COGNITIVE BIASES AND FUNCTIONAL BRAIN NETWORKS UNDERLYING
DELUSIONS IN SCHIZOPHRENIA

by

KATIE M. LAVIGNE

B.A., Concordia University, 2008
M.Sc., The University of British Columbia, 2013

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Neuroscience)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

February 2018

Abstract

Integrating evidence that contradicts a belief is a fundamental aspect of belief revision and is closely linked to delusions in schizophrenia. In this research, we examined the cognitive and brain mechanisms underlying disconfirmatory evidence integration, their relation to delusions and the bias against disconfirmatory evidence (BADE) in schizophrenia, as well as associations between changes in delusion severity, BADE, and functional brain activity related to evidence integration. Across three neuroimaging studies, two functional brain networks emerged as central to disconfirmatory evidence integration: a visual attention network (VsAN) including dorsal anterior cingulate cortex and bilateral insula; and a cognitive evaluation network (CEN) involving rostrolateral/orbitofrontal cortex, inferior frontal gyrus, and inferior parietal lobule. In study 1, these networks showed sequential activation and increased activity during disconfirmatory evidence integration, suggesting they were involved in distinct evidence detection and integration processes. In study 2, we found that activity in these networks was differentially associated with delusions, with delusional schizophrenia patients showing VsAN hyperactivity and CEN hypoactivity relative to controls. Subsequent analyses examining associations between activity in these networks and behaviour revealed that BADE was positively associated with VsAN activity during confirmatory evidence integration, and negatively associated with CEN activity during disconfirmatory evidence integration. These findings indicate that VsAN hyperactivity underlies the focus on confirmatory evidence, and CEN hypoactivity the avoidance of disconfirmatory evidence, that contributes to impaired evidence integration and delusion maintenance in schizophrenia. Finally, in study 3, we demonstrated that poorer evidence integration over time was related to greater hyperactivity in the VsAN and

hypoactivity in the CEN, from time 1 to time 2, and that improved positive symptoms (including delusions) were associated with normalization of activity in the CEN, showing that activity in these networks fluctuates as a function of changes in behavioural BADE evidence integration and symptoms. This research represents the first comprehensive study of the cognitive and brain mechanisms underlying disconfirmatory evidence integration and behavioural BADE in schizophrenia patients with delusions, and highlights brain networks underlying cognitive biases related to important aspects of delusion maintenance in schizophrenia: the focus on confirmatory evidence; and the avoidance of disconfirmatory evidence.

Lay Summary

Integrating new information is a fundamental aspect of belief revision, and impairments in this process have been related to delusions (false beliefs) in schizophrenia. Across three studies, this research identified alterations in networks of brain regions underlying the focus on confirmatory evidence, and avoidance of disconfirmatory evidence, seen in schizophrenia patients with delusions. These brain networks were found to be differentially related to both delusions and cognitive biases, which are biased thought processes that can affect decision-making. Brain activity was separable from neurocognition (e.g., memory and attention), and was also found to fluctuate with changes in symptoms and cognitive biases over time. This research represents the first comprehensive study of the cognitive and brain mechanisms underlying disconfirmatory evidence integration and BADE in schizophrenia patients with delusions, and highlights brain networks underlying important aspects of delusion maintenance in schizophrenia: the focus on confirmatory evidence; and the avoidance of disconfirmatory evidence.

Preface

This dissertation is the original intellectual product of the author, Katie Lavigne, and all work was conducted in the Cognitive Neuropsychiatry of Schizophrenia Lab at the University of British Columbia, under the supervision of Dr. Todd Woodward. All studies were approved by the University of British Columbia Clinical Research Ethics Board (Certificate numbers: H07-02786, H14-01434, and H12-01968, for Chapters 2 to 4, respectively), and the research described in Chapter 4 was also registered as a Clinical Trial with the ClinicalTrials.gov Protocol Registration System (NCT01764568).

The work presented in Chapter 2 ("Functional brain networks underlying detection and integration of disconfirmatory evidence") has been published as "Lavigne, K.M., Metzak, P.D., and Woodward, T.S. (2015). Functional brain networks underlying detection and integration of disconfirmatory evidence. *NeuroImage, 112*, 138-151", and is presented in published form, with the exception of some minor changes made to incorporate new knowledge gained from this research, and to maintain continuity throughout this dissertation. I was responsible for generating the research hypotheses, analyzing and interpreting the results, and writing of the manuscript. Dr. Metzak was involved in designing and conducting the experiment, and contributed intellectually to interpretation of the results and manuscript revisions. Dr. Woodward was lead investigator on this project and contributed throughout the project in terms of design, analysis, and manuscript revisions.

The research reported in Chapter 3 ("Functional brain networks underlying evidence integration and delusions in schizophrenia") is currently under review for publication, and has been presented as part of symposia at the following international conferences: *Organization for Human Brain Mapping* (Lavigne, K.M. (June, 2017). Functional brain

networks underlying impaired disconfirmatory evidence integration in schizophrenia. 23rd Annual Meeting of the Organization for Human Brain Mapping, Oral Session: Higher Cognitive Functions, Vancouver, BC), *Japan Neuroscience Society* (Lavigne, K.M. (July, 2017). Functional brain networks underlying the bias against disconfirmatory evidence in schizophrenia. 40th Annual Meeting – Japan Neuroscience Association, Symposium "Current status of neurocognitive bias research for delusions", Chiba-city, Japan), and *Society for Neuroscience* (Lavigne, K.M. (November, 2017). Functional brain networks underlying the bias against disconfirmatory evidence and delusions in schizophrenia. 47th Annual Meeting - Society for Neuroscience, Nanosymposium "Human studies of circuits and systems in schizophrenia and first episode psychosis", Washington, DC, USA). The research presented in Chapter 4 ("Changes in BADE, delusions, and functional brain activity underlying evidence integration in schizophrenia") is currently being prepared for publication in a journal to be determined.

I was involved in generating the hypotheses of this research, conducting the experiments and statistical analyses (including developing software to batch preprocess fMRI data), designing the novel experimental task used in Chapters 3 and 4, supervising and training volunteers and research coordinators, as well as interpreting the results and writing the manuscripts and this dissertation. Dr. Woodward supervised all aspects of this work.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AN(C)OVA	Analysis of (co)variance
BA	Brodmann area
BADE	Bias against disconfirmatory evidence
BDI	Beck Depression Inventory
BOLD	Blood-oxygen-level-dependent
BP	Baseline-to-peak
CEN	Cognitive evaluation network
COWAT	Controlled Oral Word Association Test
(C)PCA	(Constrained) Principal component analysis
(d)ACC	(Dorsal) Anterior cingulate cortex
DAN	Dorsal attention network
DLPFC	Dorsolateral prefrontal cortex
DSM	Diagnostic and Statistical Manual of Mental Disorders
EPI	Echo planar imaging
EVH	Evidence-hypothesis
FA	Flip angle
FIR	Finite impulse response
(f)MRI	(Functional) Magnetic resonance imaging
FOV	Field of view
FPN	Frontoparietal network
FWHM	Full-width-at-half-maximum
GFI	Goodness of Fit index

GSCA	Generalized structured component analysis
HDR	Hemodynamic response
HSD	Honest significant difference
IFG	Inferior frontal gyrus
IQ	Intelligence quotient
ITI	Inter-trial interval
JTC	Jumping to conclusions
LNS	Letter-Number Sequencing
MINI	Mini International Neuropsychiatric Interview
MNI	Montreal Neurological Institute
OFC	Orbitofrontal cortex
PB	Peak-to-baseline
PSYRATS	Psychotic Symptoms Rating Scales
RBANS	Repeatable Battery for the Assessment of Neuropsychological Status
ROI(s)	Region(s) of interest
(r)PFC	(Rostrolateral) Prefrontal cortex
RT	Reaction time
SANS	Scale for the Assessment of Negative Symptoms
SAPS	Scale for the Assessment of Positive Symptoms
SD	Standard deviation
SEM	Structural equation modeling
SMA	Supplementary motor area
SN	Salience network

SPM	Statistical Parametric Mapping
SPQ	Schizotypal Personality Questionnaire
SPSS	Statistical Package for the Social Sciences
SRMR	Standardized Root Mean Square Residual
T	Tesla
TE	Time to echo
TMT	Comprehensive Trail-Making Test
ToPF	Test of Premorbid Functioning
TR	Time to repetition
UBC	University of British Columbia
VAN	Ventral attention network
(V)DMN	(Visual/)Default-mode network
vmPFC	Ventromedial prefrontal cortex
VsAN	Visual attention network
WASI	Wechsler Abbreviated Scale of Intelligence
YY/NN/NY/YN	Yes-Yes/No-No/No-Yes/Yes-No

To my parents, Allan Lavigne and Deborah O'Brien,
for providing me with the skills to succeed,
and the opportunity to discover my passion.

1. Introduction

1.1. Schizophrenia

Schizophrenia is a debilitating neuropsychiatric disorder characterized by a wide range of cognitive and behavioural dysfunctions that lead to severe impairments in functioning (e.g., occupational, interpersonal, self-care). The economic and societal burden of schizophrenia can be substantial, despite lifetime prevalence rates of less than 1% (McGrath, Saha, Chant, & Welham, 2008; Saha, Chant, Welham, & McGrath, 2005; Simeone, Ward, Rotella, Collins, & Windisch, 2015). In addition to the severe burden placed on patients, family members, and caregivers (Millier et al., 2014; Rossler, Salize, van Os, & Riecher-Rossler, 2005), the annual economic impact of schizophrenia has been estimated at over $6 billion in Canada (Goeree et al., 2005), which has a higher prevalence and incidence of schizophrenia than other countries, and appears to be increasing over time (Dealberto, 2013). The majority of these costs are due to productivity losses resulting from unemployment or reduced earnings in those who are able to work (Chong et al., 2016; Goeree et al., 2005; Jin & Mosweu, 2017).

The Diagnostic and Statistical Manual of Mental Disorders (5th ed., DSM-V; American Psychiatric Association, 2013) defines several symptoms that are characteristic of schizophrenia: delusions (false beliefs), hallucinations (false perceptions), disorganized thought, abnormal motor behaviour (e.g., catatonia), and negative symptoms (e.g., avolition, anhedonia). Among these, delusions and hallucinations are the hallmark symptoms of psychosis. Delusions are beliefs about oneself, one's internal or external environment that are not supported by evidence, and are generally believed by others to be untrue. Hallucinations are perceptions (e.g., visual, auditory, olfactory) that are experienced in the absence of

external stimuli, most commonly voices. Much like hallucinations can occur in different sensory modalities, the content of delusions can vary widely (American Psychiatric Association, 2013; Coltheart, Langdon, & McKay, 2011). Persecutory delusions, which involve a belief that one is being followed, threatened, or at risk of being harmed, are most prevalent. Other types include grandiose delusions (belief that one has exceptional abilities or is an influential figure), referential delusions (belief that mundane events have personal meaning), delusions of thought withdrawal and/or insertion (belief that thoughts are being taken from/put into one's mind), and delusions of control (belief that one's body/actions are being manipulated by an external force). Despite the large variability in the potential content of delusions, a common feature underlying all delusions is that, although they are unsupported by – or even contradicted by – valid evidence, they are often held with unwavering conviction. When faced with disconfirmatory evidence, individuals with delusions find ways to dismiss it as invalid, or contort the evidence in such a way that it can be incorporated into the delusional belief (Freeman, Garety, Kuipers, Fowler, & Bebbington, 2002; Garety, 1992; Joseph, 1986; Simpson & Done, 2002). This inability or unwillingness to properly integrate evidence that contradicts beliefs prevents the delusion from being discarded, and is thought to be a major contributing factor to the maintenance of delusions in schizophrenia (Freeman et al., 2007; Woodward, Moritz, Cuttler, & Whitman, 2006b).

1.2. Theories of Delusions in Schizophrenia

One of the earliest explanations put forth on the manifestation of delusions was that they were a rational way of making sense of the perceptual abnormalities (i.e., hallucinations) that commonly accompanied delusions in schizophrenia (Maher, 1974, 1988). Although hallucinations often co-occur, and may be consistent, with delusions (Coltheart et al., 2011),

hallucinations are neither necessary nor sufficient for their manifestation (Bell, Halligan, & Ellis, 2008). First, abnormal percepts/hallucinations do not always precede delusions, and are not always present in deluded individuals; second, and most relevant to this research, there is now a large literature documenting numerous cognitive biases and deficits in schizophrenia and their relation to delusions, showing that reasoning processes are impaired in schizophrenia and may contribute to delusions (Blackwood, Howard, Bentall, & Murray, 2001; Garety & Freeman, 1999, 2013; Gilleen & David, 2005). Thus, while delusion-consistent hallucinations may reinforce delusional beliefs, cognitive dysfunction likely represents a major factor in the formation and maintenance of delusions, especially in chronic psychotic disorders such as schizophrenia.

The distinction between delusion formation and delusion maintenance is an important one, as these may involve separate cognitive processes, one or both of which might be affected in schizophrenia patients with delusions. Delusion formation refers to the development of the delusional belief, including its content, whereas delusion maintenance refers to the persistence of the belief (Freeman et al., 2002). Delusion formation has been described as a search for meaning that serves to explain experience (Freeman et al., 2002; Maher, 1974), which may or may not be driven by anomalous perceptions such as hallucinations. Several theoretical models have been proposed to explain delusion formation, some of which distinguish between cognitive processes related to the content and those related to the presence of delusions. For example, delusional content may be determined by perceptual abnormalities, emotional trauma, and/or attributional biases (Coltheart et al., 2011; Freeman et al., 2002; Garety, Kuipers, Fowler, Freeman, & Bebbington, 2001). Given that most individuals with perceptual aberrations or negative life experiences do not develop

delusions, and in conjunction with the growing literature on reasoning biases related to delusions, most researchers agree that one or more additional factors (e.g., cognitive biases or impaired reasoning/belief formation) are necessary for the manifestation of delusions. One of the most widely-studied cognitive biases in schizophrenia is jumping to conclusions (JTC), which refers to a tendency to reach a decision with very little evidence (Freeman et al., 2002; Garety & Freeman, 2013). JTC may contribute to the formation of delusions in that a belief will be adopted at a lower plausibility threshold than would be observed in the normal population. Belief inflexibility and intolerance for ambiguity may also reduce the threshold for plausibility of a delusional belief (Freeman et al., 2002). Other cognitive processes, such as theory of mind (Frith, 1992, 2014) and attributional biases (Bentall, Corcoran, Howard, Blackwood, & Kinderman, 2001; Bentall, Kinderman, & Kaney, 1994), may contribute to delusion formation through biased judgments in terms of interpreting the thoughts of others and assigning external attributions to events (Garety & Freeman, 1999).

In terms of delusion maintenance, it may be driven by two distinct but related processes, namely, focusing on obtaining evidence that confirms the belief and avoiding or disregarding evidence that contradicts the belief (Broyd, Balzan, Woodward, & Allen, 2017; Freeman et al., 2002). Gathering confirmatory evidence may be influenced by common cognitive biases seen in the general population that are exacerbated in schizophrenia, such as confirmation bias, attentional biases, and self-fulfilling prophecy. Additional cognitive biases and aberrant reasoning processes may lead disconfirmatory evidence to be discarded. For example, safety behaviours, such as avoiding situations which might contradict a delusional belief, are often practiced by individuals with delusions, and prevent the disconfirmatory evidence from coming to light (e.g., "the only reason they didn't try to kill me today was because I took a

4

different way to work"; Freeman et al., 2007). Individuals with delusions also attempt to incorporate disconfirmatory evidence into their delusions in such a way that it is no longer contrary to the belief (e.g., "the only reason they didn't try to kill me today is because they want to watch me suffer"; Freeman et al., 2002; Garety et al., 2001). While these rationalizations are often observed when delusional patients are presented with evidence that disconfirms their own delusions, reasoning-based accounts of delusion maintenance in schizophrenia should also be relevant for delusion-neutral material, as is the case for biases relating to delusion formation, such as JTC (Broyd et al., 2017). One promising general cognitive bias implicated in the maintenance of delusions is the bias against disconfirmatory evidence (BADE).

1.3. BADE and Evidence Integration

BADE refers to an unwillingness to revise a belief or conviction in light of evidence that contradicts that belief, and has been shown to be exacerbated in schizophrenia patients with delusions using an experimental paradigm in which a story gradually unfolds over three pictures or sentences, after each of which four potential interpretations of the story are rated (Woodward et al., 2006b). One interpretation is true, and becomes increasingly likely after each sentence, another is absurd with low likelihood throughout the trial, and the remaining two interpretations are lures, which are initially plausible, but become less plausible as more information is provided. Schizophrenia patients with delusions demonstrate a BADE by not decreasing their ratings on lure interpretations as much as healthy individuals when these become less plausible (i.e., not modifying their interpretations in light of new, contradictory evidence).

The BADE phenomenon has been observed in schizophrenia patients with, but not without, delusions (Sanford, Veckenstedt, Moritz, Balzan, & Woodward, 2014; Speechley, Moritz, Ngan, & Woodward, 2012; Woodward et al., 2006b), as well as in healthy individuals with high delusional ideation (Buchy, Woodward, & Liotti, 2007; Woodward, Buchy, Moritz, & Liotti, 2007). Moreover, it is not prominent in other psychiatric diagnoses, such as non-delusional bipolar disorder (Speechley et al., 2012) and obsessive-compulsive disorder (Sanford et al., 2014), the latter showing similar rigidity in beliefs to delusions in schizophrenia. A recent meta-analysis (McLean, Mattiske, & Balzan, 2017) of BADE and three other cognitive biases implicated in delusions (including JTC) supported this work, suggesting that they are specific to delusions in schizophrenia and not simply the result of general psychiatric illness. Together, these findings point to BADE, in addition to other cognitive biases, as an important contributor to psychotic delusions and sub-clinical delusional ideation in the schizophrenia spectrum. BADE is computed on delusion-neutral material, suggesting that a BADE is a generalized brain bias that affects all types of information processing, not only that which is the subject of individuals' delusions.

The BADE paradigm taps into cognitive processes related to evidence integration, particularly for evidence that contradicts an initial interpretation or belief. Evidence integration is a multifaceted construct fundamental to belief revision, and serves to update one's belief system so that it corresponds to real-world experiences (Lavigne, Metzak, & Woodward, 2015a). This process allows individuals to develop an increasingly comprehensive understanding of the world around them through the gathering of novel, and often contradictory, information. When this information is valid, but conflicts with a currently-held belief (i.e., disconfirmatory evidence), the belief must be modified or

abandoned in order to assimilate the evidence and maintain a coherent belief system. The process of integrating disconfirmatory evidence involves, at minimum, two stages (see Chapter 2 and Lavigne et al., 2015a). First, the evidence must be identified as being in conflict with one or more already-established beliefs. This initial alerting or detection stage serves to identify the evidence as relevant to one or more beliefs, and, in the case of disconfirmatory evidence, notes the disparity between the evidence and the belief in question, setting in motion a series of processes aimed at resolving the conflict. Second, once the evidence is identified as being relevant to a belief (or in conflict with a belief in the case of disconfirmatory evidence), the belief must be evaluated in light of the new information in order to determine the extent to which it should be modified to integrate the evidence (given that the evidence is deemed credible). These *detection* and *integration* stages of belief revision are most prominent when the evidence contradicts a belief, but are also involved, to a lesser extent, during the integration of confirmatory evidence (i.e., evidence is identified as relevant to the belief instead of conflictual, and the belief is updated to take into account the new information). Although as of yet untested, dysfunction in one or both of these stages may contribute to BADE in schizophrenia patients with delusions.

1.4. Functional Brain Activity Underlying Evidence Integration

Although the brain mechanisms underlying BADE have yet to be established, some previous research has examined functional brain activity associated with evidence integration and belief revision, finding increased activity in several overlapping brain regions. Specifically, increased activity in dorsal anterior cingulate cortex (dACC) and adjoining medial superior frontal gyrus is observed during uncertainty in belief revision (Behrens, Woolrich, Walton, & Rushworth, 2007; Volz, Schubotz, & von Cramon, 2004; Yoshida &

Ishii, 2006), and, along with the anterior insula and other regions, the dACC has been implicated in evidence accumulation and the moment of recognition of a stimulus (Krueger et al., 2017; Liu & Pleskac, 2011; Ploran et al., 2007). In addition, right rostrolateral prefrontal cortex (rPFC) is preferentially activated when evidence contradicts an original expectation (i.e., prediction error; Corlett et al., 2004; Fletcher et al., 2001; Turner et al., 2004), and has been theorized to underlie hypothesis evaluation impairments related to delusions (Coltheart et al., 2011). Similarly, inferior frontal gyrus (IFG) has been implicated in belief formation and updating (d'Acremont, Schultz, & Bossaerts, 2013; Sharot, Korn, & Dolan, 2011), and there is evidence that disruption of left IFG via transcranial magnetic stimulation improves integration of unfavourable evidence (Sharot et al., 2012), suggesting it may play a key role in disconfirmatory evidence integration in particular. While these findings point to specific brain regions of interest (ROIs) that may underlie evidence integration processes, the field of neuroscience is moving towards a network-based, rather than region-based, understanding of brain structure and function. Using both resting-state and task-based functional magnetic resonance imaging (fMRI) along with multivariate statistical analysis techniques, researchers have now identified several interconnected networks of brain regions and have associated these with various cognitive processes and psychiatric diseases. Comparison of these networks with previous ROI task-based studies on evidence integration can inform our understanding of the cognitive functions underlying activity within these networks and how they might relate to evidence integration and BADE.

Some resting-state-derived functional brain networks likely to be of greatest relevance to evidence integration are the salience (SN), ventral attention (VAN), and frontoparietal (FPN) networks, because of their overlap with regions implicated in evidence integration in ROI-

based studies. The SN includes nodes in the dACC and bilateral insula, and is involved in detection of subjectively salient stimuli (Menon & Uddin, 2010; Seeley et al., 2007; Uddin, 2015). The VAN is a right-lateralized network including temporoparietal junction and ventral frontal cortex (Fox, Corbetta, Snyder, Vincent, & Raichle, 2006) as well as dACC and insula (Yeo et al., 2011), and is involved in stimulus-driven attention (Corbetta & Shulman, 2002). The FPN includes rPFC, IFG, and inferior parietal lobule, and is involved in initiating control processes and adjusting responses based on feedback (Dosenbach et al., 2007), as well as in integrating information from other brain networks (Vincent, Kahn, Snyder, Raichle, & Buckner, 2008). In some studies (Vincent et al., 2008; Yeo et al., 2011), the FPN also includes dACC and insula, suggesting a degree of overlap with the SN, as is observed for the VAN; however, other research has distinguished between SN and FPN configurations (Dosenbach et al., 2007; Seeley et al., 2007), suggesting they are involved in distinct salience/task demand and control processes. These discrepancies in the resting-state literature might reflect methodological differences or might be due to the SN's involvement in modulating other large-scale networks in response to salient events (Goulden et al., 2014; Menon & Uddin, 2010), which could lead it to be captured on other networks due to similar timing of activation. This may also be the case for other regions (e.g., IFG, rPFC) that show overlap between different resting-state networks, indicating they are highly-connected hub regions involved in multiple large-scale brain networks.

Importantly, the functional brain networks that emerge from task-based fMRI studies are likely to differ from those identified using resting-state brain activity, especially if they are derived from task-specific variance in blood-oxygen-level-dependent (BOLD) signal (Lavigne, Menon, & Woodward, 2016; Lavigne et al., 2015a). Nonetheless, both ROI-based

and resting-state studies can inform our hypotheses with respect to networks that might be involved in the detection and integration stages of evidence integration. First, given the involvement of the SN and VAN in salience processing and stimulus-oriented attention, an SN/VAN-like network, with nodes in the dACC and insula, would be expected to be involved in the detection of evidence. The dACC in particular is involved in alerting to requirements to adjust behaviour and change mental set (Behrens et al., 2007; Whitman, Metzak, Lavigne, & Woodward, 2013; Woodward, Metzak, Meier, & Holroyd, 2008), and both the dACC and insula have been implicated in evidence integration as described above in ROI-based studies, particularly during the moment of recognition or Aha! Moment (Ploran et al., 2007; Whitman et al., 2013; Woodward et al., 2008). Thus, an SN/VAN-like network may be specifically recruited during the evidence detection stage (immediately following evidence presentation), with increased activity during disconfirmatory relative to confirmatory evidence due to the increased salience of evidence that contradicts a belief.

Second, an FPN-like network, including rPFC and IFG, might be involved in the integration of evidence, due to the importance of these regions in the evaluation of self-generated information (Christoff, Ream, Geddes, & Gabrieli, 2003), and in the comparison of internal representations and external information (Burgess, Dumontheil, & Gilbert, 2007; Gilbert, Spengler, Simons, Frith, & Burgess, 2006a; Gilbert et al., 2006b), which could include beliefs and evidence, respectively. Regions within the FPN have also been identified separately as relating to evidence integration, most notably for IFG (Sharot et al., 2012; Sharot et al., 2011) and rPFC (Coltheart et al., 2011). Given the FPN's involvement in control and integration processes, and in line with Coltheart and colleagues (2011) proposal that rPFC is critical for hypothesis evaluation, an FPN-like network may be preferentially

recruited during the integration of evidence (following evidence presentation but peaking later than the proposed SN/VAN-like network), with greater activity for disconfirmatory relative to confirmatory information due to the increased need for evaluative and integrative processes in the former.

In order to test these network-based hypotheses, we examined functional brain networks underlying the detection and integration of evidence in healthy individuals (see Chapter 2 and Lavigne et al., 2015a), identifying three functional brain networks associated with disconfirmatory evidence integration: (1) a visual attention network including dACC and anterior insula, overlapping primarily with the SN/VAN; (2) a cognitive evaluation network (rPFC, orbitofrontal cortex (OFC), IFG) overlapping primarily with the FPN; and (3) a response network. The visual attention and cognitive evaluation networks showed staggered peaks of activation, with activity in the former preceding that of the latter, and peak activity in the response network was situated between the visual attention and cognitive evaluation networks. This suggests distinct cognitive processes underlying all three networks, due to their very different spatial configurations and peak timing patterns. Moreover, peak timing occurred following the presentation of evidence in all networks, and showed higher activity in the disconfirm relative to confirm conditions, indicating that these networks were involved in distinct aspects of evidence integration specifically. This sequential activation, and increased recruitment during disconfirmatory evidence integration, is in line with the notion that an SN/VAN-like visual attention network serves to identify the evidence as relevant (i.e., salient) to one or more beliefs, while an FPN-like cognitive evaluation network is involved in evaluating the belief in light of the new evidence and subsequently integrating that evidence and modifying the belief.

Schizophrenia patients show abnormal connectivity within and between multiple brain networks, including the SN, VAN, and FPN. For example, in fMRI studies, schizophrenia patients have been reported to show reduced functional connectivity between the anterior insula and dACC nodes of the SN as well as between these nodes and other large-scale networks, such as the central-executive network, which has common nodes with the FPN, and the well-known default-mode network (DMN; White, Joseph, Francis, & Liddle, 2010). Similarly, decreased functional connectivity within the SN has been reported to be associated with positive symptoms, as well as with increased functional connectivity between the anti-correlated DMN and central-executive network (Manoliu et al., 2014). Reduced functional connectivity has also been reported in schizophrenia patients in the VAN and FPN (Baker, Holmes, Masters, & et al., 2014). These findings are supported by structural MRI research reporting grey matter volume reductions in the anterior insula and dACC, which were related to the severity of positive symptoms in schizophrenia (Palaniyappan, Mallikarjun, Joseph, White, & Liddle, 2011). Overall, these network findings suggest that communication within the SN, and its proposed modulation of other large-scale networks, including the FPN, is compromised in schizophrenia and may contribute to positive symptoms, including delusions (Palaniyappan & Liddle, 2012). Within a BADE/evidence integration framework, these findings would suggest that disconfirmatory evidence might not be perceived as a salient event, which would be reflected as decreased activity in SN/VAN regions (e.g., visual attention network) immediately following the presentation of disconfirmatory evidence and, subsequently, decreased activity in FPN regions (e.g., cognitive evaluation network), ultimately hindering integration of evidence and the belief from being revised. However, functional brain network activity in schizophrenia patients (with or without delusions) during

evidence integration has yet to be investigated, and it is unknown whether this putative

dysfunction might be characteristic of all patients, or just those with delusions.

1.5. BADE and Neurocognition

Schizophrenia patients show widespread deficits in neurocognitive functions, including

memory, attention, and executive functioning (Aleman, Hijman, de Haan, & Kahn, 1999;

Heinrichs, 2005; Heinrichs & Zakzanis, 1998; Kahn & Keefe, 2013). Impairment in one or

more of these processes can adversely impact performance on a wide range of cognitive

tasks, including those assessing cognitive biases. However, these cognitive deficits are not

generally correlated with delusion severity (Addington, Addington, & Maticka-Tyndale,

1991; Aleman et al., 1999; Basso, Nasrallah, Olson, & Bornstein, 1998; Berman et al., 1997;

Cuesta & Peralta, 1995; Frith, Leary, Cahill, & Johnstone, 1991; Moritz, Heeren, Andresen,

& Krausz, 2001; O'Leary et al., 2000). It is, therefore, necessary to establish that a general

cognitive bias proposed to underlie delusions in schizophrenia (i.e., BADE) diverges

empirically from neurocognitive functioning, demonstrating that it is not simply an

alternative measure of one or more neurocognitive processes.

Some previous research has investigated the role of neurocognitive measures on BADE

to examine whether these are separable (for a review, see Eisenacher & Zink, 2017).

Woodward and colleagues (2007) conducted a principal component analysis on BADE items

and neurocognitive variables (i.e., IQ, memory, executive function, and attention) using data

from healthy individuals and found that BADE items form distinct components that are

orthogonal to neurocognition. In schizophrenia patients, Moritz *et al.* (2010) found no

evidence for a correlation between BADE and executive functioning or verbal learning, two

neurocognitive processes which show some of the most severe impairments in schizophrenia

(Heinrichs & Zakzanis, 1998). However, other studies have shown associations between BADE and measures of IQ, theory of mind, and executive functioning (Eifler et al., 2014; Moritz et al., 2010; Riccaboni et al., 2012), suggesting that there is some overlap between these constructs. Finally, in a study examining the effects of psychological treatment on BADE, Buonocore and colleagues (2015) found that cognitive remediation therapy, which targets neurocognitive functioning, was not sufficient to improve BADE, which only benefited from a combined treatment including cognitive remediation and metacognitive training, a cognitive intervention targeting cognitive biases underlying delusions. These and other findings were summarized in a recent review paper of the BADE literature (Eisenacher & Zink, 2017), supporting the notion that while neurocognitive functions may contribute to BADE, they cannot fully explain it, because there remains unique variance that cannot be predicted by neurocognition. Thus, the role of neurocognition on BADE remains equivocal, and additional examination of these associations using more comprehensive neurocognitive measures and more complex statistical techniques is necessary to understand the degree to which these measures overlap, and how they might relate to the symptoms of psychosis. Moreover, it is unknown whether neurocognitive abilities influence functional brain networks underlying BADE and evidence integration in schizophrenia.

