share the propensity to respond to prediction and surprise. Higher in the predictive hierarchy and above the sensory PEs, cognitive PEs in the dorsolateral prefrontal cortex (dlPFC) are sensitive to the violation of abstract rules. The activity of the dlPFC was maximal when all associations were unpredictable in a causal associative learning task in which participants were asked to predict outcome stimuli based on previously learnt associations, suggesting that PEs are a 'driving force' in associative learning (Fletcher et al., 2001; Corlett et al., 2004).

Signed PEs or the motivational/reward type of PEs reflects whether the surprising outcome is better or worse than expected, which allows an update of the value of the stimulus currently experienced. Motivational PEs are so called because they have the potential to drive behaviour as they are associated with reward and punishment and hence are key to reinforcement learning (den Ouden et al., 2012). Unlike most unsigned PEs which are usually generated in cortical areas, signed PEs are coded by subcortical areas such as the dopaminergic ventral tegmental area (VTA) for reward PEs, which have been shown to be disrupted in psychosis (Murray et al., 2007). Nevertheless, unsigned PEs have been reported in subcortical areas such as the

midbrain dopaminergic substantia nigra (for a study in primates, see Matsumoto and Hikosaka, 2009) and signed PEs in the orbitofrontal cortex for reward which are related to PEs in the VTA (Takahashi et al., 2009). Therefore it would be more helpful to view the relationship between signed and unsigned PEs as dynamic and interlinked rather than a simple and absolute cortical/subcortical divide, although such a divide may still be useful for classification purposes.

2.3 The Neurobiology and Function of Prediction Error

Due to the observation that all currently licensed antipsychotic drugs are either full antagonists or partial agonists at the dopamine D_2 receptor (although it must borne in mind that clozapine, often reserved for treatment-resistant schizophrenia, only has low affinity for D_2) and their efficacy in treating positive symptoms (Seeman and Tallerico, 1998; Nordström et al., 1993), the dominant theory for explaining delusions and hallucinations, at least in terms of neurobiology, has long been the dopamine hypothesis of schizophrenia (Howes and Kapur, 2009; Carlsson, 1988; Meltzer and Stahl, 1976). Neuropsychological theories (e.g. 'jumping to conclusions' in persecutory delusions), on the other hand, were not often viewed in conjunction with the biochemical model. However, in more recent years with the advent of the scientific discipline Cognitive Neuropsychiatry, the neurochemical and neuropsychological approaches to the understanding of delusions and hallucinations have begun to converge. According to the founders of the discipline, Cognitive Neuropsychiatry 'represents a systematic and theoretically driven approach to explain clinical psychopathologies in terms of deficits to normal cognitive mechanisms' (Halligan and David, 2001): for example, explaining delusions with semantic processing and probabilistic reasoning, and hallucinations with verbal self- or source monitoring deficits. These cognitive processes can be further extrapolated to brain (dys)function; typical corresponding brain areas involved in the previous examples would be the fronto-temporal network for delusions and superior temporal gyrus for AVHs (Halligan and David, 2001).

The Section above mentioned the brain regions from where different types of PEs arise; here the aim is to give a more detailed account as to how PEs are generated and

what functions they may serve. Figure 2.1(A-C) exemplifies the generation of PEs in the three classes mentioned above (perceptual, cognitive and motivational).

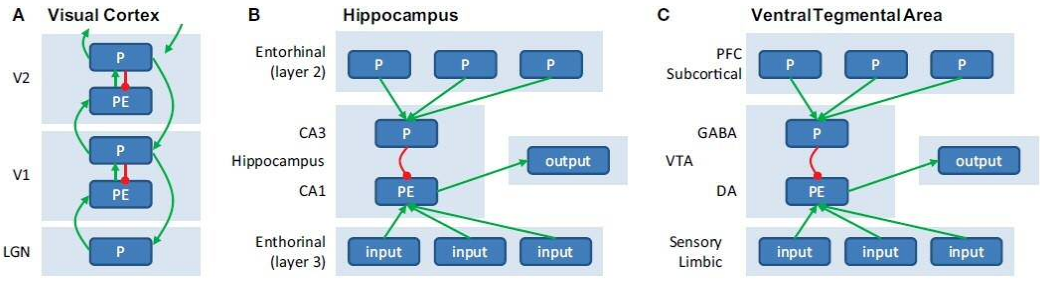

Figure 2.1(A-C). Generation of different types of prediction errors in cortical and subcortical areas. V1, primary visual cortex; V2, prestriate cortex; LGN, lateral geniculate nucleus; CA, *cornu ammonis*; PFC, prefrontal cortex; GABA, gamma-aminobutyric acid; VTA, ventral tegmental area; DA, dopamine. Adapted from den Ouden et al., 2012.

In terms of cortical PEs, it has been suggested that separate 'representation' or 'presentation' (P) units exist alongside PE units which connect between one another either within each cortical column (intrinsic) or between columns (extrinsic). The model in Figure 2.1(A) illustrates the generation of perceptual PEs in the visual cortex as the result of a mismatch between predictions (P, inhibitory, red) and input (excitatory, green), leading to the PE unit as the difference between input and prediction which will then be minimised by the activity in P units. P units can be sent forwards as input to a higher level in the hierarchy or backwards in order to update lower-level predictions and the excitatory feedback from higher-level P units activate lower-level P units. Figure 2.1(B) shows the generation of cognitive PEs in the hippocampus; predictions based on prior memories in layer 2 of the entorhinal cortex act together to drive the activity in CA3 whereas the P unit in CA3 inhibits the PE signal in CA1 (it is viewed as a general principle that P units are inhibitory to PE units). Sensory inputs (layer 3 of the entorhinal cortex) are in turn excitatory to CA1, which then computes an output based on the comparison between predictions and inputs (den Ouden et al., 2012).

Subcortical motivational PEs are thought to have arisen via similar mechanisms with those of cortical (in particular cognitive) PEs; an optogenetic study by Cohen, Haesler, Vong, Lowell, and Uchida (2012) found that activity in GABAergic inhibitory neurons signals/ indicates the delay between a reward-predicting cue and the outcome, whereas dopaminergic neurons in the VTA were active when encoding reward PEs. The actual outcome (presence or omission) of the reward played little role

in how the GABAergic neurons responded; instead, their activity was proportional to how much the cue predicted reward itself. Figure 2.1(C) shows how GABAergic neurons in the VTA inhibit the excitatory input from sensory and limbic regions (primary rewards) when the reward is expected (minimal PE). Punishment or negative reward PEs on the other hand are thought to result from the activity of lateral habenula neurons which project to VTA GABAergic neurons (Jhou et al., 2009). It is also worth noting that although some subcortical PEs are entirely contained in subcortical areas (e.g. midbrain), some may also result from cortical inputs. Area CA1 of the hippocampus (Figure 2.1B) is able to compare between the predictions from CA3 and the actual outcome from sensory inputs which sends the output to the dopaminergic neurons in the VTA (Lisman and Grace, 2005). As den Ouden et al. (2012) summarises, 'the exact content and nature of the PEs is determined by the neural circuitry in which the PEs arise'.

Unsigned and signed PEs serve different functions: for example, unsigned perceptual PEs draw attention to the surprising nature of a presence or absence in the visual scene and allows the updating of how one experiences the world, whereas signed motivational or reward PEs signal the direction of the update by informing the subject about the valance of the stimulus. Also, PEs can either directly lead to changes in perception, cognition, action and motivation by inducing short-lived postsynaptic effects or modulate the storage and updating of prior predictions by regulating synaptic strength (den Ouden et al., 2012). The brain creates a generative model in which one predicts what sensory stimuli will be next experienced or observed in the world and perceptual PEs inform the subject how good their predictions are (perceptual inference). PE minimisation or 'iterative hypothesis testing' across the cortical hierarchy gives rise to the most appropriate and energy-efficient explanation of the sensory inputs in that model and therefore shapes perception itself and helps to reduce its ambiguity unless there are very strong PE signals (Sterzer, Frith, and Petrovic, 2008). On the other hand, signed or motivational PEs play important roles in reinforcement learning such as the concept of blocking where if a cue or cues fully predict a reinforcer, it will not be associated with an additional cue even if the later cue and the reinforcer are consistently paired together.

2.4. Predictive Processing and Psychotic Experiences

One of the most influential theories about altered cognition in psychosis posits that predictive processing is altered in healthy individuals prone to psychosis-like experiences (e.g. Corlett and Fletcher, 2012; Palmer, Davare and Kilner, 2016), patients with first-episode psychosis (e.g. Corlett et al., 2007) and patients with established schizophrenia (e.g. Schlagenhauf et al., 2014).

Perceptual or sensory predictive processing is central to the monitoring and control of motor acts (Shadmehr, Smith, and Krakauer, 2010); in self-generated actions, the predicted outcome of a motor command matches the actual sensory feedback. It has been argued that this match in turn becomes our experiential marker for the sense of agency (i.e. one is the causal agent of one's action) (Sato and Yasuda, 2005). In other words, sensory input caused by self-initiated motor acts is attenuated and there is very little prediction error to minimise (Bays, Flanagan, and Wolpert, 2006; Brown, Adams, Parees, Edwards, and Friston, 2013). Of particular interest is the failure to assign agency to the self in delusion of control (Frith, 2012; Wilkinson, 2014), which is one of the 'first-rank' symptoms of schizophrenia.

Previous studies have demonstrated sensory prediction deficits in patients with established schizophrenia (Lindner, Thier, Kircher, Haarmeier, and Leube, 2005; Shergill, Samson, Bays, Frith, and Wolpert, 2005; Synofzik, Thier, Leube, Schlotterbeck, and Lindner, 2010). To date, three studies have used a nonclinical sample with schizotypal traits who were otherwise healthy (Teufel, Kingdon, Ingram, Wolpert, and Fletcher, 2010; Lemaitre, Luyat and Lafargue, 2016; Palmer, Davare and Kilner, 2016). The authors of the first study (Teufel et al., 2010) found a statistically significant negative correlation between predictive processing in the sensory-motor domain (as indexed by an overcompensation score in a force-matching task; Chapter 4) and delusional ideation as measured by PDI-21, which followed the same pattern as Shergill et al. (2005)'s finding in schizophrenia patients. Using a similar force-matching paradigm, Palmer et al. (2016) have replicated this relationship with PDI-21. Another very recent study (Lemaitre et al., 2016) used a measure of 'self-tickling' as an index of sensory prediction in a student population with high and low positive schizotypy who experienced aberrant perceptions as well as passivity-like phenomena using more specific scales; it followed the same principle that self-initiated tickling

sensations should be reduced due to the same sensory attenuation. The authors found that individuals who rated highly in positive schizotypy (as measured by the Schizotypal Personality Questionnaire) were better at tickling themselves, suggesting reduced sensory attenuation and therefore heightened prediction error signals.

The phenomenon of 'blocking' (Kamin, 1969) mentioned above in associative learning occurs when only one stimulus of a stimulus pair with a given outcome (e.g. AB+) has been previously associated with the same outcome (A+). Responses to stimulus B alone are usually attenuated ('blocked') compared to responses to stimulus A alone or if A had not been associated with the outcome. This weakening of associative strength for B, or indeed any change in the strength of association between stimulus and outcome, can be formalised as a function of a prediction error (the Rescorla-Wagner model; see Tobler, O'Doherty, Dolan, and Schultz, 2006; Haselgrove and Evans, 2010).

There is a significant amount of evidence that in patients with schizophrenia blocking is attenuated or even absent. Patients often view both cues A and B as equally salient or equally good predictors of the outcome and the associative strength for B does not change even after previous training with A+. What is more equivocal is the particular symptom dimension that is associated with this deficit. According to the predictive processing model it would be anticipated that relationships would be found between the positive dimension of schizophrenia and a decrement in blocking. This was supported by Jones, Gray, and Hemsley (1992) who found that blocking was abolished in patients in the acute phase of the disorder, where there is a preponderance of positive symptoms, but was present in those with chronic schizophrenia. Further support is provided by Corlett and colleagues who found links between neural prediction error signals and delusional symptoms (Corlett et al., 2007). However, and in contrast, other researchers have found links between reductions in blocking and negative or nonparanoid symptoms (Oades, Zimmermann, and Eggers, 1996; Bender, Müller, Oades, and Sartory, 2001; Moran, Al-Uzri, Watson, and Reveley, 2003; Moran, Owen, Crookes, Al-Uzri, and Reveley, 2008).

This situation has been mirrored when researchers have adopted a continuum model of schizophrenia and examined schizotypy. Blocking has been found to be reduced in those high in: positive (Moran et al., 2003), and the negative dimension (Haselgrove and Evans, 2010), delusions (Moore, Dickenson and Fletcher, 2011) and

the distress associated with schizotypal delusion-like beliefs (Corlett and Fletcher, 2012).

Patients with schizophrenia show a multitude of deficits in reward processing (see Gold, Waltz, Prentice, Morris, and Heerey, 2008). Previous studies have used a reversal learning paradigm in both medicated and unmedicated patients (Murray et al., 2008a, 2008b; Waltz and Gold, 2007; McKirdy, Sussmann, Hall, Lawrie, Johnstone, and McIntosh, 2009; Schlagenhauf et al., 2014; Reinen et al., 2016). The simplest design of a reversal learning task involves participants choosing between two visually presented stimuli (e.g. geometrical shapes): participants receive some kind of reward for choosing the correct stimulus and are punished (e.g. a reduction in the amount of money earned) for choosing the wrong stimulus. When a reversal happens, the rules are switched so that the previously correct stimulus becomes the wrong one, and vice versa. Of particular importance to psychosis is the observation that manipulation of dopamine levels modulates reversal learning, in which subjects with high baseline dopamine synthesis in the striatum showed relatively better reversal learning from unexpected rewards than from unexpected punishments. This pattern was reversed in those with low baseline dopamine synthesis (Cools, Frank, Gibbs, Miyakawa, Jagust, and D'Esposito, 2009). In addition, when the positive reward PE is large it will strengthen the association between reward and action and lead to a selection bias. Reward PEs hence have both immediate and direct influence on action selection and also longer term effects on learning due to a selection bias towards reinforced actions. Significant to the pathogenesis of psychosis is that abnormally large PEs result in increased salience of stimuli as top-down predictions can no longer find suitable explanations for such events (Murray et al., 2007; Heinz and Schlagenhauf, 2010; Berridge, 2007).

Current evidence (e.g. Schlagenhauf et al., 2014; Reinen et al., 2016) suggest that acutely psychotic patients display an insensitivity to positive reinforcement and increased tendency to switch regardless of reversal status which corresponds to reduced error signals in the ventral striatum. This is consistent with other studies on delusions, proneness to switching and reward insensitivity across a variety of tasks not limited to reversal learning, but also other set-shifting tasks (e.g. Cella, Dymond, and Cooper, 2009). Chronically medicated schizophrenia patients, on the other hand, tend

to show more deficits in perseveration but not switching (Elliott, McKenna, Robbins, and Sahakian, 1995).

To date, the vast majority of studies using the predictive processing framework in psychosis have been done in patients with non-affective psychoses (mostly schizophrenia); as outlined in the previous Chapter, affective psychoses are a related yet distinctly heterogeneous group within the spectrum of psychotic disorders, but have received comparatively little focus. Perhaps due to the low prevalence, patients with schizoaffective disorder only constitute a very small proportion of patients recruited in such studies and are mixed with patients with non-affective psychoses (e.g. Fogelson, Litvak, Peled, Fernandez-del-Olmo, and Friston, 2014) and there have not yet been any studies exclusively on predictive processing in patients with schizoaffective disorder. However, Gradin et al. (2011) found significant differences in reward predictive coding between patients with unipolar depression without psychotic features and those with schizophrenia: the abnormally blunted encoding of PEs contributed to anhedonia in depression whereas disruptions in PE-dependent inferences/belief updating were responsible for psychotic symptoms in schizophrenia. Therefore, the presence of significant mood symptoms could confound or complicate the predictive and/or source monitoring processes in an affective psychosis. For the purposes of the current Thesis, depressive symptoms in clinical patients were assessed to ensure a full diagnosis of affective psychosis or mood disorder was not met, and as such non-affective psychotic symptoms were the sole focus of investigation.

2.5 Delusions and Hallucinations: 'Perceiving is Believing'?

Previous research has mostly focused on delusion formation and hallucinations as if they are separate processes, with prediction error and source monitoring deficits as their 'corresponding' bases in psychopathology. However, more recent efforts (e.g. Fletcher and Frith, 2009, Griffin and Fletcher, 2017) have attempted to establish a relationship between belief and perception which can be extrapolated to a unified account for explaining delusions and hallucinations. In this approach the fundamental deficit lies within a Bayesian framework where experiences are dependent on one's beliefs and the extent to which one updates the beliefs are affected by the experiences themselves. In other words, there is significant interaction between abnormalities in

belief (i.e. delusions) and those in perception (i.e. hallucinations). Learnt associations can influence perception and delusional beliefs can alter the content of hallucinations so that they are consistent with that of the delusions (Kot and Serper, 2002).

Within this Bayesian hierarchy, the PE signal acts as an internal marker that the existing inference cannot fully account for the input (Fletcher and Frith, 2009). It has been proposed that minimisation of PE signals at the next level of the hierarchy may improve the accuracy of the current model; however, if this fails higher-level adjustments are needed in order to keep the inferences of sensory input internally consistent. According to Fletcher and Frith, what drives the learning of new associations is the fact that lower-level PEs cannot be accounted for by the current model, hence creating a heightened salience towards the newly emerged attention-grabbing stimuli.

In schizophrenia, patients' falsely generated and imprecise PE signals are sent to higher-levels of the hierarchy and are constantly demanding an update of the current inference, yet because they are 'false alarm' signals the higher-level readjustments are never enough to resolve the problem. Therefore, it leads to a vicious circle where these abnormal PE signals go higher and higher in the levels of abstraction and the higher they propagate, the more severe the disruption is to the Bayesian system. This may not only result in the formation of delusions but also their maintenance and persistence because of repeated reinforcement and reconsolidation (Corlett, Krystal, Taylor, and Fletcher, 2009).

Nevertheless, there is also a distinction between the presence of PE themselves and their precision or uncertainty (Corlett et al., 2010): as precision and uncertainty are inverse concepts to each other, delusions form because of a failure to accommodate for such uncertainty and imprecise PEs. Dopaminergic hyperfunction in the VTA-PFC circuit has been suggested to play a key role in both encoding reward PEs and modulating the level of precision (i.e. signal-to-noise ratio) of reward and non-reward PEs, whereas glutamate (NMDA receptors) signalling would be diminished. Interestingly, although this account has been primarily attributed to delusions, motivational states and reward-driven learning can also influence perceptual judgement which leads to the (related) theory of source or reality monitoring deficits in some delusions (e.g. delusions of control) and auditory-verbal hallucinations (Frith, 2012).

For example, in passivity phenomena the efference copy of motor commands which makes a prediction about their sensory outcome is either not sent by the internal forward model or is unable to cancel out the prediction of the consequences of the patient's actions (Chapter 1). Sometimes passivity and delusions of control are used interchangeably; however, there is some debate over whether the delusional elaboration is in fact phenomenologically separate and is not an intrinsic requirement for passivity experiences to arise. If this argument is indeed true it might be that PE signals can mediate both the formation of the delusion and the passivity experience itself, of which the latter can also be extrapolated to explaining certain subtypes of AVHs.

The two-factor account of delusions proposes that both perceptual and reasoning deficits are needed in order for a delusion to form (Coltheart et al., 2007; McKay et al., 2007; Davies et al., 2001): this is based on empirical findings that the sites of damage in neurological patients who also have delusions include both a perceptual site and a belief evaluation site which are responsible for perceptual abnormalities and their delusional explanations, respectively. Also, the first factor is thought to explain the content of the delusion and the second factor the maintenance of the delusion. Nevetheless, some argue that the two-factor account cannot accommodate all types of delusions, especially polythematic delusions. On the other hand, the Bayesian approach based on PE signalling combines the two-factor theory into a single deficit in Bayesian inference: abnormal perceptions originate from faulty or noisy PE signals which lead to the need for updating prior beliefs and further, delusion formation.

The key region for registering PE signals in causal learning is the right dlPFC whose abnormal activities are correlated with the severity of delusions in both endogenous and psychotogenic drug-induced (e.g. ketamine) psychosis (Corlett, Honey, and Fletcher, 2007a). Although the PE approach is sometimes viewed as a part of the damage to the belief evaluation site by two-factor theorists, it has been suggested that the former can better account for many of the abovementioned interactions between belief and perception by applying the faulty PE signal to the entire hierarchy, both bottom-up and top-down (feedforward/feedback) with deficits in Bayesian inference as the single factor that has become abnormal. Similarly however, this approach has also received criticism that aberrant PE signalling alone is insufficient

for a delusion to fully develop and does not explain the elasticity of some delusional experiences such as double-bookkeeping (Sass, 2004).

In order to solve the problems in both accounts, some of the latest theories propose that the PE approach should not be considered a rival to the two-factor account but as an 'ally' in the sense that the PE approach can in fact be integrated into both factors of the two-factor account despite many of their intrinsic differences (Miyazono et al., 2015).

2.6 Conclusion

This Chapter consists of summaries of both theoretical underpinnings and recent empirical support for the roles of prediction error in the pathogenesis of positive psychotic symptoms, namely delusions and hallucinations. Areas covered include the classification of positive symptoms and their diagnostic significance, the various types of prediction errors (perceptual, cognitive and motivational), their function and neurobiological basis. Some of the latest cognitive neuropsychiatric approaches to explaining delusions and hallucinations, in particular the Bayesian framework in relation to the link between belief and perception, are also reviewed. The theoretical validity of the roles played by predictive processing is indeed convincing and there should be little doubt about the significance of their involvement in the development of early psychosis.

Chapter Three: Source Monitoring and Psychotic Symptoms

3.1. Introduction

Having discussed how psychotic symptoms may arise from the perspective of the predictive processing framework and the concepts of prediction error, this Chapter focuses on the other main component of this Thesis: the source monitoring framework. Building upon theories of recognition memory in general, this framework is then applied to empirical studies about how deficits in source monitoring can contribute to the genesis of psychotic experiences. More specifically, how the genesis of such symptoms and experiences results from failures to attribute the correct source (as a consequence of disruptions in source monitoring) of internal mental states and processes which may be linked to the detection of prediction error. With a more integrative account in mind, the aim of this Chapter alongside with the previous one is to demonstrate that predictive processing and source monitoring are two interlinked frameworks which, when considered in unity, may have significant theoretical and practical implications for the understanding of psychosis and related symptoms.

3.2. Recognition Memory

Before introducing the source monitoring framework, it is important to offer a brief review of the most influential theories of recognition memory. Recognition memory is a subcomponent within episodic memory which, together with semantic memory, fall under the umbrella category of declarative memory.

3.2.1. Signal-Detection Theory

The signal-detection theory is in essence a single-process model and views familiarity as the unidimensional signal underlying the strength of recognition memory (Wixted, 2007) which originated from the 1960s. The classical signal-detection theory is comprised of two Gaussian distributions with equal variance (Fig. 3.1 below), one representing targets and the other representing new items, together with a single decision criterion. Any memory signal exceeding the criterion strength is deemed old

('this really happened') whereas any signal falling under the criterion threshold is deemed new ('first time encounter').

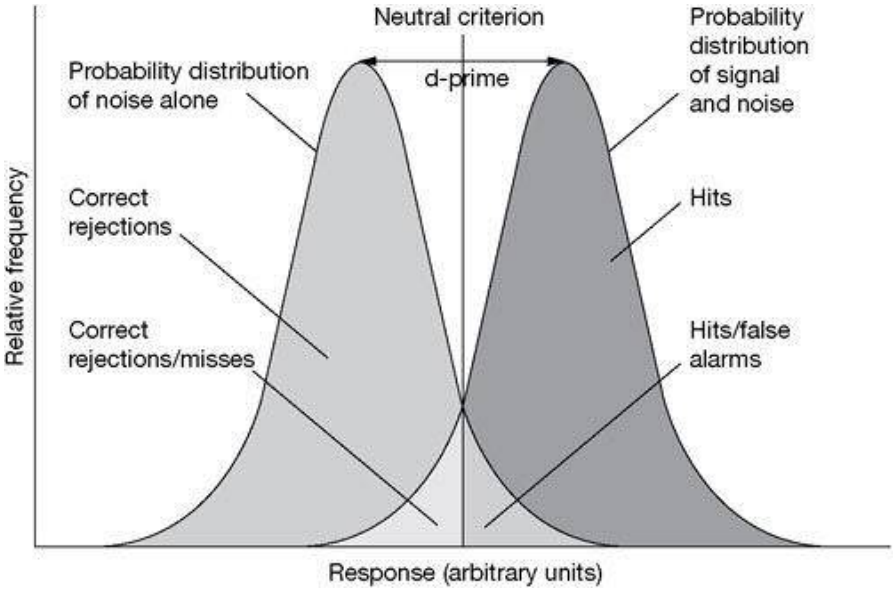

Figure 3.1. A schematic representation of signal detection theory showing two Gaussian distributions with equal variance. Adapted from Oliver et al., 2008.

Instead of the equal variance (or symmetrical) version of the theory which may initially seem more appealing, a more recent revision of the signal-detection theory suggests that the unequal variance model is more plausible (Mickes, Wixted, and Wais, 2007). For example, each item on a standard memory recall list would not have exactly the same amount of memory strength added during the study (the equal variance version assumes that targets have memory strength added to them in equal increments) and as such resulting in the unequal variance.

This type of model (whether the variance of the distributions is equal or not) proposes a single process that needs accounting for, namely familiarity. The strongest evidence for this theory perhaps comes from study findings employing receiver-operating characteristics (ROC), the function which relates the proportion of *correct* recognitions (hit rates) with the proportion of *incorrect* recognitions (false alarm rates).

3.2.2. Dual-Process Model

Whilst it is indisputable that the signal-detection theory has offered crucial contributions to the studies of memory, more contemporary researchers have instead suggested that two processes contribute to recognition memory decisions. Familiarity describes the ability to determine whether an event has happened or not but without

any contextual details, whereas recollection is the ability to retrieve specific details about the said event (Yonelinas, 2001).

Empirical evidence supporting the dual-process model is substantial, and consistently converge towards the dissociable nature of recollection and familiarity. These findings range from behavioural studies to neuropsychology, electrophysiology and neuroimaging (Rugg and Yonelinas, 2003). Studies using imaging methods (e.g. Rugg and Curran, 2007) such as electroencephalography (EEG) and functional magnetic resonance imaging (fMRI) support the notion that there are dissociable neural correlates of recollection and familiarity. For example, Woodruff, Hayama and Rugg (2006) demonstrated that recollection is not just a consequence of strong familiarity (as one variation of a single-process model would propose): in their study they found a double dissociation between a mid-frontal Event-Related Potential (ERP) associated with familiarity and a left parietal ERP which indexed recollection. Recollection and familiarity dissociated across temporal, functional and topological domains. In a follow-up study by Yu and Rugg (2010), they replicated this dissociable effect and offered strong supporting evidence for the dual-process model. This dissociation is also supported by fMRI studies (e.g. Yonelinas, Otten, Shaw, and Rugg, 2005) demonstrating separable activation patterns across different subregions of the frontal and parietal cortices when performing tasks designed to differentiate between familiarity- and recollection-based memory. Further evidence comes from studies in amnesiac patients (regardless of aetiology), who have severe and persistent deficits in episodic memory but otherwise maintain unaffected cognition in other areas. Recollection is more severely affected than familiarity in amnesia and brain injuries, indicating separable processes which are dependent on different brain regions (Yonelinas, Kroll, Dobbins, Lazzara, and Knight, 1998).

It is worth noting that there are multiple methods by which one can measure familiarity and recollection. The first of these is the ROC procedure already mentioned above, where participants rate the confidence of their recognition judgements. Hits (accurate 'old' judgements) can either be made by above-threshold familiarity in the absence of recollection or on the basis of recollection, whereas false alarms can only be driven by above-threshold familiarity. An ROC is then plotted as a function of hits relating to false alarms. A different but also well-used approach is the Remember-Know-New (RKN, Tulving, 1985) procedure where participants are asked about the

basis of their judgements, for example, report 'remember' (remembering an item as old), 'know' (familiar items without recollection, i.e. just knowing an item is old) or 'new' (neither process is involved). A third method, the process-dissociation (PD, Gruppuso, Lindsay, and Kelley, 1997) procedure, uses accuracies in relational recognition judgements (where participants are asked to retrieve specific aspects or context about the study phase) as an estimate of recollection and the conditional probability of recognising an item that is not recollected as an estimate of familiarity.

When the dual-process model by Yonelinas was first proposed it was considered in direct opposition to the signal-detection theory, given that the former introduced a second separable process of recollection. For instance, provided a participant is able to successfully retrieve information that the recall item passes the familiarity threshold and is recognised as old, the confidence response would be naturally high not because of familiarity but because the recollection of additional information surrounding that item without affecting the false alarm rates (the item is still judged as old, just with higher confidence). The signal-detection theory may explain this as an increase in the variance of the old item distribution but proponents of the dual-process model would argue that the old item distribution is right-skewed because of the factor of recollection, whereas the familiarity factor in fact does exhibit *equal* variance by itself.

More recently, Wixted (2007) put forward an alternative model which aims to reconcile the differences between the more traditional signal-detection and dual-process theories. According to these authors, old-new recognition decisions are made with increasing confidence to the extent that graded recollection of source and associative information, *together with familiarity*, is retrieved. By combining the processes of familiarity and recollection, this alternative model revises the unidimensional memory strength variable of signal-detection theory whilst incorporating the second factor of recollection as well. In other words, an individual's recognition decisions are not based on just familiarity or recollection but rather on the aggregate of the two processes. There is still much research to be completed to determine which of these models most accurately describes recognition memory but for the purposes of the current Thesis the critical factor is that there are two processes: Recollection and Familiarity.

3.3. Recollection and Familiarity in Schizophrenia

In certain psychiatric disorders such as schizophrenia, difficulties in episodic memory reflect a core cognitive impairment (Elvevag and Goldberg, 2000; Ragland et al., 2009), which is observed in young medication-naïve patients (MacDonald III et al., 2005) and healthy first-degree relatives of those with schizophrenia (Toulopoulou et al., 2003; Snitz et al., 2006). These memory impairments are largely unaffected by antipsychotic medication (Vinogradov et al., 1997). Research which elucidates the nature of the memory impairment is of vital importance because memory performance is one of the strongest predictors of functional outcome (Green, 1996; Milev et al., 2005).

Given that impairments in episodic memory are a core feature of the schizophrenic illness (Clare et al., 1993), many researchers have examined whether the deficit is in recollection and/or familiarity. In a recent quantitative review, Libby, Yonelinas, Ranganath, and Ragland (2013) concluded that deficits exist in both processes in patients with schizophrenia even after accounting for methodological differences; however, effect sizes in familiarity-based memory deficits tended to be smaller. Weiss, Goff, Duff, Roffman, and Schacter (2008), on the other hand, only found specific deficits in familiarity but not in recollection in 18 established schizophrenia patients and 18 matched controls. In the study phase of their study participants saw and heard 26 consecutive items, 13 spoken by a male and 13 by a female and identified the gender of the voice. In the test phase, participants were presented with 26 previously studied items and 26 new items, then asked to identify the source (male, female or new). Control participants were significantly more accurate than patients in old-new recognition (familiarity) but there was no evidence for difficulties associated with source memory (recollection) in patients. Weiss and colleagues also did not find any significant relationship between corrected recognition, source accuracy and psychopathology measurements.

Of the studies included in Libby et al.'s review, the majority employed the RKN model which the authors argue drove the hypothesis that recollection is the form of memory most impaired in schizophrenia, as studies using this procedure did not usually account for the non-independence between familiarity and recollection or response bias. In other words, in classic RKN models the 'remember' responses would

also rely on a certain degree of familiarity and not just the 'know' responses. It was only after a re-analysis of existing results by the RKN studies that the magnitude of deficits in familiarity-based memory increased in patients, although the effect sizes were smaller than those for recollection-related deficits. ROC and PD studies were not reanalysed and these were also the types of studies that found deficits in both retrieval processes. On the other hand, all the studies included in the review which either found unaffected familiarity (7 studies) or increased familiarity (5 studies) in patients with schizophrenia employed the RKN procedure.

A potential confound in the RKN procedure is the heavy reliance on metacognitive capacities which have been shown to be impaired in schizophrenia (Moritz, Woodward, Jelinek, and Klinge, 2008; Lysaker and Dimaggio, 2014; but also see Bacon and Izaute, 2009). This is more complex from the memory confidence ratings in the ROC approach in that it places a higher cognitive load on patients, even though some (albeit much less) metacognition is indeed required for the ROC procedure as well. Difficulties in metacognition are likely to also relate to the source judgement deficits frequently seen in patients with schizophrenia (see below), confidence and liberal acceptance over 'false' memories are found to be associated with delusional thoughts (Moritz and Woodward, 2006) compared with hallucinations. Nevertheless, such false memories may well be the consequences of hallucinated voices, for example.

As can be seen from the evidence outlined above, patients with schizophrenia frequently have larger deficits in recollection-based recognition memory than in familiarity-based recognition memory. The source monitoring framework will be described next, which relates to how one decides upon the contextual features of a memory.

3.4. The Source Monitoring Framework

In a seminal paper by Johnson, Hashtroudi, and Lindsay (1993), the authors define source monitoring as 'the set of processes involved in making attributions about the origins of memories, knowledge and beliefs' (p.3). Most importantly, the source monitoring framework posits that individuals do not retrieve some kind of abstract tag or label about the source of a certain memory record, but such a memory record is

evaluated by decision processes during remembering and eventually attributed to a particular source. 'Source' here refers to the combination of characteristics that specify the conditions under which an episode was committed to memory. The various types of sources include perceptual details (e.g. sound, colour), contextual information (when and where), semantic detail and emotional reactions (affective information) and cognitive operations. It is worth noting that although Johnson and colleagues propose that source monitoring involves decision making, such a process needs not to be deliberative and may well occur quite rapidly in the course of remembering without much conscious effort, especially when the memory record is rich with contextual information. One can arrive at a decision based upon these different characteristics heuristically or systematically. Heuristic source judgements involve criteria or thresholds, e.g. if the familiarity level exceeds X, the event is likely to have happened, or if the amount of perceptual detail is over Y, the event is likely to have been perceived externally. Most source judgements are made heuristically, but systematic processes are slower and can check the plausibility of a memory judgement which would have otherwise passed the heuristic threshold (especially in the case of highly salient perceptual information being present). On the other hand, sometimes recollections are easily accepted because they fit well with one's beliefs and knowledge, and in this case heuristic processes can ensure such memory events possess sufficient sensory detail.

Source monitoring, however, is by no means a unitary concept. Different types of source monitoring have been proposed based on the perceived internality or externality of the 'source' of a certain event to be remembered and there are three different subcategories of source monitoring:

1) Reality monitoring, defined as the act of differentiating memories of internally generated information from externally generated information, such as discriminating internal mental events (memories for imaginations and thoughts, for example) from external perceptions;
2) Internal source monitoring, defined as the act of discriminating memories between two internal sources (e.g. between what one has said or performed from memories of what one has imagined or thought); and

3) External source monitoring, defined as the act of differentiating between memories from two different external sources (e.g. statements made by person A from those by person B).

It may seem apparent that the first component, errors in reality monitoring (also called internal-external source monitoring), can in fact go both ways: one could confuse having thought of something with seeing someone else do the same thing, but also vice versa. Indeed, Johnson and colleagues (1993) also point out that one type of reality monitoring emphasises the self (such as discriminating one's own imagination from externally occurring events, like hearing someone else say something) and the other type puts the emphasis on the covert or private properties of internal mental events rather than the public qualities of other events. However, this second emphasis may entangle some instances of internal source monitoring with those of reality monitoring: for example, if one confuses what one did with what one thought, which kind of source memory error would this be? With the first type of definition it would be an internal source monitoring error but with the second type of definition it would be classed as a reality monitoring error. The authors seem to think that it is not a case of one definition being more correct than the other, but it is simply that 'one is emphasising the self versus external source as the *origin* of information whereas the other is emphasising the actual (public) versus imaginal (private) *status* of information' (p.4). However, for the current Thesis the first classification is adopted given its importance in the genesis of psychotic symptoms.

In patients with schizophrenia-spectrum psychoses, source monitoring mechanisms (in particular, reality monitoring) are thought to be severely disrupted and may act as some of the generative processes underpinning positive psychotic symptoms, especially hallucinations (Bentall et al., 1991; Johns et al., 2001). The next section will outline the evidence as to whether these individuals have deficits in certain types of source monitoring and if so, whether these are related to any symptoms.