1.6. Changes in BADE, Delusions, and Functional Brain Activity Underlying Evidence Integration in Schizophrenia

In addition to being detectable using delusion-neutral testing material, and separable from neurocognition, any cognitive bias hypothesized to underlie delusions would be expected to fluctuate with changes in delusion severity. For example, there is evidence that decreases in delusion severity are associated with decreases in the JTC bias (Sanford, Lecomte, Leclerc,

Wykes, & Woodward, 2013; Woodward, Munz, Leclerc, & Lecomte, 2009), and that reasoning processes such as JTC and belief flexibility may moderate changes in delusion severity (Broyd et al., 2017; Garety & Freeman, 2013; So et al., 2012), suggesting that JTC captures some of the cognitive processes underlying delusion severity. Less research has been conducted examining the relationship between changes in BADE and delusions, however. Cross-sectional studies indicate that BADE is exacerbated in groups with worsening delusion severity; that is, BADE shows a linear increase from healthy individuals with high delusional ideation to at-risk patients to first-episode psychosis patients (Eisenacher et al., 2016; Eisenacher & Zink, 2017). There is also evidence that cognitive interventions targeting cognitive biases improve both symptoms and cognitive biases (Kumar et al., 2010; Moritz, Veckenstedt, Randjbar, Vitzthum, & Woodward, 2011; Ross, Freeman, Dunn, & Garety, 2011), including BADE (Buonocore et al., 2015), suggesting a common underlying mechanism. However, only one study to date has specifically examined whether changes in BADE are associated with changes in delusion severity, finding no significant correlation (Buonocore et al., 2015).

In order to better understand the relation between delusions and evidence integration, it is important to examine whether changes in delusion severity correspond to changes in both behaviour and functional brain activity underlying evidence integration. If poor evidence integration contributes to delusion maintenance in schizophrenia, changes in delusion severity would be expected to correspond to changes in BADE behaviourally. Similarly, if aberrant activity in a given network (e.g., the visual attention or cognitive evaluation networks) reflects the BADE observed in schizophrenia patients with delusions, improvements in BADE and/or delusions would be expected to correspond to normalization

of activity within that network, just as further impairment in BADE and/or worsening delusion severity should correspond to exacerbation of any hyper/hypoactivity observed in delusional patients relative to controls.

1.7. Dissertation Overview

Successful integration of disconfirmatory evidence is fundamental to belief revision, and dysfunction in this process may contribute to the maintenance of delusions in schizophrenia. Although little research has been conducted on the functional brain networks underlying disconfirmatory evidence integration, it may be intimately linked with the ventral attention, salience and frontoparietal networks of resting-state studies, which have already been shown to be dysfunctional in schizophrenia. Improper detection of disconfirmatory evidence as a salient event might lead to reduced integration of that evidence, which has been observed behaviourally in schizophrenia patients, and is especially affected in those with delusions. Functional brain networks underlying BADE would also be expected to fluctuate as a function of changes in delusions, in that improvements in delusion severity should correspond to normalization of functional brain activity and BADE.

The main goals of this research were to (1) determine the cognitive and brain mechanisms underlying evidence integration; (2) identify which of these mechanisms underlies the bias against disconfirmatory evidence in schizophrenia and delusions; and to (3) examine associations between changes in BADE, delusions, and functional brain activity underlying disconfirmatory evidence integration in schizophrenia. A secondary aim of this research was to determine whether BADE and its underlying functional brain activity could be separable from neurocognitive processes, such as memory and attention.

In Chapter 2, we report on an fMRI study in healthy individuals, in which we identified distinct, sequentially-active functional brain networks underlying detection and integration of disconfirmatory evidence. This work addresses our first aim, has been published in *NeuroImage* (Lavigne et al., 2015a) and is included here in its published form, with some minor changes made to reflect additional knowledge and to maintain consistency throughout this work. Chapter 3 presents a study comparing functional brain activity during a novel evidence integration task between healthy controls and schizophrenia patients with and without delusions. This chapter focuses on expanding the findings from the previous chapter using an improved measure of evidence integration, and addresses our second aim to identify whether activity in these networks is impaired in schizophrenia, and whether it can explain the bias against disconfirmatory evidence observed in schizophrenia patients with delusions. The third aim is addressed in Chapter 4, in which we examine whether changes in BADE and delusion severity over time are associated with changes in functional brain activity within networks underlying evidence integration in schizophrenia. Chapter 5 presents an overall discussion and overview of the conclusions of this research.

2. Functional Brain Networks Underlying Detection and Integration of Disconfirmatory Evidence (Study 1)

The evaluation and integration of evidence that disconfirms a prior belief is a fundamental aspect of belief revision. Failures in evidence integration, and particularly in the ability to integrate disconfirmatory evidence, has social relevance as it can lead to resistance in modifying outdated or unhelpful beliefs (Turner & Pratkanis, 1998), and has clinical relevance as it has been linked to delusions in schizophrenia (Sanford et al., 2014; Speechley et al., 2012; Woodward, Moritz, & Chen, 2006a) and to self-regulation deficits in traumatic brain injury (Flashman & McAllister, 2002) and obsessive-compulsive disorder (Marsh et al., 2014).

Evidence integration involves multiple cognitive processes, including alerting to the piece of evidence in question, and integration of that evidence into the current belief. When evidence contradicts a currently-held belief (i.e., disconfirmatory evidence), this would increase demand for alerting and integrating processes, as the initial belief must either be revised or discarded in order to assimilate the newly-accepted evidence and maintain a coherent belief system. When the evidence is neutral, or consistent with a belief (i.e., confirmatory evidence), these cognitive processes would be expected to have a reduced role. To date, there have been few investigations into the functional brain networks underlying disconfirmatory evidence integration, and it is not known whether distinct, sequentially-active brain networks that correspond to alerting and integration processes can be measured. However, the left IFG has been implicated in disconfirmatory evidence integration, with previous studies finding improved integration following transcranial magnetic stimulation (Sharot et al., 2012). The dACC, with regard to its role in adjusting behavior and changing

mental set (Behrens et al., 2007; Whitman et al., 2013; Woodward et al., 2008) may play a role in alerting. In the current fMRI study, we used multivariate analysis methodology on two datasets to attempt to identify functional brain networks underlying different stages of disconfirmatory evidence integration.

In order to assess spatial and temporal replication of network configurations, and take advantage of spatial replication combined with temporal differences to interpret function of brain networks, two versions of an evidence integration task were run and analyzed simultaneously using constrained principal component analysis for fMRI (fMRI-CPCA; Lavigne et al., 2015b; Metzak et al., 2011; Metzak et al., 2012; Whitman et al., 2013; Woodward, Feredoes, Metzak, Takane, & Manoach, 2013). fMRI-CPCA allows observation of coordinated task-based activity of multiple distinct, sequentially-active functional brain networks based on distinct hemodynamic response (HDR) shapes and spatial distributions. fMRI-CPCA determines the degree to which each functional brain network replicates across tasks by the magnitude and pattern of the HDR shape associated with each network. When two (or more) task versions elicit the same underlying cognitive operation (e.g., evidence integration), spatial and temporal replication would be observed if HDR shapes were not distinguishable between the two task versions, and this should be the case if the timing of the cognitive operation does not differ between task versions. In contrast, spatial but not temporal replication would be observed if HDR shapes were reliably different between the two task versions, and this should be the case if the timing of the cognitive operation differs between tasks. This case (spatial but not temporal replication) provides an important scientific opportunity to use differences between tasks to help interpret the cognitive function of brain networks. Finally, if a cognitive operation is elicited by only one version of the task

19

but not the other, the version not eliciting this cognitive operation would show a flat HDR shape for that functional brain network, and therefore it could be concluded that neither spatial nor temporal replication has been observed.

In the current study, we examined the functional brain networks underlying disconfirmatory evidence integration by combining data from two versions of an evidence integration task. The main distinction between the two task versions was a persistent visual display throughout the trial in version 1, and the removal of the visual display during rating in version 2. This was expected to elicit distinct HDR shapes for visual-processing brain networks between versions, producing spatial but not temporal replication for visual-processing networks, but similar HDR shapes for evidence integration brain networks, producing spatial and temporal replication for evidence integration brain networks. This method will facilitate separation of cognitive processes underlying visual processing from those related specifically to the alerting to and integration of disconfirmatory evidence. In accordance with the two-stage process mentioned above, we hypothesized that two separable and sequentially active functional networks (viz., alerting followed by integration), would be associated with disconfirmatory evidence integration to a greater degree than confirmatory evidence integration, and would not be associated with pure visual processing.

2.1. Method

2.1.1. Participants

Participants were 39 healthy volunteers (version 1: 10 male, 10 female, mean age = 24.90, SD = 6.87; version 2: 9 male, 10 female, mean age = 26.84, SD = 7.34), most of which were native English speakers (version 1: 17 participants; version 2: 15 participants). Non-native English speakers had been using English daily for at least the past five years and

responded accurately to questions about the consent form designed to confirm their ability to read and understand English. All participants were right-handed (Annett, 1970), with the exception of one left-handed and two mixed-handed participants who completed version 2. Participants were recruited via advertisements and word-of-mouth from Vancouver, British Columbia, and participated in exchange for $10/hour and a copy of their structural brain image. All were screened for MRI compatibility, and gave written informed consent prior to participation. All experimental procedures were approved by the University of British Columbia clinical research ethics board.

2.1.2. Experimental Design

Participants completed one of two versions of a novel evidence integration task while undergoing fMRI. In version 1, each trial began with a brief (500 ms) presentation of a heavily distorted image (50 random noise, brightness -80, mosaic 8 & 8, ripple 5, 5, 50, 50; see Figure 1A) of two animals (e.g., Animal A = Bird; Animal B = Dolphin) morphed together at a ratio of 60:40 or 40:60 (Animal A/Animal B). Participants were presented with a 16-point rating scale and were asked to indicate the degree to which the image appeared to be of one animal or the other. After six seconds, or once a rating was made, a mildly distorted image (brightness -50, mosaic 8 & 8) of the same animals morphed together at a ratio of 60:40 (Animal A/Animal B) was displayed on screen for three seconds, and participants were asked to re-rate the image. This led to the design of two types of trials: confirm (image 1: 60% animal A; image 2: 60% animal A); and disconfirm (image 1: 40% animal A; image 2: 60% animal A).

Version 2 differed from version 1 primarily in the following respects (see Figure 1B): (1) removal of images during presentation of the rating scales; and (2) the addition of a

backwards mask lasting 250ms between the offset of the first image and the onset of the first

rating scale. These changes removed the requirement to visually process the images when

responding, facilitating separation of visual-processing networks from those underlying

alerting and evidence integration. In addition, (3) the morphing ratios were increased to

70:30 for image 1 and 10:90 for image 2 in an attempt to intensify the disconfirmatory

evidence presented in image 2; (4) the name of either animal A or animal B was centered

above the rating scale in version 2 rather than both names appearing at opposite ends of the

scale, which ensured greater variability in participants' responses (i.e., selecting a degree of

belief towards one animal rather than choosing between one or the other); and (5) jittered

inter-trial intervals (ITIs) of 2, 4, 6, 8, 10, and 20 seconds (rather than the 2 second ITI in

version 1) were included to optimize the deconvolution of the BOLD signal (Serences,

2004).

2.1.3. Response Conditions

For each trial, participants rated each of the two images on a 16-point scale to describe

the degree to which they believed the image depicted the queried animal(s). In order to

emphasize that participants were to revise their initial ratings after viewing the second image,

participants' ratings on the first image were preserved on the second rating scale, and ratings

were modified from that point. Assignment of all experimental conditions (for both versions)

was based on participants' rating changes from image 1 to image 2. These response-based

conditions were labeled no change, confirm, and disconfirm. The *no change* response

condition included trials in which participants' ratings changed by less than or equal to two

points on the rating scale in either direction (e.g., image 1 rating = 9, image 2 rating = 7). The

confirm response condition consisted of trials in which the initial rating was supported by the

second rating. Specifically, this refers to trials in which ratings did not cross the mid-point of the scale (8) *and* where image 2 was rated closer to the extremes of the scale (e.g., image 1 rating = 6, image 2 rating = 3; or image 1 rating = 9, image 2 rating = 14). The *disconfirm* response condition consisted of trials in which the second rating contradicted the initial rating, such that ratings *either* crossed the mid-point of the scale (8) or image 2 was rated closer to the middle of the scale (e.g., image 1 rating = 4, image 2 rating = 9; or image 1 rating = 15, image 2 rating = 11). All response conditions were created such that they were mutually-exclusive (i.e., trials with rating changes of less than two that fit under either *confirm* or *disconfirm* conditions were classified as *no change*).

2.1.4. Image Acquisition and Processing

Imaging was performed at the University of British Columbia MRI Research Centre on a Philips Achieva 3.0 Tesla (T) MRI scanner with quasar dual gradients (maximum gradient amplitude, 80mT/m; maximum slew rate, 200 mT/m/s). The participant's head was firmly secured using a customized head holder. Functional image volumes were collected using a T2*-weighted gradient-echo spin pulse sequence with 36 axial slices; thickness/gap, 3/1 mm; matrix, 80×80; repetition time (TR), 2000 ms; echo time (TE), 30 ms; flip angle (FA), 90°, field of view (FOV), 240×240 mm, effectively covering the whole brain. In version 1, between 288 and 296 images were acquired in each of 3 runs lasting approximately 9 min and 52 s each. In Version 2, 350 volumes were acquired in each of two runs lasting 11 minutes and 40 seconds each. For both versions, run order was randomly assigned for each participant in order to minimize order effects.

Functional images were pre-processed using Statistical Parametric Mapping 8 (SPM8; Wellcome Trust Centre for Neuroimaging, UK). For each participant, each functional run

was corrected for slice-timing, realigned, co-registered to their structural (T1) image, and subsequently normalized to the Montreal Neurological Institute (MNI) T1 brain template. All images were spatially smoothed with an 8mm full-width-at-half-maximum (FWHM) Gaussian filter. Runs for which motion correction exceeded 4mm or degrees were excluded from analysis. This led to the exclusion of four runs across four participants, two in each task version.

2.1.5. Data Analysis

2.1.5.1. Functional Connectivity

fMRI data analysis was carried out using constrained principal component analysis for fMRI (fMRI-CPCA) with orthogonal rotation (Lavigne et al., 2015b; Metzak et al., 2011; Metzak et al., 2012; Whitman et al., 2013; Woodward et al., 2013). The theory and proofs of CPCA are detailed in previously published work (Hunter & Takane, 2002; Takane & Hunter, 2001; Takane & Shibayama, 1991) and the fMRI-CPCA application is available on-line, free of charge (www.nitrc.org/projects/fmricpca). Briefly, fMRI-CPCA combines multivariate multiple regression analysis and principal component analysis into a unified framework to reveal multiple independent sources of poststimulus fluctuations in brain activity. fMRI-CPCA is able to (1) identify multiple functional brain networks simultaneously involved in executing a cognitive task, (2) estimate the task-related time course of coordinated BOLD activity fluctuations associated with each functional network, and (3) statistically test the effect of experimental manipulations and group differences on BOLD activity associated with each functional brain network.

2.1.5.1.1. Matrix Equations

We now present a brief summary of the logic and matrix equations for fMRI-CPCA. Broadly speaking, whole-brain BOLD activity variance was partitioned into (i.e., constrained to) task-related fluctuations using multivariate multiple regression. Orthogonal sources (components) of task-related BOLD activity fluctuations were then determined using PCA. Functional brain networks associated with each orthogonal source of BOLD variance were spatially interpreted by viewing the networks represented by voxels dominating each component, and temporally interpreted by viewing the HDR shape associated with each component.

To begin, two matrices were prepared for further analysis. The first matrix, Z, contained the intensity values for normalized and smoothed BOLD time-series of each voxel, with one column per voxel and one row per TR or scan. Subject-specific data sets were stacked vertically to produce Z. The second matrix, G, consisted of a finite impulse response (FIR) basis set, which was used to estimate the change in BOLD signal at specific poststimulus scans relative to all other scans. The value 1 is placed in rows of G for which BOLD signal amplitude is to be estimated, and the value 0 in all other rows ("mini boxcar" functions). The time bins for which a basis function was specified in the current study were the 1st to 12th scans following stimulus presentation. Since the TR for these data was 2s, this resulted in estimating BOLD signal over a 24s window, with the start of the first time bin (time = 0) corresponding to stimulus onset. In this analysis, we created a G matrix for estimating subject-and-condition specific effects by including a separate FIR basis set for each condition and for each subject. The columns in this subject-and-condition based G matrix code 12 poststimulus time bins for each of the three conditions (viz., no change, confirm, and

disconfirm) for each of the 39 subjects, totaling 1404 columns ($12 \times 3 \times 39 = 1404$). Each column of Z and G were standardized for each subject separately.

The matrix of BOLD time series (Z) and the design matrix (G) were input to group fMRI-CPCA, with BOLD signal in Z being predicted from the FIR model in G. In order to achieve this, multivariate least-squares linear multiple regression was carried out, whereby the BOLD time series (Z) was regressed onto the design matrix (G):

$$Z = GC + E, \qquad\qquad\qquad (\,1\,)$$

where $C = (G'\,G)^{-1}\,G'Z$. The C matrix represents condition-specific regression weights, which are akin to the beta images produced by conventional univariate fMRI analyses. GC represents the variability in Z that was predictable from the design matrix G, that is to say, the task-related variability in Z.

The next step used singular value decomposition (of which PCA is a special case) to extract components in GC that represented temporally orthogonal functional brain networks in which BOLD activity fluctuated coherently with experimental stimuli. The singular value decomposition of GC resulted in:

$$UDV' = GC \qquad\qquad\qquad (\,2\,)$$

where U = matrix of left singular vectors; D = diagonal matrix of singular values; V = matrix of right singular vectors. After reduction of dimensionality (discussed in more detail below) and orthogonal rotation (Metzak et al., 2011) each column of $VD/\sqrt{(m-1)}$, where m = number of rows in Z, was overlaid on a structural brain image to allow spatial visualization of the brain regions dominating each functional network. $VD/\sqrt{(m-1)}$ is referred to as a

loading matrix, and the values are correlations between the component scores (in *U*) and the variables in *GC*.

2.1.5.1.2. Predictor Weights

To interpret the functional brain networks with respect to the conditions represented in *G*, *predictor weights* in matrix *P* are produced. These are the weights that, when applied to each column of the matrix of predictor variables (*G*), create *U (U=GP)*. Thus, the *P* matrix relates each column of the *G* matrix to the component scores in *U*, and provides information about the similarity of the fluctuation of the BOLD signal over all scans to the FIR model coded into *G*. For the current analysis, this would provide 1404 values per functional brain network, one for each combination of poststimulus time (12), subject (39), and condition (3). Each subject- and condition-specific set of predictor weights is expected to take the shape of a HDR, with the highest values corresponding to the HDR peaks.

These predictor weights provide estimates of the engagement of functional networks at each point in poststimulus time, and can be submitted to an analysis of variance (ANOVA) to test for (1) reliability of each component/functional brain network over subjects, (2) differences between conditions in the activation of each network, and (3) differences between task versions in the activation of each network. These analyses were carried out as $12 \times 3 \times 2$ mixed-model ANOVAs (one for each component extracted), with the within-subjects factors of Poststimulus Time (12 whole-brain scans after the onset of each trial were estimated in the FIR model) and Response Condition (no change, confirm, and disconfirm), and the between-subjects factor of Version (version 1, version 2). Any impact of Version or Response Condition would typically be reflected by a significant interaction with Poststimulus Time for the measure of estimated HDR (i.e., the predictor weights), suggesting that the HDR

27

shape depends on Version or Response Condition, although main effects are also possible. Significant interactions were interpreted using analysis of simple main effects involving the relevant factors. Spatial and temporal replication would be indicated by a reliable HDR shape over subjects (i.e., a significant Poststimulus Time effect) and no difference between task versions (i.e., no significant Version main effect or Version × Poststimulus Time interaction effect). Spatial but not temporal replication would be indicated by a reliable HDR shape over subjects (i.e., a significant Poststimulus Time effect) and a significant difference between task versions (i.e., a significant Version main effect or Version × Poststimulus Time interaction effect). Spatial (and temporal) non-replication would be indicated by a reliable HDR shape over subjects in only one task version (i.e., a non-significant Poststimulus Time effect at one but not the other level of Version) and a difference between task versions (i.e., a significant Version × Poststimulus Time interaction effect). Tests of sphericity were carried out for all ANOVAs, and adjustment in degrees of freedom for violations of sphericity did not affect the results; therefore, the original degrees of freedom are reported.

2.2. Results

Inspection of the scree plot of singular values (Cattell, 1966; Cattell & Vogelmann, 1977) suggested that five components should be extracted. The percentages of task-related variance accounted for by each rotated component were 10.88%, 10.43%, 9.00%, 7.64%, and 6.20%, for Components 1 to 5, respectively. For Component 3[1], no main effects or interactions involving Response Condition or Version were significant (all ps > .07) so it is not discussed

---

[1] Component 3 included activations in bilateral intracalcarine cortex (BAs 17, 18), lingual gyrus (BA 19), pre- and post-central gyri (BAs 3, 6), ventromedial prefrontal cortex (BAs 9, 10), and posterior cingulate cortex (BAs 23, 30). This component showed a significant main effect of Poststimulus Time, $F(11,374) = 28.75$, $p <$ .001, but no other significant main effects or interactions were present, suggesting that although it was a biologically plausible network, activity was not related to the experimental conditions of interest.

further. Visual inspection of the predictor weights for each component confirmed a HDR shape (see Figures 2 to 5, for Components 1, 2, 4, and 5 respectively). Components 1, 2, 4, and 5 each showed a significant effect of Poststimulus Time, $F(11,374) = 47.47$, $p < .001$, $F(11,374) = 55.99$, $p < .001$, $F(11,374) = 56.30$, $p < .001$, $F(11,374) = 38.70$, $p < .001$, respectively, demonstrating detection of a biologically plausible and reliable HDR signal for each functional brain network (Metzak et al., 2011; Metzak et al., 2012; Woodward et al., 2013).

2.2.1. Anatomical Descriptions and Relations to Experimental Conditions

The brain regions associated with Components 1, 2, 4, and 5 are displayed in Figures 2A to 5A, with anatomical descriptions in Tables 1 to 4, respectively. All components showed spatial but not temporal replication, described in detail below.

2.2.1.1. Component 1: Cognitive Evaluation Network

Component 1 (Figure 2A, Table 1) was characterized by a functional network that included activations in rPFC & OFC (BAs 10, 11, 47), bilateral IFG (BAs 6, 38), right dorsolateral prefrontal cortex (DLPFC; BA 46), inferior parietal lobule (extending into angular and supramarginal gyri; BAs 2, 40), and bilateral cerebellar and occipital (BAs 17, 18, 19) regions. This network showed significant Poststimulus Time × Version, $F(11,374) = 10.52$, $p < .001$, and Poststimulus Time × Response Condition, $F(22,748) = 6.29$, $p < .001$, interactions, but no significant three-way interaction. This suggests that the HDR shape associated with Component 1 depended on Version and Response Condition, but that each could be interpreted independently. In order to interpret the Version effect, simple contrasts averaging over Response Condition were observed (Figure 2B), and revealed significant differences between versions at 9, 11, and 21 s (all ps < .005), due to higher activity for

version 2 relative to version 1. In order to interpret the Response Condition effect, simple contrasts averaging over Version were observed (Figure 2C), and revealed significantly greater activity for (1) the confirm relative to no change response conditions at 17 s, (2) the disconfirm relative to no change response conditions from 17 to 21 s, and (3) the disconfirm relative to confirm response conditions from 17 to 23 s (all ps < .005). Thus, activity in this network was highest for the disconfirm response condition after the onset of the second image, when the evidence was presented, and remained elevated throughout the remainder of the trial. Since Response Condition did not interact with Version, this pattern can be considered present in both task versions. Based on these differences between response conditions, peak timing, and the spatial distribution of the network, this network was labelled *Cognitive Evaluation Network*.

2.2.1.2. Component 2: Visual/Default-Mode Network

Component 2 (Figure 3A, Table 2) was characterized by a functional network including activations in bilateral occipital cortex (BAs 17, 18, 19), superior parietal (BA 7) regions, and bilateral middle frontal gyrus (BAs 44, 45). This network also included deactivations (negative loadings) in ventromedial prefrontal cortex (VmPFC; BAs 9, 10), precuneus and posterior cingulate gyrus (BA 23), bilateral anterior middle temporal gyrus (BA 21), and bilateral angular/supramarginal gyri (BAs 39, 40), regions commonly associated with the default-mode network (DMN; Buckner, Andrews-Hanna, & Schacter, 2008; Raichle & MacLeod, 2001). Component 2 showed significant Poststimulus Time × Version, $F(11,374)$ = 9.35, p < .001, and Poststimulus Time × Response Condition, $F(22,748)$ = 3.57, p < .001, interactions, but no significant three-way interaction. In order to interpret the Version effect, simple contrasts averaging over Response Condition were observed (Figure 3B), and

revealed significant differences between versions at 1, 3, and 7-23 s (all ps < .01). This was

attributable to greater activity in version 1 relative to version 2 across all significant time

bins. In order to interpret the Response Condition effect, simple contrasts averaging over

Version were observed (Figure 3C), and revealed significantly increased activity for (1) the

disconfirm relative to no change response conditions at 21 and 23 s, (2) the disconfirm

relative to confirm response conditions at 21 and 23 s, and for (3) the no change relative to

disconfirm response conditions at 17 s (all ps < .005). This network showed the largest

Version, rather than Response Condition, effect with greater activity across the trial for

version 1 (in which the images were continuously displayed) than version 2. Due to this

sustained activity during version 1 versus the two peaks observed in version 2, as well as to

the involvement of primary visual cortex and DMN regions, this functional network was

labelled *Visual/Default-Mode Network*.

2.2.1.3. Component 4: Response Network

Component 4 (Figure 4A, Table 3) was characterized by a functional network including

activations in bilateral cerebellum and occipital (BAs 18, 19) regions, left-dominant pre- and

post-central gyri and supplementary motor area (BAs 3, 4, 6), and left thalamus. This

network also included deactivations in VmPFC (BA 10), precuneus (BA 23), and left

posterior middle temporal gyrus (BA 21). Component 4 showed significant Poststimulus

Time × Version, F (11,374) = 5.43, p < .001, and Poststimulus Time × Response Condition,

F (22,748) = 10.19, p < .001, interactions, but no significant three-way interaction. In order

to interpret the Version effect, simple contrasts averaging over Response Condition were

observed (Figure 4B), and revealed significant differences between versions at 19 and 21 s

(ps < .005), due to increased activity in version 2 relative to version 1. In order to interpret

the Response Condition effect, simple contrasts averaging over Version were observed (Figure 4C), and revealed significantly increased activity for (1) the confirm relative to no change response conditions at 17 and 19 s ($ps < .001$), (2) the disconfirm relative to no change response conditions from 15 to 19 s ($ps < .001$), (3) the disconfirm relative to confirm response conditions at 7, 9, 17, and 19 s ($ps < .005$), and for (4) the confirm relative to disconfirm response conditions at 23 s ($p < .005$). This functional network displayed two peaks of activity, corresponding to the time at which ratings were made. Two peaks of activation would be necessary for a response network, since ratings were made for each of the two images displayed. Due to this, and the spatial distribution of the network, it was labelled *Response Network*.

2.2.1.4. Component 5: Visual Attention Network

Component 5 (Figure 5A, Table 4) was characterized by a functional network including activations in superior frontal gyrus (BA 8) extending into dACC, right lateral prefrontal cortex extending into IFG (BAs 44, 45), bilateral anterior insula (BA 47) and bilateral occipital cortex (BAs 18, 19). Component 5 showed significant Poststimulus Time × Version, F (11,374) = 5.30, $p < .001$, and Poststimulus Time × Response Condition, F (22,748) = 6.46, $p < .001$, interactions, but no significant three-way interaction. In order to interpret the Version effect, simple contrasts averaging over Response Condition were observed (Figure 5B), and revealed significant differences between versions from 5 to 9 s, and at 17 s ($ps < .01$) due to greater activity in version 2 relative to version 1. In order to interpret the Response Condition effect, simple contrasts averaging over Version were observed (Figure 5C), and revealed significantly greater activity for (1) the disconfirm relative to no change response conditions at 13 and 15 s ($ps < .005$), (2) the disconfirm

relative to confirm response conditions at 15 and 17 s (*p*s < .001), (3) the no change relative to confirm response conditions from 17 to 21 s (*p*s < .01), and for (4) the no change relative to disconfirm response conditions at 21 s (*p* < .01). This functional network peaked briefly at the onset of the second image, when the evidence was first presented, was highest in the disconfirm condition, and was present in both task versions. For this reason, and due to the spatial distribution of the network, it was labelled *Visual Attention Network.*

2.3. Discussion

In the current study, we used multivariate methodology on two datasets in an attempt to link two sequential cognitive stages involved in integrating disconfirmatory evidence to distinct functional brain networks. Three functional networks showed greater intensity (i.e., increased activations and/or increased deactivations) during integration of disconfirmatory relative to confirmatory evidence for both task versions. In order of peak timing (see Figures 5B, 4B, and 2B, respectively), these reflected (1) a visual attention network including dACC and bilateral insula; (2) a sensorimotor response-related network; and (3) a cognitive evaluation network including bilateral rPFC, OFC, inferior parietal lobule, and IFG. Activity and deactivity associated with visual processing and the DMN separated out from other networks based on HDR shape differences due to stimulus timing differences between the two versions of the task.