3.5. The Role of Source Monitoring Deficits in Psychotic Experiences

One paradigm that has been used to investigate the relationship between deficits in source monitoring and psychotic experiences in both schizophrenia and healthy schizotypy involves presenting incomplete but well known word pairs (e.g. fish and ?)

or sentences for the participants to fill in the blank in the study phase (e.g. fish and chips) and in the test phase, participants are asked to decide whether they generated the word, the word was given to them or new (for example, see Simons, Henson, Gilbert, and Fletcher., 2008; Garrison et al., 2017a, 2017c). Other studies involved filling in the blank part of a sentence (e.g. Vinogradov et al., 2008).

A great deal of research has focused on reality monitoring because it has been proposed that it may play a role in the pathogenesis of some of the positive symptoms of schizophrenia, such as hallucinations and delusions (Bentall et al., 1991; Frith, 1992; Frith and Done, 1988; Rankin and O'Carroll, 1995). Studies which have examined this capacity have typically involved presenting participants with either a complete sentence or one where they need to fill in the blank. In the test phase participants need to indicate whether they generated the word, it was given to them or is new (Vinogradov et al., 1997). Some findings suggest that individuals with schizophrenia have deficits in reality monitoring and, in particular, that they misattribute self-generated events to an external source (Johns et al., 2001; Keefe et al., 2002; Vinogradov et al., 2008). As anticipated, many of these studies found the deficit to be linked to the positive symptoms (Brébion et al., 2000, 2002). A study by Johns et al. (2010) specifically investigated verbal self-monitoring in individuals at clinical high risk of psychosis and found that at-risk individuals misattributed the source of their speech to 'other' when acoustic distortions were applied to their own speech, and these misattributions were at an intermediate level between established schizophrenia patients and healthy controls. Their paradigm was significantly different from that used in Chapter 5 as no acoustic distortions or any external interference were used in the latter study; however, their findings may still have implications that verbal self-monitoring impairments were already present in the prodromal phase.

More recently, internal source monitoring has also been examined because the distinction between imagination and reality is often blurred in schizophrenia (Mintz and Alpert, 1972; Brébion et al., 2008). A wide variety of source monitoring tasks have been utilised to study the performance of patients with schizophrenia. For example, Gawęda et al. (2012) asked patients to either imagine or actually perform an action and found that they confused the source of these actions in a subsequent test phase.

However, the overall evidence is mixed with regard to the notion that reality monitoring errors made by patients with schizophrenia, in particular those involving attributing internally generated stimuli to an external source, are associated with positive symptoms. In a study comparing hallucinating versus non-hallucinating patients with schizophrenia, Brunelin et al. (2006) investigated their performance in a 'Say-Imagine' condition and a 'Hear-Imagine' condition corresponding to internal source monitoring and reality monitoring, respectively. The authors found that hallucinating patients misattributed an internal event to an external source more frequently (reality monitoring errors) than non-hallucinating patients with schizophrenia, and the two groups did not differ significantly on internal source monitoring errors ('Say-Imagine' condition). They concluded that reality monitoring deficits could be seen as a state marker for auditory verbal hallucinations whereas internal source monitoring deficits were trait markers for a schizophrenic illness in general, and may relate to symptoms such as thought insertion and delusions of control where the patient fails to correctly monitor their own thoughts and actions.

In other studies the specificity of reality monitoring deficits to hallucinations has been brought into question. In addition to hallucinations, reality monitoring deficits have also been associated with delusional ideation (e.g. Brébion et al., 2000, 2002), with some authors reporting relationships with delusions only (Anselmetti et al., 2007). In Brébion and colleagues' study of 40 inpatients with schizophrenia and 40 matched controls, higher hallucinations scores were associated with increased tendency towards false recognition of new (non-produced) items and hallucinators were more prone to misattributing self-produced items to another source. In addition, hallucinators as well as delusional patients more prone to report spoken items as pictures. Anselmetti et al. (2007) studied 45 patients with chronic schizophrenia and 54 healthy controls using a task involving conditions of word-stem completion, viewing pictures or listening to the experimenter read out words. Both old/new item recognition and source attribution were investigated, and the authors reported that only delusions significantly related to internal-external source misattribution but not hallucinations or disorganisation.

Such findings fit well with the concept of schizophrenia as an ego-boundary disorder where patients experience a disturbed sense of agency over inner speech and thought processes (e.g. McGuire et al., 1995; Jones and Fernyhough, 2007) which then

manifest as auditory-verbal hallucinations and thought interference symptoms (e.g. thought insertion with delusional elaboration). However, measures of delusions and hallucinations do not typically focus on these subtypes and it is difficult if implausible to apply the source monitoring model to other types of delusions which do not usually involve a *lack* of agency but a *heightened* sense of the self as the central point of psychopathological experience (such as delusions of persecution and reference). Moreover, some theorists propose that hallucinations are extremely heterogeneous and cannot always be accounted for by misattributed inner speech (McCarthy-Jones et al., 2014). Further, other researchers have found no correlations between reality monitoring deficits and schizophrenia symptomatology (Henquet, Krabbendam, Dautzenberg, Jolles, and Merckelbach., 2005) even though difficulties in differentiating imagined thoughts and verbalised speech were still prominent in patients. Other researchers have found poor reality monitoring to be associated with negative symptoms (Brébion et al., 2002; Moritz et al., 2003) as well as thought disorder (Nienow and Docherty, 2004).

Source monitoring deficits in schizophrenia are not only apparent in the domain of words and speech, but also in the monitoring of actions. However, studies employing action-based source memory tasks tend to find impairments in internal source monitoring but no association with symptomatology (e.g. hallucinations). For example, Gawęda et al., (2012, 2013) found significant deficits in distinguishing imagined versus performed actions in patients with schizophrenia but such deficits were *not* more prominent in patients who reported frequent auditory hallucinations than those who did not. These findings may indicate general source monitoring deficits in schizophrenia, regardless of symptom severity; however, they could also mean that the assessment tools used were not sensitive enough to detect effects often associated with very specific forms of symptoms, for example delusions of control and passivity symptoms, as many of the commonly used symptom assessment tools do not separate out ego-boundary disturbances. These kinds of 'first-rank' symptoms are of particular interest because they involve failures in monitoring the source of one's own thoughts and actions by definition (Frith, 2005, 2012). It might just be that the effect has been diluted because different kinds of psychopathological symptoms are often 'bundled' together which would in fact require more focused assessments due to potentially

different pathogenetic mechanisms involving both diminished and heightened senses of agency (van Duppen, 2016).

Hallucination-proneness and high schizotypy in general have both been linked with deficits in internal source monitoring for action-based tasks (Collignon, van der Linden, and Larøi, 2005; Peters, Smeets, Giesbrecht, Jelicic, and Merckelbach, 2007). In Collignon and colleagues' study, 65 normal participants underwent a source monitoring task of five conditions (participant perform the action, watch the experimenter perform the action, imagine oneself perform the action, imagine the experimenter perform the action, or listen to the experimenter say the action verbally) and were tested for old-new recognition and identifying the source for old items. Participants were stratified according to their scores on the Launay-Slade Hallucinations Scale (Revised), with the top quartile and the bottom quartile corresponding to hallucination-prone and non-hallucinators, respectively. Hallucination-prone participants displayed more internal source monitoring errors but not reality monitoring errors. In Peters et al.'s study, similar patterns were found in internal source monitoring deficits (confusing participant performed and participant imagined actions) in a healthy undergraduate sample assessed by the Schizotypal Personality Scale. Further, source misattribution was related to working memory capacities.

Such findings about internal source monitoring are interesting because, in a sense, they challenge the traditional notions of first-rank symptoms (third-person auditory hallucinations, thought interference, etc.) and even the nature of disordered ego-boundary in schizophrenia: although the sense of agency is rarely measured directly alongside source monitoring, some studies on agency and schizotypy suggest that healthy individuals with schizotypal traits also experience the lack of agency over thoughts and actions because of such source monitoring difficulties (Asai and Tanno, 2007; Asai, Sugimori, and Tanno, 2008). However, these individuals are by no means psychotic at the time of study and the vast majority will never become clinically psychotic. This suggests that the presence of unusual experiences or even problems with agency are not the sole indicators of an illness, and rather than defining schizophrenia as an ego-boundary disorder, it might be more useful to focus on why some individuals, despite significant difficulties in source monitoring and feelings of agency, do not find their experiences distressing or bothersome.

However, it must be mentioned that very recent evidence (Garrison, Moseley, Alderson-Day, Smailes, Fernyhough, and Simons, 2017b) suggests that there is no reduction in both reality and internal source monitoring abilities in hallucination-prone (more specifically, hallucinations in the auditory-verbal modality) but otherwise healthy individuals. Although the authors used a word-based task rather than an action-based one, the absence of deficits in the hallucination-prone group challenged a fully dimensional model of schizotypy and provided support for a quasi-dimensional model, where the distribution of psychosis-like symptoms does not follow the normal bell-shaped curve but skews towards very few symptoms in the majority of the general population. The authors also acknowledged a potential specificity regarding impairments in the monitoring of actions, which classes the intention to speak as an action prior to the production of inner speech as described by the comparator or forward model (Chapter 1), which follows a predictive processing framework.

3.6. Conclusion

This Chapter offered an overview of source monitoring from the viewpoint of recognition memory and discussed how source monitoring processes are often impaired in both patients diagnosed with schizophrenia and healthy individuals with schizotypal personality traits. Although the current evidence is mixed as to what kind of source monitoring deficits is most relevant to specific psychotic experiences and symptoms of schizophrenia, it is widely accepted that there are clear cognitive deficits associated with source monitoring in individuals with psychosis-like experiences and more severely, in individuals with clinical schizophrenia.

Chapter Four: Aims and Hypotheses of the Current Thesis

4.1. Introduction

The previous three Chapters provided detailed background information about psychosis and review of the literature in predictive processing and source memory relevant to psychosis research. Although both frameworks have already been examined in individuals with various schizotypal traits and also in patients at different stages of psychosis (e.g. early versus chronic), there has been no research to date which combines the different domains of these frameworks in the same cohort of individuals. Given that the different types of predictive coding drive learning and inferential processes in unique ways (Chapter 2) yet are interrelated at the neural level, it is important to study these types of prediction error-based mechanisms in the same individuals, and how these processes relate to their ratings of schizotypy. For example, previous evidence has been rather inconsistent in terms of blocking and dimensions of schizotypy (positive, negative, disorganised, etc.); in contrast, findings in sensory predictive processing have been relatively consistently related to delusional ideation.

A similar line of reasoning follows with the source monitoring tasks, which despite having been extensively used in healthy individuals and patient populations in an isolated manner, these tasks have not been employed in the same individuals combining different types of source monitoring. Many previous studies have found that deficits in the source monitoring of words correlate most significantly with auditory-verbal hallucinations, for example, but again these findings vary across individuals and largely depend on the features of the tasks used and the measures taken. Therefore, it is crucial to use a battery of tasks aimed at tapping into these different processes in the same cohort.

The general aim of the current Thesis is to report the findings from two empirical studies, one in a large number of healthy individuals with varying degrees of schizotypal traits and the other in a much smaller sample of individuals displaying an early stage of psychosis and matched controls from the general population. The first study employed a battery of behavioural tasks aimed at tapping into domains of predictive processing (sensory, associative and reward) and source monitoring

(actions and words), whereas the clinical study used a subset of these behavioural tasks. Each study is divided into two Chapters according to theme (predictive processing/source monitoring), with Chapters 5 and 6 being the healthy schizotypy study and Chapters 7 and 8 the clinical study.

4.2. Study 1: Predictive Processing and Source Monitoring in Healthy Individuals with Schizotypal Traits

4.2.1. Aims

The primary aims of Study 1 were to examine the associations between schizotypy in healthy individuals and any behavioural deficits in prediction error responses and source monitoring.

For predictive processing, associations were examined between delusional ideation (measured by PDI-21), hallucinatory experiences across modalities (measured by CAPS), and general schizotypy domains (measured by O-LIFE) and different domains (sensory, associative and reward types) of predictive processing as measured by three well-validated behavioural tasks (force-matching, blocking and reversal learning).

For source monitoring, associations were examined between general positive schizotypy (as measured by O-LIFE) and old/new recognition, source memory accuracy, internal source monitoring errors and reality monitoring errors in both action-based and word pair-based tasks. It is important to note that O-LIFE was the only schizotypy measure used in the analysis of the source monitoring tasks; PDI-21 and CAPS were omitted on purpose due to the fact that only limited prior evidence was available on specific associations between hallucinations across modalities (as opposed to the auditory-verbal domain) and delusional ideation and source monitoring deficits.

The secondary aims of Study 1 were to compare results with those from previous studies using similar (but not identical) behavioural tasks and to determine the reproducibility of such findings.

4.2.2. Hypotheses

In order to avoid the inflation of family-wise errors resulting from multiple comparisons, frequentist confirmatory analyses were only carried out based on the following *a priori* hypotheses tailored to each task:

1) Force-matching task:
    a) Participants would consistently apply more force in the Finger condition (applying force directly by finger) than the Slider condition, creating an overcompensation score.
    b) Overcompensation scores would negatively correlate with the participants' PDI-21 (delusional ideation) total scores.
2) Kamin blocking task:
    a) Participants would show the blocking effect by giving stimulus B a lower rating than stimulus D.
    b) Participants who score highly on positive schizotypy (O-LIFE unusual experiences) *and* those who score highly on negative schizotypy (O-LIFE introvertive anhedonia) would both exhibit attenuated blocking by giving stimulus B a higher rating than those who have low negative schizotypy.
    c) The extent of blocking (rating of D minus that of B) would correlate negatively with introvertive anhedonia scores.
    d) The extent of blocking (rating of D minus that of B) would correlate positively with the distress subscale of PDI-21.
3) Reversal learning task:
    a) The tendency to switch after probabilistic errors would correlate positively with delusional ideation (PDI-21) total scores.
    b) The same pattern of correlation would not be seen with perseveration (post true reversal accuracy) scores.
4) Action and word-based source monitoring tasks:
    a) O-LIFE unusual experience scores would negatively correlate with old/new recognition measures in both tasks.
    b) O-LIFE unusual experience scores would negatively correlate with overall source accuracy measures in both tasks.
    c) O-LIFE unusual experience scores would positively correlate with the number of internal source monitoring errors in both tasks.

d) O-LIFE unusual experience scores would positively correlate with the number of reality monitoring errors in both tasks.

In addition to frequentist analyses, Bayesian statistics were also employed to ascertain the levels of evidence supporting each of the hypotheses in a given direction (as opposed to significance testing) for the prediction error tasks only. Exploratory frequentist analyses were also carried out for both prediction error and source monitoring tasks, but inferences were not drawn due to the abovementioned issues with multiple comparisons which were not hypothesis-driven.

4.3. Study 2: Predictive Processing and Source Monitoring in Patients with Early Psychosis

4.3.1. Aims

The aims of Study 2 were to examine the associations between psychotic symptomatology in individuals with early (first episode) psychosis and any behavioural deficits compared with healthy controls in prediction error responses of the sensory and reward domain, and source monitoring of actions, which constituted a shortened battery of behavioural tasks from those used in Study 1.

For predictive processing, associations were examined between delusions, hallucinations, and general positive psychotic symptoms (measured by SOPS) and predictive processing as measured by force-matching and reversal learning tasks.

For source monitoring, associations were examined between delusions, hallucinations, and general positive psychotic symptoms (measured by SOPS) and old/new recognition, source memory accuracy, internal source monitoring errors and reality monitoring errors in the action-based task.

4.3.2. Hypotheses

The *a priori* hypotheses for Study 2 were as follows:

1) Force-matching task:
    a) All participants would consistently apply more force in the Finger condition (applying force directly by finger) than the Slider condition, creating an overcompensation score.

b) Overcompensation scores would be significantly lower in patients than in controls, and would negatively correlate with patients' delusional symptomatology scores.

2) Reversal learning task:
    a) Patients would be significantly less accurate than controls after both true reversal and probabilistic error trials.
    b) The tendency to switch after probabilistic errors would correlate positively with general positive symptom scores.
    c) The same pattern of correlation would not be seen with perseveration (post true reversal accuracy) scores in patients.
3) Action source monitoring task:
    a) Patients would make significantly more errors than controls in old/new recognition and source accuracy.
    b) More specifically, patients would make significantly more internal source monitoring and reality monitoring errors than controls.
    c) Patients would be significantly slower than controls when making source judgements.
    d) The number of errors in internal and reality monitoring would positively correlate with general positive symptom scores.

In addition to frequentist analyses, Bayesian statistics were also employed to ascertain the levels of evidence supporting each of the hypotheses in a given direction (as opposed to significance testing) for both prediction error tasks and source monitoring tasks. Exploratory frequentist analyses were also carried out for both sets of tasks, but inferences were not drawn due to the abovementioned issues with multiple comparisons which were not hypothesis-driven.

4.4. Author's Contributions to the Current Thesis

The author of the current Thesis was responsible for the initial conceptualisation of the project, study design, gaining ethical approval (both within the Cardiff School of Psychology and from NHS/R&D organisations in England and Wales), attending research governance training, recruitment of all participants for both studies, consenting of all participants, all behavioural testing, psychometric assessments and

clinical interviews, the entirety of data entry and data analyses and writing of two publications resulting from Study 1 as first author.

Chapter Five: Dimensions of Schizotypy in Relation to Different Types of Predictive Processing

Parts of this Chapter have been published the following paper:

Humpston, C. S., Evans, L. H., Teufel, C., Ihssen, N., & Linden, D. E. (2017). Evidence of absence: no relationship between behaviourally measured prediction error response and schizotypy. *Cognitive Neuropsychiatry*, *22*(5), 373-390.
DOI: 10.1080/13546805.2017.1348289.

5.1. Abstract

The predictive processing framework has attracted much interest in the field of schizophrenia research in recent years, with an increasing number of studies also carried out in healthy individuals with nonclinical psychosis-like experiences. The current research adopted a continuum approach to psychosis and aimed to investigate different types of prediction error responses in relation to psychometrically defined schizotypy. One hundred and two healthy volunteers underwent a battery of behavioural tasks including a) a force-matching task, b) a Kamin blocking task, and c) a reversal learning task together with three questionnaires measuring domains of schizotypy from different approaches. Neither frequentist nor Bayesian statistical methods supported the notion that alterations in prediction error responses were related to schizotypal traits in any of the three tasks. These null results suggest that deficits in predictive processing associated with clinical states of psychosis are not always present in healthy individuals with schizotypal traits.

5.2. Introduction

The predictive processing model has been adduced to explain the positive symptoms of schizophrenia (delusions and hallucinations; Fletcher and Frith, 2009). This framework proposes that sensory and cognitive experiences are not simply passive events but involve the active prediction of incoming information, with the purpose of

minimising prediction errors. A prediction error occurs when there is a mismatch or discrepancy between the expectation of an experience and the actual experience itself; it has been suggested that PEs are 'a general neural coding strategy' present in the whole brain which are involved in perception, cognition and motivational control (den Ouden et al., 2012). In the present study different aspects of predictive processing were tested, namely that in the perceptual/sensory and motivational/reward domains, in relation to the same individuals' psychometrically measured schizotypal traits. The reward domain was further divided into associative learning and probabilistic (reversal) learning.

The three prediction error-based tasks (force-matching, blocking, reversal learning) were chosen because they tapped into multiple aspects of the predictive processing framework as outlined in Chapter 2 above and could potentially elicit error signals in different domains. As detailed in Chapter 4, the general hypothesis was that participants scoring highly on schizotypy measures would exhibit deficits in prediction error responses across sensory, associative and reward domains. In the force-matching task, participants rating highly on delusional ideation would exhibit reduced sensory attenuation as indexed by an overcompensation score; in the blocking task, the blocking effect would be attenuated in individuals high in general negative schizotypy; and in the reversal learning task, individuals scoring highly on delusional ideation would display increased switching tendency.

5.3.Methods

5.3.1. Task Piloting and Power Calculation

The behavioural battery of five tasks (three prediction error tasks and two source-monitoring tasks; see Chapter 5 for details of the latter) were piloted in eight female psychology undergraduates recruited from the School of Psychology online Experimental Management System in order to determine the feasibility and optimal timing of tasks, levels of performance (whether the task showed the phenomena they were designed to measure) and time intervals between the study and the test phases of source-monitoring tasks. Five participants completed all behavioural tasks whilst the first three participants only completed the source-monitoring tasks. A fixed order of tests was adopted in preference to a counterbalanced order purely due to logistic

reasons (conflicts in the timing and booking of equipment and laboratory space). All participants in the piloting phase received appropriate course credits according to the School of Psychology guidelines.

Power calculations were carried out in GPower 3.1 to determine a suitable sample size. In the sensory domain, previous work examining schizotypy by Teufel et al. (2010), Lemaitre et al. (2016) and Palmer et al. (2016) indicate effect sizes ranging from r = -0.35 to r = -0.58 (negative correlations between schizotypy and diminished sensory attenuation). Given an alpha level of .05 and a power level of 0.90, this gives an estimated sample size of 23 – 78 (two-tailed correlations). For the blocking task, previous studies by Haselgrove and Evans (2010) and Moran et al. (2003) have estimated effect sizes from Cohen's f = 0.30 to 0.39, giving a sample size of 61 - 109. In the reward domain due to a lack of studies examining schizotypy, an effect size of Cohen's f = 0.59 have been generated from schizophrenia patient datasets (Schlagenhauf et al., 2014) yielding a sample size of 22. However, it should be noted that the effects in schizotypy would be anticipated to be smaller and hence a larger sample size would be necessary. In other to maximise power, it was decided to recruit up to 120 participants (greater than the highest number estimated).

5.3.2. Participants

One hundred and fifteen healthy volunteers (25 males; mean age 22.23 years, standard deviation 3.18 years, range 18-33 years) were recruited through the Experimental Management System (EMS) as well as the electronic Noticeboard from different departments across Cardiff University (Ethical approval number: EC.14.12.09.4044R2A). Inclusion criteria were that participants must be aged between 18-35 years, have normal or corrected to normal vision and hearing and possess a high level of fluency in English. Exclusion criteria were a current diagnosis of any psychiatric illness, taking psychotropic medication, experiences of current illicit substance abuse or having mobility problems. Fifteen participants identified themselves as left-handed. Eligibility relied on self-report. All participants gave informed consent in written form and were fully debriefed after the experimental session. The majority of participants also gave written consent to potentially being contacted for future studies. All participants received 8 course credits (for those

recruited from EMS) or a single sum of £15 (for those recruited from Noticeboard) after the session as reimbursement for their time.

Of the 115 participants, all but one (who did not complete the O-LIFE) participant completed all three schizotypy measures. Thirteen participants were excluded from the current study because they failed to meet the inclusion criteria for one or more of the behavioural tasks (see below for the specific criteria and the number excluded for each task). The final 102 participants consisted of 21 males and 81 females with a mean age of 21.96 (SD = 3.14) years. Assuming the smallest effect size of 0.30, this has yielded an achieved power of 0.88.

5.3.3. Procedure

Each experimental session took 2 hours in total. Participants were tested individually and completed the tasks and questionnaires in the following order: study phase of action source-monitoring; blocking; study phase of verbal source-monitoring; all three questionnaires; reversal learning; test phase of verbal-source-monitoring; force-matching; test phase of action source-monitoring.

5.3.4. Force-matching Task

This procedure (Teufel et al., 2010) focused on the perceptual/sensory type of prediction error. Participants were asked to place their left index finger under a lever attached to a torque motor, which then applied four different levels of forces (1.0N, 1.5N, 2.0N and 2.5N) in a random order. Participants were asked to match the presented force in two conditions, which were always counterbalanced across participants. In the 'Finger' condition, participants matched the force by directly pressing down onto the tip of the lever using their right index finger in order to exert the same (perceived) level of force on their left index finger. In the 'Slider' condition, participants matched the force indirectly by moving a linear potentiometer up and down which controlled the torque motor. The gain of the slider was 0.5N/cm.

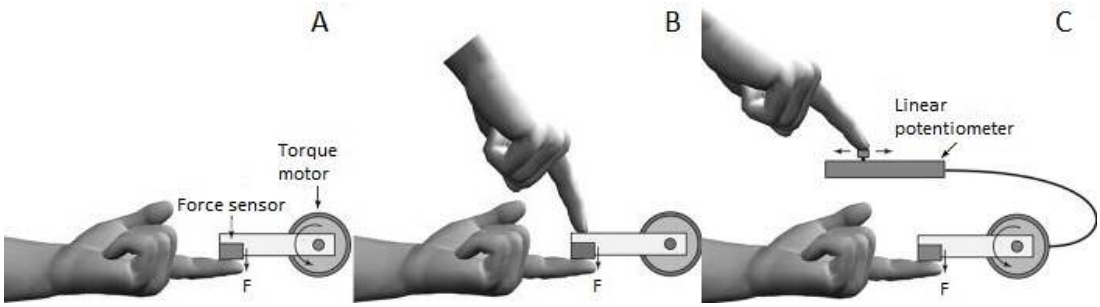

Figure 5.1. The force-matching task. A) A force is applied on the participant's left index finger; B) Participant matches the force in the Finger condition; C) Participant matches the force in the Slider condition. Adapted from Teufel et al. (2010).

Participants were predicted to exert more force (hence less accurate matching/overcompensation of the applied force) than the actual applied force in the Finger condition due to sensory attenuation (i.e. efference copy of motor command matching the outcome of motor act) so that the sensation of the force was perceived as weaker. The Slider condition by contrast involved an unusual relationship between the action and its sensory consequences and as such, participants' performance were expected to be more accurate with much less or no overcompensation of the applied force. Participants received training of the task in the form of a practice session (8 trials) of both conditions before progressing to testing sessions of 32 trials each. Five outliers (differences in applied forces deviating two standard deviations from the mean) were removed in the final analysis according to the same criteria by Teufel et al. (2010).

5.3.5. Kamin Blocking Task

This task focused on the associative type of prediction error and used the same paradigm as that by Haselgrove and Evans (2010). Participants were asked to play the role of a hospital inspector and evaluate whether certain food items and pairings of food items caused food poisoning by putting in numbers with the keyboard between 1 (completely safe to eat) and 9 (completely dangerous) with number 5 as being uncertain. Then they would click on the 'Results' button to find out whether the food item(s) actually led to food poisoning. The task was programmed in VisualBasic and ran under Microsoft Windows XP. There were 12 types of food items (e.g. salmon, lamb, peas, potatoes, etc.) which were randomly assigned as cues A to L. Stage 1 contained 10 trials whereas Stage 2 (which followed from Stage 1 with no indication of a break) had 5 trials, both in a block randomised manner. In the test stage

participants were asked to give their final ratings of cues B, D, F and K without any feedback. If there is blocking present the participants' ratings of B would be smaller than those for D or F; in other words, blocking occurs because the associative strength for B from compound AB+ is attenuated due to prior association with stimulus A+. Table 5.1 below details the task design. Data from nine participants were excluded due to a failure to learn or failure to respond with appropriate keys (remaining $N = 106$). The blocking effect was denoted as a final rating of D minus B.

Stage 1	**Stage 2**	**Test**
A+	AB+	B
	CD+	D
E-	EF+	F
	K+	K
GH+	L-	
IJ-	IJ-	

Table 5.1. Design of Kamin blocking task. Cues A to L indicate each food item, either associated with the outcome of food poisoning (+) or not (-). GH+. L- and IJ- are filler trials.

5.3.6. Reversal Learning Task

This task aimed to tap into the reward/motivational type of prediction error and used Lancaster et al. (2015) and Ihssen et al. (2016)'s baseline reversal learning paradigm (i.e. the private condition). Participants were asked to choose between two coloured squares, blue and green, which were displayed on the same screen side by side. Participants earned 1 penny (reward; +1p) if they chose the correct colour and lost 1 penny (punishment; -1p) if they chose the wrong colour. Participants were informed that they would not lose any money if their total was negative at the end and their final payment would remain the same (£15) due to ethical considerations. The experiment was programmed in Presentation and ran under Microsoft Windows 8 on a laptop. A practice block was offered prior to the actual experimental session.

At the beginning the colour blue was set to be the correct colour; however, after a variable number (between 7 and 15) of trials the reward/punishment contingencies (0.8/0.2) were reversed (true reversal) so that blue became the wrong colour and was punished, whereas green became the corrected colour and was rewarded. Feedback was given immediately after each choice in the form of a smiley face (reward) or a sad face (punishment). Probabilistic errors were also included between two true reversal trials, whose numbers were again variable (between 1 and 3). Such errors meant that participants were unexpectedly punished even though they chose the correct colour (i.e. 'wrong feedback'). Participants were told that only one colour would be correct at one time and were aware of the existence of true reversals as well as probabilistic error trials, but did not know when they would occur. The task contained 132 trials with an average of 11 true reversals in total. Fig. 5.2 below shows a diagrammatical representation of the task design. No participants were excluded on the basis of performance on this task.

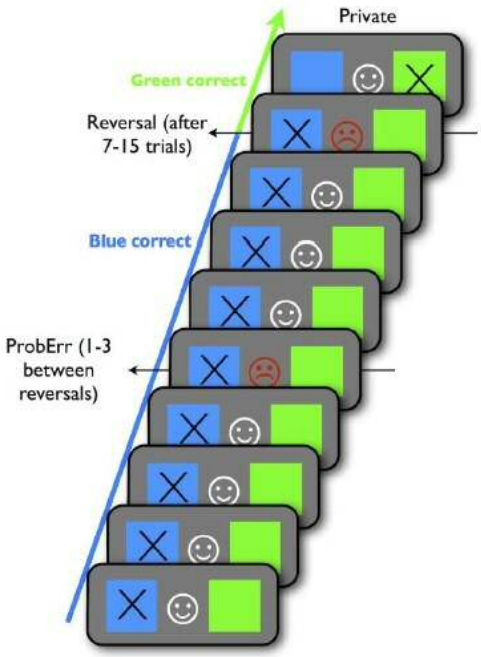

Figure 5.2. Design of reversal learning task showing one reversal episode. X refers to choices made by a hypothetical participant. Adapted from Ihssen et al. (2016). ProbErr, Probabilistic errors.

5.3.7 Questionnaires

Participants were asked to complete three questionnaires on the different dimensions of schizotypy: the 21-item Peters et al. Delusions Inventory (PDI-21; Peters et al.,

2004), the 34-item Cardiff Anomalous Perceptions Scale (CAPS; Bell et al., 2006) and the Oxford-Liverpool Inventory of Feelings and Experiences (Mason et al., 1995). PDI-21 and CAPS were in pen and paper format whereas the O-LIFE was computerised running under Microsoft Windows XP on a laptop.

The PDI-21 contains distress, preoccupation and conviction subscales and was the closest to a clinical scale for delusional ideation (in other words, akin to schizophrenic psychopathology). The O-LIFE on the other hand measured different dimensions of schizotypy with a much stronger focus on the individual-difference/personality trait approach rather than a pathological one, and was therefore chosen over the Schizotypal Personality Questionnaire (SPQ) which was based on DSM-III criteria for schizotypal personality disorder. The O-LIFE contained four subscales: unusual experiences (positive domain), cognitive disorganisation (disorganised domain), introvertive anhedonia (negative domain) and impulsive nonconformity (behavioural domain). The three questionnaires used in the present study tapped into of schizotypal traits from different approaches, from the most benign and everyday experiences as purely a personality construct (as measured by O-LIFE) to mild sensory distortions and hallucinatory phenomena (as measured by CAPS) to the more clinical delusion-like beliefs (as measured by PDI-21). Based on findings from previous studies, these scales have been examined in relation to the various types of prediction error responses and so were included in this study in order to fully replicate previous study procedures.

5.4. Data Analysis

5.4.1. Checks for Normality and Reliability

Five percent of the questionnaire scores in pen and paper format (PDI-21 and CAPS) were re-scored by another rater in order to ensure there were no systematic errors in scoring. All three schizotypy questionnaires including the fully computerised O-LIFE and their subscales were also subjected to a split-half reliability test, yielding a Spearman-Brown coefficient of 0.940. This indicates that the psychometric scores obtained in this experiment were highly reliable.

Both questionnaire scores and task performance data were checked by Shapiro-Wilk tests for normal distribution. All schizotypy questionnaire scores except O-LIFE

impulsive nonconformity (which was not of interest in the current study) subscale yielded statistically significant results (test statistics above 0.900, $p < 0.001$), which indicated that they did not follow a normal distribution. Q-Q plots further confirmed this observation. Log transform was considered but rejected due to the fact that some of the scores were zeroes. Consequently, nonparametric correlations (Spearman's rho) were chosen as the main method for correlational analyses.

5.4.2. Analysis of Behavioural Data

A parallel analysis strategy was employed in which both frequentist (Null Hypothesis Significance Testing) and Bayesian approaches were used; Bayes factors were calculated in order to explore the strength of evidence or the confidence with which the null hypotheses are supported. It has also been suggested that Bayesian approaches are resistant to multiple comparison problems (Dienes, 2011). All frequentist data analyses were carried out using SPSS 23 (IBM Corp.) and all correlations were two-tailed; all Bayesian analyses (Bayesian Correlation Pairs) were carried out in JASP Version 0.8.0.0 (https://jasp-stats.org/).

Measures or proxy measures of prediction error-based behavioural responses are as follows: in the force-matching task, an overcompensation score was calculated for each participant by subtracting the mean difference between active (applied by the participant) and passive (original force applied by the machinery) forces in the Slider condition from that in the Finger condition. In the Kamin blocking task, the extent of blocking was calculated by the final rating for cue D minus the final rating for cue B. Data (i.e. participants' ratings) for each learning stage are plotted as line graphs to ensure that learning occurred. Total accuracy, post-probabilistic error accuracy (an index of switching or 'switchiness') and post-true reversal accuracy (an index of perseveration) were entered in the analysis as dependent variables for the reversal learning task.

For all three tasks, the main outcome measures were correlated with corresponding schizotypy scales (the same as those used in frequentist statistics) by using a Bayesian Correlation Pairs analysis in JASP; for the force-matching task, this was the overcompensation score and the grand total score of PDI-21; for Kamin blocking, this was the blocking score and the unusual experiences score of O-LIFE; and for the reversal learning task the correlation was done between switching tendency

and PDI-21 total score. Bayesian factors in the form of BF_{01} (null over alternative) were calculated from *a prori* hypotheses regarding the direction of the correlation together with robustness checks to reflect the strength of evidence. For the force-matching and Kamin blocking tasks, the direction of the correlations was hypothesised to be negative whereas for reversal learning, the direction of the correlation was hypothesised to be positive. Beta* (stretched beta) prior width for these correlations was set to a relatively conservative value of 0.5.

5.5. Results

5.5.1. Schizotypy Questionnaire Scores

Descriptive data for the three scales are shown in Table 5.2 below. The mean total scores of PDI-21 and CAPS were consistent with those from previous research (approximately 6 and 8, respectively) by Peters et al. (2004) and Bell et al. (2006); whereas scores for the O-LIFE unusual experiences and introvertive anhedonia subscale were lower than the expected mean reported (7.1/4.8 in the current sample versus 8.8/6.4) by Mason and Claridge (2006).

	Mean (SD)	Range	Normative Mean (SD)
PDI-21 Total Y/N	5.88 (3.47)	0 – 16	6.7 (4.4)
PDI-21 Distress	15.95 (12.01)	0 – 51	15.5 (14.1)
PDI-21 Preoccupation	14.84 (11.45)	0 – 57	15.4 (14.1)
PDI-21 Conviction	17.92 (11.79)	0 – 52	20.4 (16.0)
CAPS Total Y/N	8.29 (6.03)	0 – 22	7.3 (5.8)
CAPS Distress	20.92 (18.13)	0 – 84	15.5 (14.5)
CAPS Intrusiveness	22.43 (19.09)	0 – 92	18.0 (17.0)
CAPS Frequency	17.65 (15.72)	0 - 79	14.6 (14.2)
O-LIFE UnExp	7.14 (5.44)	0 – 25	8.82 (6.16)
O-LIFE IntAn	4.79 (4.31)	0 – 22	6.38 (4.49)

Table 5.2. Descriptive data for schizotypy scales and their subscales ($N = 102$). SD, standard deviation; PDI-21, 21-item Peters et al. Delusions Inventory; CAPS, Cardiff Anomalous Perceptions Scale; O-LIFE, Oxford-Liverpool Inventory of Feelings and Experiences; UnExp, Unusual experiences; IntAn, Introvertive anhedonia.

5.5.2. Force-matching

Participants consistently applied more force in the Finger than in the Slider condition, demonstrating the overcompensation effect (Fig. 5.3). A paired-sample t-test showed that the mean difference between active and passive forces applied in the Finger condition was significantly greater than that in the Slider condition [$t\,(101) = 13.26$, $p < .001$), Cohen's d = 1.53].

Figure 5.3. Comparisons between mean applied forces in the Finger and Slider conditions.