2.3.1. Visual Attention Network (Component 5)

The dACC (e.g., 2, 28, 48) and bilateral insula (e.g., -30, 24, -4) were the dominant regions of the visual attention network (Component 5; see Figure 5A), which became most active during the onset of the second image, when the confirmatory/disconfirmatory evidence was presented. These regions correspond to the well-documented salience network, which is

involved in attending to environmentally-salient stimuli, and has been hypothesized to be responsible for switching between large-scale brain networks to allow access to relevant cognitive and sensory systems (Goulden et al., 2014; Menon & Uddin, 2010). Relating the current network to the 7-network brain parcellation derived from resting state data (Buckner, Krienen, Castellanos, Diaz, & Yeo, 2011; Choi, Yeo, & Buckner, 2012; Yeo et al., 2011), the dACC (e.g., 2, 28, 48), prefrontal (e.g., -46, 10, 32), caudate (subthreshold; e.g., 16, 14, 10), and parietal activations (e.g., -46, -40, 48) were all located on the frontoparietal network, and the occipital activations (e.g., 18, -92, -8) on the visual network. All deactivations were located on the DMN. The dACC and insula have been implicated in the "moment of recognition" of an object during evidence accumulation (Liu & Pleskac, 2011; Ploran et al., 2007); the dACC specifically is involved in surprise and error detection, and has been suggested to play a role in the "Aha! Moment", or to alert when behavioral adjustment is required (Carter & van Veen, 2007; Egner, 2011; Walsh, Buonocore, Carter, & Mangun, 2011; Whitman et al., 2013; Woodward et al., 2008). The right lateral prefrontal cortex, also involved in this network in the current study (e.g., 48, 12, 34), has been shown to activate in response to prediction error during associative learning (Corlett et al., 2004; Fletcher et al., 2001; Turner et al., 2004), and has been identified as important to hypothesis evaluation in clinical settings (Coltheart, 2010). The significantly higher HDR peak in the disconfirm (relative to confirm) condition may be interpreted as activation due to the conflict between the initial belief (formed during the presentation of image 1) and the disconfirmatory evidence presented at image 2. Detection of conflict between a held belief and presented evidence is a crucial first step in the process of belief revision, and dysfunction in the detection of this conflict could contribute to resistance in modifying outdated beliefs.

2.3.2. Cognitive Evaluation Network (Component 1)

Like the visual attention network, the cognitive evaluation network (Component 1; bilateral rPFC/OFC, inferior parietal lobule, and IFG; see Figure 2A) distinguished between confirmatory and disconfirmatory evidence, and peaked following confirmatory/disconfirmatory evidence presentation; however, its peak was noticeably later than that of the visual attention network (19 vs. 15 seconds), suggesting sequential activation. Relating the cognitive evaluation network to the 7-network brain parcellation derived from resting state data (Buckner et al., 2011; Choi et al., 2012; Yeo et al., 2011), the rostrolateral and superior prefrontal (e.g., 2, 28, 48), caudate (subthreshold; e.g., 20, 17, 14), and cerebellar activations (e.g., -38, -64, -50) were located on the frontoparietal network, inferior frontal gyrus/pars opercularis (e.g., 56, 12, 10) and putamen (subthreshold) (e.g., 25, 4, 6) on the ventral attention network, and the occipital activations on the visual network. Parietal activations (e.g., 46, -44, 58) were located on the dorsal attention network. The rPFC is involved in the evaluation of self-generated information (Christoff et al., 2003), and has been proposed as a key region involved in the balance between self-generated and externally-generated information (Burgess, Simons, Dumontheil, & Gilbert, 2005; Gilbert et al., 2006b). Together, the IFG and posterior parietal regions (e.g., supramarginal and angular gyri), are involved in semantic processing and visual word recognition (Binder, Desai, Graves, & Conant, 2009), and are recruited during perceptual decision-making tasks, such as in the current study. The IFG has also been implicated in belief formation and updating (d'Acremont et al., 2013; Sharot et al., 2011), and there is evidence that disruption of the left IFG improves integration of unfavorable evidence (Sharot et al., 2012), suggesting it may play a key role in disconfirmatory evidence integration in particular.

Much like the visual attention network, functional brain activity in the cognitive evaluation network was higher during disconfirmatory relative to confirmatory evidence integration, but this difference was even greater in the cognitive evaluation network. Taken together with its delayed peak relative to the visual attention network, this suggests a role in evaluating the presented evidence relative to the initial belief (formed at image 1). This would also be necessary during confirmatory evidence integration, but might be expected to elicit lesser and less sustained activity than during disconfirmatory evidence integration, as was evident in the current findings (see Figure 2C). Evaluating presented evidence and comparing it to prior knowledge is another crucial aspect of evidence integration, and dysfunction in this network could also contribute to resistance in modifying beliefs.

2.3.3. Visual/Default-Mode Network (Component 2)

The visual/default-mode network showed sustained activity during version 1 and two peaks in version 2, and was characterized by deactivations in DMN regions, and activations in visual cortex regions (see Figure 3A). Relating the response network to the 7-network brain parcellation derived from resting state data (Buckner et al., 2011; Choi et al., 2012; Yeo et al., 2011), the occipital activations were all located on the visual network. The parietal (e.g., -26, -62, 46) and lateral prefrontal (e.g., 48, 36, 32) activations were located on the dorsal attention network. All deactivations were located on the DMN, but the superior temporal deactivations (which included primary auditory cortices, e.g., -58, -32, 16) were located on the somatosensory network. The auditory cortex deactivations present on this component have been shown to be sensitive to load-dependent task-related decreases in activity in working memory and source experiments that employing visual encoding (Metzak et al., 2011; Metzak, Lavigne, & Woodward, 2015; Metzak et al., 2012; Woodward et al.,

2013). This coordinated decrease in bilateral primary auditory cortex activity could relate to reduced activation during inner speech (Buchsbaum et al., 2005; Frith, Friston, Liddle, & Frackowiak, 1991), or a more general phenomenon whereby task-irrelevant primary sensory cortices (with visual cortices being task-relevant) are deactivated during task performance (Laurienti et al., 2002; Shulman et al., 1997). The fact that the bilateral primary auditory cortex deactivity emerged uniquely on the visual processing component provides support for the latter interpretation.

2.3.4. Response Network (Component 4)

Relating the response network (Figure 4A) to the 7-network brain parcellation derived from resting state data (Buckner et al., 2011; Choi et al., 2012; Yeo et al., 2011), the dACC (e.g., -4, 8, 54) and cerebellar (e.g., 32, -56, -22) activations were located on the ventral attention network, the frontal (e.g., 28, -2, 54) and parietal (e.g., -26, -60, 50) activations on the dorsal attention network, the left sensorimotor activations (e.g., -40, -20, 58) on the somatosensory network, and the occipital activations on the visual network. All deactivations were located on the DMN. The sensorimotor response network identified in the current study peaked during responses made to both the first and second images, and showed significantly greater activity for disconfirmatory evidence integration during the response to the second image. These differences were likely the result of the way in which the conditions were encoded, given that the disconfirm responses included greater rating changes between image 1 and 2 (e.g., rating change from 14 to 4) than the confirm condition (e.g., rating change from 9 to 15). Consistent with this interpretation, the no change condition (which included rating changes of two steps or fewer) showed the least activity after the onset of the second image (see Figure 4C).

It should be noted that there was a small, but significant difference between the confirm and disconfirm conditions during the response to image 1. Given that one of the parameters for the disconfirm condition was that image 2 be rated closer to the middle of the scale than image 1 (whereas the opposite was true for confirm), this unexpected finding likely reflects initial ratings closer to the extremes of the scale in the disconfirm condition. This would result in greater response-related activity in the disconfirm relative to confirm condition during image 1, since the initial point from which participants made their ratings was at the middle of the scale.

Although activity in Components 1 (cognitive evaluation network) and 4 (response network) began increasing simultaneously after evidence presentation, the cognitive evaluation network peaked later than the response network (19 vs. 17 seconds post-stimulus), and activation extended past the time at which the response network returned to baseline. This is explained by the fact that participants were able to modify their responses throughout the 6 s time window during which the response options were displayed. This earlier peak for responding versus integrating may be because participants began responding based on the Aha! moment elicited at the onset of the second image (which would guide their decision to begin either down-rating or up-rating their initial responses), and then continued to evaluate their decision during and following finalization of their responses throughout the 6 s time window. This HDR pattern (viz., the response network peaking earlier than the cognitive evaluation network) was also observed in a previous study from our lab on controlled semantic association, though with a language-based subnetwork of the cognitive evaluation network identified in the current study (Woodward et al., 2015).

2.3.5. Differences Between Task Versions

In addition to differences between the response conditions, all functional networks also demonstrated distinct activation patterns across the two versions of the task, meaning that all components showed spatial (same networks) but not temporal (different HDR shapes) replication, providing an opportunity to use the differences between tasks to help interpret the cognitive function of the networks. For example, for Component 2 (visual/default-mode network) there was a higher, sustained peak in version 1 relative to version 2, which corresponded to the sustained presentation of the images. In addition, on Components 1 (cognitive evaluation network) and 5 (visual attention network), version 2 showed an early peak not present in version 1. Although these networks showed higher activity during disconfirmatory evidence integration, the time bins on which a version effect was observed (5 to 11 s) were not the same as those on which a response conditions effect was observed (15 to 23 s). Given that both of these networks (i.e., Components 1 and 5) included brain regions related to visual processing, it is possible that these differences between versions were also driven by the visual stimuli (or attention to visual imagery during masking), as with Component 2. Finally, some late trial differences between versions 1 and 2 were present on Component 1 (21 s), Component 4 (19 to 21 s), and Component 5 (17 s), due to a higher or more sustained peak in version 2 relative to version 1 for these components. These higher, prolonged peaks on the cognitive evaluation (Component 1) and visual attention (Component 5) networks could be the result of the increased morphing ratios and improved experiment design (jittered ITIs), which presumably served to increase the magnitude of the experimental effects in version 2. In the case of the response network (Component 4), the difference between versions is likely due to the single animal name in version 2 (compared to

both names in version 1), which would lead to larger and more extreme response changes after evidence presentation. Importantly, these version effects were statistically independent from the response condition effects, due to the absence of both three-way interactions and version × response condition interactions.

Inclusion of the no change (rating changes less than 3) response condition in the current study was intended as a control; however, there was evidence of *increased* activity during evidence integration in the no change condition relative to the other response conditions within two functional networks: the visual/DMN network (Component 2) and the visual attention network (Component 5). While an unexpected finding, this might be a consequence of uncertainty about the nature of the second image on the part of participants. If participants were unable to identify the animal in the second image, they would likely have had difficulty determining whether they should respond in a clearly confirmatory or disconfirmatory manner, and might be less inclined to change their initial ratings. The increased attention associated with uncertainty could account for the heightened activity in functional networks associated with visual and cognitive attention.

2.3.6. Limitations

One limitation of this study was that this comparison was carried out between subjects. Ideally, one would conduct multi-version comparisons within-subjects (de Zubicaray, Hansen, & McMahon, 2013; Metzak, Meier, Graf, & Woodward, 2013); however, since fMRI data is expensive to collect, and testing time is limited, comparing task versions often must be carried out between subjects. Combining versions of a task with a number of differences in the experimental design (e.g., persistence of the visual stimuli, differences in response interface, etc.) facilitates powerful comparisons for interpreting network functions,

and it is the analysis of differences and commonalities between HDR shapes across task versions that is of scientific interest. In the current analysis, differences between task versions did not appear to substantially impact interpretation of the response conditions, evidenced by the absence of three-way interactions. For example, only minor variations between task versions were present on the response network, despite substantial differences in the nature of the response interface. In addition, there was an early peak present on the visual attention and cognitive evaluation networks in version 2 that was not present in version 1, which may have been driven by visual processing regions that were part of each network, or may have been due to the increased morphing ratios and other improvements implemented in version 2. In either case, these early differences between versions did not affect interpretation of the differences between response conditions, which occurred later in the trial. The jittered ITIs in version 2, designed to increase power of the manipulations, did appear to impact the magnitude of the HDR shapes for alerting or integration of disconfirmatory evidence, which were greater in version 2. However, the conclusions reached for this study would ideally be tested by conducting multi-version comparisons within-subjects.

2.3.7. Conclusion

In the current study, we examined the functional networks associated with the processing of disconfirmatory relative to confirmatory evidence. By combining data from two versions of the same task that differed primarily in terms of stimulus timing, we were able to distinguish between functional brain activity associated with detection and integration of evidence, as well as others associated with responding and visual processing. We identified three functional networks that showed increased activity during disconfirmatory relative to

confirmatory evidence integration: a visual attention network involved in detecting a mismatch between the presented evidence and the initially-formed belief, a response network, and a cognitive evaluation network involved in the integration of the evidence and in comparison of that evidence to the initial belief.

These findings highlight two distinct functional networks underlying disconfirmatory evidence integration that correspond to two important cognitive processes underlying belief revision: (1) detection of a conflict between an initial belief and a piece of (disconfirmatory) evidence; and (2) evaluation of that evidence in light of the initial belief in order to determine whether it should be integrated into the current belief system and the belief modified or dropped. In cases where an individual is consistently resistant to disconfirmatory evidence (e.g., groupthink, stereotyping), or in clinical settings (psychotic delusions), one or both of these mechanisms may play a role. For example, delusional schizophrenia patients who demonstrate a bias against disconfirmatory evidence (Sanford et al., 2014; Speechley et al., 2012; Woodward et al., 2006a) may show decreased activity in both the visual attention and cognitive evaluation networks when faced with disconfirmatory evidence. Future research is necessary to determine whether resistance to modifying beliefs when faced with disconfirmatory evidence is due to a lack of attention toward/detection of that evidence, to an inability to integrate the evidence into the current belief system, or some combination of both processes.

3. Functional Brain Networks Underlying Evidence Integration and Delusions in Schizophrenia (Study 2)

In the previous chapter, we identified two sequentially-active functional brain networks associated with evidence integration in healthy individuals: a visual attention network (VsAN) including dACC and bilateral insula, and a cognitive evaluation (CEN) network including rPFC/OFC, IFG, and inferior parietal lobule. Both networks showed increased activity during disconfirmatory relative to confirmatory evidence integration, and were interpreted as being involved in the detection and integration of evidence (particularly disconfirmatory evidence), respectively. Schizophrenia patients show important differences in the function of these networks relative to controls, including reduced functional connectivity within the SN/VAN (which has common regions with the VsAN identified in study 1), as well as its modulation of other large-scale networks, such as the FPN (Baker et al., 2014; Manoliu et al., 2014; Palaniyappan & Liddle, 2012; White et al., 2010), which shares important nodes with the CEN identified in study 1. The objective of the current study was to examine the degree to which activity in these networks is affected in schizophrenia and delusions, and whether activity in one or both networks underlies BADE.

In this study, we compared schizophrenia patients and healthy controls on a novel evidence integration task using task-based fMRI. This novel evidence integration task was developed to address some of the methodological limitations of the task used in study 1, namely, the confound between evidence integration and responding, and that two different images were presented. As in the previous study, we expected to identify sequential activation of the VsAN and CEN following evidence presentation, with higher activity during disconfirmatory relative to confirmatory evidence integration. We hypothesized that

schizophrenia patients with delusions, but not those without, would show attenuated

activation within both networks during disconfirmatory evidence integration. We also

expected that activity within these networks would be associated with BADE behaviourally,

such that individuals with a stronger bias would show decreased activity in these networks.

Finally, we hypothesized that both behavioural BADE and functional brain activity

underlying BADE would be separable from neurocognitive measures, including memory and

attention.

3.1. Method

3.1.1. Participants

Forty healthy controls and 58 schizophrenia patients completed a battery of demographic,

clinical, and neurocognitive tests, as well as a novel fMRI evidence integration task.

Schizophrenia patients were divided into delusional and non-delusional subgroups based on

the Psychotic Symptoms Rating Scales (PSYRATS; Haddock, McCarron, Tarrier, &

Faragher, 1999) delusions total score using a median split, which resulted in a sample

configuration of 40 controls, 31 non-delusional patients (PSYRATS delusions total score <

12), and 27 delusional patients (PSYRATS delusions total score > 11). Healthy participants

were recruited through local posted advertisements as well as through word of mouth. Patient

recruitment occurred primarily through presentations to outpatient mental health

organizations throughout Vancouver, British Columbia, but some were also recruited through

local posted advertisements and word of mouth. Individuals who took part in past research

and indicated interest in future research were also contacted. Participants took part in

exchange for $10/hour, reimbursement of travel expenses, and a copy of their structural brain

image. The study was approved by the University of British Columbia (UBC) Clinical Research Ethics Board.

Patients were outpatients with a DSM-IV-TR (American Psychiatric Association, 2000) diagnosis involving psychosis being followed by a doctor and/or mental health professional. Diagnoses were confirmed with the Mini International Neuropsychiatric Interview (MINI; Sheehan et al., 1998). In cases where there was a discrepancy between the reported diagnosis and that identified by the MINI, the patient's mental health care worker was consulted. When that was not possible, a consensus was determined by the clinical psychologists who were part of the research team. All patients but one were currently taking antipsychotic medication. Fifty patients were on atypical antipsychotics (e.g., clozapine, olanzapine, risperidone), while seven patients were on typical antipsychotic medication (e.g., loxapine, fluphenazine) as their primary medication. Twenty patients were taking a second antipsychotic medication (15 atypical, 5 typical), and two patients were taking a third atypical medication. Means and standard deviations per patient group on chlorpromazine dose equivalents are listed in Table 5; no significant differences were observed between delusional and non-delusional patients. Other prescribed medications included antianxiety (n = 11), antidepressant (n = 29; 1 missing data), anticonvulsant (n = 10), anticholinergic (n = 6), and other (e.g., diabetes, blood pressure; n = 16; 1 missing data) medications. Healthy controls were not currently taking any psychotropic medications.

Participants with a history of neurological infection or head injury/loss of consciousness that lasted more than 10 minutes and resulted in cognitive sequelae were excluded from the study. Participants were also given an eye test to determine their visual acuity, and those with less than 20/50 visual acuity with corrective lenses were excluded from the study; two

patients had 20/50 visual acuity, while the remaining participants had 20/30 visual acuity or better. Other exclusion criteria included an IQ of less than 80 on the Wechsler Abbreviated Scale of Intelligence II (WASI; Wechsler, 2011), current or past history of substance/alcohol dependence, and major medical illness. Patients with severe thought disorder (i.e., a score of 4 of higher on the Scale for the Assessment of Positive Symptoms (SAPS; Andreasen, 1984b) formal thought disorder global rating) were also excluded due to a high likelihood of being unable to complete the study. Healthy participants with a recurrent history of psychiatric illness or a family history of psychosis were also excluded.

3.1.2. Experimental Design

The study design consisted of two research sessions: an assessment session and an fMRI session. The assessment session lasted approximately 2.5 hours for healthy controls, and up to 4 hours for patients, depending on the extent of their symptoms and whether further assessment was required. In some cases, patient assessments were split into two sessions, with the second session usually preceding the imaging session. In the assessment session, informed consent was given, MRI compatibility was determined, and participants were assessed on symptoms, cognitive and intellectual functioning, and provided demographic information. The fMRI session lasted approximately 2 hours (30 mins for practice and preparation, 1-1.5 hours scanning time). Participants completed a novel evidence integration task (see below) as well as two other tasks not addressed in this research (a working memory task and a probabilistic reasoning task). The evidence integration task was always performed at the start of the session; structural MRI data were also collected at the end of the session.

3.1.3. Measures

The assessment session took the form of an interview, which gathered demographic information, education/occupation history, medication and substance use, and self and family psychiatric history. Handedness was determined using the Annett Handedness Questionnaire (Annett, 1970), a 12-item questionnaire describing common manual tasks (e.g., "to write a letter legibly") and requiring participants to state their preferred hand (right, left, no preference) for each activity as well as the frequency with which they would use that hand (always, usually). Each item is scored on a range of -2 (always left) to +2 (always right), with the total score ranging from -24 to +24. Right-hand preference is denoted by a score of +9 to +24, left-hand preference a score of -24 to -10 and mixed-handedness -9 to +8. One patient failed to answer one question, which prevented a total score from being calculated; however, they were coded as left-handed due to a predominance of "always left" responses as well as self-reported left-handedness. This measure showed good reliability ($\alpha = 0.86$), and was highly correlated with self-reported handedness (data available for 57 patients; $r(57) = 0.82$, $p < .001$).

3.1.3.1. Clinical

Psychiatric history was examined using the MINI (Sheehan et al., 1998), a semi-structured diagnostic interview assessing DSM-IV and International Statistical Classification of Diseases and Related Health Problems (ICD-10) criteria for primary psychiatric disorders. The following sections were included in the current study for both patients and controls: major depressive episode; dysthymia; (hypo)manic episode; panic disorder; agoraphobia; social phobia (social anxiety disorder); obsessive-compulsive disorder; alcohol use and dependence; non-alcohol psychoactive substance use disorders; generalized anxiety disorder;

and psychotic disorders parts 1 and 2 (differential diagnosis between psychotic and mood disorders). Primary psychotic disorders assessed by the MINI in patients included schizophrenia (n=29), schizoaffective disorder (n=22), schizophreniform disorder (n=1), and psychotic disorder not otherwise specified (n=6). Possible comorbid disorders identified by the MINI included major depressive episode (5 current, 24 past, 2 missing data), (hypo)manic episode (17 past, 1 missing data), panic disorder (3 current, 3 past), panic disorder without agoraphobia (1 current, 1 missing data), social anxiety disorder (2 current, 2 missing data), obsessive-compulsive disorder (2 current), non-alcohol substance dependence (2 current), and generalized anxiety disorder (7 current, 1 missing data). Nine control participants met criteria for past major depressive disorder (n = 6), past dysthymia (n = 1), alcohol abuse and dependence (n = 5), non-alcohol substance use or dependence (marijuana; n = 3), and/or generalized anxiety disorder (n = 1). These participants were determined eligible for the study following review by the research team.

The Beck Depression Inventory II (BDI; Beck, Steer, & Brown, 1961), a 21-item self-report questionnaire assessing depressive symptoms, was administered to both patients and controls. The BDI showed excellent reliability, with a Cronbach's alpha of 0.92. The total score (maximum = 63) was used in the current study. In general, the total scores are divided into ranges of depression severity as follows: 0-13 (minimal depression), 14-19 (mild depression), 20-28 (moderate depression), and 29-63 (severe depression). Four participants (3 patients and 1 control) were missing data on one or more items and a total score could not be calculated.

Patients' psychotic symptoms were assessed in detail using the Scales for the Assessment of Positive/Negative Symptoms (SAPS/SANS; Andreasen, 1984a; Andreasen, 1984b) and,

depending on whether auditory hallucinations and/or delusions were endorsed, the PSYRATS (Haddock et al., 1999) auditory hallucinations and/or delusions subscales, respectively. The SAPS and SANS assess positive (i.e., hallucinations, delusions, bizarre behavior, formal thought disorder) and negative (i.e., affective flattening, alogia, avolition-apathy, anhedonia-asociality, attention, inappropriate affect) symptoms, respectively, on a 0 to 5-point scale with 0 = absent, 1 = questionable, 2 = mild, 3 = moderate, 4 = marked, and 5 = severe. Total overall and subscale scores were used in subsequent analyses. The PSYRATS measures the severity of multiple features of hallucinations (e.g., frequency, loudness, origin) and delusions (e.g., preoccupation, duration, conviction) on a 0 to 4-point scale of increasing severity. Separate total scores for the hallucinations and delusions subscales were used in subsequent analyses. Patients who did not report hallucinations and/or delusions (and were not administered the PSYRATS) were given a score of 0 on the relevant subscale(s) to prevent missing data. In cases where the time between the assessment session and fMRI session exceeded two weeks, symptoms were re-assessed at the fMRI session, and the updated symptom scores were used. Data were missing for nine patients on the SANS, due to refusal to answer questions related to sex and intimacy, and data were missing for one patient each on the PSYRATS delusions and hallucinations subscales.

Control participants completed the Schizotypal Personality Questionnaire (SPQ; Raine, 1991) to assess the presence of subclinical schizotypal symptoms. The SPQ is a 74-item questionnaire designed to assess the nine schizotypal personality disorder criteria of the DSM-III-R in the general population. The original dichotomous (yes-no) response format was expanded to a 5-point Likert scale (1 = strongly disagree, 2 = disagree, 3 = neither agree nor disagree, 4 = agree, 5 = strongly agree) for the current study based on suggestions of

previous research (Wuthrich & Bates, 2005). In addition to the nine subscales, five higher-order factors were calculated based on Chmielewski & Watson's (2008) item-level factor analysis (i.e., social anhedonia, social anxiety, eccentricity/oddity, mistrust, and unusual beliefs and experiences), rather than the more commonly used three factors (cognitive-perceptual, interpersonal, disorganized; Raine et al., 1994) that emerge from factor analysis at the subscale level. In the current study, the SPQ total score showed excellent reliability (α = 0.96), with subscale/factor scores ranging from α = 0.75 (magical thinking subscale) to α = 0.90 (eccentricity/oddity and mistrust factors).

3.1.3.2. Neurocognitive

Intellectual functioning was assessed in both patients and controls using the Test of Premorbid Functioning (ToPF; Wechsler, 2009) and the WASI (Wechsler, 2011). The ToPF is a 70-item reading and pronunciation test of premorbid intelligence, in which participants are asked to read aloud a series of words with irregular grapheme-to-phoneme translation that become increasingly difficult as the test progresses. A score of 1 is given for each item correctly pronounced, leading to a maximum total score of 70. The age-corrected standard scores were used for the current study. The WASI is a short battery of four neurocognitive tests designed to assess general intelligence. It consists of four subscales (block design, vocabulary, matrix reasoning, and similarities), and produces two index scores (verbal comprehension and perceptual reasoning) as well as a full-scale intelligence quotient, the latter of which was used in the current study after correcting for age.

Patients also completed a battery of neurocognitive measures consisting of the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS; Randolph, Tierney, Mohr, & Chase, 1998), the Comprehensive Trail-Making Test (TMT; Reitan &

Wolfson, 1985; Reynolds, 2002), the Letter-Number Sequencing (LNS) subtest of the Wechsler Adult Intelligence Scale – Fourth Edition (Wechsler, 2014), and the Controlled Oral Word Association Test (COWAT; Benton, 1967).

The RBANS (Randolph et al., 1998) is a battery of neurocognitive tests indexing five cognitive domains: immediate memory, visuospatial/constructional abilities, language, attention, and delayed memory. Index scores are created by summing the scores for the individual tests underlying a given domain (e.g., list learning and story memory for the immediate memory domain). Scaled scores, corrected for age and standardized to a mean of 100 and standard deviation of 15, were used for each of the five cognitive domains.

The TMT (Reitan & Wolfson, 1985; Reynolds, 2002) is a brief measure requiring participants to trace a line on a piece of paper following 25 pseudorandomly-located numbers and/or letters in chronological/alphabetical order. It consists of two versions: version A involves tracing the numbers 1-25 in chronological order; version B includes both numbers and letters, and requires alternating between numbers and letters (i.e., 1-A-2-B-3-C, etc.). TMT-A assesses attention and visual scanning, whereas TMT-B introduces an additional set-shifting component. Scores for the time to completion (seconds) are derived for each version and transformed into standard scores (higher values = better performance) based on Halstead-Reitan norms (Fromm-auch & Yeudall, 1983), which are corrected for age, gender, and education level.

The LNS (Wechsler, 2014) measures working memory and mental manipulation, and involves repeating a series of numbers and letters in chronological and then alphabetical order (e.g., T-9-A-3 repeated as 3-9-A-T). The subscale includes 10 items of 3 trials each,

and one point is given for each series of numbers and letters repeated in correct order. Total scores are age-corrected and standardized to a mean of 10 and standard deviation of 3.

The COWAT (Benton, 1967) assesses verbal fluency and involves listing as many words as possible that start with a given letter (e.g., S) within 60 seconds. The letters F, A and S were used in the current study. One point is given for each correct word starting with a given letter, and these are summed to calculate a total score. Proper nouns and variations of previous words (e.g., plural versions of a previously stated word) are not counted towards the total. The total score is transformed into a standard score based on Halstead-Reitan norms (Fromm-auch & Yeudall, 1983), which is corrected for age, gender, and education level.

3.1.3.3. Behavioural BADE Task

All participants completed the behavioural BADE task (Sanford et al., 2014; Speechley et al., 2012; Woodward et al., 2007), which involves rating interpretations of a delusion-neutral story that unfolds over three sequentially-presented sentences. For each story, sentences are presented one at a time, and participants are asked to rate (on a scale of 0-100) the likelihood of four interpretations of that story. One interpretation is true, and becomes increasingly plausible as the story unfolds; two interpretations are (emotional and neutral) lures, which appear plausible at first, but become less plausible as more information is provided; and one interpretation is absurd and not very plausible throughout the trial. Participants rate each of these interpretations independently of one another after each of the three sentences is presented. Mean ratings are calculated for each of the four interpretation types (true, neutral lure, emotional lure, absurd) after each of the three sentences by summing across all trials, leading to twelve scores: true 1, 2 and 3; emotional lure 1, 2, and 3, neutral lure 1, 2, and 3; and absurd 1, 2, and 3. BADE evidence integration (degree to which disconfirmatory

evidence has been integrated; higher scores = worse evidence integration) and conservatism (a reduced willingness to rate high when justified; higher scores = less conservatism) scores are calculated by summing the relevant items derived from a principal component analysis (Speechley et al., 2012), and these scores were used in the current study.

The current version of the task included 16 trials, 12 of which were BADE trials and 4 of which were fillers. Filler trials followed the same structure as BADE trials, except that the true interpretation was evident from the start of the trial. These were included to prevent participants from relying on one type of response pattern (Speechley et al., 2012), and were pseudo-randomly interspersed between BADE trials. The order of the interpretations (true, absurd, lures) was randomized across trials. The behavioural BADE task was presented on a computer screen and participants responded with a mouse using their dominant hand. Prior to the task, the experimenter verbally explained the instructions while participants followed along on screen, emphasizing that each interpretation should be rated independently of the others, and that ratings could be changed after each sentence presentation if they felt it was warranted. Participants then completed two practice trials under the supervision of the experimenter to ensure they understood the task. Figure 6 shows an example of a typical BADE trial after all sentences are presented and ratings completed. Two patients did not complete the task due to technical issues or time constraints.

3.1.3.4. Neuroimaging BADE/Evidence Integration Task

A novel task was designed to assess evidence integration using neuroimaging, and builds upon BADE/evidence integration tasks used in previous behavioural (Woodward et al., 2007) and neuroimaging (Lavigne et al., 2015a) studies. Similar to the previous evidence integration neuroimaging task (see Chapter 2 and Lavigne et al., 2015a), the experimental

design involved rating consecutive images; however, this task addressed evidence integration and accumulation in more depth, as the images displayed were increasingly more detailed versions of a final picture. Briefly, participants were presented with a partial line drawing of common objects, food, or animals, and asked to rate whether they believed the full picture was of a word listed below it using a dichotomous (yes/no) response interface. Following a response, they were presented with a second image showing more of the full picture and asked to re-rate. Finally, the full picture was presented, which did not require a response. For the remainder of this text, the term "picture" will be used to specify the full line drawing, while the term "image" is used to describe the partial versions of the line drawings.