In terms of the relationship between the overcompensation score and delusional ideation (as measured by PDI-21 total score), a two-tailed Spearman's correlation was calculated. There was a non-significant correlation between these two variables [ρ(100) = 0.139, p = 0.163]. Table 5.3 shows exploratory correlations with other schizotypy scales.

5.5.3. Kamin Blocking

When correlated with O-LIFE unusual experience and introvertive anhedonia (positive and negative schizotypy, respectively) scores, there was no significant correlation between these variables ρ(100) = 0.028, p = 0.782 for the positive dimension and ρ(100) = -0.106, p = 0.290 for the negative dimension). In order to replicate the methodology used by Haselgrove and Evans (2010), median splits of O-LIFE unusual experiences and introvertive anhedonia scores were computed which divided the participants into high and low positive/negative schizotypy groups. The median value for unusual experiences was 6, with scores equal to these values included in the 'low' group whereas that for introvertive anhedonia was 4. Fig. 5.4 visualises the mean ratings for each stimulus in the testing stage for low and high positive (A) and negative (B) schizotypy groups. There was no significant effect of group for either positive [$F(1, 416) = 3.544$, p = 0.680, partial eta squared = 0.008] or negative [$F(1, 416) = 15.975$, p = 0.078, partial eta squared = 0.037] schizotypy. A Spearman's rho correlation was calculated between the extent of blocking and the distress subscale of

PDI-21, as previous research (Corlett and Fletcher, 2012) indicated a specific relationship; no significant relationship was found [ρ(100) = .130, p = .191].

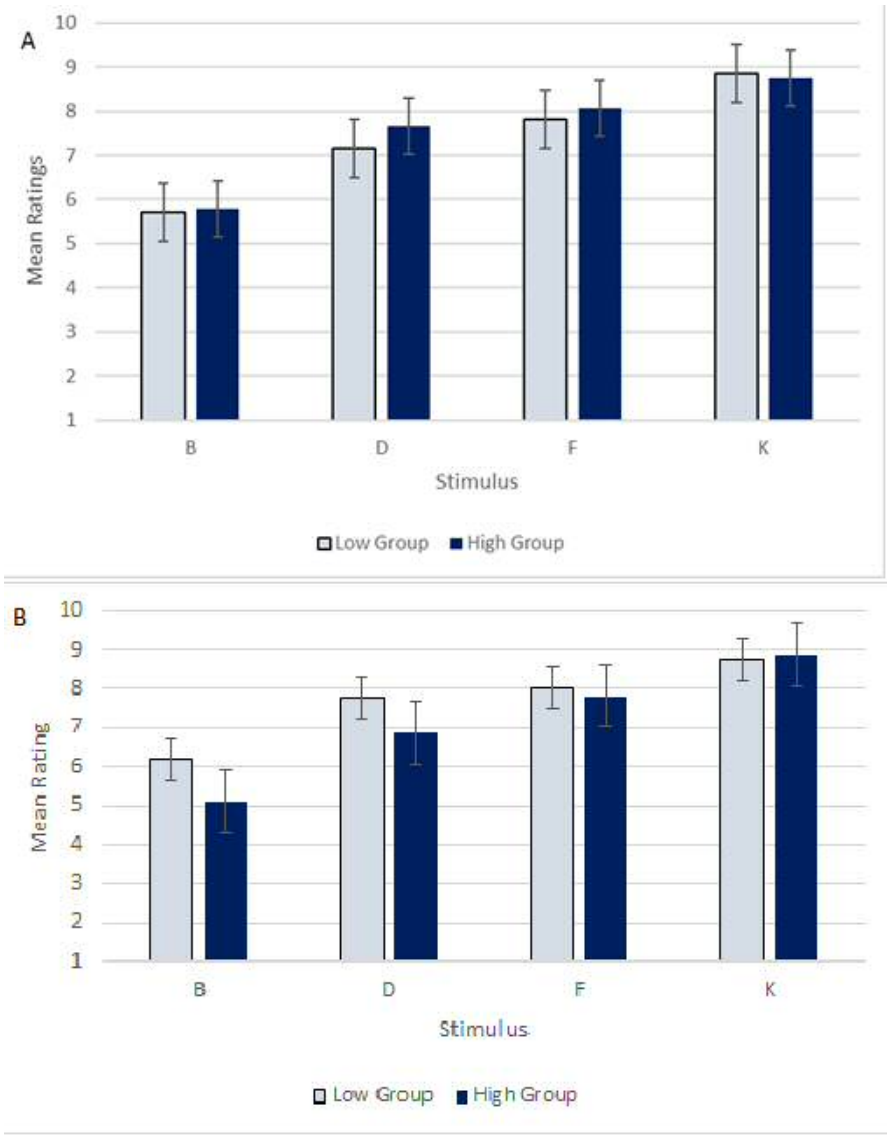

Figure 5.4 (A-B). Low and high group ratings for each stimulus in the test stage for O-LIFE unusual experiences (A) and introvertive anhedonia (B).

Further still, the learning of stimulus-outcome associations in Stages 1 and 2 were analysed to determine whether there were differences between individuals high and low in positive and negative schizotypy. There were no differences in the learning acquisition between these two groups in unusual experiences (Fig. 5.5). For Stage 1, a two-way analysis of variance (ANOVA) with factors of group (high versus low unusual experiences) and stimulus (A+ and E-) and mean ratings as dependent variable yielded a highly significant effect of stimulus [$F(1, 200) = 4013.03$, $p < 0.001$, partial

eta squared = 0.953], but no significant effect of group [$F(1, 200) = 0.715$, $p = 0.399$, partial eta squared = 0.004] or group*stimulus interaction [$F(1, 200) = 0.372$, $p = 0.543$, partial eta squared = 0.004]. An identical ANOVA carried out with stimuli GH+ and IJ- also revealed a highly significant effect of stimulus [$F(1, 200) = 1980.24$, $p < 0.001$, , partial eta squared = 0.908], no significant effect of group [$F(1, 200) = 0.688$, $p = 0.408$, partial eta squared = 0.003] but a significant group*stimulus interaction [$F(1, 208) = 4.398$, $p = 0.037$, partial eta squared = 0.021].

For Stage 2, a two-way ANOVA performed with factors of group (high/low unusual experiences) and stimulus (AB+, CD+) and mean ratings as dependent variable did yield a weak but significant effect of group [$F(1, 200) = 4.668$, $p = 0.032$, partial eta squared = 0.023], a highly significant effect of stimulus [$F(1, 200) = 41.904$, $p < 0.001$, partial eta squared = 0.173], but no significant interaction [$F(1, 200) = 0.012$, $p = 0.911$, partial eta squared = 0.00006]. An identical ANOVA carried out with stimuli EF+, K+ and IJ- also revealed a highly significant effect of stimulus [$F(2, 300) = 968.28$, $p < 0.001$, partial eta squared = 0.433], no significant effect of group [$F(2, 300) = 0.426$, $p = 0.514$, partial eta squared = 0.001] and no significant group*stimulus interaction [$F(2, 312) = 1.765$, $p = 0.173$, partial eta squared = 0.006]. There was a similar pattern of results when identical ANOVAs were performed with data split by negative schizotypy scores (as measured by introvertive anhedonia, Fig. 4.6): only stimulus type yielded significant effects whereas group status did not.

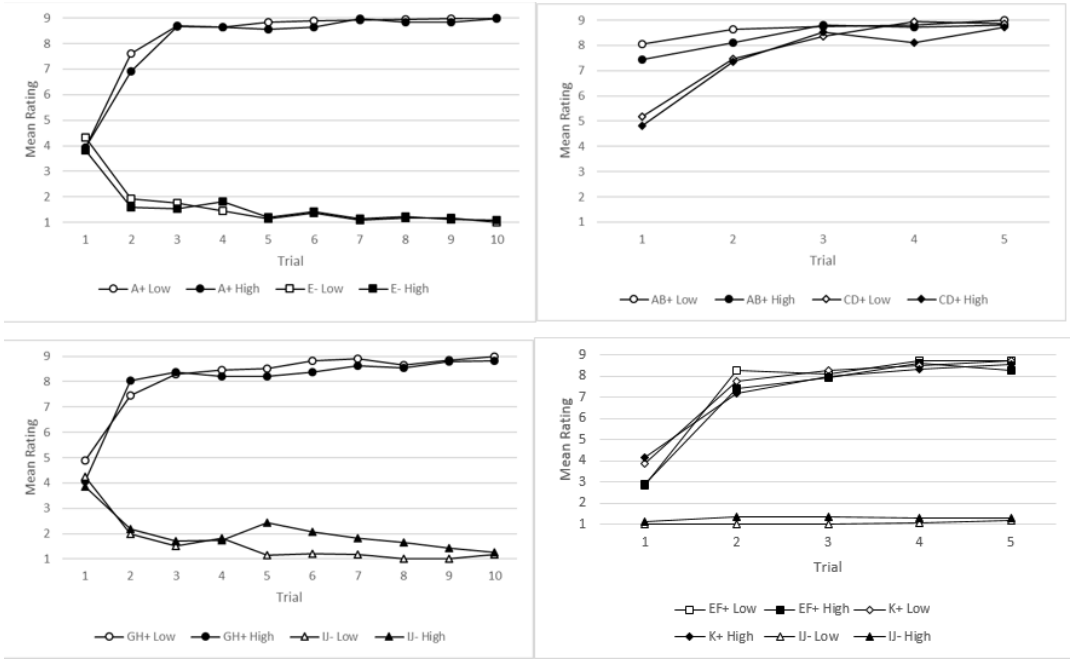

Figure 5.5. Low and high group ratings for unusual experiences across learning stages. + and – refer to the presence or the absence of the outcome, respectively.

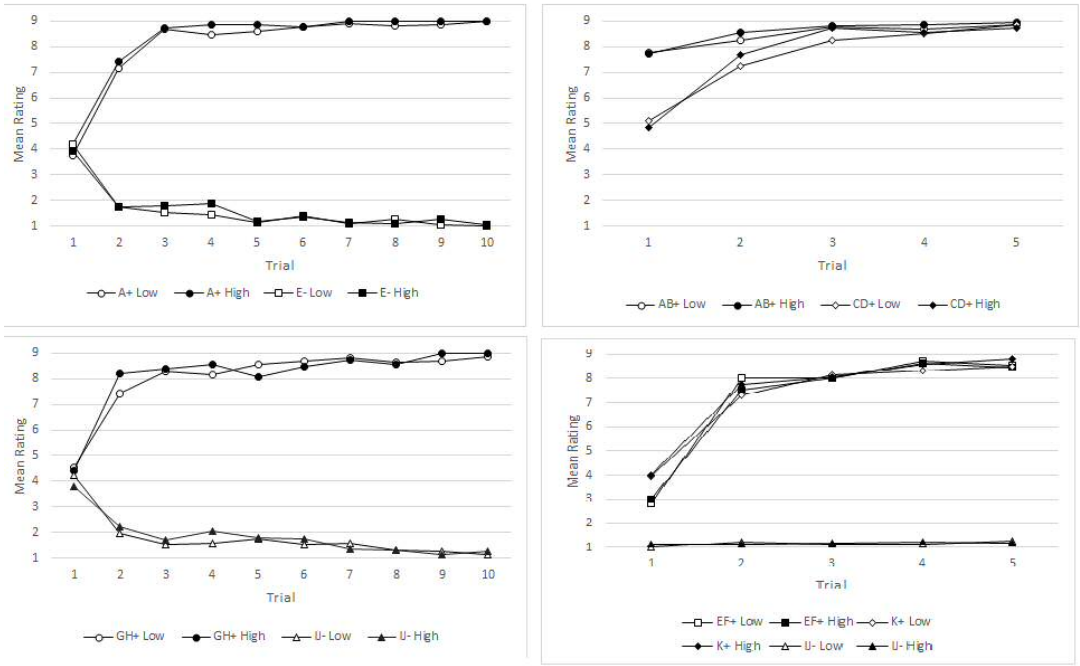

Figure 5.6. Low and high group ratings for introvertive anhedonia across learning stages. + and – refer to the presence or the absence of the outcome, respectively.

5.5.4. Reversal Learning

Fig. 5.7 shows mean accuracy data for trials surrounding true reversals and probabilistic errors; the latter was further divided into first and late (second/third)

probabilistic errors, as participants tend to display more switching in the second than in the third errors. Accuracy was greatly reduced at reversal trials from 90% to below 10% and then recovered within two trials to the pre-reversal level. Trials after late probabilistic errors demonstrated a lower accuracy than those after the first error (30% versus 40%). It required at least two further trials to restore task performance back to ceiling level in both situations. This pattern of results is compatible with other studies employing this and similar reversal learning paradigms (Ihssen et al., 2016). Switching and perseveration scores for each participant were calculated as the inverse of post-probabilistic error and post-reversal accuracies. In a subsequent correlational analysis switching score was not significantly correlated with delusional ideation as measured by PDI-21 total scores [$\rho(100) = 0.008$, $p = 0.937$].

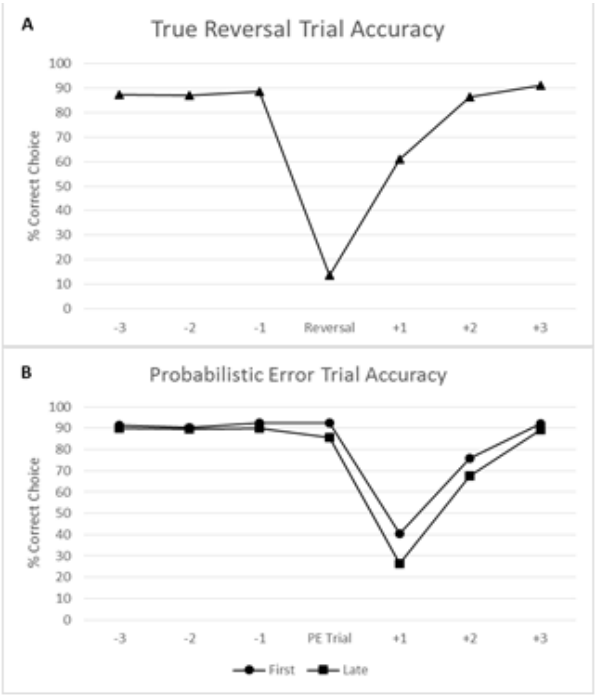

Figure 5.7 (A-B). Accuracies of true reversal trials (A) and probabilistic error trials (B).

Additional bivariate relationships in the form of a correlation matrix can be found in Table 5.3 for reference. As can be seen none of these relationships reached conventional levels of significance ($p < 0.05$) even without the application of a correlation for multiple comparisons.

	PDI-21Tot	CAPSTot	OLIFEUnExp	OLIFEIntAn
Force-matching Overcompensation	.139	.074	.127	.100
Blocking Score	.108	.130	.028	-.196
Post Reversal Perseveration	.025	-.113	-.075	.013
Post Probabilistic Error Switching	.008	.063	.089	.046

Table 5.3. Nonparametric bivariate correlation coefficients (Spearman's rho, two-tailed) between schizotypy measures and task measures (N = 102). PDI-21, 21-item Peters et al. Delusions Inventory; CAPS, Cardiff Anomalous Perceptions Scale; Tot, Total yes/no endorsements; O-LIFE, Oxford-Liverpool Inventory of Feelings and Experiences; UnExp, Unusual experiences; IntAn, Introvertive anhedonia.

5.5.5. Bayesian Analysis: Evidence of Absence

Results from Bayesian analyses are presented in Fig. 5.8. For the force-matching task, BF_{01} was estimated to be 19.623, meaning that the data provided were highly in favour of the null hypothesis (19 times the likelihood of the alternative hypothesis, in this case a negative correlation between overcompensation and PDI-21 total scores) with strong to very strong evidence, meaning that there was a significant amount of support for no effect.

For the Kamin blocking task where BF_{01} was estimated to be 11.434 for the positive dimension, which also meant that the data was in favour of the null hypothesis (10 times the likelihood of the alternative hypothesis, in this case a negative correlation between blocking and O-LIFE unusual experiences scores). BF_{01} was estimated to be 5.092 for the negative dimension, which also meant that the data provided support in favour of the null hypothesis (5 times the likelihood of the alternative hypothesis, in this case a negative correlation between blocking and O-LIFE introvertive anhedonia scores, graphs not included in Figure). Robustness checks demonstrated a moderate to strong level of evidence favouring the null hypothesis for both correlations.

For switching tendency of the reversal learning task, BF_{01} was estimated to be 11.083 which meant that the data was favouring the null hypothesis 11 times; in other

words, the alternative hypothesis was highly improbable. Robustness checks demonstrated a strong level of evidence for the null hypothesis. In terms of the correlation between perseveration and PDI-21 total score, BF_{01} values were estimated to be 8.031, which favoured the null hypothesis with strong to very strong levels of evidence (graphs not included in Fig. 5.8).

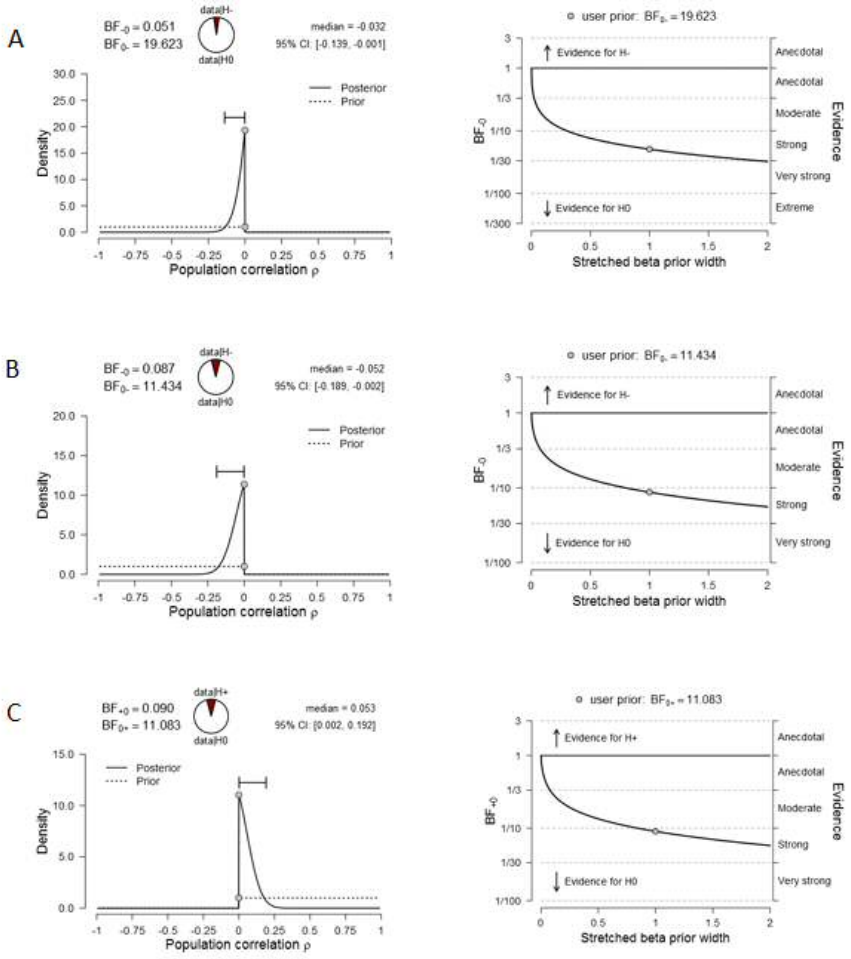

Figure 5.8 (A-C). Results from Bayesian correlation pairs analyses. Panels A, B and C show results for the force-matching, Kamin blocking and reversal learning tasks respectively. CI, Credibility Interval; BF, Bayes Factor.

5.6. Discussion

The current study investigated the relationships between different types of predictive processing and domains of psychometrically defined schizotypy in the same

individuals. There was little evidence for disrupted sensory predictive processing (as indexed by the force-matching task) in individuals with high scores of delusional ideation. Moreover, there was no significant difference in associative learning (as indexed by the blocking effect) between individuals with high and low positive or negative schizotypy or distress caused by delusion-like beliefs (as found by Corlett and Fletcher, 2012). Also, there was no evidence for alterations in switching tendency or perseveration as indexed by the reversal learning task in individuals with higher levels of delusional ideation. Importantly, the present study failed to demonstrate the same pattern of findings from previous studies which separately investigated sensory prediction, blocking phenomenon and reversal learning in relation to domains of schizotypy.

In the force-matching task, participants significantly overcompensated in the finger condition, which demonstrates the classic force-matching effect, which has been found in previous studies reviewed in Chapter 2. However sensory attenuation was not impaired in individuals with high delusional ideation. The use of PDI-21 rather than the PDI-40 may raise some concerns about the omission of items capturing delusions of control or passivity-like experiences which, by definition, have higher relevance with sensory prediction than other delusions such as paranoia. This is supported by the observation that in addition to positive schizotypy in general, Lemaitre et al. (2016) also found a significant negative correlation specifically between passivity-like experiences and the index of sensory attenuation. However, given that previous studies examining force-matching, such as those by Teufel et al. (2010) and Palmer et al (2016), used the PDI-21 and not the PDI-40, this does not explain the failure to observe this relationship in the current study. In addition, the PDI-21 was derived from the 40-item version with very similar psychometric properties (Peters et al., 2004). One methodological detail which differs between the current study and that of Teufel et al. (2010) is that in the latter study more repetitions were used to average applied and presented force (eight rather than four levels of forces). Therefore, it may be the case that the measurements were somewhat noisier in the current study because of the necessity of reducing the length of tasks to accommodate for the overall duration of testing (2 hours).

In the associative learning task, the blocking task utilised was exactly the same as that used by Haselgrove and Evans (2010). In contrast to their study, there was a

failure to find any relationships with the negative dimension of schizotypy, even when following the same analytic methods used in that study (e.g. carrying out a median split with the same median). Given it is a well-powered study it could be that the failure to find this relationship might have been affected by other factors such as smoking status which was not measured in the current study. For example, nicotine has been shown to reduce dopamine release (Zhang and Sulzer, 2004) and may attenuate the prediction error responses mediated by dopamine. Furthermore, there were no significant relationships between blocking and any other schizotypy dimensions, such as the positive dimension as previously found by Moran et al. (2003), the total PDI score (as found by Moore et al., 2011) or the distress aspect of delusional ideation (as found by Corlett and Fletcher, 2012). For these correlations the same schizotypy measures as those in previous studies were used but the blocking task and measure of this phenomenon were different. For example, Corlett and colleagues used computer-paced tasks whereas a self-paced task was used in the present study, and the former group did not use behavioural measures for blocking unlike in this study. There is some debate about whether prediction error as a latent process in associative learning is best studied by neuroimaging or behavioural methods, or perhaps a combination of both (see Griffiths, Langdon, Le Pelley, and Coltheart, 2014; Corlett and Fletcher, 2015).

In the reversal learning task, switching tendency was used as an index of reward sensitivity driven by prediction error-related learning and found no significant associations between an increased tendency to switch after probabilistic errors and delusional ideation in either frequentist or Bayesian statistical analyses. In fact, accurate responding was restored very soon after both true reversals and probabilistic errors, suggesting that participants performed the task effectively and learnt when to switch or stay relatively quickly. These findings are clearly in contrast with findings in schizophrenia (e.g. Schlagenhauf et al., 2014), but due to a lack of studies using reversal learning in healthy schizotypy, comparisons can only be made with other set-shifting tasks in individuals prone to psychosis-like experiences (e.g. Cella et al., 2009) which once again do not support current findings. Also, the reward/punishment manipulation (1p) was perhaps too weak, as participants were told they would not win or lose real money. However, this theoretically would have increased randomness in

switching, but this was not shown to be the case in the restoration of post-reversal and post-probabilistic error accuracies.

The current hypotheses focused on delusion-proneness and there were no significant correlations between hallucination measures (i.e. CAPS) and behavioural performance in the current study. Hallucinations have been recently linked with predictive coding (e.g. Horga, Schatz, Abi-Dargham, and Peterson, 2014) in established schizophrenia; however, in nonclinical groups hallucinations can also persist without causing distress or leading to a need for psychiatric care (Linden et al., 2010; Hill, Varese, Jackson, and Linden, 2012; Johns et al., 2014) in many high-functioning individuals.

In this sample, participants were all functioning relatively highly. In fact, although there were some individuals who endorsed the more 'bizarre' items such as thought echo in the schizotypy questionnaires, these were a very small minority of participants. The majority of schizotypy scores in the current sample were positively skewed towards 'normal experience' even though the means of these scores were comparable to those from previous general population studies of schizotypal traits.

However, it is also possible that there was potential disconnection between subjective experiences of schizotypy and objective measures of neurocognitive deficits, in which the subjective complaints from psychometrically-measured schizotypy do not match the magnitude of deficits seen in behavioural tasks (e.g. Chun, Minor and Cohen, 2013). This could have led to the observation that individuals rating highly in subjective reports of schizotypal (psychosis-like) experiences did not display significant deficits in behaviours supposedly affected by such experiences.

Cross-sectional studies of this kind are unable to establish causal relationships. A possibility for future research would thus be a longitudinal study with structured assessments at regular intervals in order to determine the persistence of psychosis-like experiences and any rate of transition to clinical disorders, as well as incorporating a range of methods for measuring prediction error responses (e.g. combining imaging with behavioural testing).

The present study may also have been affected by a selection bias where only participants with certain traits and interests were 'attracted' to research or motivated to take part in the study (see Martin et al., 2016, who found significant relationships between non-participation and individuals' risk factors for schizophrenia) which

would further reduce the generalisability of these findings. However, this factor would similarly apply to previous studies on this topic.

In sum, although much caution needs to be taken when interpreting the results, the present study furthers our understanding of the construct of schizotypy by employing an integrative approach to predictive processing in relation to different domains of schizotypal traits in a large sample of high-functioning individuals with no past or present psychiatric diagnosis. These null findings suggest that predictive processing mechanisms, at least in the forms of sensory, associative and reward prediction error responses, are *not* always associated with positive schizotypal personality traits in the general population.

Chapter Six: Source-Monitoring for Action and Speech in Healthy Individuals with Schizotypal Personality Traits

Parts of this Chapter have been published the following paper:

Humpston, C. S., Linden, D. E., & Evans, L. H. (2017). Deficits in reality and internal source monitoring of actions are associated with the positive dimension of schizotypy. *Psychiatry Research*, *250*, 44-49.
DOI: 10.1016/j.psychres.2017.01.063.

6.1 Abstract

Adopting a continuum approach to psychosis, the aim of the current study was to assess the relation between schizotypy and source memory of actions and word pairs in healthy volunteers. One hundred and two participants completed two source memory tasks: one which involved the completion of well-known word pairs (e.g. fish and ?) and an action based task (e.g. nod your head). At test participants needed to indicate whether the act had been performed or imagined by themselves, performed by the experimenter, or was new. The positive dimension of schizotypy was positively correlated with source errors in both reality monitoring and internal source monitoring. Individuals with high ratings of unusual experiences attributed self-performed actions and imaginations to the experimenter (reality monitoring errors), as well as confusing self-performed actions with their own imaginations (internal source monitoring errors). However, these relationships were not found in the word pair task. Such findings suggest a degree of specificity for the source monitoring of actions in a schizotypal population and may have implications for the study of clinical symptoms such as delusions of control.

6.2 Introduction

The aim of the current study was to provide a more detailed and integrated understanding of source memory and its relationship to schizotypy in a large sample of healthy volunteers. The first issue was whether individuals high in schizotypy

would display deficits in familiarity (i.e. knowing the event has previously occurred). On the basis of the review by Libby et al. (2013) it would be anticipated that a deficit in discriminating old from new items would be seen in those high in schizotypal traits. However, research findings on this issue have been mixed: Peters et al. (2007) found evidence for a deficit, whereas Collignon et al. (2005) did not. Next, source memory was investigated by assessing in the same participants reality monitoring (discriminating between internally and externally generated stimuli) and internal source monitoring (discriminating between two internally generated stimuli, such as one's own speech and imagination). Previous work reported only deficits in internal source monitoring but not in reality monitoring (Collignon et al., 2005). This is surprising given the wealth of work highlighting problems in reality monitoring in schizophrenia (Johns et al., 2001; Keefe et al., 2002; Vinogradov et al., 2008). Therefore, the overall aim was to examine whether individuals scoring high on the positive dimension of schizotypy (psychosis-like experiences) would have a deficit in both of these types of memory.

It has been argued by some researchers that the generalisability of word based paradigms to real-world situations is limited (Henquet et al., 2005; Parks, 1997) and that action based tasks might be a more naturalistic method of examining source memory. However, no study has given participants these two types of tasks and assessed whether they both lead to the same findings. Therefore in this study participants completed two source memory tasks: one where a word needed to be generated (e.g. fish and ?) and an action based task (e.g. nod your head). In both tasks, participants needed to indicate at test whether the action was i) performed, ii) imagined, iii) performed by the experimenter, or iv) was new. It was hypothesised that source memory deficits would be related to the positive dimension of schizotypy and so focused primarily on this dimension, due to the findings of previous studies in this area (e.g. Brébion et al., 2000, 2002; Collignon et al., 2005; Peters et al., 2007).

6.3 Methods

6.3.1. Participants

This study used the same pool of participants as those in Chapter 5. Five of the 115 participants did not complete the source monitoring task, and of the remaining 110,

eight participants were excluded from the study because their performance on the memory task(s) failed to exceed a threshold of 0.1 above chance i.e. less than 0.1 for corrected recognition and source accuracy of less than 0.43. Thus 102 participants (mean age 22.30 years, 80 females) were included in the study.

6.3.2. Procedure

The two source monitoring tasks were a part of a larger battery consisting of five tasks (the remaining three focused on predictive processing, whose results are reported in Chapter 5); each experimental session took a maximum of two hours in total and participants were all tested individually. All tasks were piloted first in eight psychology undergraduate students before the study began.

6.3.3. Action Source Monitoring Task

The action task involved one study-test block separated by 100 minutes. At study participants were asked to sit in a neutral position (arms and legs uncrossed) at a table opposite the experimenter. On the table were objects needed to complete some of the actions and a stack of cards with an action printed on it and above this who should complete it (Participant Perform, Participant Imagine, Experimenter Perform). Each card was turned over by the experimenter one at a time and the participant/experimenter was encouraged to complete the action in a maximum of 6 seconds. There were 75 actions with an equal number in each action condition. Approximately half required everyday objects (e.g. stretch the rubber band, staple pieces of paper together, draw a line with the ruler) and the others were actions without using objects (e.g. nod your head, stand up and sit down, look backwards). The majority of these actions were taken from Collignon et al. (2005). An additional 12 actions were used as practice trials at the start of the study and test phases. All objects were removed prior to the test phase. Here all actions presented in the study phase were randomly intermixed with 25 new actions. The action was presented on a computer screen for 2000ms. Participants were asked to recall whether they performed the action in the study phase (Participant Perform, PP), did they imagine completing the action (Participant Imagine, PI), whether they watched the experimenter perform the action (Experimenter Perform, EP) or whether the action was New.

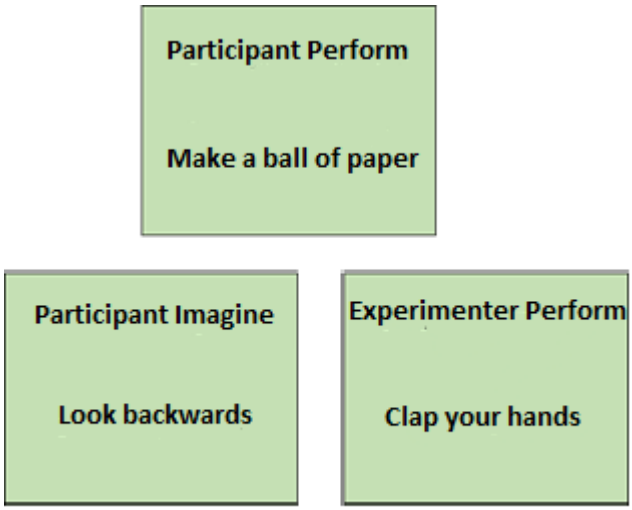

Figure 6.1. Examples from the action source-monitoring task.

6.3.4. Word Pair Source Monitoring Task

The word task also had one study-test block but with an interval of 45 minutes. It was completed on a computer. In the study phase 72 widely known but incomplete word pairs were presented in the centre of the display one at a time, e.g. Mum and ?, Bread and ? (most were taken from Simons et al., 2008) with the condition displayed directly above the incomplete word pairs. In the Participant Perform condition the participant generated the second word and said it out loud, or they imagined the second word (Participant Imagine condition) or listened to the experimenter complete the word pair (Experimenter Perform condition). After the act had been performed the participant needed to press a key to indicate which condition had just been completed. This terminated the trial and the next one commenced. An additional 12 word pairs were used as practice trials at the start of the study and test phases. Participants were asked to complete the word pairs to create a rich encoding context and to produce comparable levels of performance between the two source tasks. In the test phase all actions presented in the study phase were randomly intermixed with 24 new actions. The first word of the pair was presented until the participant made a response. Only the first word of the pair was presented because occasionally participants generate a different second word to what would normally be expected. The discrimination at test was the same as in the action task test phase. For both memory tests participants were

encouraged to respond as quickly but as accurately as they could and actions/word pairs were counterbalanced across conditions.

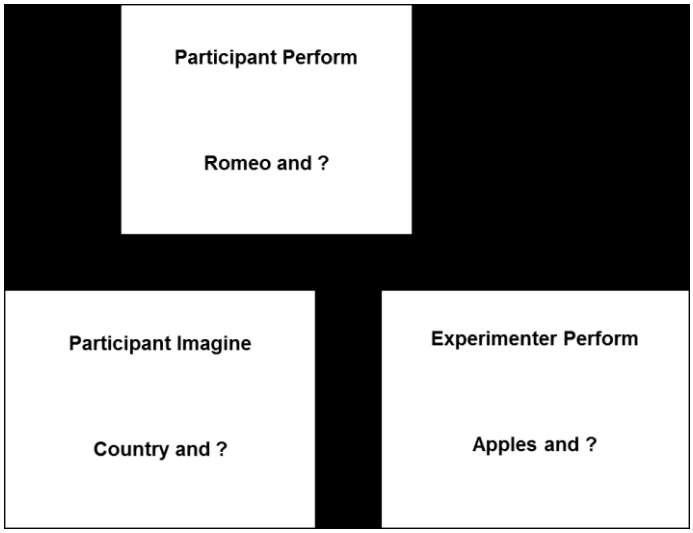

Figure 6.2. Examples from the word pair source-monitoring task.

6.3.5. Measurement of Schizotypy

The Oxford-Liverpool Inventory of Feelings and Experiences (O-LIFE; Mason et al., 1995) has four dimensions: unusual experiences, which indexes experiences akin to hallucinations and delusions; introvertive anhedonia, which describes a lack of pleasure in social or physical activities; cognitive disorganisation, which taps distractibility and disorganisation; and impulsive nonconformity which describes reckless and antisocial behaviour. There has been evidence which questions whether impulsive nonconformity is a meaningful schizotypy construct, so this dimension will not be considered further (Cochrane et al., 2010).

6.3.6. Analysis

For both action and verbal source-monitoring tasks, performance ratios (the accuracy in telling old stimuli from new, i.e. hit rate minus false alarm rate) and conditional probabilities (how accurate one's choices were of the source given the stimulus was old) were calculated for each participant. More specific source memory errors, of particular interest were errors made in the differentiation between internal and external stimuli (i.e. between participant perform/imagine and experimenter perform

conditions) or reality monitoring errors, were also calculated by dividing the number of mistakes made by the total number of stimuli in a given condition.

		Actual			
	Conditions	Participant Perform	Participant Imagine	Experimenter Perform	New
Response	Participant Perform	Source Hit	Internal SM Error	Reality Monitoring Error	False Alarm
	Participant Imagine	Internal SM Error	Source Hit	Reality Monitoring Error	False Alarm
	Experimenter Perform	Reality Monitoring Error	Reality Monitoring Error	Source Hit	False Alarm
	New	Miss	Miss	Miss	Correct Rejection

Figure 6.3. Types of source monitoring errors. SM, source monitoring.

A Shapiro-Wilk test for normality was carried out on data from both tasks and schizotypy questionnaires, which revealed that the data were not normally distributed; as a result, a non-parametric correlational analysis was used (Spearman's rho).

6.4 Results

6.4.1. Schizotypy Questionnaires

The mean schizotypy scores obtained were as follows (standard deviations in parentheses): unusual experiences, 7.31 (5.83); introvertive anhedonia, 5.05 (4.52); and cognitive disorganisation, 12.31 (5.83).

6.4.2. Overall task performance

The descriptive data from the memory tasks can be seen in Table 6.1.

	Action Task	**Word Pair Task**
Hits	0.79 (0.10)	0.75 (0.11)
False Alarms	0.20 (0.17)	0.22 (0.18)
Corrected Recognition	0.59 (0.18)	0.54 (0.17)
Source Accuracy	0.81 (0.10)	0.75 (0.10)

Table 6.1. Mean proportions with standard deviations in parentheses for each of the memory tasks.

6.4.3. Action Source Monitoring Task

Initially data were examined in terms of the proportion of actions correctly recognised as old (Hits) and the new items falsely identified as old (False Alarms). From these data a corrected recognition score can be calculated (Hits – False Alarms; Snodgrass and Corwin, 1988) which gives an index of a participant's ability to discriminate old from new items, see Table 6.1. A significant negative correlation was found between the corrected recognition score and the unusual experiences dimension of schizotypy [$\rho(100) = -0.28$, $p = 0.004$].

A measure of overall source accuracy was calculated as the total number of items correctly assigned to Participant Perform, Participant Imagine and Experimenter Perform sources divided by the number of Participant Perform, Participant Imagine and Experimenter Perform items correctly identified as old (regardless of whether the source judgment was correct). There was a negative correlation between source accuracy and scores on the unusual experiences dimension, $\rho(100) = -0.21$, $p = 0.034$. Given that source errors on this task could be due to internal source monitoring i.e. confusing Participant Imagine with Participant Perform and vice versa; or reality monitoring i.e. confusing Participant Perform/Imagine with Experimenter Perform

and vice versa, these were assessed separately. Fig. 6.4 displays the number of internal source monitoring and reality monitoring errors, which correspond to the black and white bars, respectively. The notation used in the figure and below is that the first abbreviation corresponds to the actual source and the one after is the participant's memory judgement e.g. PP/PI would be an item that the participant performed but which they thought they had imagined. A significant relationship was found between unusual experiences and total number of internal source memory errors (the sum of errors in PP/PI and PI/PP conditions, see Fig. 5.1), $\rho(100) = 0.22$, $p = 0.03$.

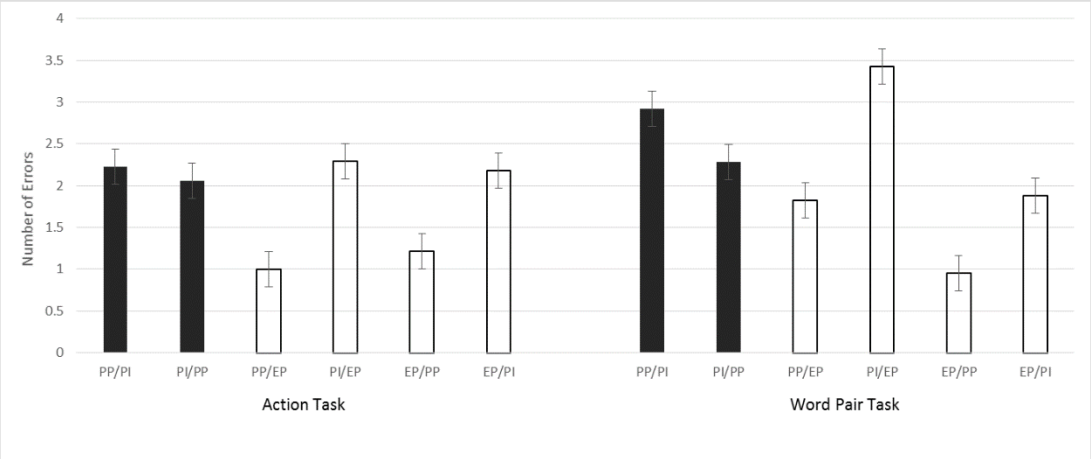

Figure 6.4. The mean number of errors produced in each memory task (action on the left, word pairs on the right) with error bars (± SEM). Internal source monitoring errors are in the filled bars and reality monitoring errors in the unfilled bars. Abbreviations are as follows: PP (Participant Perform), PI (Participant Imagine), and EP (Experimenter Perform). The first abbreviation is the actual source of the event and the second one is what the participant stated.

Moreover, there was also a significant positive correlation between unusual experiences and the overall number of reality monitoring errors (the sum of errors in PP/EP, PI/EP, EP/PP, and EP/PI conditions, see Figure 6.3), $\rho(100) = 0.24$, $p = 0.014$. There is a wealth of evidence demonstrating that reality monitoring problems in schizophrenia are in the direction of misattributing self-generated events to an external source i.e. externalising (e.g. Vinogradov et al., 1997, 2008). Therefore two additional correlations were conducted separately for two components of the reality monitoring score. There was a significant relationship between unusual experiences and errors in attributing an action that the participant performed to the experimenter (PP/EP), $\rho(100) = 0.27$, $p = 0.005$; but the same relationship was not found for imagined actions (PI/EP), $\rho(100) = 0.13$, $p = 0.19$.

6.4.4. Word Pair Source Monitoring Task

This task was analysed within the same framework as described above for the action task. There were no significant correlations between unusual experiences and corrected recognition score [$\rho(100) = -0.02$, $p = 0.81$] or overall source memory accuracy [$\rho(100) = -0.07$, $p = 0.48$]. There were also no significant associations with number of internal source memory errors [$\rho(100) = 0.15$, $p = 0.14$] or reality monitoring errors [$\rho(100) = 0.19$, $p = 0.06$]. No significant relationships were found between unusual experiences and externalising errors (ps > 0.88).

6.4.5. Other Schizotypy Dimensions

Although the focus of this study was on the unusual experiences dimension correlations were also conducted with the introvertive anhedonia and cognitive disorganisation dimensions of schizotypy to determine the specificity of the relationship. As can be seen from Table 6.2, there were no relationships with the introvertive anhedonia dimension but some with cognitive disorganisation. This might have resulted from the high degree of correlation between unusual experiences and cognitive disorganisation [$\rho(100) = 0.65$, $p < 0.001$)].

	Action Task		Word Pair Task	
	RM	**ISM**	**RM**	**ISM**
Unusual Experiences	**0.24***	0.22	0.19	0.15
Introvertive Anhedonia	0.06	0.07	-0.05	-0.01
Cognitive Disorganisation	0.14	**0.22***	**0.24***	0.13

Table 6.2. Spearman's rho correlation matrix showing coefficients between reality and internal source monitoring errors and different domains of the O-LIFE. Abbreviations are as follows: RM (reality monitoring), ISM (internal source monitoring).

6.5 Discussion

The aim of this study was to provide a more detailed understanding of the nature of the memory deficits associated with the schizotypy continuum. A significant negative correlation was found between the positive dimension of schizotypy (unusual experiences) and the corrected recognition score, indicating that participants high in unusual experiences exhibited poorer old-new discrimination. Furthermore, they were also more inclined to make errors in determining the source of the memory, even in those items correctly recalled as old. In particular, there was a positive correlation between unusual experiences and internal source monitoring errors; those participants with high scores on this dimension confused whether they had performed an act or just imagined doing it. There was also a positive relationship between the same schizotypy dimension and reality monitoring errors i.e. in determining whether the act originated from the participant (performed or imagined) or the experimenter. Consistent with previous research there was an externalising bias, such that those high in unusual experiences tended to attribute actions they had physically performed themselves to the experimenter (PP/EP errors). However, the same pattern of results was not found

for those acts the participant had just imagined (PI/EP errors). All of these relationships were only found in the action based task.

It is widely acknowledged that individuals with schizophrenia have deficits in recollection but findings on familiarity have been less consistent (Achim and Lepage, 2003; Libby et al., 2013; Ranganath et al., 2008). This is also true in schizotypy work, for example Peters et al. (2007) found evidence for deficits in old-new recognition, whereas Collignon et al. (2005) did not. It is possible that the particular measure of schizotypy used may be important. Collignon et al. (2005) used a measure that specifically assessed hallucinatory proneness (Launay and Slade Hallucinations Scale; Launay and Slade, 1981), whereas Peters et al. (2007) used the Schizotypal Personality Questionnaire (Claridge and Broks, 1984) and the current study used the unusual experiences dimension of the O-LIFE (Mason et al., 1995). These latter questionnaires index positive symptoms more widely and, for example, also encompass distortions in sensory experiences and psychotic-like delusional ideation. Thus it would appear to be the case that difficulties in making old-new discriminations are related to positive symptom-like experiences more broadly, or a specific aspect of these, but not hallucinations.

The finding of more internal source errors being related to high unusual experiences is consistent with the work of Collignon et al. (2005) and Peters et al. (2007). However, this finding was extended to include reality monitoring errors being associated with the positive dimension of schizotypy as well, which was not found by Collignon et al. (2005). There are methodological differences between the current study and that by Collignon et al. (2005) which might explain this. Firstly, in the latter study there were more conditions for participants to differentiate between; they had the added conditions of the participant imagining the experimenter performing the action and the experimenter verbalising the action (but not performing it). Secondly, the way the test response was made was quite different with Collignon et al. (2005) requiring participants to make a four-stage response at test compared to just one-stage in this study. Finally, their participants made very few errors (mean of < 1) in some of the conditions, particularly those relevant to reality monitoring, such as participant performed and experimenter performed. These floor effects might have precluded relationships being found with hallucinatory proneness by Collignon et al. (2005).

The Source Monitoring Framework (Johnson et al., 1993; See Chapter 3) offers a useful way of understanding the errors that people make when trying to retrieve the source of a piece of information. According to this framework there are no specific memory 'tags' or markers on events indicating where they originated. Instead, various attributes of the memory encoded at the time it happened later serve as the basis for making the decision as to its origin. These attributes include qualities like perceptual, semantic, spatial, temporal, sensorimotor and affective details and records of cognitive operations that created them (Johnson et al., 1993; Johnson and Raye, 1981). For example, a memory that is rich in perceptual detail, with substantial contextual information but a lack of consciously remembered details of the cognitive operations which might have generated it would likely be judged as having been perceived, whereas the opposite profile would be associated with imagined experiences. Therefore, anything which increases the similarity of these memory attributes from different sources will decrease source accuracy. For example, if imagination was particularly vivid and detailed this could be confused with an event that was actually experienced. This is pertinent because there has been a wealth of research demonstrating that people with schizophrenia (Mintz and Alpert, 1972; Rasmussen and Parnas, 2015) and those high in schizotypy (Winfield and Kamboj, 2010; Currie, 2000) tend to have more active and vibrant imaginations (Oertel et al., 2009; Sack et al., 2005). In future research it might be useful to include a measure of how well participants feel they are able to imagine completing acts as this could mediate the relationship between schizotypy/schizophrenia and memory performance.

The novel finding from this study is that significant relationships were found between memory measures and unusual experiences in the action task but not the word pair one, indicating specificity of deficits in monitoring actions. No previous work has examined both action and verbal source monitoring concurrently in the same cohort of participants, although deficits in both domains have been found separately in either different participants or unrelated tasks. The same direction of result was found in the word pair task, between schizotypy and internal and reality monitoring errors, but these did not reach statistical significance. This suggests that the action task might have greater utility in examining relationships with symptoms or experiences. Due to the well-known enactment effect (Cohen, 1989; Madan and Singhal, 2012) the study-test interval for the action task was longer (100 minutes) than for the word pair task

(45 minutes). This was done to ensure that performance was not at ceiling in the action task and both tasks were broadly comparable in terms of participant performance. As can be seen from Fig 5-3 the profile of errors between tasks is similar. Moreover, the errors also exhibit a similar profile as to what might be anticipated. For example, there is less overall confusion between Participant Perform and Experimenter Perform than between Participant Imagine and Experimenter Perform. This is likely due to the fact that when the participant performs the act there is movement as well as afferent feedback but this is not present when they imagine the act or watch the experimenter perform it, which makes the former two conditions more distinctive than the latter two.

The action memory task has been used in a number of studies both in schizophrenia and schizotypy (Collignon et al., 2005; Gawęda et al., 2012; Peters et al., 2007) and there is substantial evidence that people with schizophrenia have abnormalities in the awareness of motor actions (Frith et al., 2000; Blakemore et al., 2002). Computational models of motor control have been developed and these have been applied to schizophrenia, particularly the forward model (Wolpert, 1997). According to this account, whenever a motor command is initiated a parallel efference copy is also generated (Von Holst, 1954). This can be used to make predictions about the sensory consequences of an action, which can be compared with the actual sensory feedback of a movement. If the predicted action and the sensory input match then the action would be considered to be self-generated.

In schizophrenia, it is thought that there may be deficits in the generation of the efference copy and/or in the comparison between predicted and actual action which results in certain positive symptoms (Frith, 2005, 2012; Synofzik et al., 2010). Importantly, this would produce externalising errors, which have been found in a number of studies (for a review, see Brookwell et al., 2013), because no efference copy or a mismatch between prediction and reality would suggest an external source. In the current study, the only relationship found was between schizotypy and one type of externalising error: an act physically performed by the participant being attributed to the experimenter and not when the act had only been imagined by the participant. One potential explanation for this is that perhaps the forward model, and the hypothesised deficits that individuals within the schizophrenia spectrum have with aspects of this, can only be applied to overt actions and not internal mental events such as thinking and imagining. Indeed, this model was adapted and used by Frith and

colleagues to explain such phenomena as delusions of control and anarchic hand (e.g. Blakemore and Frith, 2003; Frith et al., 2000). A number of arguments have been raised about the possibility of extending this model to covert forms of behaviour, such as thinking. Gallagher (2004) argues that using the forward model makes sense for overt actions because one needs to know if one's actions are internally or externally caused and if one's action is not going to achieve its goal, this needs to be known in advance so that adjustments can be made. However, these reasons do not make sense when applied to thoughts. All of one's thoughts are internally generated, so there is never any possibility of having to work out whether it was oneself who thought something or someone else in normal circumstances. Thus there is currently a great deal of debate around whether Frith's forward model can be applied to internal mental states (for other work on this issue see Seal, Aleman, and McGuire., 2004; Stephens and Graham, 2000; Vicente, 2014).

To conclude, these results demonstrate that there is a negative relationship between scores on the positive dimension of schizotypy, unusual experiences, and the ability to correctly identify the source of memory information. Furthermore, the correlational analyses indicated that individuals with high scores on unusual experiences have deficits in distinguishing between actions they performed versus i) imagined and ii) those the experimenter performed. These relationships were only found in the action based task and further research is now needed to determine if a similar set of results would be found in people with schizophrenia.

Chapter Seven: Sensory and Reward Predictive Processing in Early Psychosis

7.1. Abstract

This study aimed to investigate behaviourally measured prediction error responses in individuals experiencing early psychotic symptoms compared with healthy controls. Ten patients with early psychosis and ten controls matched for gender, age and years of education took part in a force-matching task and a reversal learning task which assessed sensory and reward domains of predictive processing. Patients underwent a clinical interview for early psychosis symptoms whereas controls were screened for neuropsychiatric disorders and completed a questionnaire for schizotypal personality traits. In the force-matching task, patients did not demonstrate higher accuracy (i.e. less sensory attenuation and better matching) than controls, and their level of attenuation was not significantly correlated with their delusional ideation scores. In the reversal learning task, patients made significantly more errors than controls in general, and took longer to restore accuracy after the first post-probabilistic error trial. However, switching tendency was not significantly correlated with their positive symptoms. Bayesian methods have on the other hand shown anecdotal levels of evidence for the correlations in the directions predicted. In this study there was no evidence of a deficit in predictive processing in early psychosis patients and there were no correlations between performance and symptomatology. However, these must be taken only as preliminary findings given the small sample size.

7.2. Introduction

As outlined in Chapters 2 and 5, previous studies have provided strong, albeit not universal, support for the predictive processing model for the pathogenesis of psychotic symptoms. In patients with established schizophrenia, deficits have been found across the sensory, cognitive and reward domains of prediction error responses by both behavioural and neuroimaging methods (e.g. Lindner et al., 2005; Shergill et al., 2005; Jones et al., 1992; Corlett et al., 2007). In addition, such deficits often correlate with symptom severity. Although these studies do not offer causal

explanations due to the cross-sectional and correlational nature, they at least demonstrate significant associations between schizophrenic symptomatology and neural/behavioural markers of disturbed prediction error signalling. Studies of this kind are mostly done in either patients with schizophrenia or healthy individuals with schizotypal traits (e.g. Corlett and Fletcher, 2012; Teufel et al., 2010) such as the study described in Chapter 5. However, there are few studies conducted in individuals in the earlier stages of psychotic disorders (i.e. the stage after latent psychosis vulnerability/prodromal symptoms and before fully diagnosed schizophrenia; see Fusar-Poli, Yung, McGorry, and van Os, 2014 for a staging model), which may present a gap in the understanding of predictive processing deficits from a continuum model of psychosis.

In the current Thesis, individuals who have experienced a first brief psychotic episode as well as those with attenuated psychotic symptoms were chosen over patients with an established diagnosis of schizophrenia. This was mainly due to the higher level of positive symptoms (i.e. symptoms of interest for the current project) in the early phases of psychosis compared to chronic schizophrenia. Although many of the patients were already treated with medication and entered symptomatic remission, they had not been taking antipsychotics for over twelve months so that any potential behavioural manifestations of antipsychotic side-effects (e.g. parkinsonian symptoms) would be minimal.

Originally only individuals who were considered as ultra-high risk were the target population for recruitment, however this was later expanded to patients with first episode psychosis as long as they have not received a firm diagnosis of schizophrenia, schizoaffective disorder or substance-induced psychosis (in fact, there was no report of current illicit drug misuse at all in both patients and controls), and have not been receiving treatment from Early Intervention Services for over twelve months. A first brief psychotic episode is not always an indicator for later schizophrenia and can follow a variety of different trajectories (Marneros, Pillmann, Haring, Balzuweit, and Blöink, 2003; Singh, Burns, Amin, Jones, and Harrison, 2004). By focusing on a more extended early period of psychosis, one can maximise potential participant pool and also expand the concept of early psychosis more broadly, in order to capture a wider variability in symptomatology. In fact, in the current project the chosen assessment tool was one for prodromal syndromes (the Structured Interview for Prodromal

Syndromes, SIPS) and not one for established schizophrenia such as the Positive and Negative Syndrome Scale (PANSS), as the latter would not have been sensitive enough to detect the subtleties in the symptomatology of early psychosis. As a consequence, no formal diagnoses were made in patients. Nevertheless, it is worth mentioning that the patients in the current study have all been referred to clinical services due to functional impairment and/or distress caused by their anomalous experiences, and it is still possible that they have been experiencing symptoms for a prolonged period of time before deciding to seek help. As such, enquiries were made about the approximated date of first symptom onset during assessments, and retrospective scores of symptom severity at peak were also recorded.

The tasks chosen in the clinical study were derived from those used in healthy volunteers with schizotypal traits (Chapter 5), which consisted of the force-matching task and the reversal learning task alongside a source-monitoring task for actions (see Chapter 8). The Kamin blocking task was omitted for the clinical study mainly because it did not correlate with schizotypy in the previous study and has some conceptual overlap with the reversal learning task. Additional reasons for shortening the duration of testing were to avoid overloading the patients given the time required (at least 90 minutes) for symptom assessment and other tasks such as assessment for IQ and substance use. Although none of the previous tasks significantly correlated with schizotypy, force-matching task was included as it is the only task measuring sensory prediction error responses, and reversal learning was chosen as the task measuring associative and reward predictive processing. The hypotheses were that individuals with early psychosis would display significant resistance to the force-matching illusion compared to controls as demonstrated by lower overcompensation and thus better performance. Patients would also show a heightened tendency to switch in the reversal learning task as indexed by their post-probabilistic error accuracies. Moreover, it was predicted that overcompensation (an index of sensory attenuation) would correlate negatively with their delusional ideation scores, whereas a tendency to switch would correlate positively with general positive symptom scores.

7.3. Methods

7.3.1. Power Calculation

Power calculations were carried out in order to determine the desired sample size. Estimates of effect sizes were derived from three studies using somewhat similar tasks: 1) a study on the sense of agency and intentional binding in prodromal psychosis patients yielded an effect size of r = 0.45 (Hauser et al., 2011); 2) a study on associative learning in individuals at risk for psychosis yielded an effect size of partial eta squared = 0.17 (Orosz et al., 2010); and 3) a study using a very similar reversal learning paradigm in unmedicated schizophrenia patients yielded an effect size of Cohen's d = 1.37 (Schlagenhauf et al., 2014). Based upon these figures, if one wanted an 80% probability (1-beta) of detecting a difference between patients and controls while setting a conventional alpha level of 0.05, between 28 and 31 individuals with early psychosis and at least the same number of matched controls would be needed. Unlike the healthy volunteer study, power was set at 80% in the clinical study instead of 90% to avoid potentially yielding unrealistic target sample sizes while still maintaining an acceptable level of power.

7.3.2. Patients with Early Psychosis

Despite continued efforts, only 10 individuals who have experienced an initial brief episode of psychotic symptoms were recruited from secondary care (e.g. early intervention service clinics) within the specified timeframe of 10 months across two NHS Boards in Wales (Cardiff and Vale University Health Board and Aneurin Bevan University Health Board), an NHS mental health Trust in south-west England (Avon and Wiltshire Mental Health Partnership NHS Trust) and databases from other research studies in progress at Cardiff University.

Initially members of the research team approached healthcare professionals at relevant clinics, asking them to give out invitation letters to patients who may be interested. This method of recruitment keeps the identity of the individual unknown to the researcher unless they are interested in taking part. If participants were interested in taking part they were provided with an Information Sheet (also given to the healthcare professionals) and asked to phone, e-mail or send a reply slip to the research

team in prepaid envelopes if they wanted further information or to book an appointment (or their care coordinators could do this). Follow-up calls with the healthcare professional were carried out after a minimum of one week if a participant had not already made contact. Participants were given as much time as they needed to read and consider these documents and consult with relatives or carers if they wished. Participants were contacted a maximum of three times after an expression of interest with regard to participating in this study.

Following expressions of interest in taking part in the study, a meeting was arranged with the potential participant to establish their suitability for the study. This screening session reviewed inclusion and exclusion criteria as well as asking for the participant's written consent to contact their care coordinator to review issues of suitability and risk. They were also given the opportunity to ask any questions about the study. If the participant was eligible, a mutually convenient day and time was arranged for them to come into Cardiff University to take part in the study.

Inclusion criteria for all patients were:

- Aged 18-60 years old;
- High fluency in the English language, as indicated on the demographics questionnaire (first language or bilingual proficiency);
- Normal or corrected-to-normal vision and hearing;
- Attendance at secondary mental health services, including early intervention for psychosis services and community mental health teams, provided by the organisations listed above; and
- The ability to give informed consent to take part in the study.

Exclusion criteria for all patients were:

- Current presentation or a history of a clinically significant neurological condition (e.g. Migraine, stroke, traumatic brain injury, epilepsy, space occupying lesions, multiple sclerosis, Parkinson's disease, dementia, etc); and
- Current presentation or a history of clinically significant substance misuse/alcohol dependence (not just recreational use).

A larger age range was chosen (18 – 60 instead of 18 – 35, which is the peak age group for psychosis onset and cut-off for referrals to most early intervention services

in the NHS) because of recent research pointing towards the importance of early intervention in adults aged above 35 (Greenfield et al., 2016). The average duration of symptoms at the time of referral was 7.2 months (SD = 3.85 months). Medication status was not an exclusion criterion; however, patients must not have been receiving antipsychotic treatment for over 12 months. This criterion was chosen because of potential adverse effects on behaviour and brain physiology following long-term antipsychotic treatment (e.g. Ho, Andreasen, Ziebell, Pierson, and Magnotta, 2011). Of the ten recruited patients, seven were prescribed low-dose second-generation antipsychotics (risperidone 2mg/d, olanzapine 10mg/d and aripiprazole 15mg/d), three were prescribed antidepressants (mirtazapine 30mg/d, citalopram 20mg/d and sertraline 50mg/d) in addition to their antipsychotics and two were prescribed a benzodiazepine (lorazepam 2mg) for use when needed (PRN) in addition to their antipsychotics. The remaining three were not taking any form of psychotropic medication at the time of testing.

7.3.3. Controls

Ten healthy controls matched for gender, age, handedness and years of education were recruited from the Cardiff University School of Psychology Community Panel, other existing databases and through advertisement. In the initial invitation letter it was clearly stated that controls should have no present psychiatric illness or a history of psychiatric illness, and that they should also fulfil the first three inclusion criteria outlined above (age, language, normal vision and hearing). All interested controls received a Participant Information Sheet prior to the testing session Although there was a screening interview aimed to ensure the absence of psychiatric illnesses (the M.I.N.I. Neuropsychiatric Interview), this only took place after the control participant had given consent to ensure that no study procedures would be undertaken without prior consent being obtained.

7.3.4. Procedure

Ethical approval for the current study was obtained from Wales Research Ethics Committee 2 (Reference: 16/WA/0039) which covered the recruitment of both patients and controls. All participants gave informed consent before any of the study

procedures commenced and this was documented on the Informed Consent Form. Participants completed all the measures outlined below. For patients, the study session took approximately 3.5 hours and this was reduced in controls to 3 hours due to a shorter clinical assessment. The session for both patients and controls began with a demographics questionnaire, followed by behavioural tasks (action source monitoring – study phase, force-matching, reversal learning), IQ measure and substance use questionnaire, and lastly the test phase of the source monitoring task. For patients, the clinical assessment was at the very end of the session whereas the symptom screening was at the beginning of the session for controls. Patients also completed an additional questionnaire on depression and anxiety before the substance use questionnaire. Regular breaks were arranged for both groups and all participants were able to ask for a pause in the session whenever they felt necessary. Every participant was paid for their time (£35 for patients and £30 for controls) and travel expenses were also reimbursed.

7.3.5. Force-matching Task

This was the same force-matching task used in Chapter 5, adopted from Teufel et al. (2010).

7.3.6. Reversal Learning Task

This was the same reversal learning task used in Chapter 5, identical to the 'private condition' used by Ihssen et al. (2016).

7.3.7. Assessments

Both patients and controls were asked to provide basic demographic information and information on smoking status, recreational drug use and/or alcohol consumption. The levels of substance use must fall below the threshold for a clinically relevant diagnosis of substance dependence/abuse, otherwise they would not meet the inclusion criteria. A brief IQ measure (WASI; Wechsler Abbreviated Scale of Intelligence, Second Edition, Two-subtest form) were also given to both patients and controls which contained items on vocabulary and matrix reasoning, yielding an overall score of general cognitive ability.

Each patient underwent one clinical interview: the Structured Interview for Prodromal Syndromes (SIPS, with the companion Scale of Prodromal Symptoms, SOPS). Of particular interest for the current study is the SOPS, which has 19 items under 4 subscales (positive, negative, disorganised and general symptoms) which are rated on a 7-point Likert scale from absent (0) to severe (6). The SIPS includes a checklist for Schizotypal Personality Disorder and a research version of the Global Assessment of Functioning (GAF) scale which allows the measurement of general levels of functioning. The SIPS and SOPS have good predictive validity (Miller et al., 2003) and in the present study, scores were agreed between two raters who were both present at the assessment.

The Hospital Anxiety and Depression Scale (HADS; Zigmund and Snaith, 1983) were given to all participants in the case group. The HADS is a 14-item short questionnaire which asks about the individual's current state of anxiety and depression levels on a Likert scale from 0 (symptom absent) to 3 (severe symptom).

Only control participants underwent the M.I.N.I. International Neuropsychiatric Interview (Sheehan et al., 1998) in order to screen for any psychiatric disorders before completing the rest of the session, as the major criterion for being in the control group was the absence of any mental illness. No control was excluded on the basis of the M.I.N.I. interview. In addition, control participants were given the Oxford-Liverpool Inventory of Feelings and Experiences (O-LIFE; Mason et al., 1995), a questionnaire designed to assess schizotypal personality traits in healthy individuals from a non-clinical perspective.

7.4. Data analysis

Both frequentist and Bayesian methods, using SPSS 23 and JASP 0.8.1.2, respectively, were used in the current study. Although it was found that scores from symptom assessment in the current study followed a normal distribution after inspecting Q-Q plots, and Shapiro-Wilk tests showed no significant deviation from normality (all ps > 0.20), this was not the case for all of the measures. Therefore, nonparametric statistical tests were used (i.e. Spearman's rho) when correlating behavioural outcomes with symptom assessment scores, but parametric t-tests were used within tasks where the measures followed a normal distribution.

7.5. Results

7.5.1. Sample Characteristics

Table 7.1 below shows the characteristics of both patients and controls:

Characteristics	Patients	Controls	
Demographics			
Age [Mean years (SD)]	28.40 (11.96)	29.20 (11.10)	$t(18) = -0.155$
Gender (Male: Female)	9: 1	9: 1	
Handedness (% Right)	100%	100%	
WASI-II [Mean (SD)]	84.80 (16.28)	96.80 (7.80)	$t(18) = -2.102$*
Years of education	15.20 (3.05)	15.90 (1.45)	$t(18) = -0.656$
SOPS rating: Current [Mean (SD)]			
Total positive score	8.60 (5.94)		
Delusional ideas	2.40 (1.77)		
Suspiciousness	1.90 (1.45)		
Grandiosity	1.10 (1.37)		
Hallucinations	2.80 (2.04)		
Disorganisation	0.40 (0.70)		
Total negative score	12.30 (8.46)		
Total disorganised score	3.50 (2.88)		
Total general score	6.00 (4.29)		
SOPS rating: Peak [Mean (SD)]			
Total positive score	15.80 (5.12)		
Delusional ideas	4.40 (1.26)		
Suspiciousness	3.90 (1.29)		
Grandiosity	1.90 (1.85)		
Hallucinations	4.10 (2.08)		
Disorganisation	1.50 (1.43)		
Total negative score	17.10 (5.88)		
Total disorganised score	5.60 (2.67)		
Total general score	11.50 (4.28)		
Substance use score	0.00	0.00	
HADS rating [Mean (SD)]			
Anxiety	9.10 (4.28)		
Depression	7.90 (5.59)		
O-LIFE score [Mean (SD)]			
Unusual experience		1.40 (1.77)	
Introvertive anhedonia		4.70 (3.23)	
Alcohol use and smoking			
Mean weekly alcohol intake (UK units)	1.3	3.7	
Current smokers	70%	30%	

*$p < 0.05$.

Table 7.1. Sample characteristics for both patients and controls. WASI-II, Two-subset form of the Wechsler Abbreviated Scale for Intelligence, Adult version; SOPS, Scale of Prodromal Symptoms; HADS, Hospital Anxiety and Depression Scale; O-LIFE, Oxford-Liverpool Inventory of Feelings and Experiences.

Demographic characteristics (age, gender, handedness) for controls were well-matched to those of patients; although IQ was significantly higher in controls, the years of formal education in the two groups were comparable. In terms of symptomatology at the time of testing, positive symptom ratings were relatively low on average. For example, hallucinations score was 2.8 (mild to moderate) out of a potential 6.0, which would indicate current psychotic syndrome). This was not surprising given the early and brief nature of these patients' psychotic symptoms as well as the fact most were on antipsychotic treatment already which would have ameliorated acute symptoms. Nevertheless, compared to current symptoms, peak symptom levels were much higher on average, potentially indicating effective treatment response. In a validation study for the SIPS by Woods et al. (2009), early psychosis patients scored an average of 11.9 for positive symptoms, compared with 8.60 in the current study. Although symptoms scores were indeed considerably higher during the first episode onset (positive symptom score of 15.80 in the current sample) at the time of first presentation and referral to mental health services, these were difficult to ascertain retrospectively and accurately. Therefore, the peak symptom scores were not entered into statistical analyses.

7.5.2. Force-matching Task

Figure 7.1 shows the overall performance of patients and controls in the Finger and Slider conditions. Two patients had to be excluded due to exerting forces as high as 11N and thus suggesting an inability to understand task instructions. No control participant was excluded on the basis of performance. It is clear that both patients and controls exerted much higher matching forces in the Finger condition, indicating the overcompensation effect. On the other hand, both patients and controls were much more accurate at matching forces in the Slider condition (where participants match the presented force with a slider). This shows that the task itself worked as expected. However, patients did not demonstrate lower overcompensation compared to controls and the levels of sensory attenuation in patients and controls were in fact very similar: a simple t-test revealed no significant differences between the forces exerted by patient and controls in the Finger condition (where participants directly match the presented force with their own right index finger), [$t(16) = 0.70$, $p = 0.508$, Cohen's $d = 1.80$].

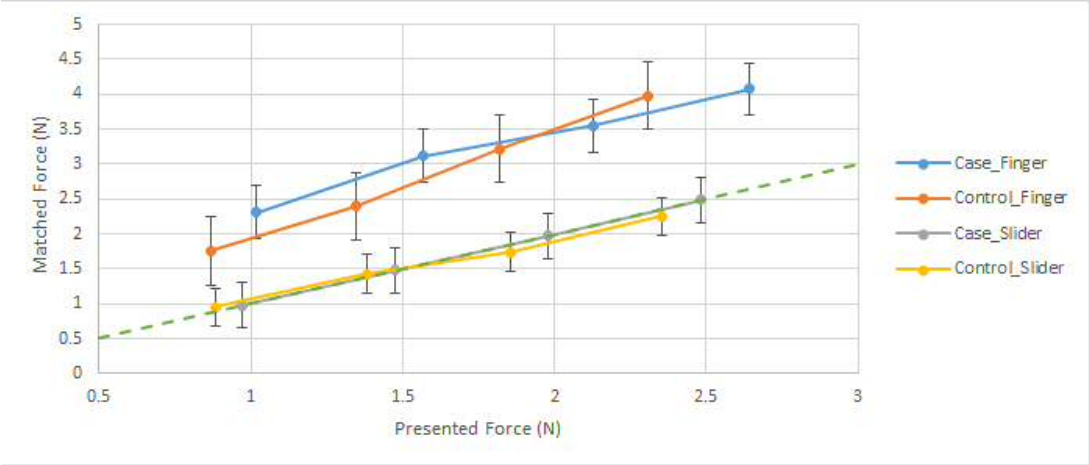

Figure 7.1. Comparisons between mean applied force in the Finger and Slider conditions in patients and controls. Dotted line indicates perfect performance.

Further, in order to replicate the analytic methods employed by Shergill et al. (2005), a mixed ANOVA (patients versus controls and Finger versus Slider conditions) was carried out, showing no significant interaction between clinical status (patient/control) and force condition (Finger/Slider), [$F(3, 48) = 0.216$, $p = 0.885$, partial eta squared = 0.004], but only a main effect of force condition [$F(3, 48) = 32.353$, $p < 0.001$, partial eta squared = 0.223]. Further, the overcompensation score was not significantly correlated with SOPS delusion subscale [$\rho(6) = 0.424$, $p = 0.295$].

7.5.3. Reversal Learning Task

Figures 7.2 and 7.3 display the true reversal error accuracies as an index of perseveration, and probabilistic error accuracies as an index of switching tendency, respectively. In both patients and controls, accuracy restored quickly after the first trial post-error and no participants were excluded on the basis of task performance.

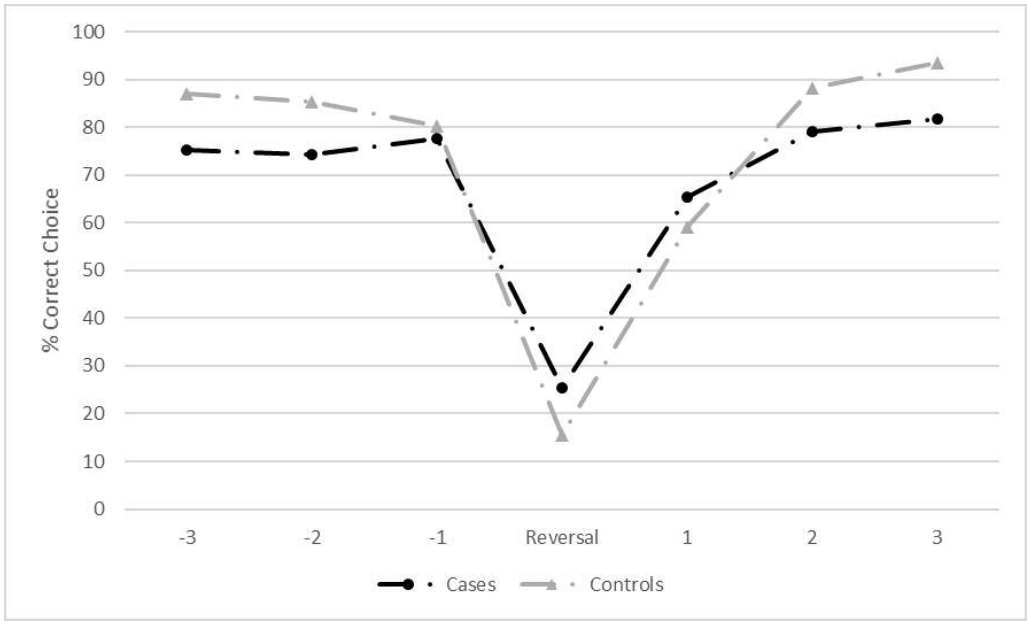

Figure 7.2. True reversal accuracies for patients (black rounded marker) and controls (grey triangular marker).