3.1.3.4.1. Stimulus Creation

376 grayscale line drawings of common objects, food, and animals were either created by hand or found via Google Image Search (https://images.google.com). Each picture was edited in the GNU Image Manipulation Program (GIMP 2; http://gimp.org) by erasing portions of the picture to identify potential lure images within them (e.g., umbrella as a lure image of bat; see Figure 7). If more than one lure was identified, each was created and noted for later discussion within the research team. Approximately 80 pictures underwent pilot testing, whereby each was split into a series of 6 to 7 images of increasing detail, and potential lures were listed. Eight to eleven colleagues, family and friends of the research team not involved in image creation were asked to rate the degree to which each image looked like each word beside it, with the highest-rated lures and images being selected for that item. Due to time constraints, the remaining pictures were not piloted and lures were determined via consensus of the research team.

3.1.3.4.2. Condition Selection

Contrary to previous iterations of the BADE/evidence integration tasks, which were rated on a scale, this novel task was rated using a dichotomous (yes/no) response interface, in order to simplify response-related brain activity and reduce movement artefacts, as well as to remove the confound of differences between conditions possibly being related to responding, as was observed in the previous study (see Chapter 2 and Lavigne et al., 2015a, Component 4). This response interface led to 4 possible response patterns given two responses per trial (one each to image 1 and image 2): Yes-Yes (YY), No-No (NN), No-Yes (NY), and Yes-No (YN). YY and NN are "confirm" conditions, since the initial rating is supported by the second rating, whereas NY and YN are "disconfirm" conditions, since the second rating contradicts the first rating.

The stimuli and prompt words for these conditions were based on the expected response patterns (see Figure 8 for examples of images for each condition). For the YY condition, the prompt word was the true word (i.e., that which matched the full picture), and both images looked like the picture, with each image showing increasingly more of the picture. For the NN condition, the prompt word was an absurd word that did not look like either image 1 or 2, nor the full picture. For the NY condition, the prompt word was the true word; the first image did not look like the prompt word, but the second image showed more of the picture such that it did look like the prompt/true word. For the YN condition (which was the condition of interest), the prompt word was a lure, which looked like the first image. However, the second image was designed to look more like the full picture, and as such no longer looked like the lure word. The true word was always shown below the full picture at the end of the trial. Two images were created for each condition for each of the 376 images, leading to 3008 total

images. The most compelling lures/images were selected first to create the YN condition as it was the condition of interest, and the remaining conditions were determined by selecting the strongest conditions via consensus of the research team.

3.1.3.4.3. Procedure

Participants were presented with a partial image and asked to rate (on a dichotomous yes-no response interface) whether they believed the full picture (shown at the end of the trial) was of the word listed below the image. They were then presented with a second image showing more of the picture and asked to re-rate. Finally, a third image was displayed showing the full picture. Figure 9 shows the timing of the experimental paradigm using an example disconfirm (YN) trial with the picture "bat" and lure word "umbrella". Each trial proceeded as follows (presentation times in milliseconds are listed in parentheses): image 1 & rating (4000ms) → image 2 & rating (4000ms) → image 3 (1000ms). The first two images and response options were displayed for 4000ms regardless of whether a response was made, and participants were allowed to modify their responses during this time window. If multiple responses were made, the last response was selected as the final response for that image, and other responses were discarded. The full picture (image 3) was displayed for 1000ms and any responses during that time were not recorded. Jittered ITIs of 2500ms, 3500ms, 5000ms, 10000ms, and 20000ms were included to optimize deconvolution of the BOLD signal (Serences, 2004).

Eighty images were included in the current study; the remaining images were used for practice trials, the longitudinal study described in Chapter 4, as well as for accompanying EEG data collection not addressed in this research. The 80 images were split into two runs of 40 trials per run, lasting approximately 11 minutes each. Participants performed a practice

session approximately 15 minutes prior to the experimental session. The practice session consisted of 16 trials (4 of each condition) using images not present in the experimental runs. The first 4 practice trials had no time limit for responding and were included to familiarize participants with the task requirements. The remaining 12 trials followed the experimental timing (4 seconds each to respond to images 1 and 2) to allow participants to become comfortable with the timing of the task.

Unlike version 2 of the evidence integration task in study 2 (Chapter 2; Lavigne et al., 2015a), the images in this task were displayed during rating. Although this might confound brain activity related to responding and visual presentation, the combination was deemed necessary to ensure the second image would clearly be adding additional information to the first image (if image 1 is removed during responding and image 2 is then displayed from a black screen, the notion of evidence accumulation is not as salient).

3.1.4. fMRI Acquisition and Preprocessing

MRI data were collected at the UBC MRI Research Centre, which uses a Phillips Achieva 3.0 T MRI scanner (maximum gradient amplitude, 80 mT/m; maximum slew rate, 200 mT/m/s) with a single shot echo planar imaging (EPI) sequence. The participant's head was secured using a customized head holder. Functional MRI volumes were collected using a T2*-weighted gradient-echo spin pulse sequence with 35 axial slices: slice thickness/gap, 3mm/1mm; reconstruction matrix, 96 × 96; TR, 2000ms; TE, 30ms; FA, 90°; FOV, 288mm; voxel size, 3mm^3, effectively covering the whole brain. Up to 325 volumes were collected in each of the two runs, which lasted 10 min and 50 s each. For three patients, the MRI technician ran the wrong scanning protocol, which led to 310 instead of 325 volumes being collected for the second run. Although the number of volumes collected differed for these

few participants, they were not excluded from the analysis because the majority of the run was completed. A high resolution anatomical MRI image of the entire brain was collected at the end of the scanning session with the following parameters: voxel size, 1 mm^3; TR, 8.1ms; TE, 3.7ms; FA, 8°; FOV, 256 × 256 × 190; acquisition matrix, 256 × 250; number of slices, 190; sagittal orientation. Three participants (2 controls, 1 patient) did not have an anatomical MRI image due to time constraints during the scanning session or the presence of artefacts.

fMRI preprocessing was performed with SPM8 (Wellcome Trust Centre for Neuroimaging, UK). Anatomical images were first visually inspected for artefacts and manually reoriented to the anterior-posterior commissure line. One patient was excluded due to an incidental finding at this stage. For each subject and run, functional images were then corrected for slice timing acquisition (reference slice was in the middle of the brain), realigned to the mean image, and co-registered to the participant's anatomical MRI (if obtained). Structural images were segmented into grey matter, white matter, and cerebrospinal fluid, and corrected for intensity non-uniformity if required. Functional scans with an accompanying structural image were then normalized to the native T1 SPM template; those without structural images were normalized to the EPI template (voxel size = 2mm^3). Finally, functional scans were smoothed using a 6mm FWHM Gaussian kernel. Realignment parameters were then examined separately for each participant and run to determine the extent of head motion. Participants/runs with head motion exceeding 2 voxels (4mm translation or 4° rotation for a voxel size of 2mm^3) were excluded from analysis. This led to the removal of 26 runs; four participants (2 controls, 2 patients) showed excessive head motion during both runs and were excluded entirely. Excluded participants are taken into account in the sample sizes listed previously.

3.1.5. Statistical Analysis

3.1.5.1. Behavioural

Group differences between controls, non-delusional, and delusional patients on demographic variables (age, education), IQ (premorbid and current) and depressive symptoms (BDI) were examined using one-way ANOVA, and Tukey's honest significant difference (HSD) test was applied post-hoc to determine specific differences between each group pair. Differences on gender and handedness were examined using Chi-square. T-tests were used to compare the non-delusional and delusional groups on clinical (chlorpromazine equivalents, PSYRATS, SAPS, and SANS) and neurocognitive (RBANS, TMT, LNS, COWAT) measures.

For the behavioural BADE task, we conducted a PCA with varimax rotation using IBM SPSS Statistics version 20 on the mean ratings (averaged across BADE trials, excluding filler trials) to identify components (i.e., evidence integration and conservatism) underlying the twelve behavioural BADE items as has been done previously (Sanford et al., 2014; Speechley et al., 2012). Composite component scores were calculated by summing the relevant items for each component, and differences between controls, non-delusional, and delusional patients were assessed using a one-way ANOVA. Correlations were also computed between composite behavioural BADE scores and demographic measures, premorbid (ToPF) and current (WASI) IQ, clinical (SPQ for controls; PSYRATS/SAPS/SANS for patients), and neurocognitive (i.e., RBANS, TMT, LNS, and COWAT for patients only) measures. All correlations were computed using Pearson's r, except for those with the categorical variables of gender and handedness, which were calculated via Spearman's rho.

In order to further examine the influence of variables of non-interest (e.g., demographic, IQ, neurocognitive measures) on behavioural BADE components, significant correlations were followed up with an analysis of covariance (ANCOVA) to determine whether significant group differences remained after controlling for these variables. To investigate the role of neurocognition on BADE behaviourally, we submitted all neurocognitive measures and the behavioural BADE composite scores to a principal component analysis, as was done in previous research at the item-level in healthy controls (Woodward et al., 2007). We expected the behavioural BADE components to load most strongly onto a component orthogonal from those driven by neurocognitive measures, supporting the notion that they represent separate and unique cognitive processes.

For the neuroimaging evidence integration task, overall and condition-specific (confirm and disconfirm) accuracy and mean reaction time (RT) scores were calculated. Accuracy scores reflected the percentage of trials that matched the expected responses per condition (correct responses to both images were required for each trial in order to qualify the trial as a match) divided by the number of trials (i.e., number of matches/number of trials). Match scores (e.g., a Yes response followed by a No response to a YN trial) were first calculated for each run separately (maximum = 40 for 40 trials) and then averaged across runs in order to equate subjects with a different number of runs. The average number of matches across both runs was then divided by the number of trials (40 for a single run). RT scores reflected the time (in ms) from the onset of a given image to the final response (within the 4000 ms time window) ignoring any initial responses if multiple responses were given. Therefore, RTs were calculated for each of the two images in each of the 40 trials per run, and averaged across all images, trials and runs. Overall scores were created by averaging across all four

conditions (YY, NN, NY, YN) and trials (10 for each condition) in both runs, and condition-specific scores were created by averaging across trials and runs for the confirm (YY and NN) and disconfirm (NY and YN) conditions separately. The overall number of missing responses (maximum = 80 given two responses per 40 trials) was also calculated, summed across all conditions and averaged across runs. One-way ANOVAs were used to compare the groups (control, non-delusional, delusional) on measures of fMRI accuracy, RT, and missing responses.

3.1.5.2. fMRI-CPCA

fMRI data were analyzed using group constrained principal component analysis for fMRI (group fMRI-CPCA; Metzak et al., 2011; Woodward et al., 2013). fMRI-CPCA is a multivariate statistical technique that combines multivariate multiple regression and PCA to identify task-specific functional brain networks, that is, networks that vary as a function of task timing. This is achieved by computing a PCA on the predicted scores resulting from a multivariate multiple regression in which the criterion variables are fMRI BOLD signal (for each subject, scan and run for each voxel of the brain) and the predictor variables are stimulus onsets (for each subject and condition over time). The initial regression step serves to separate the variance in overall brain activity into that which can be predicted by stimulus timing (i.e., predicted scores) and that which cannot (i.e., residual scores). A PCA is then computed on these predicted scores, resulting in functional brain networks that are predictable from stimulus timing. fMRI-CPCA has the advantage of deriving networks that are inherently dependent on stimulus timing, rather than correlating networks resulting from overall brain activity with stimulus timing *ad hoc*, which is often the case with other network-based techniques (e.g., independent component analysis). When applied to multiple

groups, group fMRI-CPCA allows for visualization of common brain networks and comparison of activity across groups, through observation of spatial and temporal replication across groups within each network (Lavigne et al., 2015a; Ribary et al., 2017).

Preprocessed fMRI data from all subjects are concatenated into a data matrix, with a single row for each subject, run, and scan, and a single column for each voxel. For the current study, this led to a 57495 (98 subjects × up to 2 runs × up to 325 scans) × 585390 (number of voxels for entire MRI image subsampled to 2mm^3) matrix. This data matrix then undergoes several transformations to improve data quality; a mask is applied excluding all non-brain areas (which reduced the number of voxels to 265507 in the current study), linear and quadratic effects are regressed out, head motion is regressed out using the six realignment parameters (i.e., x, y and z translation, and pitch, roll, and yaw rotation) derived from SPM, and finally, the data is standardized to a mean of 0 and standard deviation of 1. This transformed data matrix is then regressed onto a FIR-based design matrix, which reflects the timing of the experimental stimuli for each subject, condition, and for the number of post-stimulus time bins modeled (10 TRs, or 20 seconds, for the current study). The rows of the design matrix mirror those of the data matrix, with a single row for each subject, run, and scan. There is also a single column for each subject, condition, and post-stimulus time bin combination. A value of 1 is placed in cells where hemodynamic response is to be estimated, and a value of 0 is placed everywhere else, creating mini-boxcar functions. This led to a 57495 (98 subjects × up to 2 runs × up to 325 scans) × 3920 (98 subjects × 4 conditions × 10 poststimulus time bins) design matrix, which was also standardized to a mean of 0 and standard deviation of 1 prior to the regression analysis.

The multivariate multiple regression of fMRI BOLD signal onto task timing produces a matrix of predicted scores, which reflects BOLD signal that can be predicted by the experimental timing, and produces a matrix of residual scores reflecting BOLD signal unrelated to experimental timing. The matrix of predicted scores is then subjected to a PCA, and the scree plot (Cattell, 1966; Cattell & Vogelmann, 1977) is examined to determine the number of components to extract. The components are rotated using an hrfmax orthogonal rotation (Metzak et al., 2011), which fits the component solution to user-defined HDR shapes derived from the experimental timing. An in-depth description of the calculations underlying fMRI-CPCA, which are identical to those involved in group fMRI-CPCA, is listed in the previous chapter.

3.1.5.2.1. Relation to Experimental Conditions

fMRI-CPCA provides estimates of both spatial (dominant brain regions) and temporal (task-based HDR shapes) activation for each component (i.e., functional brain network) identified. Predictor weights, which reflect estimates of HDR for each component for each combination of subject, condition, and poststimulus time bin, are the weights that would be applied to the design matrix to create the component scores. These predictor weights can be analyzed statistically to determine whether each network shows a reliable HDR shape (and does not simply vary around zero), and whether there are any differences between conditions and/or groups over poststimulus time. These analyses were carried out using repeated-measures ANOVA in SPSS, one for each component/functional brain network, with the within-subjects factors of Poststimulus Time (10 whole-brain scans after the onset of each trial were estimated in the FIR model) and Condition (YY, NN, NY, YN), and the between-subjects factor of Group (control, non-delusional, delusional). Significant interactions were

followed up with simple contrasts comparing each level of the relevant factors to one another (Howell & Lacroix, 2012). Tests of sphericity were carried out for all ANOVAs, and adjustment for violations of sphericity did not affect the results; therefore, the original degrees of freedom are reported below.

3.1.5.3. Associations Between Averaged HDR and Behaviour

3.1.5.3.1. Confirmatory and Disconfirmatory Evidence Integration

Associations between brain activity underlying dis/confirmatory evidence integration and behavioural measures were first examined using correlations, by averaging over HDR in the relevant conditions (YY and NN for confirmatory evidence integration, and NY and YN for disconfirmatory evidence integration) and all poststimulus time bins in each network separately. This produced two averaged HDR values in each network (one each for activity during confirmatory and disconfirmatory evidence integration), which were then correlated with behavioural BADE, fMRI performance, neurocognition, and symptoms. In order to determine in more detail the complex relationships between these variables, three behavioural CPCAs were also conducted, with the six averaged HDR values (2 for each network) as the dependent variables, and behavioural measures as independent variables. The behavioural measures were divided into three categories, leading to three behavioural CPCAs: (1) behavioural BADE/overall fMRI performance (n = 56; patients only); (2) neurocognition (n = 58); and (3) symptoms (n = 55). The behavioural BADE (evidence integration, conservatism) and fMRI performance (overall accuracy and RT) measures were combined as they all assessed evidence integration behaviourally.

Calculation of the behavioural CPCAs was identical to group fMRI-CPCA, except that Z consisted of one row per subject (instead of subject $\times$ scan $\times$ run) and one column per fMRI

(averaged HDR) variable (instead of voxel), and G consisted of one row per subject (instead of subject × scan × run), and one column per behavioural variable (instead of subject × condition × poststimulus time bin). Z is standardized so each column has a mean of 0 and standard deviation of 1, then is regressed onto G, producing a matrix of predicted scores, which are then submitted to a PCA with varimax rotation. Unlike in group fMRI-CPCA, G was not standardized prior to the regression analysis in order to retain the information contained in the different metrics used by the included measures. The PCA produces components that reflect combinations of functional brain network variables (e.g., averaged HDR underlying confirmatory and disconfirmatory evidence integration) that are predictable by the independent variables of interest (i.e., behavioural BADE/fMRI performance, neurocognition, or symptoms). The calculations underlying the behavioural CPCAs are identical to those involved in fMRI-CPCA, which can be found in the previous chapter.

The Z matrices were similar for each of the three behavioural CPCAs (except for the number of rows due to different numbers of subjects who completed the behavioural tasks, namely 56 for behavioural BADE/fMRI performance, 58 for neurocognition, and 55 for symptoms), and consisted of six columns reflecting estimated HDR averaged over poststimulus time for each network and condition type (i.e., VsAN confirm, VsAN disconfirm, visual/default-mode network (VDMN) confirm, VDMN disconfirm, CEN confirm, CEN disconfirm). The G matrices included the same number of rows as the corresponding Z matrices, and the columns reflected scores for the relevant behavioural variables. Specifically, the behavioural BADE/fMRI performance G matrix consisted of four columns (BADE evidence integration, BADE conservatism, overall fMRI accuracy, and overall fMRI RT), the neurocognition G matrix consisted of nine columns (RBANS

immediate memory, RBANS visuospatial abilities, RBANS language, RBANS attention, RBANS delayed memory, TMT A, TMT B, LNS, and COWAT), and the symptoms *G* matrix consisted of 11 columns (PSYRATS hallucinations, PSYRATS delusions, SAPS hallucinations, SAPS delusions, SAPS bizarre behaviour, SAPS formal thought disorder, SANS affective flattening, SANS alogia, SANS avolition-apathy, SANS attention, and SANS inappropriate affect). Both the PSYRATS and SAPS hallucinations and delusions variables were included due to their distinct assessment of the presence of versus the phenomenology of hallucinations and delusions, respectively. The SANS anhedonia-asociality subscale was excluded from this analysis due to the relatively large number of missing data. The number of components extracted for each solution and analysis was determined by inspection of the scree plots.

3.1.5.3.2. Baseline-Peak Activations

In previous research (Lavigne et al., 2016), we demonstrated that functional brain activity from baseline-to-peak (BP) and from peak-to-baseline (PB) of the HDR in each network is associated with distinct cognitive processes, and that examining these portions of the hemodynamic response separately can provide additional insight into the cognitive operations involved. Correlating these BP and PB values within and between networks also provides a better understanding of how these networks are interacting with one another. BP and PB values are calculated by averaging predictor weights occurring from the start of the trial until the peak of the response (for BP), and from the time bin immediately following the peak (when the HDR begins to fall back to baseline) to the end of the trial (for PB). This results in two values for each network (BP and PB), which can be intercorrelated to examine associations within and between networks, and may also be correlated with behavioural

measures of interest. For example, in the current study, we examined correlations between these measures and behavioural BADE, fMRI performance, neurocognition, and symptoms, as was done in the previous section for brain activity underlying dis/confirmatory evidence integration. In addition, we computed three behavioural CPCAs (dependent variables: BP/PB values for each network; independent variables: (1) behavioural BADE/overall fMRI performance; (2) neurocognition; (3) symptoms) to examine multivariate associations between BP/PB measures that are directly predictable from behaviour. The behavioural CPCA procedure was identical to that described in the previous section except that BP/PB measures were placed in the Z matrices rather than average HDR during confirmatory and disconfirmatory evidence integration. The Z matrices were the same size as the previous behavioural CPCAs (behavioural BADE/fMRI performance = 56 × 6, neurocognition = 58 × 6, and symptoms = 55 × 6) since the number of BP/PB variables (2 per network averaged over confirm/disconfirm conditions) equalled that of the confirm and disconfirm variables above (two per network collapsed over time). The G matrices were identical to those in the previous behavioural CPCAs. The number of components extracted was determined for each solution and analysis by inspection of scree plots.

3.2. Results

3.2.1. Behavioural

Descriptive information is displayed separately for each group in Table 5 for demographic variables (gender, handedness, age, education), premorbid (ToPF) and current (WASI) IQ, depressive symptoms (BDI), and chlorpromazine equivalents, Table 6 for clinical measures (PSYRATS, SAPS, SANS for patients, and SPQ for controls), Table 7 for

neurocognitive measures (RBANS, TMT, LNS, COWAT; patients only), and Table 9 for behavioural BADE components and fMRI performance.

Significant group differences were observed on education, $F(2,95) = 11.13$, $p < .001$, and current IQ, $F(2,94) = 15.80$, $p < .001$, but not for age, gender, handedness, premorbid IQ, or chlorpromazine dose equivalents (all ps $> .08$). Tukey's HSD post-hoc tests revealed that both non-delusional and delusional patients had fewer years of education and lower current IQ scores than controls (ps $< .005$), but did not differ from one another (ps $> .52$). There was also a significant difference between the groups on the BDI total score, $F(2,91) = 25.33$, $p < .001$, due to greater depressive symptoms in both non-delusional and delusional patients relative to controls, and in delusional patients relative to non-delusional patients (all ps $< .01$). Although these group differences emerged, the severity of depressive symptoms was mild or minimal in both patient groups as can be seen from the mean scores (maximum = 63; see Table 5). T-tests comparing the non-delusional and delusional groups on clinical measures showed that delusional patients had significantly higher ratings on PSYRATS hallucinations, $t(55) = -2.11$, $p < .05$, and delusions, $t(55) = -11.88$, $p < .001$, as well as on SAPS total, $t(55) = -5.27$, $p < .001$, hallucinations, $t(55) = -2.69$, $p < .01$, delusions, $t(56) = -6.15$, $p < .001$, and bizarre behaviour, $t(56) = -2.07$, $p < .05$. There were no significant differences between non-delusional and delusional patients on SAPS formal thought disorder or negative symptoms, nor on neurocognitive variables (all ps $> .09$).

3.2.1.1. Behavioural BADE

The scree plot (Cattell, 1966; Cattell & Vogelmann, 1977) and Kaiser-Guttman criterion (eigenvalues > 1; Kaiser, 1991) converged on the two-component solution reported in previous studies (Sanford et al., 2014; Speechley et al., 2012). The eigenvalues were 7.21,

2.69, 0.74, 0.56, 0.19, 0.16, 0.14, 0.11, 0.06, 0.06, 0.05, and 0.03, and the first two components accounted for 82.47% of the total variance. Based on the rotated sums of squared loadings, the first component accounted for 44.72% of the overall variance, and the second component accounted for 37.76% of the overall variance. The rotated component loadings are listed in Table 8. Bolded values correspond to those items identified within each component in previous studies (Sanford et al., 2014; Speechley et al., 2012), and the two components that emerged strongly replicated those of previous research, namely conservatism (a reduced willingness to rate high when justified; higher scores = less conservatism) and evidence integration (degree to which disconfirmatory evidence has been integrated; higher scores = less evidence integration). The conservatism component was dominated by the first and second ratings of both lure interpretations, as well as all three true interpretations, whereas the evidence integration component was primarily captured by the three absurd interpretations, the latter two ratings of both lure interpretations, and the third rating for the true interpretation (reversed).

Composite scores for each component were calculated by summing the relevant items (i.e., neutral and emotional lure ratings 1 and 2, and all three true ratings for BADE conservatism, and all three absurd ratings, neutral and emotional lure ratings 2 and 3 and true rating 3 (reversed) for BADE evidence integration), as was recommended in previous research (Sanford et al., 2014). Table 9 displays means and standard deviations for each BADE component as a function of group (control, non-delusional, delusional; patient groups based on median split of PSYRATS delusions score). There was a significant difference between groups on both BADE evidence integration, $F(2,93) = 6.65$, $p < .005$, and BADE conservatism, $F(2,93) = 4.75$, $p < .05$. Tukey's HSD post-hoc tests revealed that delusional

patients had significantly higher BADE evidence integration scores than controls and non-delusional patients (ps < .05), who did not differ from one another (p > .88, see Figure 10A). These results suggest that delusional patients had significantly higher ratings on the absurd items and later ratings of the lure items, as well as lower ratings on the final true item, thus they were impaired on evidence integration. For BADE conservatism, delusional patients had significantly greater scores than non-delusional patients (p < .01), but not controls (p > .12), who were situated between non-delusional and delusional patients (see Figure 10B). Although they did not differ significantly from controls (p > .38), these findings suggest that differences on BADE conservatism were driven by lower overall ratings (more conservatism) in non-delusional schizophrenia patients on true and early lure statements.

3.2.1.2. fMRI Performance

Table 9 displays overall and condition-specific mean fMRI accuracy (responses matching expected response pattern; in %), RT (in ms), as well as overall missing responses (out of 80 possible responses given two responses per 40 trials averaged across runs) for each group for the neuroimaging evidence integration task. For fMRI accuracy, significant group differences were observed for overall, $F(2,95) = 4.41$, p < .05, and disconfirm, $F(2,95) = 3.78$, p < .05, but not confirm (p > .12), scores. Tukey's HSD post-hoc tests revealed that delusional patients were significantly less accurate (Yes/No responses matched the designed conditions less frequently) than controls for overall accuracy (p < .05), and that non-delusional patients were significantly less accurate than controls for the disconfirm conditions (p < .05), with no other significant differences between groups (p > .05). For fMRI RT, significant group differences were observed for overall RT, $F(2,95) = 12.11$, p < .001, as well as for RT in the confirm, $F(2,95) = 13.17$, p < .001, and disconfirm, $F(2,95) = 10.79$, p < .001, conditions.

Tukey's HSD post-hoc tests revealed that both patient groups were significantly slower at responding than controls for all three measures of RT (ps < .005), with no significant differences between patient groups (p > .29). There were no significant group differences on missed responses (p > .26).

3.2.1.3. Correlations with Behavioural BADE

Correlations (see Table 10) indicated that BADE evidence integration was significantly negatively associated with WASI (poorer evidence integration associated with decreased current IQ), $r(96) = -0.32$, $p < .005$, and positively associated with BDI (poorer evidence integration associated with increased depressive symptoms), $r(93) = 0.26$, $p < .05$. BADE conservatism was significantly negatively associated with gender (greater conservatism in females relative to males), $r_s(96) = -0.23$, $p < .05$, and positively associated with BDI (greater conservatism associated with higher depressive symptoms), $r(93) = 0.22$, $p < .05$. No significant correlations were observed between either BADE conservatism or evidence integration and age, education, handedness, or premorbid IQ (ps > .08). A one-way ANCOVA was performed for each BADE component to determine whether significant group differences remained after controlling for these associated variables. For BADE evidence integration, the effect of Group remained significant after controlling for WASI and BDI, $F(2,88) = 3.26$, $p < .05$. The effect of Group on BADE conservatism also remained significant after controlling for BDI and gender, $F(2,88) = 3.15$, $p < .05$.

In terms of psychotic symptoms, as was found previously (McLean et al., 2017; Sanford et al., 2014; Speechley et al., 2012; Woodward et al., 2006b) BADE evidence integration was significantly positively associated with the PSYRATS total delusions score, $r(55) = 0.34$, $p < .05$, SAPS total score, $r(56) = 0.32$, $p < .05$, and formal thought disorder, $r(56) = 0.34$, $p <$

.05, in patients. BADE conservatism was also significantly associated with the PSYRATS

total delusions score, $r(55) = 0.29$, $p < .05$. No significant correlations were observed

between either BADE evidence integration or conservatism and negative symptoms (all ps >

.18). Also replicating past work (Buchy et al., 2007; Woodward et al., 2007), in healthy

individuals, BADE evidence integration was significantly positively associated with the SPQ

unusual beliefs and experiences factor, $r(40) = 0.35$, $p < .05$, which combines subclinical

traits of delusions (i.e., magical ideation) and hallucinations (i.e., perceptual aberrations), as

well as with the magical ideation subscale, $r(40) = 0.39$, $p < .05$. BADE conservatism was

not significantly associated with any of the SPQ factors or subscales (all ps > .07). For fMRI

performance, behavioural BADE evidence integration was negatively associated with overall

fMRI accuracy, $r(96) = -0.29$, $p < .005$, as well as fMRI accuracy for both confirm, $r(96) = -0.24$, $p < .05$, and disconfirm, $r(96) = -0.21$, $p < .05$, conditions. It was also positively

associated with overall fMRI RT, $r(96) = 0.29$, $p < .005$, and fMRI RT in the confirm, $r(96) = 0.28$, $p < .01$, and disconfirm, $r(96) = 0.29$, $p < .005$, conditions. Behavioural BADE

conservatism was negatively related to overall fMRI accuracy, $r(96) = -0.21$, $p < .05$, and

fMRI accuracy during the confirm conditions, $r(96) = -0.21$, $p < .05$, as well as positively

related to missed responses on the fMRI task, $r(96) = 0.24$, $p < .05$.

No significant correlations were observed between neurocognitive variables and

behavioural BADE evidence integration or conservatism (all ps > .07). In order to further

investigate the association between behavioural BADE and neurocognition, we submitted the

BADE components and neurocognitive variables (RBANS subscales, TMT A and B, LNS,

and COWAT) to a PCA. The scree plot (Cattell, 1966; Cattell & Vogelmann, 1977) and

Kaiser-Guttman criterion (eigenvalues > 1; Kaiser, 1991) suggested a 4-component solution

(eigenvalues = 3.82, 1.84, 1.14, 1.02, 0.72, 0.63, 0.54, 0.40, 0.34, 0.29, and 0.27), which

accounted for 71.09% of the overall variance. Based on the rotated sums of squared loadings,

Components 1 to 4 accounted for 28.03%, 17.01%, 15.40%, and 10.65% of the overall

variance, respectively. The first component consisted of high loadings (Table 11) from

RBANS language (0.51) and attention (0.78), TMT A (0.74) and B (0.77), as well as LNS

(0.71) and COWAT (0.70), and was labeled *Attention*. Component 2 was dominated by

RBANS immediate (0.88) and delayed (0.84) memory subscales, and was labeled *Memory*.