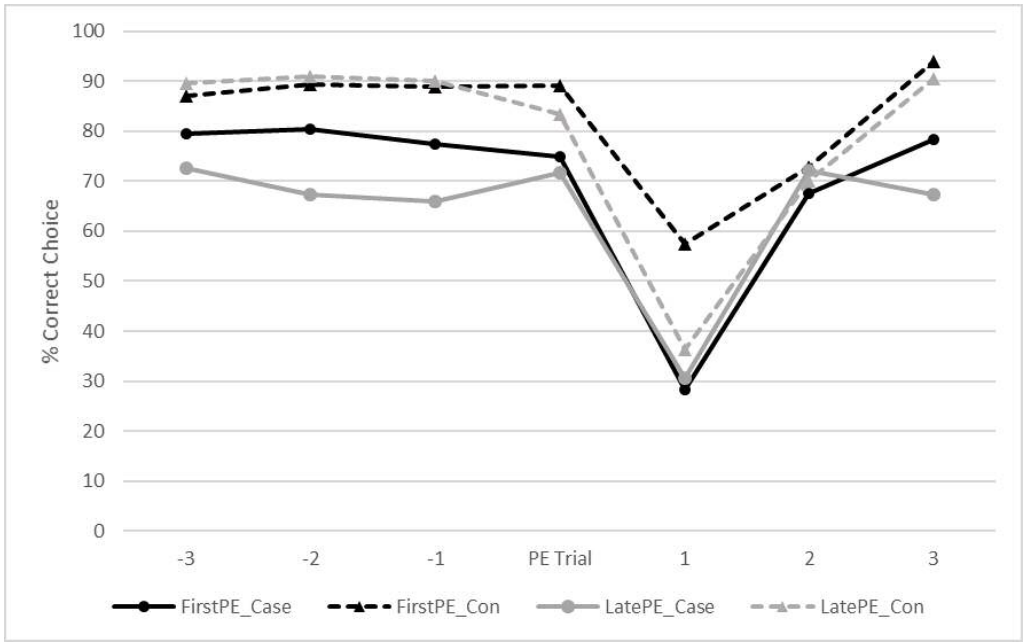

Figure 7.3. Probabilistic error (PE) accuracies for patients (black rounded marker) and controls (grey triangular marker). Solid line indicates first PE and dashed line indicates late PE. Con, Controls.

However, patients made significantly more errors overall compared with controls [$t(18) = -2.822$, $p = 0.011$, Cohen's $d = 1.26$] but the difference is not statistically significant when divided into post-reversal [$t(18) = 0.479$, $p = 0.638$, Cohen's $d = 0.21$] or post-probabilistic error trials [$t(18) = 1.867$, $p = 0.078$, Cohen's $d = 0.84$]. Further,

patients were significantly less accurate and took longer to restore accuracy after the first post-probabilistic error trial (Figure 7.3, solid versus dotted black lines) compared to controls [$t(18) = -2.773$, $p = 0.012$, Cohen's $d = 1.24$]. Patients did not show a significant relationship between switching and their positive symptom scores [$\rho(8) = -0.671$, $p = 0.867$]. However, scores for perseverative behaviour (i.e. post-true reversal accuracies) were significantly positively correlated with SOPS positive symptom scores in patients [$\rho(8) = 0.755$, $p = 0.031$]. Such an association was no longer significant when controlling for IQ [$\rho(7) = 0.659$, $p = 0.108$].

7.5.4. Exploratory Analyses with Other Symptom Domains

Further correlations were carried out for exploratory purposes only. Table 7.2 displays correlation coefficients between task performance and other specific positive symptoms (e.g. suspiciousness and hallucination) as well as overall negative symptom scores.

	Del	Sus	Hal	Neg
Force-matching Overcompensation	0.424	0.275	-0.061	0.001
Post Reversal Perseveration	0.537	0.718*	0.793*	-0.036
Post Probabilistic Error Switching	0.085	-0.025	-0.073	0.084

*, $p < 0.05$.

Table 7.2. Nonparametric bivariate correlation coefficients (Spearman's ρ, two-tailed) between SOPS symptom measures and task measures in patients. Del, delusional ideation; Sus, suspiciousness; Hal, hallucinations; Neg, negative symptoms.

Hallucination scores were found to be significantly correlated with perseveration in the reversal learning task; a similar relationship was also found with suspiciousness specifically, but not with general delusional ideation/unusual thought content. The association between suspiciousness and perseveration diminished when IQ was controlled for [$\rho(7) = 0.625$, $p = 0.133$], but the relationship with hallucination scores remained even after controlling for IQ [$\rho(7) = 0.778$, $p = 0.039$]. No other significant relationships were found with any other task measure or symptom scores. Nevertheless, inferences from these correlations can only be drawn with caution, given they would not have survived multiple comparison corrections and were only carried out as post hoc exploratory analyses.

7.5.5. Bayesian Analysis

Due to the failure to find any differences between patients and controls on the force-matching task Bayesian statistics were used to determine the level of evidence for the null/alternative hypothesis, particularly given the small sample size. A Bayesian independent samples t-test was conducted using a default Cauchy prior of 0.707, which is generally accepted (Quintana and Eriksen, 2017, accepted preprint) as a distribution of effect sizes considered realistic under the alternative hypothesis (H1), based on the consensus that mean effect sizes tend to be about half of the standard deviation of the effect. As can be seen in Figure 7.4, the Bayes Factor for the null hypothesis is 1.552 compared to that for the alternative hypothesis which is 0.644. This means that the null hypothesis (i.e. patients did *not* show less overcompensation than controls) is favoured 1.55 times over the alternative hypothesis, and this is supported by anecdotal levels of evidence under a range of priors.

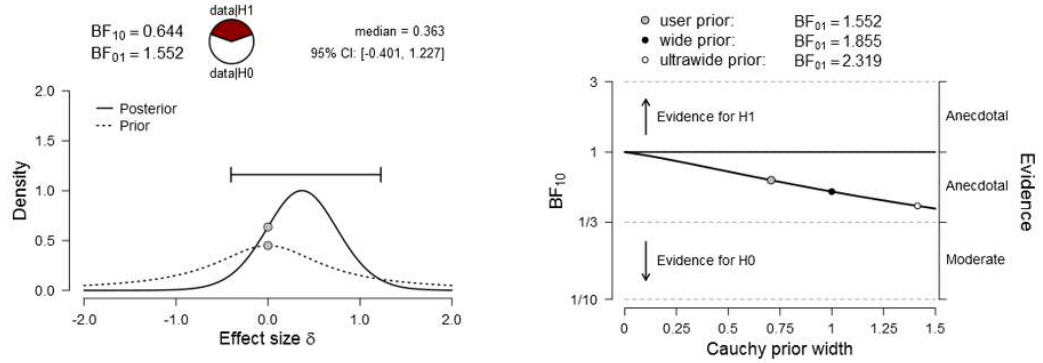

Figure 7.4. Results from Bayesian independent samples t-test. BF, Bayes Factor; CI, Credibility Interval.

Taking a similar approach to that in the first study, Bayesian correlation pairs analyses were then carried out in patients (Figure 7.5) with a conservative beta prior of 0.5. For patients, the correlations were between the force-matching overcompensation score and SOPS delusional ideation score (Figure 7.5, Panel A), and between switching tendency and SOPS overall positive symptoms score (Figure 7.5, Panel B). The relationships were hypothesised to be negative in the first correlation pair and positive in the second one. In the first correlation, Bayes Factor for the null hypothesis was favoured 2.78 times over the hypothesis that there was a

negative correlation between delusional ideation and overcompensation, with anecdotal to moderate levels of evidence. By contrast, Bayes Factor favours the hypothesis that tendency to switch was positively correlated with positive symptom scores by 1.93-fold over the null hypothesis, albeit with only an anecdotal level of evidence.

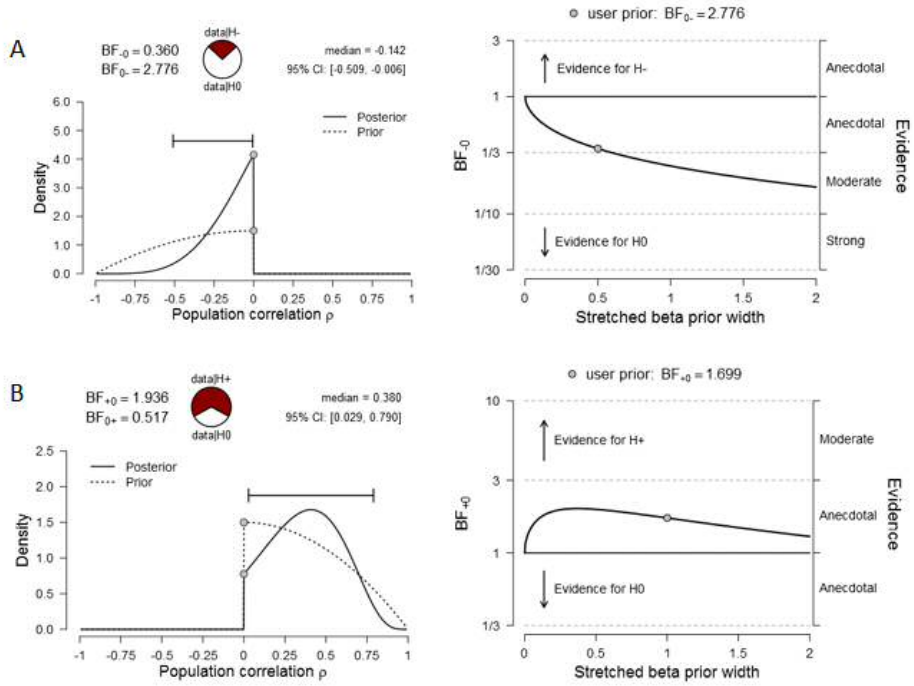

Figure 7.5. Results from Bayesian correlation pairs analyses for patients. Panels A and B show results for the force-matching, and reversal learning tasks respectively. CI, Credibility Interval; BF, Bayes Factor.

7.6. Discussion

In this Chapter, a shortened version of the tasks aimed at eliciting behaviourally measurable prediction error responses used in the first study (Chapter 5) were adopted in a case-control study. In the force-matching task, the overcompensating effect was evident in both patients and controls. Frequentist statistical methods found no significant differences between patients and controls in overcompensating for the force applied by the machine in the Finger condition, and Bayesian methods favoured the null hypothesis with anecdotal evidence. This did not replicate Shergill et al. (2005)'s behavioural findings of less overcompensation and therefore less sensory attenuation in the Finger condition in patients. Relationships with positive symptoms were not investigated in Shergill et al.'s original study in established schizophrenia patients, but failures to predict self-initiated actions may be most relevant to 'first-

rank' symptoms such as delusions of control and passivity phenomena. However, in a follow-up study, Shergill et al. (2014) found significant negative relationships between current hallucination severity and sensory attenuation using a different task and method (functional MRI). Individuals experiencing auditory-verbal hallucinations often do not recognise their voices as coming from their own mind but attribute them to an external agent as well, although the reasons for such an externalising attribution are multi-layered which may involve metacognitive beliefs and mechanisms other than basic disruptions in self-monitoring of inner speech (Jones and Fernyhough, 2007; Georgieff and Jeannerod, 1998; Morrison, Haddock and Tarrier, 1995). Bayesian methods, on the other hand, demonstrated the highest Bayes Factor supporting the null hypothesis that overcompensation and delusional ideation were not negatively correlated, with moderate levels of evidence favouring the null hypothesis.

In the reversal learning task, although patients made significantly more errors overall than controls, this did not reach the threshold for statistical significance when the errors were divided into post-probabilistic errors and post-true reversal errors. Switching tendency or sensitivity for post-probabilistic errors can be indexed by accuracies after such errors, and has been found to be more pronounced in unmedicated schizophrenia patients (Schlagenhauf et al., 2014) experiencing a relatively high level of positive symptoms. However, this relationship was not significant in the current sample of patients, who, despite never receiving a firm diagnosis of schizophrenia, were mostly under antipsychotic treatment (although less than 12 months) and with low levels of active symptoms.

A significant correlation between perseverative behaviour and hallucinatory symptoms was found in patients even after controlling for IQ (a factor which is thought to heavily influence reversal learning performance; Wolff, 1967). Although much caution needs to be borne in mind when attempting to interpret results from post-hoc exploratory analyses (Table 7.2), this relationship seemed to have been driven specifically by hallucinations rather than general positive symptoms of the first-rank. Perseverative behaviours, interestingly, were not associated with negative symptoms. Evidence for 'mapping' cognitive-behavioural deficits onto specific symptom domains has only been weak; in other words, the strengths of relationships between observed deficits in behaviour and symptom severity are not consistently high (Nieuwenstein, Aleman, and de Haan, 2001; Rund et al., 2004). Many other factors

including symptom levels at testing, the exact measurement tools used, the specific cognitive tasks employed could all affect the correlations between symptomatology and behavioural deficits. First episode psychosis patients might be 'further along' the psychosis continuum and antipsychotic medications may increase perseverative behaviour while making little improvement for decision-making (however, evidence is again inconsistent and may depend on a specific medication's receptor binding profile; see Meltzer, Thompson, Lee, and Ranjan, 1996; Meltzer and McGurk, 1999), but are often effective against perceptual distortions such as hallucinations.

There is a small possibility that some patients had difficulties understanding relatively demanding written and verbal instructions, despite that high fluency in English was an inclusion criterion. On average, patients' IQ was significantly lower than that of controls even though years spent in formal education were comparable. Low intellectual ability has been shown as a risk factor for psychosis in longitudinal studies (David, Malmberg, Brandt, Allebeck, and Lewis, 1997; Zammit et al., 2004) and this could have had an impact on how well the patients understood instructions. In the force-matching task, one of the patients excluded had the lowest IQ of 57 and the other was consistently exerting forces as high as 11N despite having an average IQ. This complicates the interpretation that low IQ is necessarily a barrier to following instructions accurately.

It is also rather unlikely that patients were distracted by cognitive intrusions (e.g. auditory verbal hallucinations) as none of the patients appeared to be responding or attending to external stimuli and their reported positive symptom levels were somewhat low. It may be the case that they had grown 'used to' their voices and delusions; alternatively, active non-reporting of symptoms and a discrepancy between actual experience and patients' description is another possibility. The latter explanation is unlikely however, as rapport was visibly established in every interview and patients would have already been assessed and spoken about their symptoms on multiple occasions to their clinical teams.

The current study was under-powered due to practical constraints in patient recruitment within the specified timescale Thus the results outlined must be viewed as preliminary. However, Bayesian statistical methods are thought to be better at overcoming the problems usually posed by small sample sizes (Dienes, 2011). The current study has demonstrated moderate levels of evidence in at least some of the

Bayesian correlation pairs (i.e. reversal learning), indicating that there is likely to be a true relationship detectable between symptom measures and task performance. Such relationships would likely emerge and reach frequentist statistical significance if the study had been fully powered.

Relationships with other symptom domains, such as disorganisation, general/affective symptoms and global assessment of functioning (GAF) scores, were not explored. Even though such data were available anyway just from conducting the whole SIPS, this was due to deliberately starting from *a priori* hypotheses and limiting the number of correlations, Bayesian or otherwise, carried out in the analyses. A higher number of correlations (i.e. 'correlating everything with everything') would inflate familywise error rates and hence false positives, which would make any significant finding uninterpretable. As such, confirmatory and exploratory analyses were deliberately set apart. As mentioned in Chapter 5, although Bayesian methods are considered more resistant to multiple comparison problems, unnecessary correlations are at best questionable, especially given the specific predictions and hypotheses set out from the very beginning.

Chapter Eight: Source Monitoring of Actions in Early Psychosis

8.1. Abstract

This study aimed to investigate source monitoring abilities of actions in individuals experiencing the early signs of psychosis. Ten early psychosis patients and ten matched healthy controls took part in an action source memory task where the participants were asked to perform, imagine, or watch the experimenter perform various simple actions. At test, the participants were asked to remember whether they performed the action, imagined the action, if the experimenter performed the action or if the action was new. When compared to controls, patients were not significantly more impaired in task performance in any measures of memory. No statistically significant correlations were found when patients' levels of current psychotic symptoms were correlated with the numbers of both reality monitoring and internal source monitoring errors. However, patients' current hallucination ratings were significantly correlated with false alarm rates, and patients were also significantly slower at making a source judgement. Given the study was under-powered, any findings must be interpreted with caution. These mostly null findings can however inform future, larger studies in the early psychosis population.

8.2. Introduction

It is widely accepted that cognitive difficulties associated with established disease states, in particular alterations in memory functions, are present at least modestly in the prodromal phases of psychosis (Simon et al., 2007). As described in Chapters 3 and 6, patients with psychosis have persistent problems with memory across different domains, but source monitoring is one particular aspect of recognition memory with which patients struggle significantly. Deficits in source memory have consistently been associated with positive symptoms at least in clinical populations, especially auditory-verbal hallucinations and verbal source monitoring (although evidence is less consistent with healthy schizotypy; see Chapter 6). However, studies on source

monitoring in individuals experiencing early phases of psychosis are relatively scarce in comparison to those in established schizophrenia.

Given the importance of targeted and effective intervention in early psychosis, it is crucial to study specific cognitive deficits in this population. The current study acts as a follow-up to that described in Chapter 6 and investigates source monitoring of actions in individuals who have experienced a first episode of psychosis. The word pair source monitoring task has been omitted due to the observation that no significant relationships were found between source memory measures and schizotypy in the previous study with this task, but they were found with the action memory task.

In Chapter 6, the positive dimension of schizotypy (i.e. psychosis-like experiences) was significantly positively associated with errors in reality monitoring as well as errors in internal source monitoring. Thus, the same types of deficits are hypothesised to be present in the early psychosis patients too. Consistent with the majority of previous literature, patients would be more impaired than controls in recollection-based task performance measures such as overall source accuracy but not in familiarity-based task performance measures. Hypotheses relating source monitoring performance and symptomatology were that 1) the number of reality monitoring (in particular, errors involving misattributing participant performed actions to an external source i.e. the experimenter); and 2) internal source monitoring errors (confusing participant performed actions with participant imagined actions) made by early psychosis patients would significantly correlate with their current scores in hallucinations severity and more generally, positive symptom scores.

8.3. Methods

8.3.1. Participants

This study was conducted with the same matched patients and controls as those in Chapter 7 (10 in each group). All patients and controls performed to a satisfactory level, exceeding a threshold of 0.1 above chance (i.e. more than 0.1 for corrected recognition and source accuracy of more than 0.43).

8.3.2. Procedure

Similar to the source memory task used in Chapter 6, the present task consisted of a study phase (approximately 8 minutes) and a test phase (varying between 5 – 10 minutes), with a 100-minute interval in-between. The study phase for the source memory task was the first to be carried out in the battery (after signing consent forms and filling out demographic details) and the test phase was the last behavioural component of the battery (before clinical interview and debriefing).

8.3.3. Action Source Monitoring Task

This task was modified from that used in Chapter 6 in two ways. First, the total number of actions in the study phase was reduced from 75 (25 in each condition) to 45 (15 in each condition) and the number of new items added in the test phase was reduced from 25 to 15. Approximately half of the excluded actions involved using actual items whereas the other half were without items. Second, unlike the previous task where participants had a maximum of 2000ms to make a response in the test phase, items in the current task were shown indefinitely and would only progress to the next item once the participant had made a choice. The reasons for these two modifications were to prevent a high number of missing data as patients tended to respond much slower, and to lessen the potential burden or cognitive load on patients so that they would not consistently perform at floor levels.

8.3.4. Assessment of Symptoms

SOPS and O-LIFE were the measures for early psychosis symptoms and schizotypal experiences for patients and controls, respectively. Details about sample characteristics and mean scores can be found in the previous Chapter (Table 7.1).

8.4. Data Analysis

Similar to the previous study on source monitoring (Chapter 6), most of the data obtained in the current task did not follow a normal distribution (Shapiro-Wilk test, ps < 0.05). Therefore, nonparametric tests were chosen for all statistical analyses, including Mann-Whitney's U for comparing performance between patients and controls, Spearman's rho (ρ) for correlations, and the Friedman's test (nonparametric

two-way ANOVA). All frequentist statistics was carried out in SPSS 23. In addition to frequentist statistical methods, a Bayesian independent samples t-test was carried out in JASP 0.8.0.0. to assess the levels of evidence for differences in source memory errors between patients and controls.

8.5. Results

8.5.1. Overall Task Performance

The same measures for hits, false alarms, corrected recognition (hits minus false alarms) and source accuracy were calculated for patients and controls. Table 8.1 below summarises these measures.

	Patients	**Controls**
Hits	0.79 (0.18)	0.82 (0.08)
False Alarms	0.11 (0.12)	0.04 (0.03)
Corrected Recognition	0.68 (0.22)	0.78 (0.10)
Source Accuracy	0.86 (0.13)	0.89 (0.08)

Table 8.1. Mean proportions with standard deviations in parentheses for patients and controls in the action source memory task.

On first inspection, controls did perform better than patients in every measure as demonstrated by higher scores of hits, corrected recognition and source accuracy in controls, as well as nearly 1/3 of the false alarms rate as compared with patients. In other words, false alarm rates in patients were approximately three times higher than those in controls, at least on a numerical level. However, Mann-Whitney's U tests revealed no statistically significant difference between the any measure of performance in familiarity or recollection between patients and controls (all ps > 0.200). In terms of the errors made in each condition, patients did not significantly

make more errors than controls in reality monitoring or internal source monitoring when the numbers of errors for each type of error were added together (Mann-Whitney's U, all ps > 0.350). For example, although numerically patients made more errors than controls by misattributing participant performed (PP) to participant imagined (PI), and once again attributing experimenter performed (EP) to participant performed (PP), these differences diminish when pooled together with other errors in internal source monitoring and reality monitoring, respectively, Figure 8.1. displays the number of errors made in each condition.

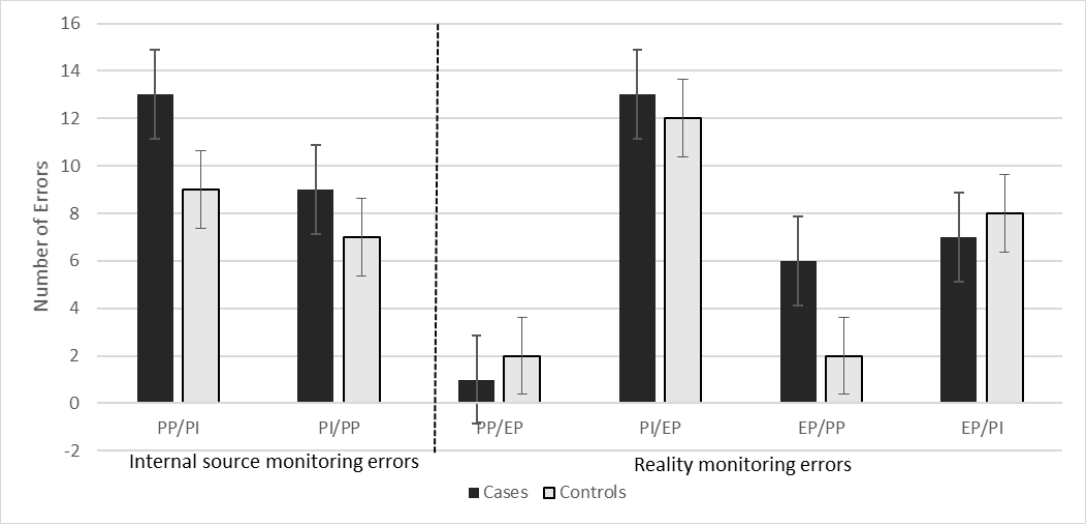

Figure 8.1. The mean number of errors produced by patients (black bars) and controls (light grey bars) with error bars (± SEM). Abbreviations are as follows: PP (Participant Perform), PI (Participant Imagine), and EP (Experimenter Perform). The first abbreviation is the actual source of the event and the second one is what the participant stated.

In fact, the number of reality monitoring errors that were of the most theoretical interest (i.e. misattributing participant perform to experimenter perform, or PP/EP errors) were very low across patients and controls. The number of internal source monitoring errors (PP/PI and PI/PP), on the other hand, were numerically much higher in patients and to a lesser degree in controls, although there was no statistically significant difference in the number of this kind of source memory errors made between the two groups. In terms of source monitoring errors, patients did not display a tendency to misattribute internally generated events to an external source; if anything, patients made more errors in the other direction (attributing external events internally) in the current sample.

8.5.2. Bayesian Statistics

Bayesian independent samples t-tests was conducted in order to compare the numbers of reality monitoring errors (Panel A) and internal source monitoring errors (Panel B) made between patients and controls, using a default Cauchy prior of 0.707 (which is generally accepted as a distribution of effect sizes considered realistic under the alternative hypothesis, H1). As can be seen in Figure 8.2 (A), the Bayes Factor for the null hypothesis is 2.516 compared to that for the alternative hypothesis which is 0.397. This means that the null hypothesis (i.e. patients did *not* make more reality monitoring errors than controls) is favoured 2.5 times over the alternative hypothesis, which is supported by anecdotal to moderate levels of evidence under a range of priors. For internal source monitoring errors a similar pattern was found: the Bayes Factor for the null hypothesis is 2.475 compared to that for the alternative hypothesis which is 0.404, indicating that patients did not make more internal source monitoring errors than controls (also with anecdotal to moderate levels of evidence).

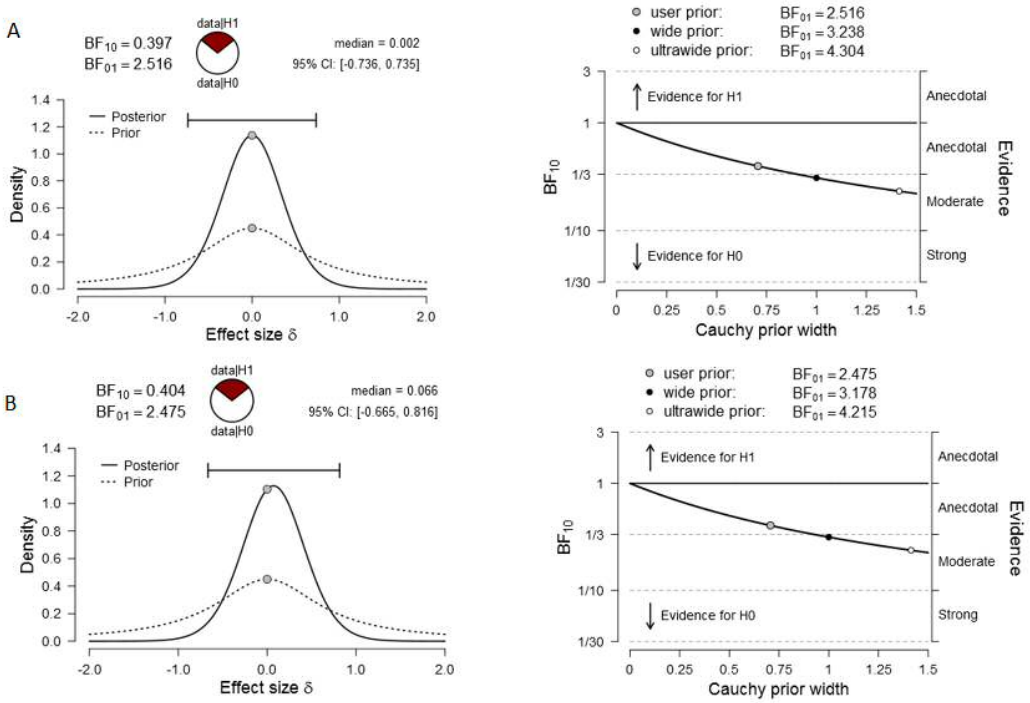

Figure 8.2. Results from Bayesian independent samples t-test for source monitoring in patients and controls. Panel A shows the results for reality monitoring errors and Panel B for internal source monitoring errors. BF, Bayes Factor; CI, Credibility Interval.

8.5.3. Source Memory Errors and Symptomatology

Two-tailed Spearman's rho correlations were carried out between positive symptoms (delusion and hallucination scores as measured by the SOPS in patients) and numbers of source memory errors (reality monitoring and internal source monitoring) made. No significant correlations were found (ps > 0.089) for either measure. As an exploratory analysis, hallucination scores in patients were significantly positively correlated with false alarm rates [$\rho(8) = 0.640$, $p = 0.046$] Nevertheless, this significant result between hallucination score and false alarm rate would not have survived multiple comparison corrections. Delusion scores did not correlate significantly with any task measure (ps > 0.440).

8.5.4. Relationship between Clinical Status and Reaction Time

Further analyses were carried out to investigate whether there might have been a speed-accuracy trade-off and whether this may better account for the absence of any statistically significant difference in task performance between patients and controls. This was particularly important given the removal of the 2000ms response window, which meant participants could in theory take as much time as they felt necessary even though the instructions clearly stated 'as quickly and as accurately as possible' when making a response. The need to respond quickly and not to *over*think about the correctness of responses was further emphasised by the experimenter prior to the test phase.

Inspecting raw reaction time data across conditions revealed that patients were in general slower at responding than controls, however, this was skewed by a handful of very long responses (8000 – 12000ms). Before reaction time data was divided into the four conditions (Participant Performed, Participant Imagined, Experimenter Performed and New) 5 responses which took longer than 8000ms were excluded from the dataset. Fig. 8.3 shows the mean reaction times for different types of source hits in patients and controls after excluding outliers. Reaction time data were also divided into different types of errors after removing 8 responses longer than 8000ms (Figure 8.4).

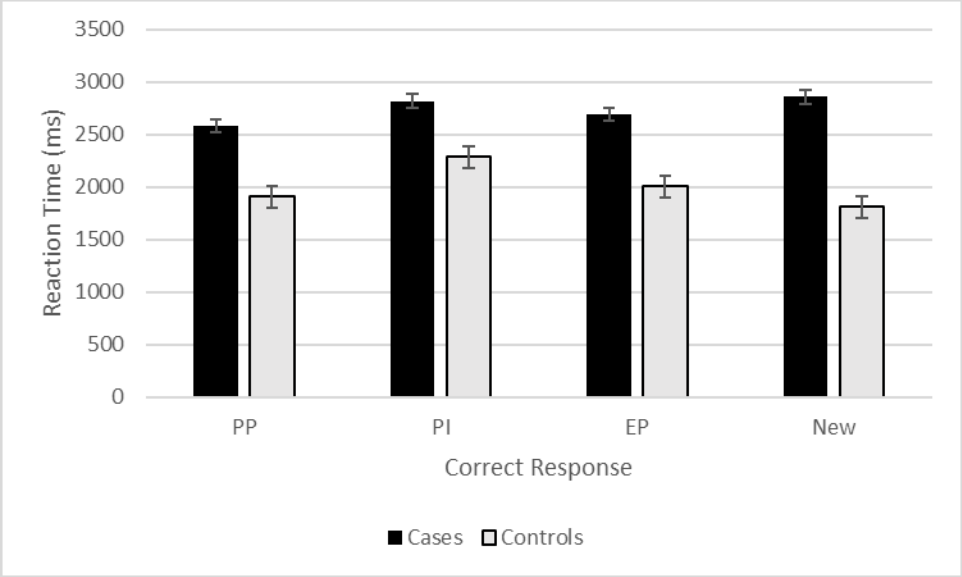

Figure 8.3. The mean reaction times for correct responses produced by patients (black bars) and controls (light grey bars) with error bars (± SEM) after removing 5 outliers (> 8000ms). Abbreviations are as follows: PP, Participant Perform; PI, Participant Imagine; EP, Experimenter Perform.

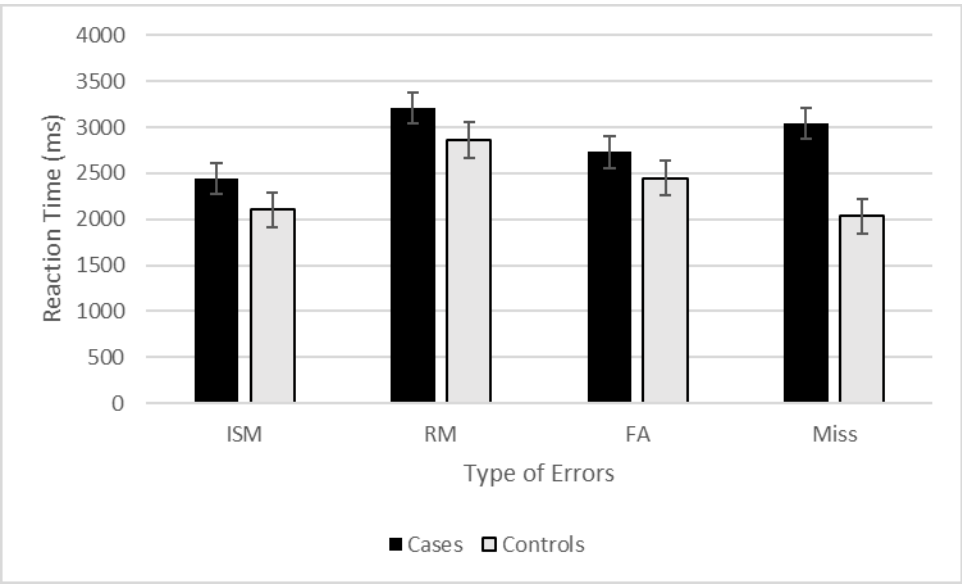

Figure 8.4. The mean reaction times for errors produced by patients (black bars) and controls (light grey bars) with error bars (± SEM) after removing 8 outliers (> 8000ms). Abbreviations are as follows: ISM, internal source monitoring errors; RM, reality monitoring errors; FA, false alarm errors.

A Friedman nonparametric two-way ANOVA was performed with reaction time as the dependent variable, clinical status (patient/control) and correct response type (participant perform, participant imagine, experimenter perform and new) as fixed factors yielded a highly significant main effect of clinical status [$\chi^2(7) = 81.10$,

$p = 0.022$, $p < 0.001$]. A second two-way ANOVA was carried out with reaction time as the dependent variable, clinical status (patient/control) and error type (internal source monitoring error/reality monitoring error/false alarms/misses) as fixed factors; however, this did not yield a significant main effect of clinical status [$\chi^2(7) = 6.22$, $p = 0.514$].

8.6. Discussion

The current study investigated source monitoring of actions in early psychosis patients and healthy controls matched for gender, age and years spent in education. Patients were able to perform the task to a satisfactory level, and did not demonstrate more impairments in familiarity- and recollection-based task performance when compared with controls at a statistically significant level. Patients did not make significantly more reality monitoring or internal source monitoring errors either, and none of the types of source monitoring errors were significantly related to psychotic symptomatology. From exploratory analyses, the only significant association was in fact between current hallucination scores and false alarm rates in patients. There was a potential speed-accuracy trade-off in patients as they were significantly slower in making a correct source judgement, but not when making errors, compared with controls across the four conditions.

The current study was under-powered. As shown in the power calculation in Chapter 7, based on previous studies the minimum number of patients required to detect an effect was 28. Therefore, the largely null findings were far more likely to be due to a lack of power rather than study design, for example, as significant associations were found in a much larger sample (Chapter 6) of healthy volunteers using a very similar design.

It was interesting that false alarm rates were related to hallucinations. Schizophrenia patients have been shown to display greater confidence in incorrect source memory responses (Moritz, Woodward, and Ruff, 2003) although this has not been related to positive symptoms. False memories have also been implicated in the pathogenesis of delusions (Moritz and Woodward, 2002; Bhatt, Laws, and McKenna, 2010). Nevertheless, the current study did not directly measure memory confidence

and as such, the relationship between false alarm rates and hallucinations could only be interpreted speculatively.

Taken together, the null findings in the current study are most likely the consequences of lacking statistical power and low symptom levels as indicated in the previous Chapter. Had more early psychosis patients been recruited in the timeframe, power would have been increased and Type II error rates (false negative findings) reduced. It should, however, serve as an indicator of the numbers of patients one may need for future studies, perhaps even in collaboration with other Universities to maximise recruitment. Still, some of the observations in the current study may point towards interesting new directions of inquiry (e.g. confidence in false memory and how it may be related to psychopathology). Longitudinal studies may also reveal how deficits in source monitoring 'evolve' as some of the early psychosis individuals transition to florid and more persistent states of psychosis.

Chapter Nine: General Discussion

9.1. Summary of Findings

The current Thesis examined behaviourally measured prediction error responses and source monitoring processes in healthy individuals with schizotypal traits as well as in individuals experiencing the early signs of psychosis. In the first study (healthy schizotypy) prediction error responses were measured in the sensory, associative and reward domains whereas the source monitoring tasks focused on that of action and words. In the second study (clinical group) the behavioural task battery was downsized to sensory and reward prediction and action source monitoring only, based on results from the first study.