The behavioural BADE evidence integration (0.89) and conservatism (0.89) scores loaded

onto the third component, which was labeled *BADE*. Finally, the RBANS visuospatial

domain (0.90) made up the fourth component, and was labeled *Visuospatial*. These results

indicate that the behavioural BADE components are orthogonal to the included

neurocognitive measures, as hypothesized based on past work (Eisenacher & Zink, 2017;

Moritz et al., 2010; Woodward et al., 2007).

3.2.2. fMRI-CPCA

Inspection of the scree plot suggested that three components best explained the variance

in fMRI BOLD signal predictable from task timing. Components 1, 2 and 3 accounted for

16.06%, 12.41%, and 4.54% of the predictable variance, respectively. All three components

showed a significant main effect of Poststimulus Time, $F(9,855) = 167.49$, $p < .001$, $F(9,855)$

$= 157.12$, $p < .001$, and $F(9,855) = 188.23$, $p < .001$, for Components 1 to 3, respectively,

and inspection of the HDR shapes (Figures 12C, 14B, and 16C) indicated that that each

network showed a biologically plausible and reliable hemodynamic response shape (Lavigne

et al., 2015a; Metzak et al., 2011; Metzak et al., 2012).

3.2.2.1. Component 1: Visual Attention Network

3.2.2.1.1. Anatomical Description

The brain regions associated with Component 1 are displayed in Figure 11, with anatomical descriptions listed in Table 12. This network was characterised by activations in bilateral occipital cortex (BAs 18, 19), supramarginal gyrus (BA 40), middle frontal gyrus (BAs 6, 45), dACC (BA 32), left motor cortex (BA 4), as well as bilateral insula, thalamus, and cerebellum. These activations spread across the visual, somatosensory, and dorsal attention (DAN) networks of Yeo and colleagues' 7-network resting-state brain network parcellation (Buckner et al., 2011; Choi et al., 2012; Yeo et al., 2011), including additional recruitment of the FPN and VAN with regard to the ACC and insula activations. Given the spatial configuration of this network, and the pattern of activation over time (see below), we labeled it the *Visual Attention Network* (VsAN).

3.2.2.1.2. Relation to Experimental Conditions

Statistical analysis of the estimated hemodynamic response shape of this network (Figure 12) via ANOVA revealed significant main effects of Condition, $F(3,285) = 3.99$, $p < .01$, and Group, $F(2,95) = 5.59$, $p < .005$, as well as a significant interaction between Condition and Poststimulus Time, $F(27,2565) = 3.21$, $p < .001$. Simple contrasts exploring the main effect of Condition (Figure 12A) revealed significantly greater activation in YN relative to YY, $F(1,95) = 7.40$, $p < .01$, and NN, $F(1,95) = 8.87$, $p < .005$. The main effect of Group (Figure 12B) was the result of significantly greater activity in delusional patients relative to controls, $F(1,95) = 9.51$, $p < .005$, and non-delusional patients relative to controls, $F(1,95) = 6.07$, $p < .05$, with no difference observed between delusional and non-delusional patients ($p > .50$). Simple contrasts examining each level of Condition at each level of Poststimulus Time

(Figure 12C) revealed significantly greater activity in (1) YY versus NN at 1s, 5s, (2) NN versus YY at 13s, (3) NY versus YY at 9s, 13s, and 15s, (4) YY versus YN at 1s, (5) YN versus YY at 13s, 15s, and 17s (6) NY versus NN at 5s, (7) YN versus NN at 5s, 7s, 13s, 15s, 17s, and 19s, (8) NY versus YN at 9s, and (9) YN versus NY at 13s and 15s (all ps < .05). These findings indicate that activity was highest in the VsAN during disconfirm conditions (particularly YN) late in the trial, and that both delusional and non-delusional patients showed hyperactivity in this network regardless of condition, with non-delusional patients situated between controls and delusional patients, suggesting a linear progression from controls to non-delusional to delusional patients in terms of intensity of activation in this network. This interpretation was supported by a significant linear trend (coefficients: 1 0 -1 for controls, non-delusional, and delusional patients, respectively; p < .005) and a non-significant quadratic trend (coefficients: 1 -2 1; p > 0.34).

3.2.2.2. Component 2: Visual/Default-Mode Network

3.2.2.2.1. Anatomical Description

The brain regions associated with Component 2 are displayed in Figure 13, with anatomical descriptions listed in Table 13. This network was characterised by activations in bilateral occipital cortex (BAs 17, 18, 19) and superior parietal lobule (BA 7), as well as deactivations in bilateral precuneus (BAs 18, 23), anterior/posterior cingulate cortex (BAs 32/23), lateral temporal cortex (BAs 21, 22, 41, 42), and ventromedial prefrontal cortex (BAs 9, 10). Relating this network to the resting-state 7-network parcellation (Buckner et al., 2011; Choi et al., 2012; Yeo et al., 2011), the activations corresponded to the visual network, extending into the DAN for parietal regions. The deactivations reflected the DMN, extending

into visual, somatosensory, and ventral attention networks. We labeled this component *Visual/Default-Mode Network* (VDMN).

3.2.2.2.2. Relation to Experimental Conditions

Statistical analysis of the estimated hemodynamic response shape of this network (Figure 14) via ANOVA revealed a significant main effect of Group, $F(1,95) = 4.45$, $p < .05$, as well as a significant interaction between Condition and Poststimulus Time, $F(27,2565) = 2.38$, $p < .001$. Delusional patients showed significantly less intensity (i.e., lesser activations and deactivations) on this network relative to controls (Figure 14A), $F(1,95) = 8.40$, $p < .005$, with no other group comparisons reaching significance (ps > .06). The Condition × Poststimulus Time interaction (Figure 14B) was the result of increased intensity in this network for the YN condition at peak. Specifically, YN showed significantly higher activity than (1) YY at 11s, and 13s, (2) NN at 9s, 11s, and 13s, and (3) NY at 11s (all ps < .05). There was also significantly increased intensity in YY relative to NN at 3s and 5s, and NY relative to YY at 13s (all ps < .05). Thus, there was increased intensity (i.e., greater deactivations in DMN regions and greater activations in visual regions) in this network during disconfirmatory relative to confirmatory evidence integration at the peak of the HDR, and delusional schizophrenia patients showed less intensity relative to controls, regardless of condition. As with the VsAN, significant linear ($p < .005$), and nonsignificant quadratic ($p > .29$), contrasts indicated progressively decreased activation from controls to non-delusional patients to delusional patients.

3.2.2.3. Component 3: Cognitive Evaluation Network

3.2.2.3.1. Anatomical Description

The brain regions associated with Component 3 are displayed in Figure 15, with anatomical descriptions listed in Table 14. This network was characterised by activations in bilateral orbitofrontal cortex (BA 38), dorsolateral prefrontal cortex (BAs 46, 47), inferior frontal gyrus, pars triangularis (BA 45), angular gyrus (BAs 39, 40), posterior middle temporal gyrus (BA 21), occipital cortex (BAs 18, 19), and cerebellum. Deactivations were also present in left precentral gyrus (BA 3) and bilateral supplementary motor area (SMA; BA 6). The activations on this network corresponded to the visual and frontoparietal networks of the 7-network parcellation (Buckner et al., 2011; Choi et al., 2012; Yeo et al., 2011), with some activations extending into the VAN, DAN, and DMN. The deactivations in precentral gyrus and SMA corresponded to the somatosensory network. Given the similarity between this network and that found in our previous study, as well as the pattern of activity across time and conditions (see below), we labeled this network the *Cognitive Evaluation Network* (CEN).

3.2.2.3.2. Relation to Experimental Conditions

Statistical analysis of the estimated hemodynamic response shape of this network (Figure 16) via ANOVA revealed a significant main effect of Condition, $F(3,285) = 42.10$, $p < .001$, as well as a significant two-way interaction between Condition and Poststimulus Time, $F(27,2565) = 11.83$, $p < .001$, and a significant three-way interaction between Condition, Poststimulus Time and Group, $F(54,2565) = 1.78$, $p < .001$. The main effect of Condition (Figure 16A) was the result of significant differences between all condition pairs (all ps < .001 other than YY vs. NN $p < .05$), except for YY and NY ($p > .16$). The Condition ×

Poststimulus Time interaction (Figure 16C) was the result of increased intensity in this network for the disconfirm relative to confirm conditions. Specifically, YN showed significantly greater activation than (1) YY 9s, 11s, 13s, 15s, 17s, and 19s, (2) NN at 1s, 3s, 7s, 9s, 11s, 13s, 15s, 17s, and 19s, and (3) NY at 5s, 7s, 9s, 11s, 13s, 15s, 17s, 19s (all ps < .05). Similarly, NY showed significantly higher activity than (1) YY at 5s, 7s, 9s, 13s, 15s, and 17s, (2) NN at 1s, 11s, 13s, 15s, 17s, and 19s (all ps < .05). There was also significantly increased activity in YY relative to NN at 1s, 7s, 9s, and 11s (all ps < .05). Simple contrasts exploring the three-way interaction revealed significant differences between (1) controls and non-delusional patients for YY at 7s, 9s, and 11s; (2) controls and delusional patients for YY at 11s, for NN at 1s, 11s, 13s, and 19s, for NY at 9s, 11s, and 13s, and for YN at 5s; and (3) between delusional and non-delusional patients for YY at 3s, and NN at 1s, 3s, and 19s (all ps < .05). These findings show that disconfirmatory relative to confirmatory evidence integration led to higher activation in this network late in the trial, and that schizophrenia patients (particularly delusional patients during disconfirmatory evidence integration) showed hypoactivity relative to controls approximately mid-way through the trial.

3.2.3. Associations Between Averaged HDR and Behaviour

3.2.3.1. Confirmatory and Disconfirmatory Evidence Integration

3.2.3.1.1. Correlations

Given the lack of significant overall differences between the two confirm (YY and NN) and between the two disconfirm (NY and YN) conditions, the predictor weights (estimated HDR) were averaged across all poststimulus time bins for the confirm and disconfirm conditions separately, in order to examine associations between brain activity underlying confirmatory and disconfirmatory evidence integration in each network and behavioural

measures (i.e., behavioural BADE, fMRI performance, neurocognitive measures, and symptoms). The correlation matrix is displayed in Table 15. For the VsAN, activity during confirmatory evidence integration was significantly negatively associated with overall fMRI accuracy, $r(98) = -0.22$, $p < .05$, but not with accuracy in the confirm or disconfirm conditions individually nor any fMRI RT measures (ps > .07). VsAN activity during disconfirmatory evidence integration was negatively associated with RBANS language, $r(58) = -0.29$, $p < .05$, and positively associated with SANS affective flattening, $r(58) = 0.27$, $p < .05$. In the VDMN, activity during confirmatory evidence integration was positively associated with COWAT, $r(58) = 0.29$, $p < .05$, and activity during disconfirmatory evidence integration was negatively associated with SANS total, $r(49) = -0.31$, $p < .05$, SANS affective flattening, $r(58) = -0.30$, $p < .05$, and SANS alogia, $r(58) = -0.31$, $p < .05$. For the CEN, activity during confirmatory evidence integration was positively associated with fMRI accuracy in the disconfirm conditions, $r(98) = 0.20$, $p < .05$, and negatively associated with RBANS attention, $r(58) = -0.29$, $p < .05$. CEN activity during disconfirmatory evidence integration was positively associated with fMRI accuracy both overall, $r(98) = 0.25$, $p < .05$, and for disconfirm, $r(98) = 0.27$, $p < .05$, and was negatively associated with fMRI RT overall, $r(98) = -0.25$, $p < .05$, as well as for both confirm, $r(98) = -0.27$, $p < .05$, and disconfirm, $r(98) = -0.24$, $p < .05$, conditions. Finally, CEN activity during disconfirmatory evidence integration was also negatively associated with BADE evidence integration, $r(96) = -0.24$, $p < .05$.

3.2.3.1.2. Behavioural CPCAs

In order to further investigate associations between brain activity underlying confirmatory and disconfirmatory evidence integration and behaviour, we performed three

behavioural CPCAs with the averaged-over-time confirm and disconfirm HDR values as the dependent variables, and behavioural measures (1: behavioural BADE and overall fMRI performance, 2: neurocognitive variables, and 3: symptoms) as the predictor variables. This type of multivariate analysis allows us to identify combinations of functional brain networks (and condition types) that are best predicted by our behavioural measures of interest that may not be readily observed through correlations alone. Scree plots (Cattell, 1966; Cattell & Vogelmann, 1977) were inspected to determine the number of components to extract for each solution and each analysis separately. Two components were extracted for the overall (i.e., overall variance in average estimated HDR prior to regressing onto behavioural measures), predicted (i.e., variance predictable from behavioural measures), and residual (i.e., variance not predictable from behavioural measures) solutions for all analyses.

3.2.3.1.2.1. Behavioural BADE/fMRI Performance

For the behavioural BADE and fMRI performance CPCA, Table 16 shows the percentage of variance accounted for by each component for each solution, and Table 17 lists the rotated component loadings (dependent variables), as well as the correlations between the independent variables and the components extracted from the predicted solution (i.e., predictor loadings). Similar components were identified for the overall and residual solutions. The first component included strong loadings from VsAN and VDMN (reversed) for both confirm (0.79 and -0.78) and disconfirm (0.74 and -0.83) conditions. This indicates that in people for whom the VsAN activated strongly, the VDMN deactivated weakly, and vice-versa. That is to say, either people activate the VsAN or deactivate the VDMN, but tend to not do both simultaneously, evidenced by the positive VsAN and negative VDMN loadings within components across the solutions, and confirmed via inspection of

scatterplots. The second component was dominated by loadings from the CEN confirm (0.92) and disconfirm (0.90) conditions, which indicates that CEN activity is largely independent from VsAN/VDMN activity, due to its low loadings on Component 1; however, the relatively high loadings for VsAN confirm (-0.38) and disconfirm (-0.37) on Component 2 show that there is a negative association between activity in the VsAN and CEN, though not as strong as for the VsAN and VDMN.

For the predicted solution, the first component included the highest loadings from VsAN confirm (0.16), VDMN confirm (-0.17) and CEN disconfirm (-0.35), while the second component was dominated by CEN confirm (0.17) and VsAN confirm (-0.16). The predictor weights showed that increased behavioural BADE evidence integration (0.93), overall fMRI RT (0.54), and reduced overall fMRI accuracy (-0.47) were associated with greater activity in VsAN confirm, and decreased activity in VDMN confirm and CEN disconfirm. In contrast, greater activity in CEN confirm (and less activity in VsAN confirm) was associated with increased fMRI accuracy (0.50) and decreased BADE conservatism (-0.81). These findings are in line with the correlational findings described above, where behavioural BADE evidence integration was associated (-0.24, Table 15) with disconfirmatory evidence integration in the CEN; however, they also demonstrate that behavioural BADE evidence integration is related to hyperactivity in the VsAN (and less intensity in the VDMN) during confirmatory evidence integration. This aberrant activity (hyperactivity in the VsAN during confirmatory evidence integration, hypoactivity in the VDMN, hypoactivity in the CEN during disconfirmatory evidence integration) was not only associated with behavioural BADE evidence integration, but poorer performance on the fMRI task in terms of both accuracy and reaction time.

Thus, BADE evidence integration showed differential associations with the VsAN and VDMN/CEN, namely, it was positively related to the VsAN (meaning poorer evidence integration was associated with more VsAN activation), and negatively related to the VDMN/CEN (meaning better evidence integration was associated with more CEN activation and greater VDMN intensity). Given the higher loadings for confirmatory evidence integration on Component 1 for the predicted solution (0.16 and -0.17 for VsAN confirm and VDMN confirm, respectively), and for disconfirmatory evidence integration for the CEN (-0.35), these results support the correlational analyses, but also show that the associations with behavioural BADE evidence integration are present for brain activity underlying both confirmatory and disconfirmatory evidence integration, but in opposite directions. That is, poorer evidence integration behaviourally was associated with more VsAN activation (and this effect was stronger for confirmatory evidence integration), and greater evidence integration was associated with more CEN activation (and this effect was stronger for disconfirmatory evidence integration). Finally, behavioural BADE conservatism showed a similar pattern to behavioural BADE evidence integration; however, the association with VsAN confirm (evident on both Components 1 and 2, 0.42 and -0.81) was stronger, and the association with the CEN disconfirm (Component 1, 0.42) weaker, compared to BADE evidence integration (0.93 for Component 1 only). Since the conservativism construct is caused by a number of people who are unwilling to rate high (when justified) on the first rating (Speechley et al., 2012), this finding implies that increased activity on VsAN is part of appropriately high ratings (when justified) on the first picture.

3.2.3.1.2.2. Neurocognition

The variance table and rotated component loadings (dependent variables) as well as the predictor loadings (independent variables) for each solution for the behavioural CPCA on neurocognition and brain activity underlying dis/confirmatory evidence integration are displayed in Table 18 and Table 19, respectively. The overall and residual solutions were similar to the behavioural BADE/fMRI performance analysis, with Component 1 reflecting VsAN and VDMN (reversed) confirm and disconfirm conditions, and Component 2 reflecting CEN confirm and disconfirm conditions. The predicted solution also distinguished between confirmatory (Component 1) and disconfirmatory evidence integration (Component 2), except that CEN confirm loaded negatively on the first component (-0.24). In terms of the predictor loadings (see Table 19, Predicted Solution – Independent Variables), Component 1 was negatively related to RBANS immediate memory (-0.12), TMT A (-0.47) and COWAT (-0.53) and positively related to LNS (0.27), whereas component 2 was negatively associated with RBANS visuospatial (-0.20), language (-0.42), attention (-0.22), and TMT B (-0.37), and positively related to RBANS delayed memory (0.41). In combination with the behavioural BADE/fMRI performance results above, this set of results implies that VsAN hyperactivity and CEN hypoactivity (which were related to behavioural BADE evidence integration and impaired fMRI performance) is also associated with decreased immediate memory, visual scanning, and fluency, and increased working memory and mental manipulation.

3.2.3.1.2.3. Symptoms

Table 20 and Table 21 show the variance table and rotated component loadings (dependent variables) as well as the predictor loadings (independent variables) for each

solution for the behavioural CPCA on symptoms and brain activity underlying dis/confirmatory evidence integration. The overall and residual solutions mirrored those of the behavioural BADE/fMRI performance analysis, and the predicted solution also followed a similar pattern, with Component 1 including VsAN and VDMN confirm (0.22 and -0.34) and disconfirm (0.22 and -0.35), and Component 2 including CEN confirm (0.21) and disconfirm (0.34). As in the behavioural BADE/fMRI performance analysis, the CEN disconfirm loading was stronger than the confirm loading for Component 2 (0.34 vs. 0.21, respectively); however, the VsAN confirm and disconfirm loadings were the same on Component 1 (0.22 and 0.22, respectively). Interestingly, there was a clear dissociation between symptoms and each of the components, with Component 1 showing positive associations with several negative symptoms (e.g., avolition-apathy, alogia; predictor loadings ranging from 0.22 to 0.52), indicating that greater negative symptoms were associated with increased VsAN activity, and Component 2 showing negative associations with positive symptoms (e.g., hallucinations, delusions, formal thought disorder; predictor loadings ranging from -0.14 to -0.62), indicating that greater positive symptoms were associated with decreased CEN activity.

3.2.3.1.2.4. Summary of Associations Between Dis/Confirmatory Evidence Integration HDR and Behaviour

The multivariate behavioural CPCAs on brain activity underlying dis/confirmatory evidence integration clarified and extended the correlational findings. In terms of BADE, we found that poorer behavioural BADE evidence integration (and to a lesser extent conservatism) was associated with hyperactivity in the VsAN (especially during the confirm conditions) and hypoactivity in the CEN (especially during the disconfirm conditions). This

indicates that hyperactivity in a visual attention network (observed in schizophrenia patients

with delusions) during *confirmatory* evidence integration, as well as hypoactivity in a

cognitive evaluation network (also observed in schizophrenia patients with delusions) during

disconfirmatory evidence integration, is associated with poorer evidence integration

behaviourally. These alterations in functional brain activity were also associated with poorer

performance on the fMRI evidence integration task, in terms of both decreased accuracy and

longer reaction times. The neurocognition CPCA followed a similar pattern, showing that

VsAN hyperactivity during confirmatory evidence integration was associated with greater

attention on the RBANS, and CEN hypoactivity (especially during disconfirmatory evidence

integration) was associated with poorer memory. Finally, the behavioural CPCA on

symptoms demonstrated that greater positive symptoms (including delusions) were

associated with hypoactivity in the CEN (more so during disconfirmatory evidence

integration) and hyperactivity in the VsAN (more so during confirmatory evidence

integration), mirroring the behavioural BADE/fMRI performance findings. In contrast,

negative symptoms, which were more strongly related to VsAN and VDMN than CEN, were

equally predictive of confirmatory and disconfirmatory evidence integration.

3.2.3.2. Baseline-Peak Activations

Computation of BP and PB values was determined by inspection of the HDR in each

network averaged over groups and conditions. This involved averaging over poststimulus

time bins that reflected the portion of the HDR shape from baseline (first poststimulus time

bin) to the peak of the response for BP, and from the time bin following the peak to the return

to baseline (final poststimulus time bin) for PB. The BP and PB values for each network

were as follows (see Figures 12C, 14B, and 16C): VsAN BP = 1s – 11s, PB = 13s – 19s;

VDMN BP = 1s – 11s, PB = 13s – 19s; CEN BP = 1s – 15s, PB = 17s – 19s. The resulting

six values (VsAN BP, VsAN PB, VDMN BP, VDMN PB, CEN BP, and CEN PB) were (1)

intercorrelated to examine within- and between-network associations, (2) correlated with

behavioural BADE, fMRI performance, neurocognition, and symptoms, to determine

associations with behaviour; and (3) submitted to three behavioural CPCAs (1: behavioural

BADE/overall fMRI performance, 2: neurocognition, 3: symptoms) to examine these

associations in more depth.

3.2.3.2.1. Intercorrelations

Intercorrelations between BP and PB values within and between networks (Table 22)

revealed a significant association between VsAN BP and VsAN PB, $r(98) = .20$, $p < .05$, a

predictable finding showing that higher activity from baseline to peak is associated with

higher activity from peak to baseline in this network. Between networks, we found

significant negative correlations between the VsAN and VDMN BP, $r(98) = -0.75$, $p < .001$,

and PB, $r(98) = -0.62$, $p < .001$, values, but not when examining BP versus PB across

different networks, with the exception of CEN BP and VsAN PB, for which there was a

negative correlation, $r(98) = -0.37$, $p < .001$. These findings support the dis/confirmatory

evidence integration behavioural CPCAs, which indicated that individuals who strongly

activate the VsAN weakly deactivate the VDMN and vice-versa. The negative association

between the VsAN PB and CEN BP is distinct, in that it involves different aspects of the

HDR, and suggests that as the VsAN decreases from peak to baseline the CEN increases

from baseline to peak. This can be observed averaged over subjects by comparing Figures

12C (VsAN) and 16C (CEN) from approximately 10s-15s poststimulus.

3.2.3.2.2. Correlations with Behaviour

In terms of correlations with behaviour (Table 23), VsAN BP was positively associated with SAPS total, $r(57) = 0.28$, $p < .05$, SANS total, $r(49) = 0.34$, $p < .05$, and SANS affective flattening, $r(58) = 0.35$, $p < .05$. In contrast, no significant correlations emerged with VsAN PB (all ps $> .08$). For VDMN, BP values were negatively associated with SANS total, $r(58) = -0.30$, $p < .05$, and SANS alogia, $r(98) = -0.29$, $p < .05$, and VDMN PB was positively associated with fMRI accuracy during the disconfirm conditions, $r(98) = 0.22$, $p < .05$. CEN BP was positively associated with fMRI accuracy for the disconfirm conditions only, $r(98) = 0.22$, $p < .05$, as well as negatively associated with behavioural BADE evidence integration, $r(96) = -0.22$, $p < .05$, overall fMRI RT, $r(98) = -0.21$, $p < .05$, and fMRI RT for both the confirm, $r(98) = -0.21$, $p < .05$, and disconfirm, $r(98) = -0.21$, $p < .05$, conditions. Finally, CEN PB was negatively associated with SANS anhedonia-asociality, $r(98) = -0.34$, $p < .05$, and positively associated with TMT A, $r(58) = 0.26$, $p < .05$. These results suggest that hyperactivity during VsAN BP primarily underlies the previous findings related to symptoms in patients, and that hypoactivity during CEN BP drives the behavioural BADE/fMRI performance findings.

3.2.3.2.3. Behavioural CPCAs

In order to further investigate associations between BP and PB HDR and behaviour, we performed three behavioural CPCAs with the BP/PB values as the dependent variables, and behavioural measures (1: behavioural BADE and overall fMRI performance, 2: neurocognitive variables, and 3: symptoms) as the predictor variables. As with the behavioural CPCAs on activity underlying dis/confirmatory evidence integration, this multivariate analysis allows for the identification of combinations of functional brain

networks (and BP/PB aspects) that are best predicted by our behavioural measures of interest that may not be observed through correlations alone. In addition, comparison of the following results with the behavioural CPCAs above (Tables 17, 19, and 21) can provide a more in-depth understanding of the aspects of the HDR that drive the previous findings on brain activity underlying dis/confirmatory evidence integration. Scree plots (Cattell, 1966; Cattell & Vogelmann, 1977) were inspected to determine the number of components to extract for each solution and each analysis separately. Two components were extracted for the overall (i.e., overall variance in average estimated HDR prior to regressing onto behavioural measures), predicted (i.e., variance predictable from behavioural measures), and residual (i.e., variance not predictable from behavioural measures) solutions for all analyses.

3.2.3.2.3.1. Behavioural BADE/fMRI Performance

For the behavioural CPCA on behavioural BADE/fMRI performance and baseline-peak HDR, four components were extracted for the overall and residual solutions, and two were extracted for the predicted solution. Table 24 shows the distributions of variance for each component and solution, and Table 25 shows the rotated component loadings (dependent variables) as well as the predictor loadings (independent variables) for each solution. The overall and residual solutions had similar loadings across analyses: Component 1 was dominated by VsAN BP and VDMN BP (reversed), Component 2 by VsAN PB and VDMN PB (reversed), Component 3 by CEN BP, and Component 4 by CEN PB. These results support the intercorrelations reported above, which showed the strongest associations between VsAN and VDMN BP and PB values separately, and little association with CEN values. The second highest loading on Component 3 (dominated by CEN BP) was VsAN PB, which is also in line with the intercorrelations reported above.

In terms of the predicted solution, the first component captured VsAN PB (0.16) VDMN PB (-0.24), and CEN BP (-0.27), while the second component included VsAN BP (0.10) VDMN BP (-0.13), and CEN PB (-0.23), suggesting distinct associations between CEN BP/PB values and VsAN/VDMN PB/BP values. The predictor loadings supported this interpretation, as Component 1 was more strongly related to behavioural BADE evidence integration and conservatism, while Component 2 was more strongly related to overall fMRI accuracy and RT. Specifically, both behavioural BADE evidence integration (0.88) and conservatism (0.89) were positively associated with Component 1. In contrast, fMRI accuracy showed a negative association with both components (-0.48 and -0.53 for Components 1 and 2, respectively), and RT was positively associated to Component 2 only (0.85). Overall, this suggests that behavioural BADE components are better predictors of VsAN PB, VDMN PB, and CEN BP, and that fMRI performance measures are better predictors of VsAN BP, VDMN BP, and CEN PB. This clarifies the previous behavioural BADE/fMRI performance CPCA on brain activity underlying dis/confirmatory evidence integration. For example, the decreased CEN activity (especially during disconfirmatory evidence integration) that was related to poorer fMRI accuracy and behavioural evidence integration is primarily associated with CEN BP. In contrast, increased VsAN activity (especially during confirmatory evidence integration) that was associated with impaired evidence integration and reduced fMRI accuracy in the previous analysis appears to be driven by VsAN PB.

3.2.3.2.3.2. Neurocognition

Tables 26 and 27 show the variance distribution and rotated component loadings (dependent variables) as well as the predictor loadings (independent variables) for the CPCA

on neurocognitive measures and baseline-peak HDR. For the predicted solution, Component 1 consisted of VsAN BP (0.41), VsAN PB (0.27), and VDMN BP (-0.30); Component 2 included VDMN PB (-0.32), CEN BP (0.23) and CEN PB (-0.23). Thus, the predicted solution distinguished between brain networks more than BP/PB values, except that VDMN BP and PB were split across Components 1 and 2, respectively. Component 1 was negatively related to several neurocognitive measures, including RBANS language (-0.49), TMT A (-0.57) and B (-0.33), and COWAT (-0.40). Component 2 was negatively related to RBANS attention (-0.59), TMT A (-0.63), TMT B (-0.37) and RBANS immediate memory (-0.26). These findings differ somewhat from the CPCA on activity during dis/confirmatory evidence integration and neurocognition, suggesting that neurocognitive measures predict functional brain activity underlying BP/PB values differently from that underlying dis/confirmatory evidence integration.

3.2.3.2.3.3. Symptoms

The variance table and rotated component loadings (dependent variables) as well as the predictor loadings (independent variables) for the CPCA on symptoms and baseline-peak HDR are displayed in Tables 28 and 29. Component 1 of the predicted solution included VsAN BP (0.39) and VDMN BP (-0.38), whereas Component 2 included CEN BP (-0.27) and CEN PB (0.31), as well as VDMN PB (0.31). Positive symptoms (e.g., hallucinations, delusions, formal thought disorder; predictor loadings ranging from 0.19 to 0.48) were positively associated with Component 2, meaning that higher positive symptoms were associated with decreased activity in CEN BP (which was found to underlie the BADE findings in the behavioural BADE/fMRI performance analysis; Table 17) and increased intensity in VDMN BP and CEN PB. Component 2 was also positively associated with

SANS inattentiveness (0.32) and negatively associated with SANS avolition-apathy (-0.40),

while the SANS affective flattening (0.44) and alogia (0.41) were more strongly related to

Component 1. These findings support the behavioural CPCA on brain activity underlying

dis/confirmatory evidence integration and symptoms described above (Table 21), in which

CEN disconfirm was associated with positive symptoms, and VsAN confirm/disconfirm was

associated with negative symptoms. However, in combination with the behavioural

BADE/fMRI performance CPCA on BP/PB values, these results further demonstrate that

hypoactivity in the BP aspect of the CEN underlies the results related to both impaired

evidence integration and increased positive symptoms.