In healthy individuals, there was no marked perturbation in predictive processing in those with high schizotypal traits (Chapter 5). This was demonstrated by normal levels of sensory attenuation in the force-matching task, no evidence for a reduction of the blocking effect in the Kamin blocking task, and no evidence for increased switching tendency in the reversal learning task. Performance levels in these tasks were not significantly correlated with various domains of schizotypy (delusional ideation, hallucinatory experiences, general positive and negative schizotypy) as measured by different psychometric scales. On the other hand, in the same healthy individuals, source monitoring tasks of action but not of words elicited deficits significantly associated with high positive schizotypy (Chapter 6). Such deficits were observed in reality monitoring (misattributing participant performed actions to the experimenter) as well as internal source monitoring (confusing participant performed actions with participant imagined actions). These findings add to the increasing body of research that deficits in cognition often seen in clinical states of schizophrenia manifest in individuals with no need for care, and support the continuum model of psychosis vulnerability.

Nevertheless, due to the difficulties encountered in participant recruitment only a very small sample of patients was recruited in the clinical study (Chapters 7 and 8). As a result, the under-powered nature of the second study meant that it was unable to detect significant differences in performance in a subset of the behavioural tasks

(force-matching, reversal learning and action source monitoring) between individuals with early/first episode psychosis and healthy controls matched for gender, age, handedness and years of education. Performance in patients was not generally associated with current symptom levels (which were relatively low in the current sample), however, some patterns have emerged. In the reversal learning task, perseverative behaviour was significantly correlated with the level of positive symptoms (combined score of delusional ideation, suspiciousness, perceptual abnormalities, grandiosity and disorganisation) but switching tendency did not correlate significantly with symptomatology.

9.2. Interpretation and Implication of Findings

In Study 1, the main hypotheses for the predictive processing tasks were that scores from three different schizotypy scales would correlate with domains of behaviourally-measured prediction error responses. In the force-matching task, although participants did indeed consistently apply more force in the Finger condition which demonstrated the overcompensation effect due to sensory attenuation of self-generated actions, this effect was not negatively correlated with their PDI-21 scores. Similarly in the Kamin blocking task, although the blocking effect was present (evidenced by a lower rating for stimulus B than that of stimulus D given by participants) this was not significantly correlated either with O-LIFE positive or negative schizotypy dimension, and did not demonstrate the positive association with PDI-21 distress subscale first shown by Corlett and Fletcher (2012). In addition, after replicating Haselgrove and Evans (2010)'s methodology and performing a median split, blocking effect was not significantly different from high or low schizotypy across neither the positive nor the negative dimension. In the reversal learning task, PDI-21 scores were correlated with neither post-probabilistic error switching nor post-true reversal perseveration tendency. Bayesian analyses further provided supportive evidence for the null effects as reflected by Bayes Factors calculated from each of the correlation pairs.

The other part of Study 2, namely those focusing on the source monitoring of actions and words, did find statistically significant associations between the numbers of reality monitoring/internal source monitoring errors and O-LIFE unusual experiences scores and such effects were specific to the action task. In addition, high

levels of positive schizotypy were associated with poorer old/new recognition and overall source accuracy measures, again only limited to the action task.

Study 1 was undoubtedly better powered than most previous studies of this kind; it also had the advantage of employing a full battery of tasks in the same individuals. However, previous studies in prediction error response and schizotypy did find significant relationships despite having a lower power. Such null findings from the present study may indicate that the concept of healthy schizotypy is noisier than previously thought, or even due to publication biases where only positive findings were published in the literature. The specificity of deficits in the action source monitoring in relation to positive schizotypy may also mean that impairments in action source monitoring were more pronounced and perhaps less latent than prediction error responses. However, this would not explain why no associations were found in the source monitoring of word pairs. One possible explanation would be that O-LIFE unusual experience as a general schizotypy measure was not sensitive enough for word pair monitoring deficits, which might only be 'picked up' by more specific measures of hallucinatory experiences in the auditory-verbal modality, for example. Again, this points towards the noisiness and non-specificity in at least some of the schizotypy measures currently used.

Confirmatory analyses in Study 2 did not find specific associations between symptomatology measures and behavioural responses to predictive processing or source monitoring of actions either; however, it must be emphasised that this study was severely under-powered, which was the most likely explanation for the failure to demonstrate significant associations. Another likely explanation was that most patients were already in remission due to antipsychotic treatment, which was evidenced by a higher rating in all SOPS measures at peak/first onset. Patients did perform significantly worse across the reversal learning and source monitoring tasks, and the association between perseveration and hallucination scores from exploratory analyses remained significant even after controlling for IQ, where in the latter scenario patients seemed to demonstrate speed-accuracy trade-off when making correct source judgements. However, due to the fact that nonparametric ANOVA was used, it was not feasible to assess interactions, hence although interpretation of main effects indicated a speed-accuracy trade-off effects, interaction effects could have also played

a role. This further complicated the evaluation of the current findings, which were already limited because of the lack of power.

Despite the under-powered nature of the clinical study, findings from the two studies in the current Thesis nevertheless shed some new light on the cognitive mechanisms in two different stages across the psychosis continuum, namely healthy schizotypy and early signs/first episode of a psychotic disorder. The results point towards differential manifestations of behavioural deficits in predictive processing and source monitoring, in that source monitoring deficits are more salient than perturbations in prediction error responses in these individuals. Although it could be seen that the lack of evidence for the latter is a kind of counterevidence for the psychosis continuum, the concept itself is not at odds with the observation that nonclinical and subclinical psychosis-like experiences are distributed in healthy individuals without a need for clinical care. It is true that such a distribution is heavily skewed towards 'normal experience', but the continuum approach does not need to be fully dimensional. An absence of related cognitive-behavioural deficits in terms of prediction error responses does not necessarily challenge the notion of a continuum (or even continua) of psychosis-like experiences, given that deficits in source monitoring were clearly demonstrated in the same group of individuals.

The findings of the present research are also of clinical relevance, especially for the development of potential tools for psychological interventions such as psychoeducation. Psychoeducation for schizophrenia has been shown to be at least moderately beneficial for treatment (especially antipsychotic medication) adherence, relapse prevention and reducing subsequent hospitalisation (Xia, Merinder, and Belgamwar, 2011). One particular application of the current findings in a psychoeducational framework might be focused on the continuum nature of psychotic experiences (that psychosis lies within a spectrum of human experiences, and healthy individuals could have attenuated symptoms too), thus potentially alleviating the psychological burden and stigma associated with being diagnosed with schizophrenia. The very observation that patients are not categorically different deviations from 'normality' may just provide hope and reassurance that they will recover in the future and return to the other, healthier end of the continuum.

Another application of the findings to psychoeducation may be that mental illnesses are not the sole products of either the brain or the environment, but a

combination of many heterogeneous and highly complex factors. For example, a deficit in source monitoring or an externalising bias may be the result of a genetic predisposition for difficulties with episodic memory, but may also be the consequences of dealing with stressors such as dissociating oneself from negative emotions or events in life. It is crucial to understand and accept that there is no single theory or framework that can fully explain every aspect of psychosis, neither is there a treatment (biological or psychosocial) that works for every patient. As such, facilitation of personalised medicine will not only help with symptom reduction but will also greatly improve the outlook of recovery. Sadly, an unhealthy competition for some kind of political supremacy persists between a minority of practitioners in biological psychiatry and in clinical psychology (to use one example). Such unnecessary friction between therapeutic professions is not only without scientific bases but also highly detrimental, if not unethical, when providing care for vulnerable patients. To acknowledge the coexistence of both the physical and the psychological - and to relinquish the dualist notion as if the brain was not a part of the whole person – may just be a simpler way forward.

A key question for consideration for future research would be whether source monitoring is at least related to the detection of prediction error, as the two concepts clearly have some overlap (Griffin and Fletcher, 2017). Theoretically, although the predictive processing approach is fundamentally reductionist and may not be able to capture the full complexity, nature and subjective reality of the psychotic experience (Corlett et al., 2010), it (alongside other accounts) does provide important insight into at least some aspects of how delusions and hallucinations may be generated. Figure 9.1. below shows a tentative diagrammatic representation for this relationship:

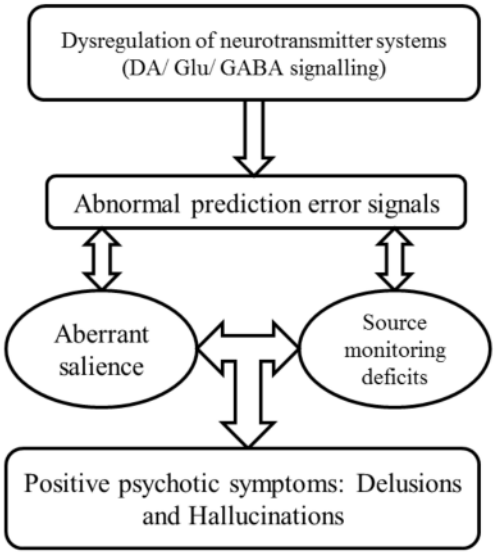

Figure 9.1. The relationship between prediction error and downstream cognitive factors. DA, dopamine; Glu, glutamate; GABA, gamma-aminobutyric acid.

As the diagram shows, aberrant prediction error signalling is the centre stage which is the result of dysregulated neurotransmitter systems (in particular dopamine and glutamate). Faulty and imprecise prediction error signalling then leads to two related processes, abnormally heightened salience and deficits in source monitoring which can feedback to update the prediction error signal. The basis for the relationship between salience and source monitoring is that aberrant salience (originating from dopaminergic hyperfunction mediated via error signalling) drives new yet incorrect associations about all the incoming stimuli to be formed, including the (mis)attribution and allocation of stimuli to their perceived source which is in essence also an association between the representation of the stimulus and its (perceived) origin. With these two factors combined, delusions and hallucinations may occur as the consequence of one's often futile efforts to seek meaning in a confusing and threatening world. As mentioned previously, some delusions and hallucinations can 'feed' into each other (e.g. the persecutory delusional content and threatening voices) to form a seemingly inescapable vicious circle in which the faulty prediction error signals are sent higher and higher up the hierarchy of abstraction, and a recent model of 'circular inference' has supported this (Jardri and Denève, 2013; Denève and Jardri, 2016; Jardri, Duverne, Litvinova, and Denève, 2017).

Alternatively, alternations in source monitoring especially when the sense of agency is in question could be a consequence of (or at least related to) disrupted

prediction error signalling which is again not always a pathological phenomenon. When viewed through the lens of the predictive processing framework, problems with agency as seen in source monitoring paradigms may be at least partly explained by inadequate detection and minimisation of prediction error. This theorisation may be particularly important for explaining the faulty monitoring of self-generated actions seen in both patients with schizophrenia and healthy individuals with high levels of schizotypal traits.

Furthermore, not all types of delusions and hallucinations can necessarily be explained by the prediction error approach. Therefore, future research should perhaps allow the fine-tuning or subtyping of symptoms in order to delineate the deeper mechanisms. In a recent review by Griffith, Langdon, Le Pelley, and Coltheart (2014), the authors have written extensively on how the prediction error signalling approach best explains delusions of control and (some) hallucinations, linking predictions with source monitoring as both are important to the discrimination of self- versus other-generated stimuli whilst providing a careful critique and re-examination of some of the findings by Corlett and colleagues. Further, the authors stress that it is not always the stimuli themselves that matter but the meaning, associations and consequences they create for the patient which give such stimuli personal significance. In fact, Corlett et al., (2014) also support the interdependent relationship between source or reality monitoring and prediction error signalling: the former is thought to be modulated by dopaminergic transmission as well, most likely as a result of aberrant prediction error signalling, and a dysconnection in the frontostriatal circuit has been implicated in both prediction error signalling and source monitoring deficits.

In the current Thesis, neither the schizotypy psychometric scales nor the symptoms assessment in patients with early psychosis differentiated subtypes of delusions and hallucinations beyond broadly paranoid, first-rank and grandiose type delusions and sensory modalities (rather than content) of hallucinations. Relationships between different domains of prediction error responses and positive symptoms were not found in healthy volunteers with schizotypy or early psychosis patients, but significant associations were revealed in the source monitoring of actions in healthy participants between positive schizotypy and internal as well as reality monitoring deficits. It is possible that because most patients were treated with anti-psychotic medications (dopamine antagonists), deficits in prediction error signalling have

diminished leading to a downstream restoration of source monitoring abilities, whereas in healthy (i.e. non-medicated) volunteers, any deficits seen were not affected by dopaminergic modulation. Nevertheless, if prediction error signalling is truly upstream from source monitoring and there is some kind of cascading relationship between the detection of prediction error and reality monitoring for example, deficits in both processes would have been found in healthy volunteers with positive schizotypy. This could have been the case in the clinical study had it been fully powered, but even in the well-powered healthy volunteer study there did not appear to be significant relationships between performance in the prediction error tasks and various psychosis-like experiences.

It has been suggested that auditory-verbal hallucinations come from strong top-down priors where the patient believes that voices will occur (Powers, Mathys and Corlett, 2017). With regards to hallucinations however, subtyping (Garwood, Dodgson, Bruce, and McCarthy-Jones, 2013; McCarthy-Jones et al., 2014) may be a useful approach in trying to apply the prediction error signalling framework to some of the AVHs where source monitoring theories alone may not be sufficient. Source monitoring deficit accounts posit that faulty monitoring of self-generated inner speech leads to the external misattribution of these internally produced stimuli as sensory percepts: for example, whilst traditional source monitoring accounts are able to account for 'inner speech hallucinations' where attention is inwardly directed, 'hypervigilance hallucinations' are not so well accounted for as badly monitored inner speech due to their outwardly directed attention.

Perhaps a complementary approach can utilise prediction error signalling theories in which hypervigilance hallucinations are viewed as excessive error signals arising from real external stimuli (e.g. traffic noise) which drive the update of prior beliefs about these stimuli in order to minimise the error signals. The result of this updating process is that the conscious percept becomes voices or at least voice-like, although there is still debate over whether AVHs are truly auditory in nature. Indeed, the existence of such hypervigilance hallucinations itself defies the traditional definition of a 'true' AVH as the hallucination must occur spontaneously without any stimulus in external reality which adds to the complex heterogeneity of the psychotic experience.

In some studies of source monitoring, participants are asked to rate the probability or confidence with which they think a certain event is of endogenous of exogenous origin (Moritz, Woodward, and Ruff 2003; Mitchell and Johnson, 2000). Gerrans (2014) posits that such paradigms are 'congenial to the Bayesian model, which treats source monitoring as the detection of prediction error' (p. 93), where the probability ratings form priors for the Bayesian belief updating system. Nevertheless, the definition of source monitoring here is perhaps less based on memory but on the determination of agency (self versus other) according to, for example, the forward model (see Chapter 1). Indeed, according to the Bayesian framework for predictive processing (Chapter 2), the misattribution or externalisation of agency may well be the consequence of detecting prediction errors when there should be none. Such error signals are then propagated to the higher hierarchies of the Bayesian iterative model which call for an update of the current inferences, resulting in erroneous judgements about the origin of a self-generated occurrence. This model has been used extensively to explain ego-boundary disorders in schizophrenia; most prominently, delusion of alien control (Blakemore, Oakley, and Frith, 2003; Frith, 2005, 2012) where the patient misattributes self-generated actions to an external (often malign) force controlling their behaviour.

However, can the same model be applied to explain disorders of thought interference, which are diagnostically meaningful 'first-rank' symptoms of schizophrenia? For example, thought insertion is the phenomenon where instead of overt actions, internal thoughts and mental events are misattributed to an external source which often has the power to control the patient's mental processes (rather than direct behaviour). It is widely acknowledged that 'normal' or one's own thoughts are immune to error through misidentification relative to the first-person pronoun (the immunity principle): that is, 'when a speaker uses the first-person pronoun ('I') to refer to him or herself, she cannot make a mistake about the person to whom she is referring' (Gallagher, 2000; Shoemaker, 1968). Hence the same efference copies which would have been generated for actions (according to the forward model) would not work in the same way for thoughts because in normal circumstances at least, all thoughts are internally generated and one would not have to work out who did the thinking. As such, even if a thought appears alien or strange, a top-down process should be initiated so that any prediction error signals generated by bottom-up

information will be minimised if not cancelled out. In fact, many theorists do not agree that the comparator model can be extended to explain thinking (e.g. Vosgerau and Newen, 2007; Vicente, 2014) whereas others are firm proponents of the 'thoughts as motor acts' model where the comparator or forward model does apply (e.g. Campbell, 1999) thus generating a great deal of debate.

Still, another aspect of agency misattribution seen in schizophrenia, namely that of exaggeration or internalisation errors, does not seem explainable by the source monitoring framework without taking into account its (assumed) Bayesian nature. For example, delusions of reference where the self takes the position of heightened centrality in receiving external stimuli which are imbued with significance and self-relevance is phenomenologically the opposite to passivity symptoms where the self loses control over one's own actions and the patient is convinced that other agents can influence their self to the extent one's sense of agency diminishes. Traditional source monitoring accounts would fail to explain how within the same person, the sense of agency can be both inflated and compromised sometimes simultaneously; however, in a very recent paper, Asai (2016) shows for the first time that the over- and under-attribution of agency where, depending on the S/N ratio (self/other ratio) of the stimuli, only self-generated signal would need to be detected by one's sensorimotor system. The S/N ratio might also be determined by the "embodiedness" of the action: embodied action (self) would produce a higher S/N ratio, while disembodied (other) action would produce a lower one in terms of motor prediction. Further, the author points out that the type of attribution would be different in schizophrenia versus schizotypy: under-attributed agency should be observed for people with schizophrenia or schizotypy only within the range of embodied action, whereas in disembodied actions a pattern of over-attribution for patients and underattribution for schizotypy should be observed. Figure 9.2 below summarises the model proposed by Asai (2016) in Bayesian probability distribution terms:

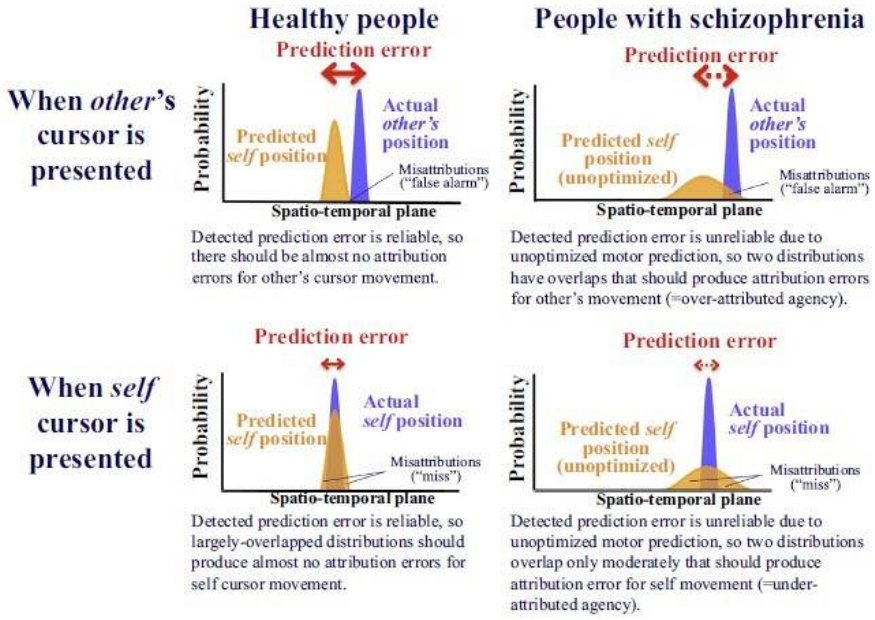

Figure 9.2. Diagrammatical representations of attribution errors. Adapted from Asai (2016).

Nevertheless, even Asai's recent account may be insufficient to explain the misattribution of agency for thinking, for the same reason that thoughts are by definition internal (self) so the S/N ratio would be maximised to the 'signal' or self condition. But the very act of misattributing agency implies volition and deliberation which borders on the realm of the 'judgement' or agency (Synofzik et al., 2008) and not just a 'sense' which is often far more minimal and basic, sometimes even possessing sensory qualities: some patients report knowing the exact location and feeling the associated tactile sensation of external thoughts entering their head (Mullins and Spence, 2003). Therefore, it is clear that further research is urgently needed to elucidate the relationship between thought processes and motor acts if such paradoxical observations are ever to be reconciled; no matter how much of a parsimonious explanation the source monitoring framework (when viewed in Bayesian terms) can offer, there are still discrepancies yet to be resolved.

9.3. Limitations

Study 1 was carried out in healthy and high-functioning individuals recruited from a University population. Like many other schizotypy studies done in healthy volunteer samples, the issue of selection bias cannot be overlooked. Although specific efforts

were in place to expand the participant pool beyond a psychology undergraduate sample and indeed younger members of staff across different departments were also recruited, in the end the vast majority of participants were still female undergraduates. Raine (1992) found significantly higher endorsement rates on the positive dimensions (odd beliefs, ideas of reference in the Schizotypal Personality Questionnaire) in females but higher endorsements on the negative dimension (e.g. constricted affect) in males. In clinical patients with schizophrenia, it has been reported that males tend to have an earlier age of onset, more acute manifestation and more impairments from negative symptoms, whereas females may retain better social functioning (Angermeyer and Kühnz, 1988; Shtasel, Gur, Gallacher, Heimberg, and Gur, 1992). Nevertheless, in the current sample a post-hoc Mann-Whitney U test found no significant differences across the three schizotypy scales *with the exception of* O-LIFE impulsive nonconformity, where males had a higher endorsement rate ($p = 0.006$). It could be that because all participants were extremely high-functioning and as a consequence all schizotypy scales were heavily positively skewed, gender effects on schizotypal traits have been diluted. Another possibility is that due to the social desirability effect, participants may tend to under-report what they consider as abnormal experiences at odds with their healthy and high-functioning nature (Pedregon, Farley, Davis, Wood, and Clark, 2012).

A major confounder, which was not measured objectively or actively controlled for, was substance misuse or even dependence. The exclusion of substance use was purely by self-report, and no saliva or blood samples were taken to be examined. Potential non-disclosure or misinformation from the participants could have implications on the schizotypy ratings, as it is generally accepted that individuals troubled by psychosis-like experiences misuse illicit drugs such as cannabis to self-medicate, which instead worsens their psychotic symptoms (Schiffman, Nakamura, Earleywine, and LaBrie, 2005), and that incidence of positive schizotypy is higher in cannabis users even without drawing temporal links (Nunn, Rizza, and Peters, 2001). However, some evidence suggests that in clinical patients, a history of cannabis use was associated with fewer symptoms and prior hospitalisations (Mueser et al., 1990), perhaps depending on the dominant active ingredient of cannabis used.

Study 2 had more detailed questionnaires on substance use, even though it was also by self-report on the testing day. However, prior to each testing session the

patient's past history of substance use was discussed with the clinician or care coordinator in charge of the patient and anyone meeting a criterion of dependence was excluded. Still, there was the potential of under-reporting if not actively downplaying of active psychotic symptoms. This could be due to the fact that most patients were in remission following antipsychotic treatment and they no longer felt 'unwell', that lack of insight was a hallmark of psychotic disorders by definition, or that self-stigmatisation prevented the patients from disclosing again what they might think as 'crazy' or 'abnormal' experiences.

Contrasting with Study 1, Study 2 (despite a small sample) almost exclusively consisted of male participants. This inevitably limits the generalisability of Study 1 to clinical populations (or even to the wider general population), not only because the sheer discrepancy in the level of functioning and IQ between patients and controls (for example, most patients were not in employment, education or training and only two had undergraduate degrees), but also due to many of the gender differences mentioned above. In fact, the only female patient scored exceptionally low on all symptom domains in the SOPS compared with her male counterparts. The very different demographics between participants in Study 1 and 2 perhaps reflects wider issues in healthy schizotypy versus clinical psychosis research, where patients with a first episode of psychosis and high-functioning individuals exhibiting mild psychosis-like experiences have little in common apart from matching age and gender. In this sense, such observations also reduce the full dimensionality of the psychosis continuum as a whole as mentioned in the previous Section.

There are also a number of conceptual issues the current Thesis did not directly address. First, the sense of agency was not investigated in either study. Although proxy measures such as sensory attenuation may be seen as a marker for agency, the concept is two-fold. These are the feeling of agency and the judgement of agency (Synofzik et al., 2008). In fact, the account proposed by Synofzik and colleagues directly challenges Frith's comparator model. Other recent studies such as those by Asai (2016, 2017) did not explicitly differentiate between feeling and judgement of agency and it can sometimes be ambiguous as to what 'attribution of agency' means. Theoretically, the feeling of agency is much more bottom-up and is largely based on multisensory integration whereas judgement of agency relies on higher-order cognitive mechanisms. Further, it would be extremely difficult to accurately measure 'levels' of agency at a

given time using psychometric assessments or questionnaires, even though there have been recent attempts (e.g. Lamaitre et al., 2016). The main reason why the two studies in the current study did not address the issue of agency was the inherent difficulties and 'noisiness' of potential measures. Practically, however, it would also have been too broad a research question if agency had been incorporated as another variable. Although agency may be intrinsic to sensory predictive processing, for example, it is less relevant for reward prediction. In addition, given the low level of agency-related symptoms (thought interference, delusions of control) in both healthy and early psychosis populations, it would be unlikely to be able to detect deficits in agency. This is because agency-related symptoms have nosological superiority in the diagnosis of schizophrenia (i.e. first-rank symptoms) and would signal transition to frank psychosis (see Marshall et al., 2017).

Third, the relationship *between* prediction error tasks and source monitoring tasks was not investigated. In the previous Section it is posited that source monitoring may be at least related to the detection of certain types of prediction error (most likely in the sensory domain), but the question remains as to how to directly measure prediction error in a source memory task, for example. It would be interesting to correlate errors in reality monitoring of actions with the overcompensation score in the force-matching task, as both involve sensory prediction processes. A negative correlation would be predicted, however, this method of correlation across tasks may be somewhat crude. It has been proposed that source monitoring acts at a higher level of explanation than predictive processing (Griffin and Fletcher, 2017), however, the precise mechanisms remain unclear and go well beyond the remit of the current Thesis.

Fourth, although there was data on the CAPS as a general hallucination measure in healthy volunteers, potential associations between this scale and verbal source monitoring (the word pair task) was not investigated. As the *a priori* hypotheses focused on general positive schizotypy in the source monitoring tasks, and no predictions with CAPS (which measures hallucinatory experiences in more than one modality) was made, such correlations were deemed unnecessary. However, specific scales on auditory-verbal hallucinations may have been useful in relating to performance in the word pair task, as the latter directly tapped into the auditory-verbal modality and would have been affected by hallucinations in the same modality.

Lastly, one conceptual consideration is how easily detectable these prediction error responses are by behavioural measures. As Corlett and Fletcher (2015) points out, prediction error responses could be much more latent than previously thought and behavioural measures alone might not be sensitive enough to study them. This is perhaps why many studies on prediction error have employed neuroimaging methods, focusing on specific brain regions such as the ventral striatum (Murray et al., 2008b). Nevertheless, in the sensory domain at least, such prediction error responses are evidently present as demonstrated by the force-matching illusion. Similarly, Kamin blocking and reward learning have been subjected to a wide range of behavioural investigations, long before neuroimaging methods were introduced. Also, using imaging would not easily have allowed one to study different domains of predictive processing in one experimental session. Perhaps another avenue of inquiry would be computational psychiatric modelling of behavioural data, and this is indeed a rapidly emerging field (e.g. Adams, Stephan, Brown, Frith, and Friston, 2013). There are many frameworks available for modelling reinforcement learning (e.g. Q-learning, actor-critic algorithms) but optimal model selection and potential over-fitting are issues commonly encountered. However, finely-tuned computational modelling has the ability to assist with disentangling complex and deeply latent processes.

9.4. Future Directions

There are many ways in which the current findings can inform future research. The most basic would be to aim towards much larger sample sizes for patients, perhaps by collaborating with other research organisations across multiple recruitment sites. Certainly, to require patients to travel long distances would be logistically difficult (or even ethically concerning), but with local early intervention services becoming more widespread and well established, participant recruitment of prodromal and first episode psychosis patients should become less effortful in the near future. Indeed, the current project emphasised the importance of early intervention by offering clear evidence for the presence of an early stage of psychosis and support for a continuum model of vulnerability to schizophrenia-spectrum psychoses. However, it also demonstrated that effect sizes were small in the early psychosis/first episode group. As mentioned above, recruiting acutely ill psychotic patients is ethically questionable;

still, a useful window for recruitment might be soon after a first episode of psychosis while the patient continues to experience significant positive symptoms is but not behaviourally disturbed. As described in Chapter 1, the prodrome is often a retrospective concept and it can be difficult to demarcate the exact point where an individual crossed the line into florid psychosis (but see Marshall et al., 2017). Given that less than half of the prodromal patients would actually transition to clinical schizophrenia longitudinal studies would be needed to 'track' who actually transition and then check their behavioural performance retrospectively. Alternatively, individuals at genetic high risk (e.g. first degree siblings) might be a useful avenue to recruit and follow-up as a comparison group to first episode psychosis patients.

Another group of potential interest are healthy hallucinators without a need for care or any diagnosis of a psychotic disorder, namely individuals claiming to be 'psychics', for example (Powers, Kelley, and Corlett, 2016). This is again a heterogeneous group, but has the clear advantage of hearing (sometimes persistent) voices without the confounds of medication. Conversely, it would be very interesting to also recruit patients with psychosis who do not hallucinate, as this would tease apart the relationships between hallucinations per se and other psychotic symptoms (such as delusions). In the current Thesis, it was not logistically possible to sub-group patients or controls according to the status of specific symptoms, which could have increased the noise in symptom assessments and relations with behavioural tasks.

Methodologically, another potential avenue would be to artificially induce anomalous experiences by administering psychoactive substances (mainly serotonergic agonists such as psilocybin) in healthy volunteers, even though this kind of studies is often very time-consuming and has more risks associated with the substances used. These kinds of studies have allowed researchers to closely control, capture and monitor such anomalous experiences in otherwise nonclinical samples, and the delusions, hallucinations and ego-dissolution experiences induced by psychoactive substances mimic those seen in psychosis (hence the name psychotomimetic for such substances). Using neuroimaging methods is another possibility, whether combined with a drug study or not; indeed, many of the recent studies on predictive processing and source monitoring involved some kind of imaging (functional MRI being the most popular choice) but this depended largely on the research questions being asked. Nevertheless, follow-up studies with both behavioural

and neuroimaging methods would certainly be useful in delineating the potential neural mechanisms of such processes, especially if using a paradigm that tapped into both prediction error responses and source monitoring.

In conclusion, the studies in the current project offered further consolidating evidence for a continuum model of psychosis from a cognitive neuropsychiatric framework. Although not all the behavioural alterations were demonstrated in individuals at various points of the psychosis continuum, deficits in the source monitoring of actions remained salient. In addition, the integrated approach taken by the current project (i.e. studying different aspects of predictive processing and source monitoring in the same cohort of participants) and the combination of frequentist and Bayesian statistical methods are novel. Findings from the current project could potentially inform cognitive therapies targeting the rehabilitation of self-agency and psychoeducation for the patient and their carers. In addition, it contributed to disentangling the different domains of neuropsychological processes such as prediction error responses and types of source monitoring, and the utility of behaviourally focused assessments.

References

Achim, A. M., & Lepage, M. (2003). Is associative recognition more impaired than item recognition memory in Schizophrenia? A meta-analysis. *Brain and Cognition*, *53*(2), 121-124.

Adams, R. A., Stephan, K. E., Brown, H. R., Frith, C. D., & Friston, K. J. (2013). The computational anatomy of psychosis. *Frontiers in Psychiatry*, *4*: 47.

Allen, P., Freeman, D., Johns, L., & McGuire, P. (2006). Misattribution of self-generated speech in relation to hallucinatory proneness and delusional ideation in healthy volunteers. *Schizophrenia Research*, *84*(2), 281-288.

Allen, P., Modinos, G., Hubl, D., Shields, G., Cachia, A., Jardri, R., ... & Hoffman, R. (2012). Neuroimaging auditory hallucinations in schizophrenia: from neuroanatomy to neurochemistry and beyond. *Schizophrenia Bulletin*, *38*(4), 695-703.

Andreasen, N. C., & Carpenter Jr, W. T. (1993). Diagnosis and classification of schizophrenia. *Schizophrenia Bulletin*, *19*(2), 199-214.

Andreasen, N. C. (1984). *Scale for the Assessment of Positive Symptoms (SAPS)*. Iowa City: University of Iowa.

Angermeyer, M. C., & Kühnz, L. (1988). Gender differences in age at onset of schizophrenia. *European Archives of Psychiatry and Neurological Sciences*, *237*(6), 351-364.

Anselmetti, S., Cavallaro, R., Bechi, M., Angelone, S. M., Ermoli, E., Cocchi, F., & Smeraldi, E. (2007). Psychopathological and neuropsychological correlates of source monitoring impairment in schizophrenia. *Psychiatry research*, *150*(1), 51-59.

Appelbaum, P. S., Robbins, P. C., & Roth, L. H. (1999). Dimensional approach to delusions: comparison across types and diagnoses. *American Journal of Psychiatry*, *156*(12), 1938-1943.

Arnal, L. H., Morillon, B., Kell, C. A., & Giraud, A. L. (2009). Dual neural routing of visual facilitation in speech processing. *The Journal of Neuroscience*, *29*(43), 13445-13453.

Arnal, L. H., Wyart, V., & Giraud, A. L. (2011). Transitions in neural oscillations reflect prediction errors generated in audiovisual speech. *Nature Neuroscience*, *14*(6), 797-801.

Asai, T., Sugimori, E., & Tanno, Y. (2008). Schizotypal personality traits and prediction of one's own movements in motor control: What causes an abnormal sense of agency?. *Consciousness and Cognition*, *17*(4), 1131-1142.

Asai, T., & Tanno, Y. (2007). The relationship between the sense of self-agency and schizotypal personality traits. *Journal of Motor Behavior*, *39*(3), 162-168.

Asai, T. (2016). Self is "other", other is "self": poor self-other discriminability explains schizotypal twisted agency judgment. *Psychiatry Research*, *246*, 593-600.

Asai, T. (2017). Know thy agency in predictive coding: Meta-monitoring over forward modeling. *Consciousness and Cognition*, *51*, 82-99.

Bacon, E., & Izaute, M. (2009). Metacognition in schizophrenia: processes underlying patients' reflections on their own episodic memory. *Biological Psychiatry*, *66*(11), 1031-1037.

Bajorek, T., & Stockmann, T. (2012). Psychiatric Assessment. In N. Burton (Ed.), *Psychiatry* (p. 14). London: JP Medical Ltd.

Barnes, M. P., Saunders, M., Walls, T. J., Saunders, I., & Kirk, C. A. (1986). The syndrome of Karl Ludwig Kahlbaum. *Journal of Neurology, Neurosurgery & Psychiatry*, *49*(9), 991-996.

Bassiony, M. M., Steinberg, M. S., Warren, A., Rosenblatt, A., Baker, A. S., & Lyketsos, C. G. (2000). Delusions and hallucinations in Alzheimer's disease: prevalence and clinical correlates. *International Journal of Geriatric Psychiatry*, *15*(2), 99-107.

Bayne, T., & Pacherie, E. (2005). In defence of the doxastic conception of delusions. *Mind & Language*, *20*(2), 163-188.

Bays, P. M., Flanagan, J. R., & Wolpert, D. M. (2006). Attenuation of self-generated tactile sensations is predictive, not postdictive. *PLoS Biology*, *4*(2), e28.

Beavan, V., Read, J., & Cartwright, C. (2011). The prevalence of voice-hearers in the general population: a literature review. *Journal of Mental Health*, *20*(3), 281-292.

Bell, V., Halligan, P. W., & Ellis, H. D. (2006a). Explaining delusions: a cognitive perspective. *Trends in Cognitive Sciences*, *10*(5), 219-226.

Bell, V., Halligan, P. W., & Ellis, H. D. (2006b). Diagnosing delusions: a review of inter-rater reliability. *Schizophrenia Research*, *86*(1), 76-79.

Bell, V., Halligan, P. W., & Ellis, H. D. (2006c). The Cardiff Anomalous Perceptions Scale (CAPS): a new validated measure of anomalous perceptual experience. *Schizophrenia Bulletin*, *32*(2), 366-377.

Bender, S., Müller, B., Oades, R. D., & Sartory, G. (2001). Conditioned blocking and schizophrenia: A replication and study of the role of symptoms, age, onset-age of psychosis and illness-duration. *Schizophrenia Research*, *49*(1), 157-170.

Benetti, S., Mechelli, A., Picchioni, M., Broome, M., Williams, S., & McGuire, P. (2009). Functional integration between the posterior hippocampus and prefrontal cortex is impaired in both first episode schizophrenia and the at risk mental state. *Brain*, *132*(9), 2426-2436.