3.2.3.2.3.4. Summary of Associations Between Baseline-Peak HDR and Behaviour

The behavioural CPCAs on baseline-to-peak and peak-to-baseline activity revealed that

the previous findings (hyperactivity in VsAN/hypoactivity in CEN being related to poorer

behavioural BADE evidence integration and fMRI performance, and greater positive

symptoms) were associated with distinct aspects of the HDR in these networks. Specifically,

poorer behavioural BADE evidence integration was associated with hyperactivity during

VsAN PB and hypoactivity in CEN BP (as well as VDMN PB). This indicates that the

behavioural BADE findings are driven by the rise from baseline to peak in the CEN (early

HDR) and the fall from peak to baseline in the VsAN and VDMN (late HDR). The negative

correlation observed between CEN BP and VsAN PB suggests that hyperactivity in the

VsAN may lead to hypoactivity in the CEN, though this should be tested further with

confirmatory or effective connectivity techniques. In terms of performance on the fMRI task,

accuracy was more strongly predicted by PB values, and RTs more strongly predicted by BP

values, especially for the VsAN/VDMN. In addition, neurocognitive measures were better predictors of networks rather than BP/PB values, supporting our interpretation that behavioural BADE and neurocognition are separable in terms of their relation to functional brain activity underlying evidence integration. We also found that attention measures were better predictors of the VsAN, and memory measures of the CEN. Finally, in terms of symptoms, greater positive symptoms (including delusions) were stronger predictors of the CEN than the other two networks, but showed differential associations with CEN BP (greater symptoms associated with hypoactivity) and CEN PB (greater symptoms associated with hyperactivity). This latter finding could indicate hyperactivity in the CEN at peak in schizophrenia patients with delusions, but given the group differences were observed earlier in the trial, another interpretation is that the fall from baseline to peak occurred more slowly in patients with greater positive symptoms.

3.3. Discussion

The purpose of this study was to examine the cognitive and brain mechanisms underlying BADE and its relation to delusions and functional brain activity involved in evidence integration in schizophrenia. Behaviourally, we replicated previous findings (Buchy et al., 2007; Sanford et al., 2014; Speechley et al., 2012; Woodward et al., 2007; Woodward et al., 2006b) showing impaired evidence integration in schizophrenia patients with delusions and in healthy individuals with high levels of delusional ideation, and replicated previous research (Woodward et al., 2007) showing that BADE is separable from neurocognitive functioning, extending this latter finding to a schizophrenia patient sample. Three functional brain networks were associated with evidence integration: VsAN (occipital cortex, supramarginal gyrus, middle frontal gyrus, ACC, insula), VDMN (deactivations in

ventromedial PFC, precuneus, posterior cingulate cortex, and activations in occipital cortex),

and CEN (rostrolateral and orbitofrontal cortex, inferior frontal gyrus, inferior parietal

lobule). Although all networks were preferentially recruited during disconfirmatory evidence

integration and showed differences between delusional patients and the other groups, only

the VsAN and CEN were associated with behavioural BADE evidence integration

(preferentially during confirmatory and disconfirmatory evidence integration, respectively).

These results suggest that the CEN underlies the evaluation and integration of evidence and

that hypoactivity in this network during disconfirmatory evidence integration underlies the

bias against disconfirmatory evidence observed in schizophrenia patients with delusions. In

addition, hyperactivity during confirmatory evidence integration in the VsAN in delusional

patients may reflect the focus on confirmatory evidence theorized to co-occur with BADE in

delusions (Broyd et al., 2017; Freeman et al., 2002). These findings highlight functional brain

networks associated with two important cognitive processes underlying delusion

maintenance in schizophrenia: the focus on confirmatory evidence and avoidance of

disconfirmatory evidence.

3.3.1. Network Comparisons Between Studies 1 and 2

3.3.1.1. Visual Attention Network

The VsAN that emerged in study 2 (Component 1 Figure 11) most closely resembled the

VsAN identified in study 1 (Component 5 Figure 5A), with partially overlapping nodes in the

dACC and bilateral insula, as well as activations in bilateral occipital cortex, dorsolateral

prefrontal cortex, and posterior parietal cortex. The latter two regions are part of the dorsal

attention network described in previous resting-state research (Yeo et al., 2011), and showed

more extensive recruitment in the VsAN of the current study. In comparison to the VsAN

identified in study 1, the current network included additional activations in somatosensory regions involved in response processes, such as bilateral thalamus, left motor cortex, and cerebellum. These regions correspond to the response network that emerged in study 1 (Component 4 Figure 4A). Thus, the VsAN was similar to that of study 1, but included regions involved in other cognitive processes (e.g., attention, task demand, responding) occurring at the same time as evidence detection. This is apparent when examining the HDR shape associated with this component, which shows an earlier peak relative to study 1 (9s vs. 15s), and is sustained throughout the trial.

3.3.1.2. Visual/Default-Mode Network

The VDMN identified in study 2 (Component 2 Figure 13) was similar to that which emerged in study 1 (Component 2 Figure 3A), including deactivations in ventromedial prefrontal cortex, precuneus, posterior cingulate cortex, and bilateral lateral temporal cortex, as well as activations in occipital cortex, and posterior parietal cortex. The deactivations on this network were more extensive than study 1, in which DMN deactivations were spread across several components. This may be the result of a smaller number of components extracted in the current study, or may be due to differences in task timing across the studies, such that the DMN deactivations were not as temporally coordinated with the other networks. This latter interpretation is supported by the less extensive occipital activations on the VDMN in study 2 relative to study 1, as well as the different HDR shape for this network relative to the two other networks identified (VsAN and CEN), as it peaked midway through the trial (instead of early in the trial like the VsAN or later in the trial like the CEN) and was not as sustained as the VsAN.

94

3.3.1.3. Cognitive Evaluation Network

The CEN in study 2 (Component 3 Figure 15) was highly similar to the network of the same name in study 1 (Component 1 Figure 2A), with activations in rostrolateral/orbitofrontal cortex, dorsolateral prefrontal cortex, inferior frontal gyrus, and inferior parietal lobule. This is striking given the use of different samples and different evidence integration tasks across the two studies. However, these regions were more extensively and more bilaterally recruited in study 2, suggesting greater recruitment of evaluation and integration processes with the novel task. There were also more extensive activations in left middle temporal gyrus (subthreshold in study 1), suggesting greater involvement of language and semantic processing (McDermott, Petersen, Watson, & Ojemann, 2003; Woodward et al., 2015) not captured as strongly on this network in study 1. Interestingly, unlike study 1, the configuration of the CEN in the current study also included deactivations in somatosensory regions (i.e., SMA, left motor cortex), which, given its late peak, may reflect a suppression of response processes during the presentation of the third and final image, which was not present in study 1. The HDR shape of this network also closely resembled that of the previous study, with a late peak at the end of the trial after visual attention and response activations as well as greater activity during disconfirmatory relative to confirmatory evidence integration.

3.3.2. Behavioural BADE and Delusions

In line with previous behavioural studies, delusional schizophrenia patients demonstrated an unwillingness to integrate evidence relative to non-delusional patients and healthy controls. Previous research has shown that BADE is a cognitive bias specific to delusions in schizophrenia and not a result of general psychiatric illness (for a review, see Broyd et al.,

2017), in that it is observed in delusional patients relative to individuals with other psychiatric diagnoses, such as non-delusional bipolar disorder (Speechley et al., 2012), obsessive-compulsive disorder (Sanford et al., 2014), and post-traumatic stress disorder (Moritz & Woodward, 2006). In combination with ours and others' (Buchy et al., 2007; Woodward et al., 2007) research finding that behavioural BADE evidence integration is also increased in healthy individuals with high levels of delusional ideation, these results confirm that poor evidence integration is specific to delusional thinking across the schizophrenia spectrum. In line with this interpretation, there is evidence that BADE gradually increases from at-risk to first-episode, and has been suggested as a cognitive marker for delusions (Eisenacher et al., 2016; Eisenacher & Zink, 2017).

In examining potential brain mechanisms underlying behavioural BADE evidence integration in delusions, we found differential associations between VsAN/CEN activity and behavioural BADE evidence integration/delusions in schizophrenia. Specifically, delusional patients showed hyperactivity in the VsAN (Figure 12), which was associated with poorer behavioural BADE evidence integration during confirmatory evidence integration on the fMRI task (Table 17). In contrast, the CEN was found to be hypoactive in delusional patients relative to controls (Figure 16), and this decreased activity was associated with poorer behavioural BADE evidence integration, particularly during disconfirmatory evidence integration on the fMRI task (Tables 15 and 17). The behavioural CPCA of behavioural BADE and fMRI performance on activity underlying confirmatory and disconfirmatory evidence integration (Table 17) also showed that hyperactivity in the VsAN during confirmatory evidence integration, and hypoactivity in the CEN during disconfirmatory evidence integration was associated with poorer performance on the fMRI task in terms of

decreased accuracy and increased reaction time. This suggests that aberrant activity in the VsAN and CEN represent distinct aspects of evidence integration that are differentially related to both behavioural BADE evidence integration and delusions in schizophrenia, and contribute to impaired performance on an fMRI evidence integration task.

First, the VsAN is involved in attending to salient stimuli, including the initial image and evidence presented. This is apparent in this network's early and sustained peak (~ 9-13 seconds; Table 15), which covers the majority of the trial, and is also supported by the heightened activity in this network during disconfirmatory evidence integration later in the trial, due to the increased salience of evidence that contradicts a belief. This interpretation is in line with our previous study (Chapter 2; Lavigne et al., 2015a), even though the current VsAN includes additional response and task-positive regions. The importance of the VsAN and its comprising regions to salience detection has been reported in ROI-based studies on evidence integration (e.g., Liu & Pleskac, 2011; Ploran et al., 2007), as well as network-based studies on the function of the ventral attention (Corbetta & Shulman, 2002) and salience (Seeley et al., 2007; Uddin, 2015) networks, which share important nodes with the VsAN identified here (viz., dACC, bilateral insula). In addition, we found that this network was hyperactive in delusional schizophrenia patients (Figure 12)[2], and that hyperactivity in the VsAN during confirmatory evidence integration was associated with poorer behavioural

---

[2] Delusional patients' hyperactivity in this network, which peaked prior to evidence presentation, was contrary to our hypothesis of decreased activity in a VsAN that peaked following evidence presentation. This difference in peak timing likely underlies this unexpected finding. Specifically, the VsAN begins to rise early in the trial (3s) and peaks at 9s, only one second following evidence presentation in real time. Given the delay observed between a stimulus and BOLD signal activation, and the sustained activity in this network across the entire trial, the current configuration of this network likely does not (only) reflect the detection of evidence as discussed in our previous study (Lavigne et al., 2015a; also see Chapter 2). The more extensive recruitment of attention- and response-related regions on this network suggests a complex combination of additional cognitive processes (e.g., responding, attention, task demand) being represented on this network that show similar timing patterns to, and are grouped with, evidence detection by the PCA.

BADE evidence integration (Table 17). Hyperactivity in this visual attention network during confirmatory evidence integration may reflect a focus on evidence that confirms a belief, which corresponds to a hypersalience of evidence-hypothesis (EVH) matches, a cognitive bias in which hypotheses with supporting evidence are seen as more plausible than those for which supporting evidence is not provided (Speechley, Whitman, & Woodward, 2010; Whitman, Menon, Kuo, & Woodward, 2012). Behaviourally, hypersalience of EVH matches has been associated with delusion severity in schizophrenia (Speechley et al., 2010; Whitman et al., 2012), and has been suggested to underlie both JTC (Speechley et al., 2010) and BADE (Broyd et al., 2017; Sanford et al., 2014; Speechley et al., 2012). Moreover, previous research on functional brain activity underlying hypersalience of EVH matches in healthy individuals identified a brain network highly similar to the VsAN found in the current study, and demonstrated that this network is hyperactive when accepting a hypothesis with an evidence match relative to an evidence mismatch (Whitman et al., 2013). These findings highlight a cognitive bias and corresponding brain network underlying the focus on confirmatory evidence that sustains delusion maintenance in schizophrenia (Freeman et al., 2002), and show that this hypersalience is associated with poorer evidence integration.

Following salience detection in the VsAN, activity in the CEN serves to evaluate the evidence and integrate it into the belief system. This interpretation is supported by the network's late-peaking HDR (approximately 15 seconds after the start of the trial; see Figure 16), which shows that activity occurs following the presentation of evidence and lasts until the end of the trial. Moreover, this network is more strongly recruited for disconfirmatory relative to confirmatory evidence (Figure 16), which would be expected for a network involved in evaluation and integration processes. These findings are in line with our previous

study (Chapter 2; Lavigne et al., 2015a), despite a different sample including schizophrenia patients and the use of a novel evidence integration task. In addition, the evaluating and integrating functions of the CEN (and regions comprising the CEN) are supported by both ROI-based research on evidence integration (Corlett et al., 2004; d'Acremont et al., 2013; Sharot et al., 2012) and network-based studies describing the functions of the FPN (Dosenbach et al., 2007; Vincent et al., 2008), which shares important nodes with the CEN (e.g., rPFC, IFG). Our novel findings that schizophrenia patients show hypoactivity in the CEN (Figure 16) during an evidence integration task, and that hypoactivity in this network (particularly during disconfirmatory evidence integration) is associated with poorer behavioural BADE evidence integration (Tables 15 and 17), suggests that hypoactivity in the CEN captures the bias against disconfirmatory evidence observed in schizophrenia patients with delusions.

Taken together, these results suggest that, in delusional schizophrenia patients, there is a focus on confirmatory evidence (hypersalience of EVH matches; hyperactivity in the VsAN), which is associated with poorer evidence integration (BADE; hypoactivity in the CEN). A hypersalience of EVH matches might result in the initial hypothesis/evidence (i.e., first picture/response in the fMRI task, and first sentence/lure interpretations in the behavioural BADE task) becoming hypersalient, hindering the integration of new (disconfirmatory) evidence. This corresponds to theoretical accounts of delusion maintenance in schizophrenia, which suggest important roles for both hypersalience of confirmatory evidence and avoidance of disconfirmatory evidence (Broyd et al., 2017; Freeman et al., 2002). Thus, this research shows, for the first time, differential activation of two functional brain networks (VsAN hyperactivity, CEN hypoactivity) in delusional schizophrenia patients that underlies

two cognitive biases related to delusions in schizophrenia (hypersalience of EVH matches, BADE), and corresponds to important processes involved in delusional maintenance in schizophrenia (focus on confirmatory evidence, avoidance of disconfirmatory evidence).

3.3.3. Baseline-Peak Activations

Our investigation of associations between baseline-peak aspects of the hemodynamic response revealed intercorrelations within networks, as well as a negative association between the VsAN and VDMN, indicating that individuals who hyperactivate the VsAN tend to recruit the VDMN less. However, we also identified a significant negative association between CEN BP and VsAN PB (-0.37, see Table 22), the only cross-network correlation involving different aspects of the hemodynamic response that reached significance. This suggests a reciprocal relationship between the VsAN and CEN, particularly, a suppression of CEN below baseline when evaluative and integrative processes are less relevant, and a decrease in the VsAN back to baseline as the CEN is recruited for evidence integration processes. Both the correlations and the behavioural CPCAs on BP/PB values and behaviour found that CEN BP drove the associations between CEN activity during disconfirmatory evidence integration and delusions, behavioural BADE, and fMRI performance. Given the reciprocal relationship between VsAN PB and CEN BP, these findings support our interpretation that a hypersalience of EVH matches to the first image/interpretation may lead to decreased activity in the CEN from baseline-to-peak, and hinder the evaluation and integration of subsequent (disconfirmatory) evidence. Future research should test this hypothesis using effective connectivity techniques, such as dynamic causal modeling.

3.3.4. Visual/Default-Mode Network

While the VDMN also showed increased (de)activity during disconfirmatory evidence integration as well as differences between delusional patients and healthy controls (Figure 14), it was not correlated with behavioural BADE evidence integration (Tables 15 and 23), although there was some evidence of association from the behavioural CPCAs. Previous research on the DMN and schizophrenia has consistently reported DMN hyperactivity (or task-related hypodeactivation) and its relation to both impaired task performance, cognitive deficits, and positive symptoms (Buckner et al., 2008; Garrity et al., 2007; Whitfield-Gabrieli & Ford, 2012; Zhou et al., 2016), as was observed in the current study. Given the DMN's involvement in stimulus-independent thought and self-referential processing (Buckner et al., 2008), task-related hypodeactivation could reflect an inability to stay engaged with the task, which would explain its associations with impaired performance. While patients' hyperactivity in the VsAN in the current study might seem contradictory to these findings, it was associated with greater inattentiveness, and might represent a focus on early stimuli (i.e., first picture) throughout the trial when focus should be on the subsequent images. Another interpretation of DMN task-related hypodeactivation in schizophrenia addressed in previous work (Buckner et al., 2008) is that in may indicate an impaired ability to distinguish between self and other or internal and external representations, which has been associated with positive symptoms. This is in line with the patient group differences identified in the current study and supports our interpretation that schizophrenia patients with delusions show impairments when comparing internal representations (beliefs) with external experiences (evidence). However, given the sparse associations between activity in the VDMN and behavioural BADE in the current study, and that DMN hyperactivity (and task-related

hypodeactivity) has been observed across a variety of tasks and in the resting state in schizophrenia (Whitfield-Gabrieli & Ford, 2012), aberrant DMN activity in schizophrenia is not specific to BADE and evidence integration processes, and likely reflects a more generalized impairment. Future studies are necessary to determine whether the observed VDMN differences are separable from the VsAN and CEN findings related to BADE and evidence integration.

3.3.5. BADE and Neurocognition

Our research supports others' (e.g., Moritz et al., 2010; Woodward et al., 2007) distinguishing BADE from neurocognitive functioning. First, we found that increased BADE evidence integration in delusional patients remained after controlling for related IQ (as well as depressive symptoms). Second, we replicated previous findings (Woodward et al., 2007) showing that BADE forms a component distinct from neurocognition, using a more extensive set of neurocognitive measures, and including the more recently defined behavioural BADE components (evidence integration and conservatism) rather than the items. However, there were some significant correlations between fMRI activity and neurocognitive variables, and the behavioural CPCAs showed that neurocognitive measures were associated with similar components as behavioural BADE (more so for brain activity underlying dis/confirmatory evidence integration than BP/PB). Overall, these findings suggest separable roles for behavioural BADE and neurocognitive functions not only behaviourally, but with regard to their influence on functional brain activity underlying an evidence integration task. However, the similarity between the predicted solutions across the behavioural CPCAs indicates some degree of overlap in terms of their impact on functional brain activity, and is in line with

interpretations of previous research that neurocognitive abilities partially contribute to behavioural BADE performance (Eisenacher & Zink, 2017).

3.3.6. Limitations

Although this study included a large number of schizophrenia patients, which allowed us to divide them into adequately-sized delusional and non-delusional groups, we could not sub-divide them further based on additional clinical characteristics. For example, the delusional group differed from the non-delusional group on not only delusions, but also on other positive symptoms, such as hallucinations and bizarre behaviour (which showed some associations with brain activity in the behavioural CPCAs), as well as depressive symptoms, though the latter did not appear to affect the behavioural BADE results. Moreover, it was not possible to examine the potential influence of type of delusion (e.g., persecutory, referential), which may affect evidence integration processes. Future research would also benefit from investigating other clinical characteristics, such as medication type and dosage, and illness onset and duration, as these have known effects on BADE (Andreou et al., 2015; Eisenacher & Zink, 2017). With regard to the functional brain activations, the VsAN identified in the current study, though similar to that of our previous study, included additional activations from task demand and response-related regions, likely due to these processes occurring simultaneously. Careful consideration of task design and inclusion of additional neuroimaging techniques with high temporal resolution (e.g., electroencephalography) would help distinguish between brain networks related to detection and integration of evidence and those related to responding or task demand.

3.3.7. Conclusion

In this study, we identified two functional brain networks differentially related to evidence integration and delusions in schizophrenia. We found hyperactivity in delusional schizophrenia patients in a visual attention network, including nodes of the ventral attention and salience networks identified in resting-state studies, that was associated with poorer behavioural BADE evidence integration during the integration of confirmatory evidence on an fMRI task. In contrast, hypoactivity was observed in delusional patients in a cognitive evaluation network (overlapping with nodes of the frontoparietal network from resting-state research, and regions implicated in evidence integration processes in ROI-based studies), and was associated with poorer behavioural BADE evidence integration during the integration of disconfirmatory evidence on an fMRI task. These findings highlight distinct brain mechanisms underlying two cognitive biases associated with delusions (hypersalience of EVH matches and behavioural BADE), which reflect two theorized processes fundamental to the maintenance of delusions: the focus on confirmatory evidence and avoidance of disconfirmatory evidence, respectively. Future research would benefit from focusing on more severely delusional patients, and examining these effects in antipsychotic-naïve first-episode patients. Finally, longitudinal research should be carried out to determine the degree to which the identified networks fluctuate as a function of BADE and/or delusions.

4. Changes in BADE, Delusions, and Functional Brain Activity Underlying Evidence Integration in Schizophrenia (Study 3)

In the previous two chapters, we hypothesized that two functional brain networks were instrumental in processing disconfirmatory evidence: a VsAN involved in the detection of evidence; and a CEN involved in the integration of evidence. We further demonstrated that these networks were differentially related to the behaviourally measured BADE and to delusions in schizophrenia, suggesting that hyperactivity in the VsAN in delusional patients underlies the focus on confirmatory evidence, and hypoactivity in the CEN the avoidance of disconfirmatory evidence, theorized to contribute to delusion maintenance in schizophrenia (Broyd et al., 2017; Freeman et al., 2002). Moreover, we provided evidence that both behavioural BADE and functional brain activity within these networks were separable from neurocognitive processes, such as memory and attention.

An additional feature of a generalized cognitive bias proposed to influence delusions is that it would be expected to fluctuate with changes in delusion severity. Although delusions are often described as fixed (on a hypothesis about reality), delusion severity and conviction tend to wax and wane over the course of the illness (Coltheart et al., 2011). Cross-sectional research also suggests that behavioural BADE worsens with increasing delusion severity during the early stages of psychosis (Eisenacher & Zink, 2017), which provides preliminary evidence for a relationship. Moreover, both behavioural BADE (Buonocore et al., 2015) and delusion severity (Moritz et al., 2011) improve following treatment with cognitive interventions; however, it is not clear whether changes in delusions are associated with changes in BADE (Buonocore et al., 2015), and the effects of changes in delusion severity

and/or BADE on functional brain networks underlying evidence integration have yet to be investigated.

The current study assessed associations between changes in delusions, clinical, neurocognitive, and behavioural BADE measures in schizophrenia. A subsample of patients also underwent fMRI in order to examine whether longitudinal changes in functional brain activity during an evidence integration task are associated with changes in behavioural BADE and delusions in schizophrenia. Schizophrenia patients were assessed before and after a minimum 8-week delay as part of a larger randomized controlled trial. We hypothesized that changes in delusions would be associated with changes in both behavioural BADE and functional brain activity; specifically, we expected that a decrease in delusion severity from time 1 to time 2 would be positively related to decreased BADE and normalization of the early-peaking VsAN during confirmatory evidence integration (e.g., decreased activity at time 2 relative to time 1), as well as normalization of the late-peaking CEN during disconfirmatory evidence integration (e.g., increased activity at time 2 relative to time 1).

4.1. Method

4.1.1. Participants

Seventy-two patients with a diagnosis involving psychosis participated in this study, which was part of a larger randomized controlled trial. As part of the larger study, in addition to treatment with medications, patients also received one of two cognitive interventions (cognitive remediation therapy, n = 22; or metacognitive training, n = 24) or treatment as usual (medication optimization, recreational/occupational therapy, psychoeducation, n = 26). Thirty-five patients overlapped with those who completed study 2, and 29 of those patients completed fMRI at both time 1 and time 2. The delay between time 1 and time 2 assessment

sessions was at least 8 weeks (full sample: mean = 13.65 weeks, SD = 3.58; range 8-23

weeks; fMRI subsample: mean = 14.57 weeks; SD = 3.23; range = 8-23 weeks). Exclusion

criteria were identical to those described in the previous chapter: neurological infection; loss

of consciousness; IQ less than 80 on the WASI (Wechsler, 2011); substance/alcohol

dependence; major medical illness; and presence of severe thought disorder based on the

SAPS (Andreasen, 1984b). Participants took part in exchange for $10/hour, reimbursement

of travel expenses, and a copy of their structural brain image. This study was approved by the

UBC Clinical Research Ethics Board and registered as a Clinical Trial with the

ClinicalTrials.gov Protocol Registration System (NCT01764568).

4.1.1.1. Diagnosis

Diagnoses were confirmed using the MINI (Sheehan et al., 1998). In cases where the

MINI diagnosis conflicted with the reported diagnosis, the participant's mental health care

worker was consulted. When that was not possible, a consensus was determined by the

clinical psychologists who were part of the research team. Primary psychotic disorders

assessed by the MINI were as follows: schizophrenia (n = 31), schizoaffective disorder (n =

35), psychotic disorder not otherwise specified (n = 4), mood disorder with psychotic

features (n = 1), and delusion disorder (n = 1). Comorbid diagnoses captured by the MINI

included: major depressive episode (6 current, 37 past, 3 missing data); (hypo)manic episode

(23 past, 2 missing data); panic disorder (6 current, 4 past, 3 missing data); panic disorder

with agoraphobia (2 current, 2 missing); panic disorder without agoraphobia (2 current, 2

missing); agoraphobia without history of panic disorder (1 current, 2 missing); social phobia

(4 current, 3 missing data); obsessive-compulsive disorder (4 current, 2 missing data);

alcohol abuse and dependence (3 current); non-alcohol psychoactive substance use disorders

(4 current, 2 missing data); and generalized anxiety disorder (11 current, 3 missing data).

Ratios for primary and comorbid diagnoses were similar in the fMRI subsample.

4.1.1.2. Medications

Medication use was assessed at both time 1 and time 2. All patients were taking antipsychotic medication at both time points. Sixty-two patients were taking an atypical antipsychotic (e.g., olanzapine, risperidone), and 6 patients were taking a typical antipsychotic (e.g., loxapine), as their primary medication (medication information was missing for 4 patients). In addition, 30 patients were taking a second antipsychotic medication (26 atypical, 4 typical), and 3 patients were taking a third medication (2 atypical, 1 typical). Other prescribed medications included: anti-anxiety (n = 18), anti-depressant (n = 33; 2 missing data), anti-convulsant (n = 13), anti-cholinergic (n = 6); and other (e.g., diabetes, blood pressure; n = 27). Chlorpromazine equivalent dosages were available for 58 out of 72 patients. Twenty patients had a change in their antipsychotic medications from time 1 to time 2; however, there were no significant differences between time 1 and time 2 on chlorpromazine dose equivalents ($p > .38$; see Table 32 for means).

4.1.2. Measures

4.1.2.1. Behavioural

At time 1, patients completed the same measures as in study 2 (Chapter 3); data for 35 patients overlapped with those from study 2. Premorbid and current IQ were assessed using the ToPF (Wechsler, 2009) and WASI (Wechsler, 2011), respectively. Depressive symptoms were assessed using BDI (Beck et al., 1961). Psychotic symptom assessments were carried out with the SAPS and SANS (Andreasen, 1984a, 1984b), and the PSYRATS (Haddock et

al., 1999) if hallucinations and/or delusions were endorsed. As in study 2, patients who did not endorse delusions or hallucinations, and were not administered the PSYRATS, were coded as 0 on the relevant subscale(s) in order to reduce missing data. Neurocognitive measures included the RBANS (Randolph et al., 1998), TMT (Reitan & Wolfson, 1985; Reynolds, 2002), LNS (Wechsler, 2014), and COWAT (Benton, 1967). Descriptions of these measures and scoring procedures are listed in the previous chapter. All measures except for the ToPF and WASI were administered at both time 1 and time 2.

Patients also completed the behavioural BADE task (Sanford et al., 2014; Speechley et al., 2012) at both time 1 and time 2. The behavioural BADE task was the same as that used in study 2, which involved rating four interpretations of a story that unfolds over a sequential series of three sentences (see Figure 6). As in study 2, the task included 16 trials or scenarios (12 experimental, 4 filler) and four interpretations (true, neutral lure, emotional lure, absurd) which are to be rated independently of one another after each sentence is presented. The true interpretation becomes increasingly more plausible as more evidence is provided, while the lure interpretations appear plausible at first, but become less likely after each sentence. Finally, the absurd interpretation is not very plausible throughout the trial. The behavioural BADE task was identical in procedure at both time points except that different scenarios were used at time 1 (same as study 2) and time 2. Ratings for each of the four interpretation types (true, both lures, absurd) after each of the three sentences are summed across all trials, leading to twelve scores: true 1, 2 and 3; emotional lure 1, 2, and 3, neutral lure 1, 2, and 3; and absurd 1, 2, and 3.

4.1.2.2. Neuroimaging BADE/Evidence Integration Task

Participants in the fMRI subsample (all of whom overlapped with study 2) also completed the novel evidence integration fMRI task introduced in Chapter 3, with identical experimental parameters except that different images were used for the time 1 (same as study 2) and time 2 versions. Mean overall and condition-specific fMRI accuracy and RTs, as well as overall missed responses were computed as in the previous study, except this was done for each time point separately.

MRI data were acquired at the UBC MRI Research Centre, and scanning protocols were identical to those described in Chapter 3. For four patients (3 at time 1, 1 at time 2), the MRI technician ran the wrong scanning protocol, which led to 310 instead of 325 volumes being collected for the second run. As in the previous study, these patients were not excluded from analysis. Data were preprocessed using SPM8 (Wellcome Trust Centre for Neuroimaging, UK). For each subject and run, functional images were corrected for slice timing acquisition, realigned to the mean image, and co-registered to the participants structural MRI. Structural images were segmented into grey matter, white matter, and cerebrospinal fluid, and corrected for intensity non-uniformity if required. Functional scans were then normalized to the native T1 SPM template (voxel size = $2mm^3$). Subsequently, functional scans were smoothed using a 6mm FWHM Gaussian kernel. Realignment parameters were examined separately for each participant, run, and time point to determine the extent of head motion. Participants or runs with head motion exceeding 2 voxels (4mm translation or 4° rotation for a voxel size of $2mm^3$) were excluded from analysis. This led to the removal of 16 runs (9 at time 1, 7 at time 2), and two participants who showed excessive head motion during both runs at either time point. Excluded participants are taken into account in the sample size listed previously.

4.1.3. Statistical Analysis

4.1.3.1. BADE Component Structure

In order to examine the reliability of the component structure of the behavioural BADE task from time 1 to time 2, we employed structural equation modeling (SEM), specifically, generalized structured component analysis (GSCA; Hwang & Takane, 2004) using the gesca package for R (https://www.r-project.org/; Kim, Cardwell, & Hwang, 2016). GSCA is an SEM technique that uses component analysis (as in PCA) rather than factor analysis, such that latent variables are a weighted combination of observed variables. It provides loadings for each observed variable included in the model as well as an overall measure of model fit. The model used for the current study (see Figure 17) was based on previous PCAs of the behavioural BADE task (Sanford et al., 2014; Speechley et al., 2012; see also Chapter 3) and consisted of two latent variables: evidence integration (degree to which disconfirmatory evidence has been integration; higher scores = worse evidence integration) and conservatism (a reduced willingness to rate high when justified; higher scores = less conservatism). The evidence integration variable was modeled with the following observed variables: absurd 1, absurd 2, absurd 3, neutral lure 2, neutral lure 3, emotional lure 2, emotional lure 3, and true 3. The conservatism latent variable was modeled with the observed variables neutral lure 1, emotional lure 1, true 1, true 2, and true 3. The true 3 item was represented on both latent variables due to its shared loadings (reversed for evidence integration) in previous research.

Comparison of the modeled component structure on time 1 and time 2 data was performed in two steps (as in Woodward *et al.* (2014) for comparing models of the PSYRATS symptom dimensions). The first step involved an unconstrained GSCA, in which the values of the observed variables' loadings on the latent variables are allowed to differ

between the two time points. In this step, the loadings and statistical significance values can be compared across time points. If they are similar and statistically significant, a second GSCA is performed, which is identical to the first model, except that the loadings of the observed variables are constrained across time points. That is, each variable is modeled to have equal loadings for time 1 and time 2. The fit values that emerge from this constrained GSCA can be compared to the those of the unconstrained GSCA using a t-test on a number of bootstrapped samples for each model (100 in the current study) to determine whether there is a difference in fit across the two models. A lack of significant difference indicates that the component structure does not vary from time 1 to time 2 (i.e., structural invariance), which supports using the same composite evidence integration and conservatism scores for both time points.

4.1.3.2. Time 1 Only

Given the overlap between patients who completed study 2 (Chapter 3) and those who took part in the fMRI portion of this study, we also examined behavioural differences at time 1 between patients from study 2 who did and did not complete study 3. This was achieved using independent-samples t-tests on the following variables: age; education; ToPF; WASI; chlorpromazine equivalent dosage; BDI; PSYRATS hallucinations and delusions; SAPS total, hallucinations, delusions, bizarre behaviour, and formal thought disorder; SANS total, affective flattening, alogia, avolition-apathy, anhedonia-asociality, attention, and inappropriate affect; RBANS immediate memory, visuospatial, language, attention, and delayed memory; behavioural BADE evidence integration and conservatism; and fMRI accuracy (overall, confirm, disconfirm), fMRI RT (overall, confirm, disconfirm), and overall

fMRI missed responses. Group differences on gender and handedness were examined using Chi-square.

4.1.3.3. Time 1 Versus Time 2

4.1.3.3.1. Behavioural

Examination of differences from time 1 to time 2 on clinical, neurocognitive, behavioural BADE, and fMRI performance variables was carried out using paired t-tests by pairing time 1 and time 2 scores on the following variables: chlorpromazine equivalent dosage; BDI; PSYRATS hallucinations and delusions; SAPS total, hallucinations, delusions, bizarre behaviour, and formal thought disorder; SANS total, affective flattening, alogia, avolition-apathy, anhedonia-asociality, attention, and inappropriate affect; RBANS immediate memory, visuospatial, language, attention, and delayed memory; behavioural BADE evidence integration and conservatism; and fMRI accuracy (overall, confirm, disconfirm), fMRI RT (overall, confirm, disconfirm), and overall fMRI missed responses. Associations between behaviour and delusion change were carried out as correlations between delusion severity (PSYRATS, SAPS) and the remaining behavioural measures, which were first transformed into change scores by subtracting the time 1 from time 2 values (i.e., time 2-time 1). Change scores were calculated in this manner for all relevant subsequent analyses, such that positive values indicated larger values at time 2 relative to time 1 (i.e., worse symptoms, behavioural BADE, fMRI RT and missed responses, and improved neurocognition and fMRI accuracy).

4.1.3.3.2. fMRI-CPCA

As in study 2, group fMRI-CPCA (Metzak et al., 2011; Woodward et al., 2013) was used
to identify task-based functional brain networks related to stimulus timing. Group fMRI-
CPCA was conducted in the same way as in study 2, with all subjects' scans (including both
time points) combined together to identify common functional brain networks that could then
be compared across time points and conditions over poststimulus time to examine differences
in the magnitude and timing of activations within each network. Details regarding the
mathematical calculations underlying CPCA are described in Chapter 2. Briefly,
preprocessed fMRI BOLD data for all subjects, scans and runs for both time points were
concatenated into a 34730 (29 subjects × up to 325 scans × up to 2 runs × 2 time points) ×
267763 (585390 voxels in 2mm^3 resolution with non-brain regions masked out) data matrix,
which was then regressed onto a 34730 (29 subjects × up to 325 scans × up to 2 runs × 2 time
points) × 2320 (29 subjects × 2 time points × 4 conditions × 10 time bins) design matrix
consisting of dummy-coded stimulus timing information. The predicted scores were then
submitted to a PCA to identify functional brain networks predictable by the timing of the
experimental task.

4.1.3.3.2.1. Relation to Experimental Conditions

The predictor weights, which reflect estimates of HDR for each component (i.e.,
functional brain network) for each combination of subject, condition, time point, and
poststimulus time, were submitted to a repeated-measures ANOVA for each network
separately. These ANOVAs included the within-subjects factors of Time Point (time 1, time
2), Condition (YY, NN, NY, YN), and Poststimulus Time (10 time bins after stimulus onset).
Significant interactions were followed up with simple contrasts comparing each level of the

relevant factors to one another (Howell & Lacroix, 2012). Tests of sphericity were carried out for all ANOVAs, and adjustment for violations of sphericity did not affect the results; therefore, the original degrees of freedom are reported.

4.1.3.4. Associations Between Averaged HDR and Behaviour

4.1.3.4.1. Confirmatory and Disconfirmatory Evidence Integration

As in Chapter 3, associations between functional brain activity underlying dis/confirmatory evidence integration and behaviour were investigated using correlations as well as three behavioural CPCAs, one each for (1) behavioural BADE and overall fMRI performance, (2) neurocognitive variables, and (3) symptoms. However, in this chapter, change scores were used as criterion and predictor variables. As with the behavioural data, predictor weight change scores were computed by subtracting the time 1 from time 2 values (time 2-time 1); thus, positive values indicated an increase in intensity (greater activations and greater deactivations) from time 1 to time 2. These change scores were averaged over poststimulus time for each condition type (confirm = YY and NN; disconfirm = NY and YN), resulting in a single value for each condition type in each of the three networks (VsAN confirm, VsAN disconfirm, VDMN confirm, VDMN disconfirm, CEN confirm, and CEN disconfirm). These values were used to compute correlations with behaviour, and were included in the Z matrix for the behavioural CPCAs.

Details regarding the procedure and calculations for behavioural CPCA are described in Chapter 3. Briefly, the Z matrix, consisting of change scores (time 2-time 1) of averaged estimated HDR (i.e., predictor weights) for the confirm and disconfirm conditions separately, is standardized and regressed onto the G matrix, which consists of (non-standardized) change scores (time 2-time 1) on the relevant behavioural measures. This results in a matrix of

predicted scores, which is then submitted to a PCA to identify components reflecting combinations of estimated HDR change score values that are predictable by change scores in the measures of interest (i.e., behavioural BADE/fMRI performance, neurocognition, and symptoms). The Z matrices were similar for each of the three CPCAs (except for the number of rows due to different numbers of subjects who completed the behavioural tasks, namely 28 for behavioural BADE/fMRI performance, 29 for neurocognition, and 26 for symptoms), and consisted of six columns reflecting change scores of estimated HDR averaged over poststimulus time for each network (VsAN, VDMN, and CEN) and condition type (confirm and disconfirm).

The G matrices included the same number of rows as the corresponding Z matrices, and the columns reflected change scores (time 2-time 1) for each of the relevant behavioural variables. Specifically, the behavioural BADE/fMRI performance G matrix consisted of four columns of change scores (BADE evidence integration, BADE conservatism, overall fMRI accuracy, and overall fMRI RT), the neurocognition G matrix consisted of nine columns of change scores (RBANS immediate memory, RBANS visuospatial abilities, RBANS language, RBANS attention, RBANS delayed memory, TMT A, TMT B, LNS, and COWAT), and the symptoms G matrix consisted of 11 columns of change scores (PSYRATS hallucinations, PSYRATS delusions, SAPS hallucinations, SAPS delusions, SAPS bizarre behaviour, SAPS formal thought disorder, SANS affective flattening, SANS alogia, SANS avolition-apathy, SANS attention, and SANS inappropriate affect). Both the PSYRATS and SAPS hallucinations and delusions variables were included due to their distinct assessment of the presence of versus the phenomenology of hallucinations and delusions, respectively. The SANS anhedonia-asociality subscale was excluded from this analysis due to a relatively large

number of missing data. The number of components extracted for each solution and analysis was determined by inspection of scree plots.

4.1.3.4.2. Baseline-Peak Activations

BP and PB values were computed by creating HDR change scores as above (time 2-time 1), with BP values corresponding to average activation from the start of the trial until the peak of the HDR, and PB values corresponding to the time bin immediately following the peak (when the HDR began to fall back to baseline) to the end of the trial. Averaging over these temporal aspects of the HDR results in two values (BP and PB) for each network, which were converted to change scores (time 2-time 1) and (1) intercorrelated to determine associations of change within and between networks, (2) correlated with change in behavioural BADE, fMRI performance, neurocognitive measures, and symptoms, and (3) submitted to three separate CPCAs (1: behavioural BADE and overall fMRI performance; 2: neurocognitive variables; and 3: symptoms) to examine more complex associations between fMRI BP and PB change and behaviour change. The behavioural CPCA procedure was identical to that described in the previous section except that BP/PB change scores (time 2-time 1) were used in the Z matrices. The Z matrices were the same size as the previous behavioural CPCAs (behavioural BADE/fMRI performance = 28×6, neurocognition = 29×6, and symptoms = 26×6) since the number of BP/PB change score variables was the same as the confirm/disconfirm variables used above; however, the data included in the Z matrices consisted of change scores (time 2-time 1) of averaged HDR for the BP/PB aspects of the HDR rather than for HDR during dis/confirmatory evidence integration. The G matrices were identical to those in the previous behavioural CPCAs.

4.2. Results

4.2.1. Behavioural BADE Component Structure

Table 30 displays the loadings and bootstrapped standard errors for the unconstrained (time 1: left column; time 2: middle column) and constrained (time 1 & time 2: right column) GSCA models on behavioural BADE. The overall goodness of fit for the unconstrained model was 0.662, meaning that 66.2% of the overall variance in the observed variables was accounted for by the latent variables of BADE evidence integration and conservatism. The Goodness of Fit index (GFI) and Standardized Root Mean Square Residual (SRMR) were 0.99 and 0.12, respectively. The overall goodness of fit of the constrained model was similar to the unconstrained model, with a value of 0.659, with a GFI of 0.99 and SRMR of 0.13. The t-test comparing GFI values on 100 bootstrapped samples of each model was not significant, $t(197.99) = 1.38$, $p = 0.17$, indicating that the component structure of the behavioural BADE items was consistent between time 1 and time 2. Thus, composite behavioural BADE evidence integration and conservatism scores were created in the same way for each time point, based on the specified model (BADE evidence integration = absurd 1, 2 & 3, neutral lure 2 & 3, emotional lure 2 & 3, true 3 (reversed); BADE conservatism = neutral lure 1, emotional lure 1, true 1, 2, & 3). These values were used for all subsequent analyses involving the behavioural BADE task.

4.2.2. Time 1 Only

Table 31 shows means and standard deviations on demographics, premorbid and current IQ, chlorpromazine equivalent dosage, symptoms, neurocognitive measures, behavioural BADE, and fMRI performance for patients from study 2 who did not complete (time 1 only; n = 29 out of 58 patients from study 2) and did complete (time 1 & 2; n = 29 out of 58

118

patients from study 2) study 3. There were no significant differences between the groups on gender, handedness, age, education, premorbid IQ, or chlorpromazine equivalents (all ps > .25); however, time 1 & 2 patients had higher current IQ scores than time 1 only patients, $t(56) = -3.44$, $p < .005$. No significant differences were observed on symptoms (all ps > .06), other than SANS attention, on which time 1 only patients had significantly higher inattentiveness than time 1 & 2 patients, $t(56) = 2.01$, $p < .05$. Time 1 only patients relative to time 1 & 2 patients also had significantly higher (worse) BADE evidence integration, $t(54) = 2.84$, $p < .01$, and higher (better/less) conservatism, $t(54) = 3.16$, $p < .005$, scores. There were no significant differences between the groups on neurocognitive measures or fMRI performance (all ps > .07).

4.2.3. Time 1 Versus Time 2

4.2.3.1. Behavioural

Paired t-tests comparing time 1 and time 2 behavioural scores (see Table 32 for means) revealed significant decreases on BDI, $t(64) = 4.86$, $p < .001$, SAPS total, $t(67) = 2.78$, $p < .01$, SAPS delusions, $t(68) = 2.67$, $p < .01$, SAPS formal thought disorder, $t(70) = 2.01$, $p < .05$, and BADE conservatism, $t(55) = 2.08$, $p < .05$. Significant increases from time 1 to time 2 were observed on RBANS immediate memory, $t(71) = -5.43$, $p < .001$, RBANS attention, $t(71) = -3.09$, $p < .005$, RBANS delayed memory, $t(71) = -3.10$, $p < .005$, TMT A, $t(71) = -6.58$, $p < .001$, TMT B, $t(70) = -4.89$, $p < .001$, LNS, $t(71) = -2.22$, $p < .05$, and BADE evidence integration, $t(55) = -5.35$, $p < .001$. No significant changes from time 1 to time 2 were observed for fMRI performance (all ps > .15), except for fMRI accuracy in the confirm conditions, $t(28) = -2.41$, $p < .05$, which increased from 85.95% at time 1 to 90.52% time 2. These results indicate that patients improved on depressive symptoms and positive psychotic

119

symptoms, as well as on several neurocognitive measures, and on fMRI accuracy during confirmatory evidence integration from time 1 to time 2. In terms of behavioural BADE, both conservatism and evidence integration worsened from time 1 to time 2.

Correlations between change scores (time 2-time 1) in delusions (on the PSYRATS and SAPS separately) and change scores (time 2-time 1) in behavioural measures (see Table 33) showed significant associations between PSYRATS delusion change scores and the SAPS total score, $r(68) = 0.31$, $p < .01$, and between SAPS delusion change and PSYRATS hallucinations, $r(61) = 0.31$, $p < .05$, SAPS total, $r(68) = 0.74$, $p < .001$, SAPS hallucinations, $r(69) = 0.35$, $p < .005$, and TMT B, $r(68) = -0.25$, $p < .05$. These results show that decreases in delusions from time 1 to time 2 were primarily associated with decreases in other positive symptoms, and that for SAPS delusions, decreased severity from time 1 to time 2 was related to improved TMT B scores. There were no significant associations between delusion change (on either the PSYRATS or SAPS measures) and change in behavioural BADE or fMRI performance (all ps > .30).

4.2.3.2. fMRI-CPCA

Inspection of the scree plot (Cattell, 1966; Cattell & Vogelmann, 1977) suggested a three-component solution, with Components 1, 2 and 3 accounting for 15.94%, 11.06%, and 4.88% of the variance predictable from task timing, respectively. All three components showed a significant main effect of Poststimulus Time, $F(9,252) = 64.34$, $p < .001$, $F(9,252) = 78.16$, $p < .001$, and $F(9,252) = 66.89$, $p < .001$, for Components 1 to 3, respectively, and examination of the hemodynamic response shapes (Figures 19B, 21C, and 23B) indicated that each network showed a reliable and biological hemodynamic response shape (Lavigne et al., 2015a; Metzak et al., 2011; Metzak et al., 2012).

4.2.3.2.1. Component 1: Visual Attention Network

4.2.3.2.1.1. Anatomical Description

The brain regions associated with Component 1 are displayed in Figure 18, with anatomical descriptions listed in Table 34. This network was characterized by activations in bilateral occipital cortex (BAs 18, 19), middle frontal gyrus (BA 6), inferior frontal gyrus, pars triangularis (BAs 44, 45), right supramarginal gyrus (BA 40), left precentral (BAs 4, 6) and postcentral (BAs 2, 3) gyri, bilateral supplementary motor area (BA 6) and anterior cingulate cortex (BA 32), as well as left thalamus, bilateral insula and cerebellum. As in the previous chapter, this network reflected nodes of the visual, dorsal attention, somatosensory, and ventral attention networks described by Yeo and colleagues (2011), and was labelled *Visual Attention Network*.

4.2.3.2.1.2. Relation to Experimental Conditions

There was a significant main effect of Condition (Figure 19A), $F(3,84) = 3.58$, $p < .05$, on this network, which was the result of significantly higher activity in the disconfirm conditions (NY and YN) relative to the NN condition (ps < .05), with no other contrasts reaching significance (all ps > .12). Neither the main effect of Time Point, nor any of the higher-order interactions between variables (Figure 19B: Time Point × Poststimulus Time; Figure 19C: Condition × Poststimulus Time) were significant (all ps > .10).

4.2.3.2.2. Component 2: Visual/Default-Mode Network

4.2.3.2.2.1. Anatomical Description

The brain regions associated with Component 2 are displayed in Figure 20, with anatomical descriptions listed in Table 35. This network was dominated by deactivations in

bilateral precuneus (BA 18), posterior cingulate cortex (BA 23), frontal pole (BAs 9, 10), anterior cingulate cortex (BA 24), temporal cortex (BAs 21, 22, 41, 42), as well as right precentral (BA 4) and postcentral (BAs 2, 3) gyri. It also included activations in bilateral occipital cortex (BAs 18, 19), superior parietal lobule (BA 7), precentral gyrus (BA 44), and left supramarginal gyrus (BA 40). The deactivations overlapped with the default-mode and somatosensory networks, and the activations overlapped with visual, somatosensory, and dorsal attention networks of the resting-state 7 network parcellation (Buckner et al., 2011; Yeo et al., 2011). Due to the dominance of default-mode related deactivations and activations in occipital cortex on this network, it was labeled *Visual/Default-Mode Network*.

4.2.3.2.2.2. Relation to Experimental Conditions

There were significant main effects of Condition, $F(3,84) = 3.00$, $p < .05$, and Time Point, $F(1,28) = 4.14$, $p < .05$. The main effect of Condition (Figure 21A) was due to increased intensity in the YN condition relative to all other conditions, and the main effect of Time Point (Figure 21B) was the result of decreased intensity from time 1 to time 2 in this network ($p < .05$). There was also a significant interaction between Time Point and Poststimulus Time (Figure 21C), $F(9,252) = 4.14$, $p < .05$, showing decreased intensity from time 1 to time 2 in this network at 7s, 9s, 11s, and 13s (all ps $< .05$). Finally, there was a significant interaction between Condition and Poststimulus Time (Figure 21D), $F(27, 756) = 1.93$, $p < .05$. Simple contrasts revealed that this interaction was due to significantly greater activity in (1) YN versus YY at 9s, 11s, and 13s, (2) YN versus NN at 3s, 5s, 7s, 9s, and 11s, (3) YN relative to NY at 9s, 11s, and 13s. Neither the Time Point × Condition nor the Time Point × Condition × Poststimulus Time interactions were significant (ps $> .40$).

122

4.2.3.2.3. Component 3: Cognitive Evaluation Network

4.2.3.2.3.1. Anatomical Description

The brain regions associated with Component 3 are displayed in Figure 22, with anatomical descriptions listed in Table 36. This network included activations in occipital cortex (BAs 18, 19), middle temporal gyrus, temporooccipital part (BAs 21, 37), lateral orbitofrontal cortex (BA 10, 38, 47), dorsolateral prefrontal cortex (BA 46), supramarginal and angular gyri (BAs 39, 40), and inferior frontal gyrus, pars triangularis (BA 44, 45) and pars opercularis (BA 45). Deactivations were also observed in supplementary motor area (BA 6) and precentral gyrus (BAs 4, 6). The activations of this network corresponded to visual and frontoparietal networks extending into the dorsal and ventral attention networks, and the deactivations to the somatosensory network of the 7-network parcellation described in previous research (Yeo et al., 2011). Given the similarities between the spatial configuration and timing of this network and those found in the previous chapters, it was labeled the *Cognitive Evaluation Network*.

4.2.3.2.3.2. Relation to Experimental Conditions

There was a significant main effect of Condition, $F(3,84) = 16.16$, $p < .001$, and a significant interaction between Condition and Poststimulus Time, $F(27,756) = 5.01$, $p < .001$. Simple contrasts revealed the main effect of Condition (Figure 23A) was the result of increased activity in the YN condition relative to all other conditions (all ps < .001), as well as increased activity in (1) NY versus NN, and (2) YY versus NN (ps < .05). The Condition × Poststimulus Time (Figure 23C) interaction was due to increased activity in (1) YN versus YY at 11s, 13s, 15s, 17s, and 19s, (2) YN versus NN at 7s, 9s, 11s, 13s, 15s, 17s, and 19s, and (3) YN versus NY at 7s, 9s, 11s, 13s, 15s, 17s, and 19s (all ps > .05). There was also

123

increased activity in (1) YY versus NN at 7s, 11s, and 19s, (2) YY versus NY at 9s, (3) NY versus YY at 15s and 17s, (4) and NY versus NN at 1s, 13s, 15s, 17s, and 19s (all ps > .05). The main effect of Time Point was not significant, nor were the remaining interactions (ps > .36).

4.2.3.3. Associations Between Averaged HDR and Behaviour

4.2.3.3.1. Confirmatory and Disconfirmatory Evidence Integration

4.2.3.3.1.1. Correlations

Correlations were computed between change scores (time 2-time 1) on average activity in the confirm and disconfirm conditions separately and change scores (time 2-time 1) on behavioural variables (see Table 37) in order to examine associations between changes in brain activity underlying dis/confirmatory evidence integration and behaviour. For VsAN, increased network activity during confirmatory evidence integration was associated with increased scores on behavioral BADE evidence integration (poorer evidence integration), $r(28) = 0.40$, $p < .05$, and increased scores on behavioural BADE conservatism (improved/decreased conservativism), $r(28) = 0.46$, $p < .05$, from time 1 to time 2. Greater VsAN activity during confirmatory evidence integration was also associated with decreased fMRI accuracy overall, $r(29) = -0.42$, $p < .05$, and for the disconfirm conditions, $r(29) = -0.37$, $p < .05$, decreased SAPS total scores, $r(27) = -0.48$, $p < .05$, more missed responses on the fMRI task, $r(29) = 0.51$, $p < .005$, and improved RBANS attention, $r(29) = 0.53$, $p < .005$. This indicates that increased network activity in the VsAN from time 1 to time 2 during confirmatory evidence integration was related to increased (poorer) BADE evidence integration scores and increased (better) conservatism scores, improved attention, and more fMRI missed responses, and was also related to decreases in positive symptoms and

124

decreased fMRI accuracy. These findings are in line with the hypotheses stemming from study 2 for behavioural BADE and fMRI performance (given hyperactivity in the VsAN for patients with higher (poorer) behavioural BADE evidence integration and impaired performance), but contradicts expectations based on the symptom findings (decreased VsAN activity from time 1 to time 2 during confirmatory evidence integration would be expected to related to decreased symptoms/delusions since hyperactivity was present in delusional patients in study 2, but the opposite was observed).

Greater activity in the VsAN during disconfirmatory evidence integration from time 1 to time 2 was associated with decreases in TMT A, $r(29) = -0.40$, $p < .05$, indicating further impairment in TMT A with hyperactivity in the VsAN during disconfirmatory evidence integration. For the VDMN, decreased intensity (lesser activations and deactivations) during confirmatory evidence integration from time 1 to time 2 was associated with more missed responses on the fMRI task, $r(29) = -0.50$, $p < .01$, and decreases in SAPS total, $r(27) = 0.53$, $p < .005$, SAPS delusions, $r(27) = 0.44$, $p < .05$, and SAPS formal thought disorder, $r(29) = 0.41$, $p < .05$. No significant associations were observed between behaviour and activity change during disconfirmatory evidence integration in this network. Finally, for the CEN, greater activity from time 1 to time 2 during confirmatory evidence integration was related to more missed responses on the fMRI task, $r(29) = 0.40$, $p < .05$, and improved RBANS delayed memory, $r(29) = 0.41$, $p < .05$, as well as to decreased fMRI accuracy during the confirm conditions only, $r(29) = -0.41$, $p < .05$, and decreased SAPS formal thought disorder, $r(29) = -0.37$, $p < .05$. For disconfirmatory evidence integration, greater activity in the CEN from time 1 to time 2 was associated with decreased SAPS inappropriate affect, $r(29) = -0.39$, $p < .05$. These findings are partially in line with the hypotheses stemming from study 2;

decreased intensity in the VDMN from time 1 to time 2 (during confirmatory evidence integration) was associated with poorer performance on the fMRI task (in line with study 2), but with decreases in positive symptoms (in contrast to study 2, in which greater positive symptoms were associated with less intensity in the VDMN). For the CEN, the opposite pattern emerged; increased activity from time 1 to time 2 (during confirmatory evidence integration) was associated with poorer performance on the fMRI task (in contrast to study 2, in which greater activity was associated with better performance), but with decreases in positive symptoms particularly formal thought disorder (which is in line with study 2). These findings suggest complex associations between symptoms, evidence integration and functional brain activity in these different networks, and are explored further in the following section using multivariate analysis.

4.2.3.3.1.2. Behavioural CPCAs

In order to further examine associations between activity underlying dis/confirmatory evidence integration and behaviour, we conducted three behavioural CPCAs (as in Chapter 3) predicting change scores (time 1-time 2) on functional brain activity in each network and condition type (confirm, disconfirm) from (1) behavioural BADE and overall fMRI performance change scores, (2) neurocognition change scores, and (3) symptoms change scores. The main difference between these analysis and those of the previous chapter is that change scores were used in Z, calculated by subtracting the time 1 from time 2 predictor weight values (time 2-time 1 as for the correlations above), and averaging them across poststimulus time for each condition type and network. The variance tables and rotated component loadings are displayed in Tables 38 to 43, for the three analyses respectively, with even-numbered tables showing the variance tables and odd-numbered tables showing the

126

rotated component loadings. The overall and residual solutions were similar across all three

analyses: the first component consisted of VsAN and VDMN (reversed) confirm and

disconfirm values, and the second component included CEN confirm and disconfirm. This

indicates that, when it comes to delusion change, activity in the CEN is orthogonal to that of

the VsAN and VDMN, as it was for a single time point (time 1 only, Chapter 3).

4.2.3.3.1.2.1. Behavioural BADE/fMRI Performance

For the behavioural BADE/fMRI performance analysis (Table 39), the predicted solution

(Predicted Solution – Dependent Variables) separated the confirm and disconfirm conditions,

with VsAN confirm (0.49) and VDMN confirm (-0.24) loading most strongly on the first

component, and VsAN disconfirm (0.49) dominating the second component. Both VDMN

disconfirm (0.17 and -0.16), CEN confirm (0.16 and -0.20) and disconfirm (-0.11 and -0.13)

loaded similarly on both components. The predictor loadings (Table 39, Predicted Solution –

Independent Variables) show that, for Component 1, an increase from time 1 to time 2 on

VsAN/CEN confirm (0.49 and 0.16, respectively), and a decrease in VDMN confirm (-0.24),

was associated with a decrease in fMRI accuracy (-.85), longer fMRI RTs (0.30), as well as

greater behavioural BADE evidence integration (0.85) and conservatism (0.80). In contrast,

on Component 2, an increase in VsAN disconfirm (0.49), and less so a decrease in VDMN

disconfirm (-0.16) and CEN disconfirm (-0.13), from time 1 to time 2 was associated with

improved fMRI accuracy (0.44), but still longer fMRI RTs (0.31) and higher BADE evidence

integration (0.31) and conservatism (0.50). These findings show that functional brain activity

underlying evidence integration does indeed fluctuate with changes in BADE evidence

integration, and this fluctuation fits with expectations that stem from the group differences

observed in study 2. That is, increased behavioural BADE from time 1 to time 2 was

associated with a further increase of functional brain activity in the VsAN (especially during the confirm condition) and a further decrease in CEN activity (for the disconfirm condition) relative to the hyper/hypoactivity observed in study 2. Activity in the CEN during confirmatory evidence integration was more complex, in that the portion of variance that was associated with Component 1 (which was dominated by VsAN confirm) was positively related (0.16) to behavioural BADE evidence integration (0.85), while the portion of variance that was associated with Component 2 (dominated by VsAN and CEN disconfirm) was negatively associated (-0.20) with behavioural BADE evidence integration (0.31). Given the bi-directional nature of these correlations, these results suggest that decreased/improved behavioural BADE (though not a general trend in the current study) was associated with normalization of the hemodynamic response in these networks (i.e., decreased activity in the VsAN during confirmatory evidence integration, and increased activity in the CEN during disconfirmatory evidence integration) based on the group differences identified in study 2, as hypothesized.