Bentall, R. P., Baker, G. A., & Havers, S. (1991). Reality monitoring and psychotic hallucinations. *British Journal of Clinical Psychology*, *30*(3), 213-222.

Berridge, K. C. (2007). The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology*, *191*(3), 391-431.

Berrios, G. E., & Beer, D. (1994). The notion of unitary psychosis: a conceptual history. *History of Psychiatry*, *5*(17), 13-36.

Bhatt, R., Laws, K. R., & McKenna, P. J. (2010). False memory in schizophrenia patients with and without delusions. *Psychiatry Research*, *178*(2), 260-265.

Bird, V., Premkumar, P., Kendall, T., Whittington, C., Mitchell, J., & Kuipers, E. (2010). Early intervention services, cognitive–behavioural therapy and family intervention in early psychosis: systematic review. *The British Journal of Psychiatry*, *197*(5), 350-356.

Blakemore, S. J., & Frith, C. (2003). Self-awareness and action. *Current Opinion in Neurobiology*, *13*(2), 219-224.

Blakemore, S. J., Oakley, D. A., & Frith, C. D. (2003). Delusions of alien control in the normal brain. *Neuropsychologia*, *41*(8), 1058-1067.

Blakemore, S. J., Wolpert, D. M., & Frith, C. D. (2002). Abnormalities in the awareness of action. *Trends in Cognitive Sciences*, *6*(6), 237-242.

Bleuler, E. (1950). Dementia praecox or the group of schizophrenias. Oxford, England: International Universities Press.

Brébion, G., Amador, X., David, A., Malaspina, D., Sharif, Z., & Gorman, J. M. (2000). Positive symptomatology and source-monitoring failure in schizophrenia—an analysis of symptom-specific effects. *Psychiatry Research*, *95*(2), 119-131.

Brébion, G., Gorman, J. M., Amador, X., Malaspina, D., & Sharif, Z. (2002). Source monitoring impairments in schizophrenia: characterisation and associations with positive and negative symptomatology. *Psychiatry Research*, *112*(1), 27-39.

Brébion, G., Ohlsen, R. I., Pilowsky, L. S., & David, A. S. (2008). Visual hallucinations in schizophrenia: confusion between imagination and perception. *Neuropsychology*, *22*(3), 383.

Breier, A., & Berg, P. H. (1999). The psychosis of schizophrenia: prevalence, response to atypical antipsychotics, and prediction of outcome. *Biological Psychiatry*, *46*(3), 361-364.

Brett, C., Heriot Maitland, C., McGuire, P., & Peters, E. (2014). Predictors of distress associated with psychotic like anomalous experiences in clinical and non clinical populations. *British Journal of Clinical Psychology*, *53*(2), 213-227.

Brockington, I. F., & Leff, J. P. (1979). Schizo-affective psychosis: definitions and incidence. *Psychological Medicine*, *9*(1), 91-99.

Brookwell, M. L., Bentall, R. P., & Varese, F. (2013). Externalizing biases and hallucinations in source-monitoring, self-monitoring and signal detection studies: a meta-analytic review. *Psychological Medicine*, *43*(12), 2465-2475.

Broome, M. R., Matthiasson, P., Fusar-Poli, P., Woolley, J. B., Johns, L. C., Tabraham, P., ... & McGuire, P. K. (2009). Neural correlates of executive function and working memory in the 'at-risk mental state'. *British Journal of Psychiatry*, *194*(1), 25-33.

Brown, H., Adams, R. A., Parees, I., Edwards, M., & Friston, K. (2013). Active inference, sensory attenuation and illusions. *Cognitive Processing*, *14*(4), 411-427.

Brunelin, J., Combris, M., Poulet, E., Kallel, L., D'Amato, T., Dalery, J., & Saoud, M. (2006). Source monitoring deficits in hallucinating compared to non-hallucinating patients with schizophrenia. *European Psychiatry*, *21*(4), 259-261.

Campbell, J. (1999). Schizophrenia, the space of reasons, and thinking as a motor process. *The Monist*, *82*(4), 609-625.

Carlsson, A. (1988). The current status of the dopamine hypothesis of schizophrenia. *Neuropsychopharmacology*.

Cella, M., Dymond, S., & Cooper, A. (2009). Impairment in flexible emotion-based learning in hallucination-and delusion-prone individuals. *Psychiatry Research*, *170*(1), 70-74.

Cermolacce, M., Sass, L., & Parnas, J. (2010). What is bizarre in bizarre delusions? A critical review. *Schizophrenia Bulletin, 36*(4), 667-679.

Chadwick, P., Lees, S., & Birchwood, M. (2000). The revised beliefs about voices questionnaire (BAVQ-R). *The British Journal of Psychiatry*, *177*(3), 229-232.

Cheniaux, E., Landeira-Fernandez, J., Telles, L. L., Lessa, J. L. M., Dias, A., Duncan, T., & Versiani, M. (2008). Does schizoaffective disorder really exist? A systematic review of the studies that compared schizoaffective disorder with schizophrenia or mood disorders. *Journal of Affective Disorders*, *106*(3), 209-217.

Chun, C. A., Minor, K. S., & Cohen, A. S. (2013). Neurocognition in psychometrically defined college schizotypy samples: we are not measuring the "right stuff". *Journal of the International Neuropsychological Society*, *19*(03), 324-337.

Clare, L., McKenna, P. J., Mortimer, A. M., & Baddeley, A. D. (1993). Memory in schizophrenia: what is impaired and what is preserved?. *Neuropsychologia*, *31*(11), 1225-1241.

Claridge, G. E. (1997). *Schizotypy: Implications for illness and health*. Oxford University Press.

Claridge, G., & Broks, P. (1984). Schizotypy and hemisphere function—I: Theoretical considerations and the measurement of schizotypy. *Personality and Individual Differences*, *5*(6), 633-648.

Cohen, J. Y., Haesler, S., Vong, L., Lowell, B. B., & Uchida, N. (2012). Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature*, *482*(7383), 85-88.

Cohen, R. L. (1989). Memory for action events: The power of enactment. *Educational Psychology Review*, *1*(1), 57-80.

Collignon, O., Van der Linden, M., & Larøi, F. (2005). Source monitoring for actions in hallucination proneness. *Cognitive Neuropsychiatry*, *10*(2), 105-123.

Coltheart, M. (2007). The 33rd Sir Frederick Bartlett lecture cognitive neuropsychiatry and delusional belief. *The Quarterly Journal of Experimental Psychology*, *60*(8), 1041-1062.

Coltheart, M., Langdon, R., & McKay, R. (2007). Schizophrenia and monothematic delusions. *Schizophrenia Bulletin*, *33*(3), 642-647.

Cools, R., Frank, M. J., Gibbs, S. E., Miyakawa, A., Jagust, W., & D'Esposito, M. (2009). Striatal dopamine predicts outcome-specific reversal learning and its sensitivity to dopaminergic drug administration. *The Journal of Neuroscience*, *29*(5), 1538-1543.

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. *Cognitive Neuropsychiatry*, *20*(2), 95-105.

Corlett, P. R., Aitken, M. R., Dickinson, A., Shanks, D. R., Honey, G. D., Honey, R. A., ... & Fletcher, P. C. (2004). Prediction error during retrospective revaluation of causal associations in humans: fMRI evidence in favour of an associative model of learning. *Neuron*, *44*(5), 877-888.

Corlett, P. R., Canavan, S. V., Nahum, L., Appah, F., & Morgan, P. T. (2014). Dreams, reality and memory: confabulations in lucid dreamers implicate reality-monitoring dysfunction in dream consciousness. *Cognitive Neuropsychiatry*, (ahead-of-print), 1-14.

Corlett, P. R., Frith, C. D., & Fletcher, P. C. (2009). 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. (2007). From prediction error to psychosis: ketamine as a pharmacological model of delusions. *Journal of Psychopharmacology*, *21*(3), 238-252.

Corlett, P. R., Krystal, J. H., Taylor, J. R., & Fletcher, P. C. (2009). Why do delusions persist?. *Frontiers in Human Neuroscience*, *3*.

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

Corlett, P. R., Taylor, J. R., Wang, X. J., Fletcher, P. C., & Krystal, J. H. (2010). Toward a neurobiology of delusions. *Progress in Neurobiology*, *92*(3), 345-369.

Costafreda, S. G., Brébion, G., Allen, P., McGuire, P. K., & Fu, C. H. Y. (2008). Affective modulation of external misattribution bias in source monitoring in schizophrenia. *Psychological Medicine*, *38*(06), 821-824.

Curran, T. (2000). Brain potentials of recollection and familiarity. *Memory & Cognition*, *28*(6), 923-938.

Currie, G. (2000). Imagination, delusion and hallucinations. *Mind & Language*, *15*(1), 168-183.

Cuthbert, B. N. (2014). The RDoC framework: facilitating transition from ICD/DSM to dimensional approaches that integrate neuroscience and psychopathology. *World Psychiatry*, *13*(1), 28-35.

Czigler, I. (2007). Visual mismatch negativity: Violation of nonattended environmental regularities. *Journal of Psychophysiology*, *21*(3-4), 224.

David, A. S., Malmberg, A., Brandt, L., Allebeck, P., & Lewis, G. (1997). IQ and risk for schizophrenia: a population-based cohort study. *Psychological Medicine*, *27*(6), 1311-1323.

Davies, M., Coltheart, M., Langdon, R., & Breen, N. (2001). Monothematic delusions: Towards a two-factor account. *Philosophy, Psychiatry, & Psychology*, *8*(2), 133-158.

Davis, J. M., Chen, N., & Glick, I. D. (2003). A meta-analysis of the efficacy of second-generation antipsychotics. *Archives of General Psychiatry*, *60*(6), 553-564.

Debbané, M., & Barrantes-Vidal, N. (2014). Schizotypy From a Developmental Perspective. *Schizophrenia Bulletin*, DOI: 10.1093/schbul/sbu175.

Debbané, M., Eliez, S., Badoud, D., Conus, P., Flückiger, R., & Schultze-Lutter, F. (2014). Developing Psychosis and Its Risk States Through the Lens of Schizotypy. *Schizophrenia Bulletin*, DOI: 10.1093/schbul/sbu176.

Debbané, M., Eliez, S., Badoud, D., Conus, P., Flückiger, R., & Schultze-Lutter, F. (2015). Developing psychosis and its risk states through the lens of schizotypy. *Schizophrenia Bulletin*, *41*(suppl 2), S396-S407.

Demjaha, A., Egerton, A., Murray, R. M., Kapur, S., Howes, O. D., Stone, J. M., & McGuire, P. K. (2014). Antipsychotic treatment resistance in schizophrenia associated with elevated glutamate levels but normal dopamine function. *Biological Psychiatry*, *75*(5), e11-e13.

Demjaha, A., Murray, R. M., McGuire, P. K., Kapur, S., & Howes, O. D. (2012). Dopamine synthesis capacity in patients with treatment-resistant schizophrenia. *American Journal of Psychiatry*, *169*(11), 1203-1210.

den Ouden, H. E., Kok, P., & de Lange, F. P. (2012). How prediction errors shape perception, attention, and motivation. *Frontiers in Psychology*, *3*. DOI: 10.3389/fpsyg.2012.00548.

Denève, S., & Jardri, R. (2016). Circular inference: mistaken belief, misplaced trust. *Current Opinion in Behavioral Sciences*, *11*, 40-48.

Dienes, Z. (2011). Bayesian versus orthodox statistics: Which side are you on?. *Perspectives on Psychological Science*, *6*(3), 274-290.

Dierks, T., Linden, D. E., Jandl, M., Formisano, E., Goebel, R., Lanfermann, H., & Singer, W. (1999). Activation of Heschl's gyrus during auditory hallucinations. *Neuron*, *22*(3), 615-621.

Eissler, K. R. (1951). Remarks on the psychoanalysis of schizophrenia. *The International Journal of Psychoanalysis, 32*, 139-156.

Elliott, R., McKenna, P. J., Robbins, T. W., & Sahakian, B. J. (1995). Neuropsychological evidence for frontostriatal dysfunction in schizophrenia. *Psychological Medicine*, *25*(3), 619-630.

Elvevag, B., & Goldberg, T. E. (2000). Cognitive impairment in schizophrenia is the core of the disorder. *Critical Reviews™ in Neurobiology*, *14*(1).

Fenigstein, A., & Vanable, P. A. (1992). Paranoia and self-consciousness. *Journal of Personality and Social Psychology*, *62*(1), 129-138.

Fletcher, P. C., & Frith, C. D. (2009). Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia. *Nature Reviews Neuroscience*, *10*(1), 48-58.

Fletcher, P. C., Anderson, J. M., Shanks, D. R., Honey, R., Carpenter, T. A., Donovan, T., Papadakis, N., & Bullmore, E. T. (2001). Responses of human frontal cortex to surprising events are predicted by formal associative learning theory. *Nature Neuroscience*, *4*(10), 1043-1048.

Fogelson, N., Litvak, V., Peled, A., Fernandez-del-Olmo, M., & Friston, K. (2014). The functional anatomy of schizophrenia: a dynamic causal modeling study of predictive coding. *Schizophrenia Research*, *158*(1), 204-212.

Ford, J. M., Morris, S. E., Hoffman, R. E., Sommer, I., Waters, F., McCarthy-Jones, S., ... & Cuthbert, B. N. (2014). Studying hallucinations within the NIMH RDoC framework. *Schizophrenia Bulletin*, S295-S304.

Freeman, D., Garety, P. A., Bebbington, P. E., Smith, B., Rollinson, R., Fowler, D., ... & Dunn, G. (2005). Psychological investigation of the structure of paranoia in a non-clinical population. *The British Journal of Psychiatry*, *186*(5), 427-435.

Freeman, D., McManus, S., Brugha, T., Meltzer, H., Jenkins, R., & Bebbington, P. (2011). Concomitants of paranoia in the general population. *Psychological Medicine*, *41*(05), 923-936.

Freeman, D., Pugh, K., & Garety, P. (2008). Jumping to conclusions and paranoid ideation in the general population. *Schizophrenia Research*, *102*(1), 254-260.

Frith, C. (2005). The self in action: lessons from delusions of control. *Consciousness and Cognition*, *14*(4), 752-770.

Frith, C. (2012). Explaining delusions of control: The comparator model 20years on. *Consciousness and Cognition*, *21*(1), 52-54.

Frith, C. D. (1992). *The Cognitive neuropsychology of schizophrenia*. Hove, East Sussex: Psychology Press.

Frith, C. D., & Done, D. J. (1988). Towards a neuropsychology of schizophrenia. *The British Journal of Psychiatry*, *153*(4), 437-443.

Frith, C. D., Blakemore, S. J., & Wolpert, D. M. (2000). Explaining the symptoms of schizophrenia: abnormalities in the awareness of action. *Brain Research Reviews*, *31*(2), 357-363.

Fusar-Poli, P., Bonoldi, I., Yung, A. R., Borgwardt, S., Kempton, M. J., Valmaggia, L., ... & McGuire, P. (2012). Predicting psychosis: Meta-analysis of transition outcomes in individuals at high clinical risk. *Archives of General Psychiatry*, *69*(3), 220-229.

Fusar-Poli, P., Yung, A. R., McGorry, P., & Van Os, J. (2014). Lessons learned from the psychosis high-risk state: towards a general staging model of prodromal intervention. *Psychological Medicine*, *44*(1), 17-24.

Gallagher, S. (2000). Philosophical conceptions of the self: implications for cognitive science. *Trends in Cognitive Sciences*, *4*(1), 14-21.

Gallagher, S. (2004). Neurocognitive models of schizophrenia: a neurophenomenological critique. *Psychopathology*, *37*(1), 8-19.

Garety, P. A., Bebbington, P., Fowler, D., Freeman, D., & Kuipers, E. (2007). Implications for neurobiological research of cognitive models of psychosis: a theoretical paper. *Psychological Medicine*, *37*(10), 1377-1391.

Garrison, J. R., Bond, R., Gibbard, E., Johnson, M. K., & Simons, J. S. (2017a). Monitoring what is real: The effects of modality and action on accuracy and type of reality monitoring error. *Cortex*, *87*, 108-117.

Garrison, J. R., Moseley, P., Alderson-Day, B., Smailes, D., Fernyhough, C., & Simons, J. S. (2017b). Testing continuum models of psychosis: No reduction in source monitoring ability in healthy individuals prone to auditory hallucinations. *Cortex*, *91*, 197-207.

Garrison, J. R., Fernandez-Egea, E., Zaman, R., Agius, M., & Simons, J. S. (2017c). Reality monitoring impairment in schizophrenia reflects specific prefrontal cortex dysfunction. *NeuroImage: Clinical*, *14*, 260-268.

Garwood, L., Dodgson, G., Bruce, V., & McCarthy-Jones, S. (2013). A preliminary investigation into the existence of a hypervigilance subtype of auditory hallucination in people with psychosis. *Behavioural and Cognitive Psychotherapy,* (ahead-of-print), 1-11.

Gaser, C., Nenadic, I., Volz, H. P., Büchel, C., & Sauer, H. (2004). Neuroanatomy of 'hearing voices': a frontotemporal brain structural abnormality associated with auditory hallucinations in schizophrenia. *Cerebral Cortex*, *14*(1), 91-96.

Gawęda, Ł., Moritz, S., & Kokoszka, A. (2012). Impaired discrimination between imagined and performed actions in schizophrenia. *Psychiatry Research*, *195*(1), 1-8.

Gawęda, Ł., Woodward, T. S., Moritz, S., & Kokoszka, A. (2013). Impaired action self-monitoring in schizophrenia patients with auditory hallucinations. *Schizophrenia Research*, *144*(1), 72-79.

Georgieff, N., & Jeannerod, M. (1998). Beyond consciousness of external reality: a "who" system for consciousness of action and self-consciousness. *Consciousness and Cognition*, *7*(3), 465-477.

Gerrans, P. (2001). Delusions as performance failures. *Cognitive Neuropsychiatry*, *6*(3), 161-173.

Gerrans, P. (2013). Delusional attitudes and default thinking. *Mind & Language*, *28*(1), 83-102.

Gerrans, P. (2014). *Measure of Madness: Philosophy of Mind, Cognitive Neuroscience, and Delusional Thought*. Cambridge: MIT Press.

Gilleen, J., & David, A. S. (2005). The cognitive neuropsychiatry of delusions: from psychopathology to neuropsychology and back again. *Psychological Medicine*, *35*(01), 5-12.

Gold, J. M., Waltz, J. A., Prentice, K. J., Morris, S. E., & Heerey, E. A. (2008). Reward processing in schizophrenia: a deficit in the representation of value. *Schizophrenia Bulletin*, *34*(5), 835-847.

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

Green, C. E. L., Freeman, D., Kuipers, E., Bebbington, P., Fowler, D., Dunn, G., & Garety, P. A. (2008). Measuring ideas of persecution and social reference: the Green et al. Paranoid Thought Scales (GPTS). *Psychological Medicine*, *38*(1), 101-111.

Green, M. F. (1996). What are the functional consequences of neurocognitive deficits in schizophrenia?. *The American Journal of Psychiatry*, *153*(3), 321.

Greenfield, P., Joshi, S., Christian, S., Lekkos, P., Gregorowicz, A., Fisher, H. L., & Johnson, S. (2016). First episode psychosis in the over 35 s: is there a role for early intervention?. *Early Intervention in Psychiatry*, DOI:10.1111/eip.12322.

Griffin, J. D., & Fletcher, P. C. (2017). Predictive Processing, Source Monitoring, and Psychosis. *Annual Review of Clinical Psychology*, in press. DOI: 10.1146/annurev-clinpsy-032816-045145.

Griffiths, O., Langdon, R., Le Pelley, M. E., & Coltheart, M. (2014). Delusions and prediction error: re-examining the behavioural evidence for disrupted error signalling in delusion formation. *Cognitive Neuropsychiatry*, *19*(5), 439-467.

Gruppuso, V., Lindsay, D. S., & Kelley, C. M. (1997). The process-dissociation procedure and similarity: Defining and estimating recollection and familiarity in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*(2), 259-278.

Gumley, A. I., Gillan, K., Morrison, A. P., & Schwannauer, M. (2011). The development and validation of the Beliefs about Paranoia Scale (Short Form). *Behavioural and Cognitive Psychotherapy*, *39*(1), 35-53.

Gusnard, D. A., Akbudak, E., Shulman, G. L., & Raichle, M. E. (2001). Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function. *Proceedings of the National Academy of Sciences*, *98*(7), 4259-4264.

Haddock, G., McCarron, J., Tarrier, N., & Faragher, E. B. (1999). Scales to measure dimensions of hallucinations and delusions: the psychotic symptom rating scales (PSYRATS). *Psychological Medicine*, *29*(4), 879-889.

Halligan, P. W., & David, A. S. (2001). Cognitive neuropsychiatry: towards a scientific psychopathology. *Nature Reviews Neuroscience*, *2*(3), 209-215.

Hanssen, M., Bak, M., Bijl, R., Vollebergh, W., & van Os, J. (2005). The incidence and outcome of subclinical psychotic experiences in the general population. *British Journal of Clinical Psychology*, *44*(2), 181-191.

Harrow, M., Grossman, L. S., Herbener, E. S., & Davies, E. W. (2000). Ten-year outcome: patients with schizoaffective disorders, schizophrenia, affective disorders and mood-incongruent psychotic symptoms. *The British Journal of Psychiatry*, *177*(5), 421-426.

Haselgrove, M., & Evans, L. H. (2010). Variations in selective and nonselective prediction error with the negative dimension of schizotypy. *The Quarterly Journal of Experimental Psychology*, *63*(6), 1127-1149.

Hauser, M., Moore, J. W., de Millas, W., Gallinat, J., Heinz, A., Haggard, P., & Voss, M. (2011). Sense of agency is altered in patients with a putative psychotic prodrome. *Schizophrenia Research*, *126*(1), 20-27.

Hawton, K., Sutton, L., Haw, C., Sinclair, J., & Deeks, J. J. (2005). Schizophrenia and suicide: systematic review of risk factors. *The British Journal of Psychiatry*, *187*(1), 9-20.

Hayward, M., Denney, J., Vaughan, S., & Fowler, D. (2008). The voice and you: development and psychometric evaluation of a measure of relationships with voices. *Clinical Psychology & Psychotherapy*, *15*(1), 45-52.

Heinz, A., & Schlagenhauf, F. (2010). Dopaminergic dysfunction in schizophrenia: salience attribution revisited. *Schizophrenia Bulletin*, *36*(3), 472-485.

Henquet, C., Krabbendam, L., Dautzenberg, J., Jolles, J., & Merckelbach, H. (2005). Confusing thoughts and speech: source monitoring and psychosis. *Psychiatry Research*, *133*(1), 57-63.

Hewitt, J. K., & Claridge, G. (1989). The factor structure of schizotypy in a normal population. *Personality and Individual Differences*, *10*(3), 323-329.

Hill, K., Varese, F., Jackson, M., & Linden, D. E. (2012). The relationship between metacognitive beliefs, auditory hallucinations, and hallucination related distress in clinical and non clinical voice hearers. *British Journal of Clinical Psychology*, *51*(4), 434-447.

Ho, B. C., Andreasen, N. C., Ziebell, S., Pierson, R., & Magnotta, V. (2011). Long-term antipsychotic treatment and brain volumes: a longitudinal study of first-episode schizophrenia. *Archives of General Psychiatry*, *68*(2), 128-137.

Hoffman, R. E., Hawkins, K. A., Gueorguieva, R., Boutros, N. N., Rachid, F., Carroll, K., & Krystal, J. H. (2003). Transcranial magnetic stimulation of left temporoparietal cortex and medication-resistant auditory hallucinations. *Archives of General Psychiatry*, *60*(1), 49-56.

Hoffman, R. E., Varanko, M., Gilmore, J., & Mishara, A. L. (2008). Experiential features used by patients with schizophrenia to differentiate 'voices' from ordinary verbal thought. *Psychological Medicine*, *38*(08), 1167-1176.

Horga, G., Schatz, K. C., Abi-Dargham, A., & Peterson, B. S. (2014). Deficits in predictive coding underlie hallucinations in schizophrenia. *The Journal of Neuroscience*, *34*(24), 8072-8082.

Howes, O. D., & Kapur, S. (2009). The dopamine hypothesis of schizophrenia: version III—the final common pathway. *Schizophrenia Bulletin*, *35*(3), 549-562.

Howes, O. D., & Kapur, S. (2014). A neurobiological hypothesis for the classification of schizophrenia: type A (hyperdopaminergic) and type B (normodopaminergic). *The British Journal of Psychiatry*, *205*(1), 1-3.

Howes, O. D., Shotbolt, P., Bloomfield, M., Daalman, K., Demjaha, A., Diederen, K. M., ... & Sommer, I. E. (2013). Dopaminergic function in the psychosis spectrum: an [18F]-DOPA imaging study in healthy individuals with auditory hallucinations. *Schizophrenia Bulletin*, *39*(4), 807-814.

Hugdahl, K., Løberg, E. M., Jørgensen, H. A., Lundervold, A., Lund, A., Green, M. F., & Rund, B. R. (2008). Left hemisphere lateralisation of auditory hallucinations in schizophrenia: a dichotic listening study. *Cognitive Neuropsychiatry*, *13*(2), 166-179.

Humpston, C. S., & Broome, M. R. (2016). The spectra of soundless voices and audible thoughts: Towards an integrative model of auditory verbal hallucinations and thought insertion. *Review of Philosophy and Psychology*, *7*(3), 611-629.

Humpston, C. S., Linden, D. E., & Evans, L. H. (2017). Deficits in reality and internal source monitoring of actions are associated with the positive dimension of schizotypy. *Psychiatry Research*, *250*, 44-49.

Ihssen, N., Mussweiler, T., & Linden, D. E. (2016). Observing others stay or switch–How social prediction errors are integrated into reward reversal learning. *Cognition*, *153*, 19-32.

Jardri, R., & Denève, S. (2013). Circular inferences in schizophrenia. *Brain*, *136*(11), 3227-3241.

Jardri, R., Duverne, S., Litvinova, A. S., & Denève, S. (2017). Experimental evidence for circular inference in schizophrenia. *Nature Communications*, *8*, 14218.

Jauhar, S., McKenna, P. J., Radua, J., Fung, E., Salvador, R., & Laws, K. R. (2014). Cognitive-behavioural therapy for the symptoms of schizophrenia: systematic review and meta-analysis with examination of potential bias. *The British Journal of Psychiatry*, *204*(1), 20-29.

Jhou, T. C., Fields, H. L., Baxter, M. G., Saper, C. B., & Holland, P. C. (2009). The rostromedial tegmental nucleus (RMTg), a GABAergic afferent to midbrain dopamine neurons, encodes aversive stimuli and inhibits motor responses. *Neuron*, *61*(5), 786-800.

Johns, L. C., & van Os, J. (2001). The continuity of psychotic experiences in the general population. *Clinical Psychology Review*, *21*(8), 1125-1141.

Johns, L. C., Allen, P., Valli, I., Winton-Brown, T., Broome, M., Woolley, J., ... & McGuire, P. (2010). Impaired verbal self-monitoring in individuals at high risk of psychosis. *Psychological Medicine*, *40*(9), 1433-1442.

Johns, L. C., Kompus, K., Connell, M., Humpston, C., Lincoln, T. M., Longden, E., ... & Larøi, F. (2014). Auditory verbal hallucinations in persons with and without a need for care. *Schizophrenia Bulletin*, *40*(Suppl 4), S255-S264.

Johns, L. C., Rossell, S., Frith, C., Ahmad, F., Hemsley, D., Kuipers, E., & McGuire, P. K. (2001). Verbal self-monitoring and auditory verbal hallucinations in patients with schizophrenia. *Psychological Medicine*, *31*(04), 705-715.

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review*, *88*(1), 67-85.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, *114*(1), 3-28.

Johnstone, E. C., Ebmeier, K. P., Miller, P., Owens, D. G., & Lawrie, S. M. (2005). Predicting schizophrenia: findings from the Edinburgh high-risk study. *British Journal of Psychiatry*, *186*(1), 18-25.

Jones, N., & Luhrmann, T. M. (2016). Beyond the sensory: Findings from an in-depth analysis of the phenomenology of "auditory hallucinations" in schizophrenia. *Psychosis*, *8*(3), 191-202.

Jones, S. H., Gray, J. A., & Hemsley, D. R. (1992a). Loss of the Kamin blocking effect in acute but not chronic schizophrenics. *Biological Psychiatry*, *32*(9), 739-755.

Jones, S. R. (2010). Do we need multiple models of auditory verbal hallucinations? Examining the phenomenological fit of cognitive and neurological models. *Schizophrenia Bulletin*, *36*(3), 566-575.

Jones, S. R., & Fernyhough, C. (2007). Thought as action: Inner speech, self-monitoring, and auditory verbal hallucinations. *Consciousness and Cognition*, *16*(2), 391-399.

Kamin, L. J. (1969). Predictability, surprise, attention, and conditioning. *Punishment and Aversive Behavior*, 279-296.

Kapur, S. (2003). Psychosis as a state of aberrant salience: a framework linking biology, phenomenology, and pharmacology in schizophrenia. *American Journal of Psychiatry*, *160*(1), 13-23.

Kapur, S. (2004). How antipsychotics become anti-'psychotic'–from dopamine to salience to psychosis. *Trends in Pharmacological Sciences*, *25*(8), 402-406.

Kay, S. R., Fiszbein, A., & Opfer, L. A. (1987). The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophrenia Bulletin*, *13*(2), 261-276.

Kaymaz, N., & van Os, J. (2010). Extended psychosis phenotype–yes: single continuum–unlikely. *Psychological Medicine*, *40*(12), 1963-1966.

Keefe, R. S., Arnold, M. C., Bayen, U. J., McEvoy, J. P., & Wilson, W. H. (2002). Source-monitoring deficits for self-generated stimuli in schizophrenia: multinomial modeling of data from three sources. *Schizophrenia Research*, *57*(1), 51-67.

Kelley, R., & Wixted, J. T. (2001). On the nature of associative information in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(3), 701-722.

Kendler, K. S. (1980). Are there delusions specific for paranoid disorders vs schizophrenia?. *Schizophrenia Bulletin*, *6*(1), 1-3.

Kendler, K. S., Aggen, S. H., Czajkowski, N., Røysamb, E., Tambs, K., Torgersen, S., ... & Reichborn-Kjennerud, T. (2008). The structure of genetic and environmental risk factors for DSM-IV personality disorders: a multivariate twin study. *Archives of General Psychiatry*, *65*(12), 1438-1446.

Koriat, A., Levy-Sadot, R., Edry, E., & de Marcas, S. (2003). What do we know about what we cannot remember? Accessing the semantic attributes of words that cannot be recalled. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(6), 1095-1105.

Kot, T., & Serper, M. (2002). Increased susceptibility to auditory conditioning in hallucinating schizophrenic patients: a preliminary investigation. *The Journal of Nervous and Mental Disease*, *190*(5), 282-288.

Kwapil, T. R., & Barrantes-Vidal, N. (2014). Schizotypy: Looking back and moving forward. *Schizophrenia Bulletin*, DOI: 10.1093/schbul/sbu186.

Lahti, A. C., Koffel, B., LaPorte, D., & Tamminga, C. A. (1995). Subanesthetic doses of ketamine stimulate psychosis in schizophrenia. *Neuropsychopharmacology*, *13*(1), 9-19.

Lancaster, T. M., Ihssen, N., Brindley, L. M., Tansey, K. E., Mantripragada, K., O'Donovan, M. C., ... & Linden, D. E. (2015). Associations between polygenic risk for schizophrenia and brain function during probabilistic learning in healthy individuals. *Human Brain Mapping,* DOI: 10.1002/hbm.23044.

Launay, G., & Slade, P. (1981). The measurement of hallucinatory predisposition in male and female prisoners. *Personality and Individual Differences*, *2*(3), 221-234.

Lemaitre, A. L., Luyat, M., & Lafargue, G. (2016). Individuals with pronounced schizotypal traits are particularly successful in tickling themselves. *Consciousness and Cognition*, *41*, 64-71.

Libby, L. A., Yonelinas, A. P., Ranganath, C., & Ragland, J. D. (2013). Recollection and familiarity in schizophrenia: a quantitative review. *Biological Psychiatry*, *73*(10), 944-950.

Lieberman, J. A. (2007). Effectiveness of antipsychotic drugs in patients with chronic schizophrenia: efficacy, safety and cost outcomes of CATIE and other trials. *The Journal of Clinical Psychiatry*, *68*(2), e04.

Linden, D. E., Thornton, K., Kuswanto, C. N., Johnston, S. J., van de Ven, V., & Jackson, M. C. (2010). The brain's voices: comparing nonclinical auditory hallucinations and imagery. *Cerebral Cortex*, DOI: 10.1093/cercor/bhq097.

Lindenmayer, M. D., & Khan, A. (2006). Psychopathology. In J. A. Lieberman, T. S. Stroup, & D. O. Perkins (Eds), *Textbook of Schizophrenia* (pp. 187-222). Arlington: American Psychiatric Publishing, Inc.

Lindner, A., Thier, P., Kircher, T. T., Haarmeier, T., & Leube, D. T. (2005). Disorders of agency in schizophrenia correlate with an inability to compensate for the sensory consequences of actions. *Current Biology*, *15*(12), 1119-1124.

Lisman, J. E., & Grace, A. A. (2005). The hippocampal-VTA loop: controlling the entry of information into long-term memory. *Neuron*, *46*(5), 703-713.

Lysaker, P. H., & Dimaggio, G. (2014). Metacognitive capacities for reflection in schizophrenia: implications for developing treatments. *Schizophrenia Bulletin*, *40*(3), 487-491.

MacDonald III, A. W., Carter, C. S., Kerns, J. G., Ursu, S., Barch, D. M., Holmes, A. J., ... & Cohen, J. D. (2005). Specificity of prefrontal dysfunction and context processing deficits to schizophrenia in never-medicated patients with first-episode psychosis. *American Journal of Psychiatry*, *162*(3), 475-484.

Madan, C. R., & Singhal, A. (2012). Using actions to enhance memory: effects of enactment, gestures, and exercise on human memory. *Frontiers in Psychology*, *3*: 507.

Maher, B. A. (1974). Delusional thinking and perceptual disorder. *Journal of Individual Psychology, 30*(1), 98-113.

Maher, B. A. (1999). Anomalous experience in everyday life: Its significance for psychopathology. *The Monist*, 547-570.

Maher, B. A. (2006). The relationship between delusions and hallucinations. *Current Psychiatry Reports*, *8*(3), 179-183.

Marneros, A., Pillmann, F., Haring, A., Balzuweit, S., & Blöink, R. (2003). Features of acute and transient psychotic disorders. *European Archives of Psychiatry and Clinical Neuroscience*, *253*(4), 167-174.

Marshall, C., Lu, Y., Lyngberg, K., Deighton, S., Cadenhead, K. S., Cannon, T. D., ... & Tsuang, M. T. Changes in symptom content from a clinical high risk state to conversion to psychosis. *Early Intervention in Psychiatry*, in press. DOI: 10.1111/eip.12473.

Martin, J., Tilling, K., Hubbard, L., Stergiakouli, E., Thapar, A., Smith, G. D., ... & Zammit, S. (2016). Association of Genetic Risk for Schizophrenia with Nonparticipation Over Time in a Population-Based Cohort Study. *American Journal of Epidemiology*, DOI: 10.1093/aje/kww009.

Mason, O., & Claridge, G. (2006). The Oxford-Liverpool Inventory of Feelings and Experiences (O-LIFE): further description and extended norms. *Schizophrenia Research*, *82*(2), 203-211.

Mason, O., Claridge, G., & Jackson, M. (1995). New scales for the assessment of schizotypy. *Personality and Individual Differences*, *18*(1), 7-13.

Mason, O., Startup, M., Halpin, S., Schall, U., Conrad, A., & Carr, V. (2004). Risk factors for transition to first episode psychosis among individuals with 'at-risk mental states'. *Schizophrenia Research*, *71*(2), 227-237.

Matsumoto, M., & Hikosaka, O. (2009). Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature*, *459*(7248), 837-841.

McCarthy-Jones, S., Trauer, T., Mackinnon, A., Sims, E., Thomas, N., & Copolov, D. L. (2014). A new phenomenological survey of auditory hallucinations: evidence for subtypes and implications for theory and practice. *Schizophrenia Bulletin*, *40*(1), 231-235.

McEvoy, J. P., Meyer, J. M., Goff, D. C., Nasrallah, H. A., Davis, S. M., Sullivan, L., ... & Lieberman, J. A. (2005). Prevalence of the metabolic syndrome in patients with schizophrenia: baseline results from the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) schizophrenia trial and comparison with national estimates from NHANES III. *Schizophrenia Research*, *80*(1), 19-32.

McGuire, P. K., David, A. S., Murray, R. M., Frackowiak, R. S. J., Frith, C. D., Wright, I., & Silbersweig, D. A. (1995). Abnormal monitoring of inner speech: a physiological basis for auditory hallucinations. *The Lancet*, *346*(8975), 596-600.

McKay, R., Langdon, R., & Coltheart, M. (2006). The persecutory ideation questionnaire. *The Journal of Nervous and Mental Disease*, *194*(8), 628-631.

McKay, R., Langdon, R., & Coltheart, M. (2007). Models of misbelief: Integrating motivational and deficit theories of delusions. *Consciousness and Cognition*, *16*(4), 932-941.

McKirdy, J., Sussmann, J. E. D., Hall, J., Lawrie, S. M., Johnstone, E. C., & McIntosh, A. M. (2009). Set shifting and reversal learning in patients with bipolar disorder or schizophrenia. *Psychological Medicine*, *39*(08), 1289-1293.

Meltzer, H. Y., & McGurk, S. R. (1999). The effects of clozapine, risperidone, and olanzapine on cognitive function in schizophrenia. *Schizophrenia Bulletin*, *25*(2), 233-256.

Meltzer, H. Y., & Stahl, S. M. (1976). The dopamine hypothesis of schizophrenia. *Schizophrenia Bulletin*, *2*(1), 19-76.

Meltzer, H. Y., Thompson, P. A., Lee, M. A., & Ranjan, R. (1996). Neuropsychologic deficits in schizophrenia: relation to social function and effect of antipsychotic drug treatment. *Neuropsychopharmacology*, *14*(3), 27S-33S.

Mickes, L., Wixted, J. T., & Wais, P. E. (2007). A direct test of the unequal-variance signal detection model of recognition memory. *Psychonomic Bulletin & Review*, *14*(5), 858-865.

Milev, P., Ho, B. C., Arndt, S., & Andreasen, N. C. (2005). Predictive values of neurocognition and negative symptoms on functional outcome in schizophrenia: a longitudinal first-episode study with 7-year follow-up. *American Journal of Psychiatry*, *162*(3), 495-506.

Miller, P., Byrne, M., Hodges, A., Lawrie, S. M., Owens, D. G. C., & Johnstone, E. C. (2002). Schizotypal components in people at high risk of developing schizophrenia: early findings from the Edinburgh High-Risk Study. *British Journal of Psychiatry*, *180*(2), 179-184.

Miller, T. J., McGlashan, T. H., Rosen, J. L., Cadenhead, K., Ventura, J., McFarlane, W., ... & Woods, S. W. (2003). Prodromal assessment with the structured interview for prodromal syndromes and the scale of prodromal symptoms: predictive validity, interrater reliability, and training to reliability. *Schizophrenia Bulletin*, *29*(4), 703-715.

Mintz, S., & Alpert, M. (1972). Imagery vividness, reality testing, and schizophrenic hallucinations. *Journal of Abnormal Psychology*, *79*(3), 310.
Mishara, A. L. (2010). Klaus Conrad (1905–1961): Delusional mood, psychosis, and beginning schizophrenia. *Schizophrenia Bulletin*, *36*(1), 9-13.
Mitchell, K. J., & Johnson, M. K. (2000). Source monitoring: Attributing mental experiences. *The Oxford Handbook of Memory*, 179-195.
Miyazono, K., Bortolotti, L., & Broome, M. R. (2015). Prediction-error and two-factor theories of delusion formation: competitors or allies? In N. Galbraith (Ed.), *Aberrant Beliefs and Reasoning* (pp. 34-54). Hove: Psychology Press.
Moore, J. W., Dickinson, A., & Fletcher, P. C. (2011). Sense of agency, associative learning, and schizotypy. *Consciousness and Cognition*, *20*(3), 792-800.
Moran, P. M., Al-Uzri, M. M., Watson, J., & Reveley, M. A. (2003). Reduced Kamin blocking in non paranoid schizophrenia: associations with schizotypy. *Journal of Psychiatric Research*, *37*(2), 155-163.
Moran, P. M., Owen, L., Crookes, A. E., Al-Uzri, M. M., & Reveley, M. A. (2008). Abnormal prediction error is associated with negative and depressive symptoms in schizophrenia. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *32*(1), 116-123.
Moritz, S., & Larøi, F. (2008). Differences and similarities in the sensory and cognitive signatures of voice-hearing, intrusions and thoughts. *Schizophrenia Research*, *102*(1), 96-107.
Moritz, S., & Woodward, T. S. (2002). Memory confidence and false memories in schizophrenia. *The Journal of Nervous and Mental Disease*, *190*(9), 641-643.
Moritz, S., & Woodward, T. S. (2006). Metacognitive control over false memories: a key determinant of delusional thinking. *Current Psychiatry Reports*, *8*(3), 184-190.
Moritz, S., Woodward, T. S., & Ruff, C. C. (2003). Source monitoring and memory confidence in schizophrenia. *Psychological Medicine*, *33*(01), 131-139.
Moritz, S., Woodward, T. S., Jelinek, L., & Klinge, R. (2008). Memory and metamemory in schizophrenia: a liberal acceptance account of psychosis. *Psychological Medicine*, *38*(06), 825-832.
Morrison, A. P., Haddock, G., & Tarrier, N. (1995). Intrusive thoughts and auditory hallucinations: a cognitive approach. *Behavioural and Cognitive Psychotherapy*, *23*(3), 265-280.
Morrison, A. P., Bentall, R. P., French, P., Walford, L., Kilcommons, A., Knight, A., ... & Lewis, S. W. (2002). Randomised controlled trial of early detection and cognitive therapy for preventing transition to psychosis in high-risk individuals. *The British Journal of Psychiatry*, *181*(43), s78-s84.
Morrison, A. P., Turkington, D., Pyle, M., Spencer, H., Brabban, A., Dunn, G., ... & Hutton, P. (2014). Cognitive therapy for people with schizophrenia spectrum disorders not taking antipsychotic drugs: a single-blind randomised controlled trial. *The Lancet*, *383*(9926), 1395-1403.
Mueser, K. T., Yarnold, P. R., Levinson, D. F., Singh, H., Bellack, A. S., Kee, K., ... & Yadalam, K. G. (1990). Prevalence of substance abuse in schizophrenia: demographic and clinical correlates. *Schizophrenia Bulletin*, *16*(1), 31-56.
Mulholland, C., & Cooper, S. (2000). The symptom of depression in schizophrenia and its management. *Advances in Psychiatric Treatment*, *6*(3), 169-177.
Mullins, S., & Spence, S. A. (2003). Re-examining thought insertion. *British Journal of Psychiatry*, *182*(4), 293-298.

Murray, G. K., Cheng, F., Clark, L., Barnett, J. H., Blackwell, A. D., Fletcher, P. C., ... & Jones, P. B. (2008a). Reinforcement and reversal learning in first-episode psychosis. *Schizophrenia Bulletin*, *34*(5), 848-855.

Murray, G. K., Corlett, P. R., Clark, L., Pessiglione, M., Blackwell, A. D., Honey, G., ... & Fletcher, P. C. (2008b). Substantia nigra/ventral tegmental reward prediction error disruption in psychosis. *Molecular Psychiatry*, *13*(3), 267-276.

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clinical Neurophysiology*, *118*(12), 2544-2590.

Nayani, T. H., & David, A. S. (1996). The auditory hallucination: a phenomenological survey. *Psychological Medicine*, *26*(01), 177-189.

Nelson, B., & Sass, L. A. (2009). Medusa's stare: a case study of working with self-disturbance in the early phase of schizophrenia. *Clinical Case Studies*, *8*(6), 489-504.

Nelson, B., Fornito, A., Harrison, B. J., Yücel, M., Sass, L. A., Yung, A. R., ... & McGorry, P. D. (2009). A disturbed sense of self in the psychosis prodrome: linking phenomenology and neurobiology. *Neuroscience & Biobehavioral Reviews*, *33*(6), 807-817.

Nelson, B., Thompson, A., & Yung, A. R. (2012). Basic self-disturbance predicts psychosis onset in the ultra high risk for psychosis "prodromal" population. *Schizophrenia Bulletin*, *38*(6), 1277-1287.

Nelson, B., Whitford, T. J., Lavoie, S., & Sass, L. A. (2014a). What are the neurocognitive correlates of basic self-disturbance in schizophrenia?: Integrating phenomenology and neurocognition. Part 1 (Source monitoring deficits). *Schizophrenia Research*, *152*(1), 12-19.

Nelson, B., Whitford, T. J., Lavoie, S., & Sass, L. A. (2014b). What are the neurocognitive correlates of basic self-disturbance in schizophrenia?: Integrating phenomenology and neurocognition: Part 2 (Aberrant salience). *Schizophrenia Research*, *152*(1), 20-27.

Nelson, B., Yuen, H. P., Lin, A., Wood, S. J., McGorry, P. D., Hartmann, J. A., & Yung, A. R. (2016). Further examination of the reducing transition rate in ultra high risk for psychosis samples: The possible role of earlier intervention. *Schizophrenia Research*, *174*(1), 43-49.

Nelson, B., Yung, A. R., Bechdolf, A., & McGorry, P. D. (2008). The phenomenological critique and self-disturbance: implications for ultra-high risk ("prodrome") research. *Schizophrenia Bulletin*, *34*(2), 381-392.

Nienow, T. M., & Docherty, N. M. (2004). Internal source monitoring and thought disorder in schizophrenia. *The Journal of Nervous and Mental Disease*, *192*(10), 696-700.

Nieuwenstein, M. R., Aleman, A., & de Haan, E. H. (2001). Relationship between symptom dimensions and neurocognitive functioning in schizophrenia: a meta-analysis of WCST and CPT studies. *Journal of Psychiatric Research*, *35*(2), 119-125.

Nordgaard, J., Arnfred, S. M., Handest, P., & Parnas, J. (2008). The diagnostic status of first-rank symptoms. *Schizophrenia Bulletin*, *34*(1), 137-154.

Nordström, A. L., Farde, L., Wiesel, F. A., Forslund, K., Pauli, S., Halldin, C., & Uppfeldt, G. (1993). Central D2-dopamine receptor occupancy in relation to antipsychotic drug effects: a double-blind PET study of schizophrenic patients. *Biological Psychiatry*, *33*(4), 227-235.

Northoff, G., & Bermpohl, F. (2004). Cortical midline structures and the self. *Trends in Cognitive Sciences*, *8*(3), 102-107.

Northoff, G., Heinzel, A., de Greck, M., Bermpohl, F., Dobrowolny, H., & Panksepp, J. (2006). Self-referential processing in our brain- a meta-analysis of imaging studies on the self. *Neuroimage*, *31*(1), 440-457.

Nunn, J. A., Rizza, F., & Peters, E. R. (2001). The incidence of schizotypy among cannabis and alcohol users. *The Journal of Nervous and Mental Disease*, *189*(11), 741-748.

Oades, R. D., Zimmermann, B., & Eggers, C. (1996). Conditioned blocking in patients with paranoid, non-paranoid psychosis or obsessive compulsive disorder: Associations with symptoms, personality and monoamine metabolism. *Journal of Psychiatric Research*, *30*(5), 369-390.

Oertel, V., Rotarska-Jagiela, A., van de Ven, V., Haenschel, C., Grube, M., Stangier, U., ... & Linden, D. E. (2009). Mental imagery vividness as a trait marker across the schizophrenia spectrum. *Psychiatry Research*, *167*(1), 1-11.

Ohayon, M. M. (2000). Prevalence of hallucinations and their pathological associations in the general population. *Psychiatry Research*, *97*(2), 153-164.

Oliver, R., Bjoertomt, O., Greenwood, R., & Rothwell, J. (2008). 'Noisy patients'—can signal detection theory help?. *Nature Clinical Practice Neurology*, *4*(6), 306-316.

Orosz, A. T., Feldon, J., Simon, A. E., Hilti, L. M., Gruber, K., Yee, B. K., & Cattapan-Ludewig, K. (2010). Learned irrelevance and associative learning is attenuated in individuals at risk for psychosis but not in asymptomatic first-degree relatives of schizophrenia patients: translational state markers of psychosis?. *Schizophrenia Bulletin*, DOI: 10.1093/schbul/sbp165.

Owen, M. J. (2014). New Approaches to Psychiatric Diagnostic Classification. *Neuron*, *84*(3), 564-571.

Palmer, B. A., Pankratz, V. S., & Bostwick, J. M. (2005). The lifetime risk of suicide in schizophrenia: a reexamination. *Archives of General Psychiatry*, *62*(3), 247-253.

Palmer, C. E., Davare, M., & Kilner, J. M. (2016). Physiological and perceptual sensory attenuation have different underlying neurophysiological correlates. *Journal of Neuroscience*, *36*(42), 10803-10812.

Patil, S. T., Zhang, L., Martenyi, F., Lowe, S. L., Jackson, K. A., Andreev, B. V., ... & Schoepp, D. D. (2007). Activation of mGlu2/3 receptors as a new approach to treat schizophrenia: a randomized Phase 2 clinical trial. *Nature Medicine*, *13*(9), 1102-1107.

Pedregon, C. A., Farley, R. L., Davis, A., Wood, J. M., & Clark, R. D. (2012). Social desirability, personality questionnaires, and the "better than average" effect. *Personality and Individual Differences*, *52*(2), 213-217.

Penn, D. L., Waldheter, E. J., Perkins, D. O., Mueser, K. T., & Lieberman, J. A. (2005). Psychosocial treatment for first-episode psychosis: a research update. *American journal of Psychiatry*, *162*(12), 2220-2232.

Peralta, V., & Cuesta, M. J. (2003). The nosology of psychotic disorders: a comparison among competing classification systems. *Schizophrenia Bulletin*, *29*(3), 413.

Peralta, V., & Cuesta, M. J. (2005). The underlying structure of diagnostic systems of schizophrenia: a comprehensive polydiagnostic approach. *Schizophrenia Research*, *79*(2), 217-229.

Pérez Álvarez, M., García Montes, J. M., Vallina Fernández, O., Perona Garcelán, S., & Cuevas Yust, C. (2011). New life for schizophrenia psychotherapy in the light of phenomenology. *Clinical Psychology & Psychotherapy*, *18*(3), 187-201.

Perkins, D. O., Leserman, J., Jarskog, L. F., Graham, K., Kazmer, J., & Lieberman, J. A. (2000). Characterizing and dating the onset of symptoms in psychotic illness: the Symptom Onset in Schizophrenia (SOS) inventory. *Schizophrenia Research*, *44*(1), 1-10.

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). *Schizophrenia Bulletin*, *25*(3), 553-576.

Peters, E., Joseph, S., Day, S., & Garety, P. (2004). Measuring Delusional Ideation: The 21-Item Peters et al. Delusions Inventory (PDI). *Schizophrenia Bulletin*, *30*(4), 1005-1022.

Peters, M. J., Smeets, T., Giesbrecht, T., Jelicic, M., & Merckelbach, H. (2007). Confusing action and imagination: action source monitoring in individuals with schizotypal traits. *The Journal of Nervous and Mental Disease*, *195*(9), 752-757.

Pflueger, M. O., Gschwandtner, U., Stieglitz, R. D., & Riecher-Rössler, A. (2007). Neuropsychological deficits in individuals with an at risk mental state for psychosis—working memory as a potential trait marker. *Schizophrenia Research*, *97*(1), 14-24.

Postmes, L., Sno, H. N., Goedhart, S., van der Stel, J., Heering, H. D., & de Haan, L. (2014). Schizophrenia as a self-disorder due to perceptual incoherence. *Schizophrenia Research*, *152*(1), 41-50.

Powers, A. R., Kelley, M. S., & Corlett, P. R. (2017). Varieties of voice-hearing: psychics and the psychosis continuum. *Schizophrenia Bulletin*, *43*(1), 84-98.

Powers, A. R., Mathys, C., & Corlett, P. R. (2017). Pavlovian conditioning–induced hallucinations result from overweighting of perceptual priors. *Science*, *357*(6351), 596-600.

Qin, J., Raye, C. L., Johnson, M. K., & Mitchell, K. J. (2001). Source ROCs are (typically) curvilinear: comment on Yonelinas (1999). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(4), 1110-1115.

Quintana, D., & Eriksen, J. A. (2017). Bayesian alternatives for common null-hypothesis significance tests in psychiatry: A non-technical guide using JASP. Accepted preprint, available at https://osf.io/preprints/wun5v/.

Ragland, J. D., Laird, A. R., Ranganath, C., Blumenfeld, R. S., Gonzales, S. M., & Glahn, D. C. (2009). Prefrontal activation deficits during episodic memory in schizophrenia. *American Journal of Psychiatry*, *166*(8), 863-874.

Raine, A. (1991). The SPQ: a scale for the assessment of schizotypal personality based on DSM-III-R criteria. *Schizophrenia Bulletin*, *17*(4), 555-564.

Raine, A. (1992). Sex differences in schizotypal personality in a nonclinical population. *Journal of Abnormal Psychology*, *101*(2), 361-364.

Ranganath, C., Minzenberg, M. J., & Ragland, J. D. (2008). The cognitive neuroscience of memory function and dysfunction in schizophrenia. *Biological Psychiatry*, *64*(1), 18-25.

Rankin, P. M., & O'Carroll, P. J. (1995). Reality discrimination, reality monitoring and disposition towards hallucination. *British Journal of Clinical Psychology*, *34*(4), 517-528.

Rasmussen, A. R., & Parnas, J. (2015). Pathologies of imagination in schizophrenia spectrum disorders. *Acta Psychiatrica Scandinavica*, *131*(3), 157-161.

Reinen, J. M., Van Snellenberg, J. X., Horga, G., Abi-Dargham, A., Daw, N. D., & Shohamy, D. (2016). Motivational Context Modulates Prediction Error Response in Schizophrenia. *Schizophrenia Bulletin*, DOI: 10.1093/schbul/sbw045.

Rotello, C. M., Macmillan, N. A., & Reeder, J. A. (2004). Sum-difference theory of remembering and knowing: a two-dimensional signal-detection model. *Psychological Review*, *111*(3), 588-616.

Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Sciences*, *11*(6), 251-257.

Rugg, M. D., & Yonelinas, A. P. (2003). Human recognition memory: a cognitive neuroscience perspective. *Trends in Cognitive Sciences*, *7*(7), 313-319.

Rund, B. R., Melle, I., Friis, S., Larsen, T. K., Midbøe, L. J., Opjordsmoen, S., ... & McGlashan, T. (2004). Neurocognitive dysfunction in first-episode psychosis: correlates with symptoms, premorbid adjustment, and duration of untreated psychosis. *American Journal of Psychiatry*, *161*(3), 466-472.

Sack, A. T., Van De Ven, V. G., Etschenberg, S., Schatz, D., & Linden, D. E. (2005). Enhanced vividness of mental imagery as a trait marker of schizophrenia?. *Schizophrenia Bulletin*, *31*(1), 97-104.

Sass, L. A. (2004). Some reflections on the (analytic) philosophical approach to delusion. *Philosophy, Psychiatry, & Psychology*, *11*(1), 71-80.

Sass, L. A., & Parnas, J. (2003). Schizophrenia, consciousness, and the self. *Schizophrenia Bulletin*, *29*(3), 427-444.

Sato, A., & Yasuda, A. (2005). Illusion of sense of self-agency: discrepancy between the predicted and actual sensory consequences of actions modulates the sense of self-agency, but not the sense of self-ownership. *Cognition*, *94*(3), 241-255.

Schiffman, J., Nakamura, B., Earleywine, M., & LaBrie, J. (2005). Symptoms of schizotypy precede cannabis use. *Psychiatry Research*, *134*(1), 37-42.

Schlagenhauf, F., Huys, Q. J., Deserno, L., Rapp, M. A., Beck, A., Heinze, H. J., ... & Heinz, A. (2014). Striatal dysfunction during reversal learning in unmedicated schizophrenia patients. *Neuroimage*, *89*, 171-180.

Seal, M., Aleman, A., & McGuire, P. (2004). Compelling imagery, unanticipated speech and deceptive memory: Neurocognitive models of auditory verbal hallucinations in schizophrenia. *Cognitive Neuropsychiatry*, *9*(1-2), 43-72.

Seeman, P., & Tallerico, T. (1998). Antipsychotic drugs which elicit little or no parkinsonism bind more loosely than dopamine to brain D2 receptors, yet occupy high levels of these receptors. *Molecular Psychiatry*, *3*(2), 123-134.

Shadmehr, R., Smith, M. A., & Krakauer, J. W. (2010). Error correction, sensory prediction, and adaptation in motor control. *Annual Review of Neuroscience*, *33*, 89-108.

Shakeel, M. K., & Docherty, N. M. (2012). Neurocognitive predictors of source monitoring in schizophrenia. *Psychiatry Research*, *200*(2), 173-176.

Shapiro, D. A., Renock, S., Arrington, E., Chiodo, L. A., Liu, L. X., Sibley, D. R., ... & Mailman, R. (2003). Aripiprazole, a novel atypical antipsychotic drug with

a unique and robust pharmacology. *Neuropsychopharmacology, 28*(8), 1400-1411.

Shawyer, F., Ratcliff, K., Mackinnon, A., Farhall, J., Hayes, S. C., & Copolov, D. (2007). The voices acceptance and action scale (VAAS): Pilot data. *Journal of Clinical Psychology*, *63*(6), 593-606.

Sheehan, D., Lecrubier, Y., Sheehan, K. H., Sheehan, K., Amorim, P., Janavs, J., Weiller, E. H. T., Hergueta, T., Baker, R. D. G., & Dunbar, G. (1998). Diagnostic Psychiatric Interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry*, *59*, 22-33.

Shergill, S. S., Brammer, M. J., Williams, S. C., Murray, R. M., & McGuire, P. K. (2000a). Mapping auditory hallucinations in schizophrenia using functional magnetic resonance imaging. *Archives of General Psychiatry*, *57*(11), 1033-1038.

Shergill, S. S., Bullmore, E., Simmons, A., Murray, R., & McGuire, P. (2000b). Functional anatomy of auditory verbal imagery in schizophrenic patients with auditory hallucinations. *American Journal of Psychiatry*, *157*(10), 1691-1693.

Shergill, S. S., Samson, G., Bays, P. M., Frith, C. D., & Wolpert, D. M. (2005). Evidence for sensory prediction deficits in schizophrenia. *American Journal of Psychiatry*, *162*(12), 2384-2386.

Shergill, S. S., White, T. P., Joyce, D. W., Bays, P. M., Wolpert, D. M., & Frith, C. D. (2014). Functional magnetic resonance imaging of impaired sensory prediction in schizophrenia. *JAMA Psychiatry*, *71*(1), 28-35.

Shinn, A. K., Heckers, S., & Öngür, D. (2013). The special treatment of first rank auditory hallucinations and bizarre delusions in the diagnosis of schizophrenia. *Schizophrenia Research*, *146*(1), 17-21.

Shoemaker, S. S. (1968). Self-reference and self-awareness. *The Journal of Philosophy*, *65*(19), 555-567.

Shtasel, D. L., Gur, R. E., Gallacher, F., Heimberg, C., & Gur, R. C. (1992). Gender differences in the clinical expression of schizophrenia. *Schizophrenia Research*, *7*(3), 225-231.

Simon, A. E., Cattapan-Ludewig, K., Zmilacher, S., Arbach, D., Gruber, K., Dvorsky, D. N., ... & Umbricht, D. (2007). Cognitive functioning in the schizophrenia prodrome. *Schizophrenia Bulletin*, *33*(3), 761-771.

Simons, J. S., Henson, R. N., Gilbert, S. J., & Fletcher, P. C. (2008). Separable forms of reality monitoring supported by anterior prefrontal cortex. *Journal of Cognitive Neuroscience*, *20*(3), 447-457.

Singh, S. P., Burns, T., Amin, S., Jones, P. B., & Harrison, G. (2004). Acute and transient psychotic disorders: precursors, epidemiology, course and outcome. *The British Journal of Psychiatry*, *185*(6), 452-459.

Snitz, B. E., MacDonald, A. W., & Carter, C. S. (2006). Cognitive deficits in unaffected first-degree relatives of schizophrenia patients: a meta-analytic review of putative endophenotypes. *Schizophrenia Bulletin*, *32*(1), 179-194.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: applications to dementia and amnesia. *Journal of Experimental Psychology: General*, *117*(1), 34-50.

Sommer, I. E., Diederen, K. M., Blom, J. D., Willems, A., Kushan, L., Slotema, K., ... & Kahn, R. S. (2008). Auditory verbal hallucinations predominantly activate the right inferior frontal area. *Brain*, *131*(12), 3169-3177.

Stagg, C., Hindley, P., Tales, A., & Butler, S. (2004). Visual mismatch negativity: the detection of stimulus change. *Neuroreport*, *15*(4), 659-663.

Startup, M., Startup, S., & Sedgman, A. (2008). Immediate source-monitoring, self-focused attention and the positive symptoms of schizophrenia. *Behaviour Research and Therapy*, *46*(10), 1176-1180.

Stefanis, N. C., Hanssen, M., Smirnis, N. K., Avramopoulos, D. A., Evdokimidis, I. K., Stefanis, C. N., ... & Van Os, J. (2002). Evidence that three dimensions of psychosis have a distribution in the general population. *Psychological Medicine*, *32*(02), 347-358.

Stephens, G. L., and Graham, G., 2000. When self-consciousness breaks: Alien voices and inserted thoughts. Cambridge, MA: MIT Press.

Sterzer, P., Frith, C., & Petrovic, P. (2008). Believing is seeing: expectations alter visual awareness. *Current Biology*, *18*(16), R697-R698.

Stroup, T. S., McEvoy, J. P., Swartz, M. S., Byerly, M. J., Glick, I. D., Canive, J. M., ... & Lieberman, J. A. (2003). The national institute of mental health clinical antipsychotic trials of intervention effectiveness (CATIE) project. *Schizophrenia Bulletin*, *29*(1), 15-31.

Sugimori, E., & Asai, T. (2015). Attribution of movement: Potential links between subjective reports of agency and output monitoring. *The Quarterly Journal of Experimental Psychology*, *68*(5), 900-916.

Sugimori, E., Asai, T., & Tanno, Y. (2010). Sense of agency over speech and proneness to auditory hallucinations: The reality monitoring paradigm. *Quarterly Journal of Experimental Psychology*, *11*, 1–17.

Sugimori, E., Asai, T., & Tanno, Y. (2011). Sense of agency over thought: External misattribution of thought in a memory task and proneness to auditory hallucination. *Consciousness and Cognition*, *20*(3), 688-695.

Synofzik, M., Thier, P., Leube, D. T., Schlotterbeck, P., & Lindner, A. (2010). Misattributions of agency in schizophrenia are based on imprecise predictions about the sensory consequences of one's actions. *Brain*, *133*(1), 262-271.

Synofzik, M., Vosgerau, G., & Newen, A. (2008). Beyond the comparator model: a multifactorial two-step account of agency. *Consciousness and Cognition*, *17*(1), 219-239.

Takahashi, Y. K., Roesch, M. R., Stalnaker, T. A., Haney, R. Z., Calu, D. J., Taylor, A. R., ... & Schoenbaum, G. (2009). The orbitofrontal cortex and ventral tegmental area are necessary for learning from unexpected outcomes. *Neuron*, *62*(2), 269-280.

Teufel, C., Kingdon, A., Ingram, J. N., Wolpert, D. M., & Fletcher, P. C. (2010). Deficits in sensory prediction are related to delusional ideation in healthy individuals. *Neuropsychologia*, *48*(14), 4169-4172.

Tobler, P. N., O'Doherty, J. P., Dolan, R. J., & Schultz, W. (2006). Human neural learning depends on reward prediction errors in the blocking paradigm. *Journal of Neurophysiology*, *95*(1), 301-310.

Toulopoulou, T., Rabe-Hesketh, S., King, H., Murray, R. M., & Morris, R. G. (2003). Episodic memory in schizophrenic patients and their relatives. *Schizophrenia Research*, *63*(3), 261-271.

Tsuang, M. T. (1991). Morbidity risks of schizophrenia and affective disorders among first-degree relatives of patients with schizoaffective disorders. *The British Journal of Psychiatry*, *158*(2), 165-170.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology/Psychologie Canadienne*, *26*(1), 1-12.
van Duppen, Z. (2016). The phenomenology of hypo-and hyperreality in psychopathology. *Phenomenology and the Cognitive Sciences*, *15*(3), 423-441.
van Os, J., Hanssen, M., Bijl, R. V., & Ravelli, A. (2000). Strauss (1969) revisited: a psychosis continuum in the general population?. *Schizophrenia Research, 45*(1), 11-20.
van Os, J., Linscott, R. J., Myin-Germeys, I., Delespaul, P., & Krabbendam, L. (2009). A systematic review and meta-analysis of the psychosis continuum: evidence for a psychosis proneness–persistence–impairment model of psychotic disorder. *Psychological Medicine*, *39*(02), 179-195.
Verdoux, H., & van Os, J. (2002). Psychotic symptoms in non-clinical populations and the continuum of psychosis. *Schizophrenia Research*, *54*(1), 59-65.
Vicente, A. (2014). The comparator account on thought insertion, alien voices and inner speech: some open questions. *Phenomenology and the Cognitive Sciences*, *13*(2), 335-353.
Vinogradov, S., Luks, T. L., Schulman, B. J., & Simpson, G. V. (2008). Deficit in a neural correlate of reality monitoring in schizophrenia patients. *Cerebral Cortex*, *18*(11), 2532-2539.
Vinogradov, S., Willis-Shore, J., Poole, J. H., Marten, E., Ober, B. A., & Shenaut, G. K. (1997). Clinical and neurocognitive aspects of source monitoring errors in schizophrenia. *American Journal of Psychiatry*, *154*(11), 1530-1537.
Vollema, M. G., Sitskoorn, M. M., Appels, M. C. M., & Kahn, R. S. (2002). Does the Schizotypal Personality Questionnaire reflect the biological–genetic vulnerability to schizophrenia?. *Schizophrenia Research*, *54*(1), 39-45.
Von Holst, E. (1954). Relations between the central nervous system and the peripheral organs. *The British Journal of Animal Behaviour*, *2*(3), 89-94.
Vosgerau, G., & Newen, A. (2007). Thoughts, motor actions, and the self. *Mind & Language*, *22*(1), 22-43.
Waltz, J. A., & Gold, J. M. (2007). Probabilistic reversal learning impairments in schizophrenia: further evidence of orbitofrontal dysfunction. *Schizophrenia Research*, *93*(1), 296-303.
Wang, L., Metzak, P. D., & Woodward, T. S. (2011). Aberrant connectivity during self–other source monitoring in schizophrenia. *Schizophrenia Research*, *125*(2), 136-142.
Waters, F., & Jardri, R. (2014). Auditory Hallucinations: Debunking the Myth of Language Supremacy. *Schizophrenia Bulletin*, sbu166.
Waters, F., Allen, P., Aleman, A., Fernyhough, C., Woodward, T. S., Badcock, J. C., ... & Larøi, F. (2012). Auditory hallucinations in schizophrenia and nonschizophrenia populations: a review and integrated model of cognitive mechanisms. *Schizophrenia Bulletin*, *38*(4), 683-693.
Weiss, A. P., Goff, D. C., Duff, M., Roffman, J. L., & Schacter, D. L. (2008). Distinguishing familiarity-based from source-based memory performance in patients with schizophrenia. *Schizophrenia Research*, *99*(1), 208-217.
Wilkinson, S. (2014). Accounting for the phenomenology and varieties of auditory verbal hallucination within a predictive processing framework. *Consciousness and Cognition*, *30*, 142-155.
Wiltink, S., Velthorst, E., Nelson, B., McGorry, P. M., & Yung, A. R. (2015). Declining transition rates to psychosis: the contribution of potential changes in

referral pathways to an ultra–high risk service. *Early Intervention in Psychiatry*, *9*(3), 200-206.
Winfield, H., & Kamboj, S. K. (2010). Schizotypy and mental time travel. *Consciousness and Cognition*, *19*(1), 321-327.
Wixted, J. T. (2007). Dual-process theory and signal-detection theory of recognition memory. *Psychological Review*, *114*(1), 152.
Wixted, J. T., & Stretch, V. (2004). In defense of the signal detection interpretation of remember/know judgments. *Psychonomic Bulletin & Review*, *11*(4), 616-641.
Wolff, J. L. (1967). Concept-shift and discrimination-reversal learning in humans. *Psychological Bulletin*, *68*(6), 369-408.
Wolpert, D. M. (1997). Computational approaches to motor control. *Trends in Cognitive Sciences*, *1*(6), 209-216.
Wood, S. J., Yung, A. R., McGorry, P. D., & Pantelis, C. (2011). Neuroimaging and treatment evidence for clinical staging in psychotic disorders: from the at-risk mental state to chronic schizophrenia. *Biological Psychiatry*, *70*(7), 619-625.
Woodruff, C. C., Hayama, H. R., & Rugg, M. D. (2006). Electrophysiological dissociation of the neural correlates of recollection and familiarity. *Brain Research*, *1100*(1), 125-135.
Woods, S. W., Addington, J., Cadenhead, K. S., Cannon, T. D., Cornblatt, B. A., Heinssen, R., ... & McGlashan, T. H. (2009). Validity of the prodromal risk syndrome for first psychosis: findings from the North American Prodrome Longitudinal Study. *Schizophrenia Bulletin*, *35*(5), 894-908.
Woodward, T. S., Menon, M., & Whitman, J. C. (2007). Source monitoring biases and auditory hallucinations. *Cognitive Neuropsychiatry*, *12*(6), 477-494.
Xia, J., Merinder, L. B., & Belgamwar, M. R. (2011). Psychoeducation for schizophrenia. *Cochrane Database of Systematic Reviews*, *6*, CD002831.
Yonelinas, A. P. (1994). Receiver-operating characteristics in recognition memory: evidence for a dual-process model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(6), 1341.
Yonelinas, A. P. (1997). Recognition memory ROCs for item and associative information: The contribution of recollection and familiarity. *Memory & Cognition*, *25*(6), 747-763.
Yonelinas, A. P., Kroll, N. E., Dobbins, I., Lazzara, M., & Knight, R. T. (1998). Recollection and familiarity deficits in amnesia: convergence of remember-know, process dissociation, and receiver operating characteristic data. *Neuropsychology*, *12*(3), 323.
Yonelinas, A. P. (1999). The contribution of recollection and familiarity to recognition and source-memory judgments: A formal dual-process model and an analysis of receiver operating characterstics. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(6), 1415-1434.
Yonelinas, A. P. (2001). Components of episodic memory: the contribution of recollection and familiarity. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *356*(1413), 1363-1374.
Yonelinas, A. P., Otten, L. J., Shaw, K. N., & Rugg, M. D. (2005). Separating the brain regions involved in recollection and familiarity in recognition memory. *Journal of Neuroscience*, *25*(11), 3002-3008.
Yu, S. S., & Rugg, M. D. (2010). Dissociation of the electrophysiological correlates of familiarity strength and item repetition. *Brain Research*, *1320*, 74-84.

Yung, A. R., & McGorry, P. D. (1996). The prodromal phase of first-episode psychosis: past and current conceptualizations. *Schizophrenia Bulletin*, *22*(2), 353-370.

Yung, A. R., Buckby, J. A., Cotton, S. M., Cosgrave, E. M., Killackey, E. J., Stanford, C., ... & McGorry, P. D. (2006). Psychotic-like experiences in nonpsychotic help-seekers: associations with distress, depression, and disability. *Schizophrenia Bulletin*, *32*(2), 352-359.

Yung, A. R., Nelson, B., Stanford, C., Simmons, M. B., Cosgrave, E. M., Killackey, E., ... & McGorry, P. D. (2008). Validation of "prodromal" criteria to detect individuals at ultra high risk of psychosis: 2 year follow-up. *Schizophrenia Research*, *105*(1), 10-17.

Yung, A. R., Yuen, H. P., McGorry, P. D., Phillips, L. J., Kelly, D., Dell'Olio, M., ... & Buckby, J. (2005). Mapping the onset of psychosis: the Comprehensive Assessment of At Risk Mental States. *Australian and New Zealand Journal of Psychiatry*, *39*(11 12), 964-971.

Zammit, S., Allebeck, P., David, A. S., Dalman, C., Hemmingsson, T., Lundberg, I., & Lewis, G. (2004). A longitudinal study of premorbid IQ score and risk of developing schizophrenia, bipolar disorder, severe depression, and other nonaffective psychoses. *Archives of General Psychiatry*, *61*(4), 354-360.

Zhang, H., & Sulzer, D. (2004). Frequency-dependent modulation of dopamine release by nicotine. *Nature Neuroscience*, *7*(6), 581-582.

Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica*, *67*(6), 361-370.