4.2.3.3.1.2.2. Neurocognition

Similar to study 2, the predicted solution for the neurocognition analysis (Table 41, Predicted Solution – Dependent Variables) showed a greater distinction between networks than conditions. Specifically, Component 1 consisted of VsAN confirm (0.60) and disconfirm (0.52) and VDMN confirm (-0.17), and Component 2 included CEN confirm (0.40) and disconfirm (0.50), and VDMN disconfirm (-0.28). The predictor loadings (Table 41, Predicted Solution – Independent Variables) for Component 1 showed that increases in VsAN (confirm and disconfirm) and decreases in VDMN (confirm) from time 1 to time 2 were associated with improved RBANS attention (0.75) and decreased TMT A (-0.32), and

RBANS visuospatial abilities (-0.30). In contrast, for Component 2, increases in CEN (confirm and disconfirm) and decreases in VDMN (disconfirm) were associated with improvements in RBANS immediate (0.18) and delayed (0.54) memory, as well as decreases on RBANS language (-0.31), TMT B (-0.49), and LNS (-0.34). These findings are highly similar to those from the neurocognition CPCA on activity underlying dis/confirmatory evidence integration in study 2, because they suggest that changes in neurocognition predict changes in the same functional brain networks that they predict at a single time point. Specifically, an increase in attention-related neurocognitive measures (which was observed behaviourally in the current study) was associated with normalization of the VsAN (viz., decreased activity) regardless of condition (and normalization of VDMN confirm). However, the RBANS attention variable shows the opposite pattern, likely due to the importance of the VsAN in orienting attention. Careful consideration of the differences in aspects of attention assessed by, for example, RBANS attention and TMT A, may clarify these apparently contradictory findings. In contrast to attention measures, we found that an increase in memory-related measures (also observed behaviourally in this study) was associated with normalization (increased activity) of the CEN for both conditions (and further deactivation in VDMN disconfirm). Unexpectedly, a decrease in performance in RBANS language, TMT B (set shifting) and LNS (processing speed) was also associated with normalization (increased activity) of the CEN, suggesting the cognitive processes underlying these neuropsychological tasks may be hindered by increased activity in the CEN in schizophrenia patients, although they enhance evidence integration.

4.2.3.3.1.2.3. Symptoms

Similar to the behavioural BADE/fMRI performance analysis, in terms of symptoms (Table 43), the predicted solution (Predicted Solution – Dependent Variables) distinguished between confirm and disconfirm conditions, except that the confirm conditions (especially VsAN (0.57) and VDMN (-0.79)), dominating the first component, accounted for more variance than the disconfirm conditions (VsAN = 0.15, VDMN = -0.46, CEN = 0.62), which dominated the second component (41.46% vs. 38.47% of predictable variance, see Table 42). One exception to this analogous pattern was that the CEN confirm condition loaded more strongly onto the second component (0.61 for Component 2 vs. 0.11 for Component 1). The predictor loadings (Table 43, Predicted Solution – Independent Variables) showed that greater VsAN activity and lesser VDMN intensity during confirmatory evidence integration (Component 1) from time 1 to time 2 was associated with decreased delusions (-0.21 for PSYRATS, -0.58 for SAPS), SAPS formal thought disorder (-0.48), and SANS affective flattening (-0.37). In contrast, greater CEN (confirm and disconfirm) and VsAN (disconfirm) activity and lesser VDMN intensity (disconfirm) was associated with decreased hallucinations (-0.15 for PSYRATS, -0.11 for SAPS), SANS alogia (-0.69) and SANS inappropriate affect (-0.35), as well as greater SANS avolition-apathy (0.40). Thus, Component 1 (VsAN/VDMN confirm) was more strongly related to positive symptoms, while Component 2 (VsAN/VDMN disconfirm and CEN confirm/disconfirm) was better predicted by negative symptoms, though not exclusively so. This differs from the study 2 findings, for which CEN activity (in both confirm and disconfirm conditions) was more strongly related to positive symptoms, and suggests that changes in delusions are a stronger predictor of activity during confirmatory evidence integration relative to disconfirmatory

evidence integration, particularly for VsAN and VDMN. However, the direction of these findings and how it relates to change over time deviates from what might be expected from the study 2 results. Specifically, a decrease in symptom severity from time 1 to time 2 was associated with an increase in VsAN and CEN activity, and a decrease in VDMN activity. This reflects normalization of the HDR for the CEN (in line with expectations), but an increase in the hyperactivity/hypoactivity in the VsAN/VDMN identified in study 2. These complex findings may have been influenced by the paradoxical increase in behavioural BADE evidence integration, coupled with decreases in symptoms observed in this study, which is addressed further in the discussion section below.

4.2.3.3.1.2.4. Summary of Associations Between Dis/Confirmatory Evidence Integration HDR and Behaviour

The behavioural CPCAs on activity change scores (time 2-time 1) underlying dis/confirmatory evidence integration revealed that functional brain activity fluctuates as a function of behavioural evidence integration, and to a lesser extent, symptoms, in a way that would be largely be expected based on the findings of study 2. Specifically, poorer behavioural BADE evidence integration and fMRI performance from time 1 to time 2 was associated with greater activity in the VsAN (more so for the confirm condition), and less activity in the CEN (for both confirm and disconfirm) from time 1 to time 2. This indicates that poorer behavioural evidence integration is associated with greater aberrations of the differences in functional brain activity (hyperactivity in the VsAN during confirmatory evidence integration, and hypoactivity in the CEN during disconfirmatory evidence integration) identified in study 2, though the VsAN associations were stronger in the current study. Neurocognitive measures were better predictors of networks' activity than activity

131

specifically related to confirmatory or disconfirmatory evidence integration, and improvements in attention and memory were associated with increases/decreases in the VsAN (depending on the measure) and CEN activity, respectively. In terms of symptoms, the findings were in line with our hypotheses for the CEN (viz., decreased positive symptoms were associated with normalization of the CEN), but were contrary to expectations for the VsAN and VDM (viz., decreased positive symptoms were associated with further hyperactivity in the VsAN and hypoactivity in the VDMN). This is addressed further in the discussion section below.

4.2.3.3.2. Baseline-Peak Activations

Computation of BP and PB values was determined by inspection of the HDR in each network averaged over conditions and time points. The BP and PB values for each network were as follows (see Figures 19B, 21C, and 23B): VsAN BP = 1s – 9s, PB = 11s – 19s; VDMN BP = 1s – 11s, PB = 13s – 19s; CEN BP = 1s – 15s, PB = 17s – 19s. BP and PB values were computed from change scores (time 2-time 1) in estimated HDR and the resulting six values (VsAN BP, VsAN PB, VDMN BP, VDMN PB, CEN BP, and CEN PB) were (1) intercorrelated to examine within- and between-network associations, (2) correlated with behavioural BADE, fMRI performance, neurocognition, and symptoms, to determine associations with behaviour, and (3) submitted to three behavioural CPCAs (1: behavioural BADE/overall fMRI performance, 2: neurocognition, 3: symptoms) to examine these associations in more depth.

4.2.3.3.2.1. Intercorrelations

Intercorrelations between BP and PB values across the three networks (Table 44) revealed that increased BP and PB values from time 1 to time 2 were positively correlated

within the same network for VsAN, $r(29) = 0.38$, $p < .05$, and VDMN, $r(29) = 0.45$, $p < .05$, but not CEN ($p > 0.69$). Between networks, increases in VsAN and decreases in DMN BP, $r(29) = -0.73$, $p < .001$, and PB, $r(29) = -0.52$, $p < .005$, were also significantly correlated. Significant associations between BP and PB change score values across networks were observed for VDMN BP and CEN PB (increased activity in VDMN BP was associated with increased activity in CEN PB from time 1 to time 2), $r(29) = 0.41$, $p < .05$, and CEN BP and VDMN PB (increased activity in CEN BP was associated with decreased activity in VDMN PB from time 1 to time 2), $r(29) = -0.41$, $p < .05$. As with study 2, these findings indicate that CEN activity is partially separable from VsAN and VDMN activity, whereby changes in the latter two networks were associated with changes in the other for the same aspect of the HDR, which was not the case for the CEN. Interestingly, changes in different aspects of the HDR for the CEN and VDMN fluctuated with one another, a finding not present in study 2. This suggests that a CEN BP increase from time 1 to time 2 was associated with decrease in the PB of the VDMN, and a CEN PB increase from time 1 to time 2 was associated with an increase in the BP of the VDMN. This supports the notion that the two aspects of the CEN HDR represent distinct cognitive processes, as they are differentially associated with other networks (VDMN in the current study, and VsAN in study 2).

4.2.3.3.2.2. Correlations with Behaviour

Correlations between BP/PB values and behaviour (Table 45) revealed that increased activity in VsAN BP from time 1 to time 2 was associated with increased (improved/less conservative) behavioural BADE conservatism, $r(28) = 0.38$, $p < .05$, greater fMRI missed responses, $r(29) = 0.50$, $p < .01$, and improved RBANS attention, $r(29) = 0.47$, $p < .01$. Similarly, increased activity in VsAN PB from time 1 to time 2 was associated with

increased (improved/less conservative) behavioural BADE conservatism, $r(28) = 0.41$, $p <$ .05, and longer fMRI RTs overall, $r(29) = 0.52$, $p < .01$, and for both the confirm, $r(29) =$ 0.46, $p < .05$ and disconfirm, $r(29) = 0.48$, $p < .01$, conditions. For the VDMN, decreased intensity in BP from time 1 to time 2 was associated with more missed responses, $r(29) = -$ 0.44, $p < .05$, and decreased SAPS total, $r(27) = 0.39$, $p < .05$. Decreased intensity in VDMN PB was associated with longer fMRI RTs during disconfirmatory evidence integration, $r(29)$ $= -0.39$, $p < .05$, decreased SAPS total, $r(27) = 0.41$, $p < .05$, and SAPS delusions, $r(27) =$ 0.38, $p < .05$. Finally, for the CEN, increased BP from time 1 to time 2 was associated with greater RBANS delayed memory, $r(29) = 0.39$, $p < .05$, and increased PB with poorer COWAT, $r(29) = -0.40$, $p < .05$. These results demonstrate that changes in the BP/PB aspects of the hemodynamic response were not related to behavioural BADE evidence integration (only conservatism), and were only related to positive symptoms for the VDMN. In study 2, associations with symptoms were also minimal, but were observed for evidence integration (CEN BP) and positive symptoms for VsAN BP (but not with delusions; see Table 23). Associations between BP and PB in each network and behavioural measures were examined in more detail using behavioural CPCA.

4.2.3.3.2.3. Behavioural CPCAs

As with activity changes underlying dis/confirmatory evidence integration, three behavioural CPCAs were conducted to examine associations between BP/PB change scores (time 2-time 1) and (1) behavioural BADE and overall fMRI performance change scores, (2) neurocognition change scores, and (3) symptom change scores. The variance tables and rotated component loadings are displayed in Tables 46 to 51, respectively, with even-numbered tables for the variance tables and odd-numbered tables for the rotated component

134

and predictor loadings. The overall solutions were similar across all three analyses:

Component 1 included VsAN and VDMN (reversed) BP change scores; Component 2

VsAN, VDMN (reversed), and CEN PB change scores; and Component 3 VDMN (reversed)

PB and CEN BP change scores. However, unlike the overall solutions, and the residual

solutions of the previous analyses, the residual solutions differed across the three behavioural

CPCAs, likely due to differences in the configurations of their respective predicted solutions,

and that they accounted for different amounts of the overall variance (see Tables 46, 48, and

50).

For behavioural BADE/fMRI performance (Table 47), the residual solution consisted of

two components: for Component 1, VsAN BP (0.73), VDMN BP (-0.90) and CEN PB (-

0.66); and for Component 2, VsAN PB (0.51), VDMN PB (-0.78), and CEN BP (0.68). For

the behavioural CPCA on neurocognition, three components were extracted for the residual

solution (Table 49). The first component was similar to the first component in the

behavioural BADE/fMRI performance BP/PB analysis (VsAN BP = 0.78, VDMN BP = -

0.84, CEN PB = -0.49), but the latter's second component was split across the final two

components in the neurocognition CPCA (Component 2: VsAN PB (0.64), VDMN (-0.78);

Component 3: CEN BP (0.74)). This division may have been the result of extracting an

additional component in the neurocognition residual solution. Finally, for behavioural CPCA

on symptoms (Table 51), two components were extracted in the residual solution, which

differed from the two previous analyses (behavioural BADE/fMRI performance and

neurocognition). The first component included PB values for VsAN (0.75), VDMN (-0.39),

and CEN (0.68), and the second component included BP values for VsAN (0.78), VDMN (-

0.64), and CEN (0.24). We focus below on the predicted solutions, with two components extracted across all three analyses, and described in detail below for each analysis separately.

4.2.3.3.2.3.1. Behavioural BADE/fMRI Performance

The first component of the predicted solution (Table 47, Predicted Solution – Dependent Variables) included VsAN PB (0.61), VDMN PB (-0.27), CEN BP (-0.26), and CEN PB (0.21), and the second component consisted of VsAN BP (0.43) and VDMN BP (-0.15). The predictor loadings (Table 47, Predicted Solution – Independent Variables) showed that increases in behavioural BADE evidence integration (0.54 and 0.68) and conservatism (0.57 and 0.79) from time 1 to time 2 were associated with increased activity in VsAN BP and PB, as well as CEN PB, and decreased activity in VDMN BP and PB, and CEN BP. This is in line with the hypotheses stemming from study 2, as it demonstrated that increased BADE evidence integration from time 1 to time 2 was associated with aberrant network activity reflected in aspects of the hemodynamic response that were underlying poor evidence integration in the previous study (i.e., hyperactivity in VsAN BP and PB, and hypoactivity in CEN BP). Thus, changes in functional brain networks (and aspects of the HDR) underlying behavioural BADE evidence integration in study 2 were correlated with changes in behavioural BADE in the current study. Decreased fMRI accuracy was related to both components (-0.19 and -0.47, respectively), but more strongly associated with VsAN BP increases and VDMN BP decreases (Component 2), suggesting that this increase/decrease in VsAN/VDMN BP was related to poorer fMRI accuracy, whereas increase/decrease in fMRI RT was related to VsAN BP increases over time (-0.35) and VsAN PB increases over time (0.85), respectively, with the reversed associations emerging for VDMN.

4.2.3.3.2.3.2. Neurocognition

Similar to the neurocognition CPCA examining activity underlying dis/confirmatory evidence integration (Table 41, Predicted Solution – Dependent Variables), the predicted solution for the neurocognition change CPCA (Table 49, Predicted Solution – Dependent Variables) showed a greater distinction between changes in functional brain networks than changes in BP/PB values. Component 1 showed the strongest loadings from VsAN BP (0.32), VsAN PB (0.65), and CEN PB (0.59), whereas Component 2 was dominated by VDMN BP (-0.34), VDMN PB (-0.26), and CEN BP (0.43). Thus, changes in neurocognitive measures were better predictors of networks than BP/PB values, as was the case for the previous neurocognition CPCA on activity underlying dis/confirmatory evidence integration (Table 41, Predicted Solution – Dependent Variables). The one exception for this configuration was the CEN, where BP change loaded with the VDMN (Component 2), and PB change loaded with the VsAN (Component 1). As in the study 2 CPCA examining dis/confirmatory evidence integration (Table 27), the predictor loadings split between attention (related to Component 1) and memory (related to Component 2) measures, though the results were not as clear as for the study 2 analysis on neurocognition (Table 27). The predictor loadings (Table 49, Predicted Solution – Independent Variables) show that greater VsAN (BP and PB) and CEN PB activity from time 1 to time 2 (Component 1) was associated with increased RBANS immediate memory (0.29) and attention (0.45), and decreased RBANS visuospatial abilities (-0.42), TMT A (-0.49), and COWAT (-0.49). In contrast, increases in CEN BP and decreases in VDMN (BP and PB) from time 1 to time 2 (Component 2) were associated with increased RBANS delayed memory (0.44), as well as decreased RBANS language (-0.39), TMT B (-0.15), and LNS (-0.46). One difference

between this and the previous neurocognition CPCA on activity underlying dis/confirmatory evidence integration (Table 41, Predicted Solution – Dependent Variables) is that RBANS immediate memory change values loaded more strongly with VsAN (and CEN PB) than VDMN (and CEN BP), the latter of which was observed in the previous analysis (Table 41). Nonetheless, these change value findings are in line with the previous neurocognition CPCAs from both the current study (Table 41) and study 2 (Table 27) indicating that neurocognition is a better predictor of network activity than activity underlying different aspects of the HDR (i.e., dis/confirmatory evidence integration or BP/PB). Moreover, these results emphasize the importance of the VsAN for attentional and the CEN for memory processes.

4.2.3.3.2.3.3. Symptoms

For symptom change values (Table 51; Predicted Solution – Dependent Variables), the predicted solution was similar to that of the behavioural BADE/fMRI performance analysis (Table 47; Predicted Solution – Dependent Variables); Component 1 included VsAN PB (0.29), VDMN PB (-0.69, and CEN BP (0.65), and Component 2 included VsAN BP (0.48), VDMN BP (-0.62), and CEN PB (-0.53). Some exceptions included CEN BP reversing sign and CEN PB flipping from Component 1 to Component 2 and reversing sign. The predictor loadings (Table 51, Predicted Solution – Independent Variables) showed that increases in VsAN PB and CEN BP (and decreases in VDMN PB) from time 1 to time 2 (Component 1) were associated with decreased delusions (-0.24 for PSYRATS, -0.22 for SAPS), hallucinations (-0.22 for PSYRATS, -0.20 for SAPS), and SAPS formal thought disorder (-0.41), as well as some negative symptoms, including decreased SANS alogia (-0.64) and increased SANS avolition-apathy (0.45). In contrast, increases in VsAN BP and decreases in VDMN BP and CEN BP from time 1 to time 2 were associated with decreased delusions

138

(SAPS only, -0.42), SAPS bizarre behaviour (-0.38), SANS affective flattening (-0.21), and SANS attention (-0.24). Although these findings indicate that functional brain activity in these networks fluctuates as a function of symptoms (including delusions), the direction of these associations was not entirely in line with expectations, as was the case with the dis/confirmatory evidence integration CPCA on symptoms above (Table 43; Predicted Solution – Independent Variables). Decreased symptoms from time 1 to time 2 were associated with further hyperactivation of the VsAN/hyperdeactivation of the VDMN (contrary to hypotheses) but normalization (increased activation) of the CEN (as hypothesized) relative to the group differences observed in study 2. This is discussed further below. Finally, the PSYRATS and SAPS delusions change scores showed differential associations with these two components (more strongly predicting Components 1 and 2, respectively), suggesting that changes in the presence (SAPS) versus phenomenology (PSYRATS) of delusions are better at predicting PB and BP aspects of the hemodynamic response, respectively (and vice-versa for CEN).

4.2.3.3.2.3.4. Summary of Associations Between Baseline-Peak HDR and Behaviour

As in Chapter 3 (study 2), the behavioural CPCAs on baseline-peak values clarified which aspects of the HDR drove the findings regarding activity underlying dis/confirmatory evidence integration. Change from time 1 to time 2 to poorer behavioural evidence integration (increased BADE, decreased fMRI accuracy, and longer fMRI RTs) was associated with greater activity in VsAN BP and PB, and CEN PB, as well as decreased activity in CEN BP and VDMN BP and PB, from time 1 to time 2. These findings indicate that, as in study 2, the change in behavioural BADE/fMRI performance findings were driven

by increases in VsAN PB (and to a lesser extent BP in the current study) and decreases in the CEN BP aspects of the HDR. Interestingly, the BP and PB aspects of the CEN showed differential associations with behavioural BADE evidence integration and fMRI accuracy (viz., poorer behavioural evidence integration associated with decreased CEN BP activity and increased CEN PB activity from time 1 to time 2), suggesting that changes in the BP and PB aspects of the HDR represent changes in distinct cognitive operations. As with the previous neurocognition analyses (Tables 27 and 41), neurocognitive measures better predicted functional brain networks than BP/PB values, and once again showed that increased attention and memory from time 1 to time 2 were associated with greater activity in the VsAN and CEN, respectively. Finally, a decrease in positive symptoms from time to time 2 was associated with increased activity in VsAN PB and decreased activity in VDMN BP and PB (contrary to hypotheses), as well as an increase in CEN BP (in line with hypotheses). This is discussed further below.

4.3. Discussion

In the current study, we examined associations between changes in delusion severity, behavioural BADE, and functional brain networks underlying evidence integration in schizophrenia. From time 1 to time 2 (approximately 3 months later), patients improved on measures of symptoms and neurocognition, but behavioural BADE evidence integration worsened. The functional brain networks identified were similar to those found in the studies 1 and 2, namely, VsAN, VDMN, and CEN. All networks showed increased activity during disconfirmatory evidence integration, as in the previous studies (Chapters 2 and 3). Comparisons of overall brain activity in each network from time 1 to time 2 showed a decrease in intensity in the VDMN only; however, more in-depth examination of changes in

network activity underlying changes in confirmatory and disconfirmatory evidence integration and baseline-peak aspects of the hemodynamic response in each network revealed more complex associations. Specifically, poorer behavioural BADE evidence integration from time 1 to time 2 was associated with increased activity in the VsAN and decreased activity in the BP aspect of the CEN. This is the pattern (hyperactivity in VsAN and hypoactivity in CEN, especially BP) that drove the group differences identified in study 2. Interestingly, changes in the BP and PB aspects of the CEN showed differential associations with changes in behavioural BADE evidence integration and symptoms, suggesting that BP and PB represent distinct cognitive processes underlying evidence integration. For example, CEN BP might involve initiation of evidence integration processes (impaired in schizophrenia patients, especially those with delusions), while CEN PB could reflect efficient inhibition (also impaired in patients to a lesser degree) of these processes once integration has been achieved, as we observed in a previous study in a response network using an auditory oddball task (Lavigne et al., 2016). In addition, improvements in neurocognitive functioning from time 1 to time 2 were generally associated with normalization of functional brain activity in these networks, but were better predictors of brain networks than activity underlying confirmatory or disconfirmatory evidence integration or BP/PB (i.e., behavioural CPCA predicted solutions were more often clustered by networks than conditions or BP/PB values). Finally, poorer behavioural BADE evidence integration and improved delusions from time 1 to time 2 showed opposite associations with functional brain activity in the VsAN (poorer BADE associated with increased VsAN activity, as hypothesized; improved delusions also associated with increased VsAN activity, contrary to hypotheses), which implies a complex interplay between changes in BADE, delusion severity, and functional

brain activity underlying evidence integration. These findings indicate that poorer

behavioural BADE evidence integration across time is associated with further hyperactivity

in the VsAN and hypoactivity in the CEN during an fMRI evidence integration task, and that

improvements in symptoms (including delusions) lead to normalization of the CEN, which

underlies the bias against disconfirmatory evidence observed in schizophrenia patients with

delusions.

4.3.1. Behavioural

Behaviourally, both symptoms and neurocognitive measures improved from time 1 to

time 2 (see Table 32); however, BADE evidence integration scores worsened from time 1 to

time 2, suggesting that patients rated later lures and absurd interpretations higher, and the

final true item lower, at time 2. Although unexpected, these results are in line with previous

research suggesting that BADE may worsen during the course of psychosis (subclinical to at-

risk to first-episode to chronic; Eisenacher et al., 2016; Eisenacher & Zink, 2017). In

addition, there was no significant correlation between behavioural BADE score increases

from time 1 and time 2 and delusion change. Although this supports the only previous study

to test this association explicitly (Buonocore et al., 2015), we expected that decreases in

delusion severity (observed in the current study) would be associated with decreased

behavioural BADE evidence integration (i.e., improvements in evidence integration) due to

the wealth of previous research reporting a BADE in schizophrenia patients with delusions

(McLean et al., 2017; Moritz & Woodward, 2006; Sanford et al., 2014; Speechley et al.,

2012; Woodward et al., 2006b). However, we failed to observe a correlation between

changes in behavioural BADE evidence integration and delusion severity, and in fact

observed an *increase* in BADE from time 1 to time 2. A possible explanation for these

unexpected findings concerns the fact that behavioural BADE evidence integration scores

were quite low at time 1 in the current study, which would limit the degree to which BADE

scores could improve over time. There was also evidence that both improvements in

delusions and impairment in BADE evidence integration over time affected functional brain

activity underlying our evidence integration task (though in opposite directions for the

VsAN, see below), which highlights the need for additional research. Future research might

consider selecting for individuals with high behavioural BADE scores and employing more

directed confirmatory analysis techniques to investigate these associations.

4.3.2. Neuroimaging

As in the previous studies, we identified three functional brain networks associated with

evidence integration: VsAN (occipital cortex, dorsal and ventral attention networks,

somatosensory regions), VDMN (deactivations in ventromedial PFC, precuneus, posterior

cingulate cortex; activations in visual cortex), and CEN (rostrolateral prefrontal/orbitofrontal

cortex, inferior frontal gyrus, inferior parietal lobule). These networks corresponded closely

to those found in both studies 1 and 2 in separate samples, and again, each network showed

increased activity during disconfirmatory relative to confirmatory evidence integration.

Network differences in study 3 relative to study 2 included: (1) less extensive activation of

the right insula and thalamus in the VsAN; (2) absence of activation in the left precentral

gyrus for the VDMN activations; (3) additional deactivation of the right pre- and postcentral

gyri in the VDMN; and (4) more anterior activation of middle temporal cortex, which did not

extend into the posterior parietal cortex cluster in the CEN as in the previous study.

However, these differences in functional brain network configurations were minor

considering the difference in samples (no overlap to study 1, and 25% overlap to study 2) and fMRI task (different from study 1, and same as study 2, but with different stimuli for time 2).

4.3.3. Evidence Integration

Averaging across all participants and conditions, only activity in the VDMN showed a significant decrease from time 1 to time 2. However, investigation of associations between changes in functional brain activity in each network and behaviour using behavioural CPCA revealed more complex associations. Specifically, examination of associations with changes in network activity underlying confirmatory and disconfirmatory evidence integration showed that greater impairment in BADE evidence integration from time 1 to time 2 was associated with increased VsAN activity (especially confirm) and with decreased CEN activity (especially disconfirm) from time 1 to time 2 (Table 39), which were the networks/conditions that drove our single time point findings in study 2. Since patients' behavioural BADE evidence integration scores increased overall (poorer performance), this indicates that activation was further removed from normal levels in comparison to the patient versus control results described in the previous chapter; however, given the correlational nature of these findings, they also suggest that improvement in behavioural BADE evidence integration is associated with normalization of the hyperactivity in the VsAN and hypoactivity in the CEN. Interestingly, CEN activation changes in the confirm condition followed the pattern of the VsAN, supporting our interpretation in the previous chapter that increased BADE is associated with hypersalience to EVH matches, or a focus on confirmatory evidence. Activity in the VDMN also showed this distinct confirm/disconfirm pattern, but in reverse, likely due to the anti-correlated nature of the default-mode network and task-positive activations (Fox et al., 2005).

We also examined associations between behaviour and the BP and PB components of the hemodynamic response, which are known to index different cognitive operations (Lavigne et al., 2016). Increased behavioural BADE evidence integration was associated with increases in VsAN BP/PB and CEN PB, and with decreases in VDMN BP/PB and CEN BP from time 1 to time 2 (Table 47). These findings are in line with study 2, in which we showed that BADE evidence integration was associated with greater activity in the VsAN and decreased activity in the CEN, the latter being driven by CEN BP values. Like the previous study, investigation of these different aspects of the hemodynamic response revealed a distinction between the BP and PB values in the CEN, suggesting they represent different cognitive processes underlying evidence integration, likely related to the rate of initiation and suppression of the cognitive processes involved in evidence integration. Specifically, the BP aspect of the CEN is suppressed during activation of the visual attention network when it is not needed. In patients with delusions, as demonstrated in Chapter 3, there is a hypersuppression of the CEN around the same time that there is hyperactivation in the visual attention network, both of which were found to underlie impairments in evidence integration. The current findings suggest that activity in these aspects of the hemodynamic response in these networks fluctuates as a function of BADE evidence integration in the direction predicted by the patients versus control findings of the previous study (i.e., poorer behavioural BADE evidence integration leads to further hyperactivity in the VsAN and hypoactivity in CEN BP), providing further support for their role in evidence integration processes.

4.3.4. Symptoms

Decreased symptoms were related to decreases in VDMN intensity and increases in VsAN and CEN activation from time 1 to time 2, with positive symptoms (including delusions) more strongly predicting activity in the confirm conditions, and negative symptoms the disconfirm conditions (Table 43). These findings suggest that decreased symptoms lead to normalization of functional brain activity in the CEN, but lead to further VsAN hyperactivity/VDMN hypoactivity. This exacerbation at time 2 in the VsAN/VDMN following *decreased* symptoms (including delusions) is an unexpected finding, given that delusions were associated with hyperactivity in this network in the previous study, and decreased from time 1 to time 2. The BP/PB analysis was similar, in that decreases in symptoms were associated with decreased VDMN BP and PB and CEN PB activity, and increased VsAN BP and PB, and CEN BP activity from time 1 to time 2 (Table 51). As with the conditions analysis, the CEN (BP) findings were in line with expectations, but the VsAN/VDMN findings were not. One explanation for these contradictory results might concern the paradoxical increase in behavioural BADE evidence integration relative to decrease in delusions observed in this study. Since both BADE and symptoms were associated with functional brain activity, but increased BADE was more strongly related to activity in these networks than decreased delusions (0.81 and -0.58, respectively, see Tables 39 and 43), it may be influencing this result. This suggests a complex interplay between changes in symptoms and activity in these networks that may be affected by changes in behavioural BADE evidence integration. Future research should investigate the potential moderating role of BADE on symptoms and functional brain activity underlying evidence

integration in order to determine how BADE and delusions/symptoms interact to exert differential effects on the CEN and VsAN/VDMN.

4.3.5. Neurocognition

We also examined associations between changes in neurocognition and changes in functional brain activity across the three networks, focusing on the confirm and disconfirm conditions and baseline-peak values separately. Improvements in neurocognition were generally associated with normalization of functional brain activity, but were better at predicting changes in functional brain networks as a whole regardless of confirmatory/disconfirmatory evidence integration or BP/PB aspect of the hemodynamic response (Tables 41 and 49, respectively). This supports our findings from study 2 that behavioural BADE is separable from neurocognitive functioning, and that BADE is not simply a proxy for one or more neurocognitive processes, since BADE showed differential associations with activity underlying dis/confirmatory evidence integration and the baseline-peak aspects of the hemodynamic response, unlike neurocognition. Moreover, the current study demonstrated a general trend where improvements in attention led to increased activation in the VsAN (though this differed depending on the measure used), and improvements in memory led to increased activation in the CEN, which is in line with the cognitive operations expected to underlie these networks, and the stages of evidence integration (detection and integration) associated with these networks in study 1. Finally, although neurocognition and BADE predicted different configurations of functional brain networks than behavioural BADE across both types of analyses (HDR underlying dis/confirmatory evidence integration, and BP/PB), there were some similarities, indicating

that neurocognitive processes may partially contribute to behavioural BADE evidence integration and underlying functional brain activity, as was also noted in the previous study.

4.3.6. Limitations

One limitation of this study includes the low behavioural BADE evidence integration scores observed in patients at time 1, which would restrict the degree to which BADE could improve from time 1 to time 2. In addition, the small sample size (n=29 patients tested across two time points) limited our ability to create groups based on changes in delusions across time points. Similarly, our study may have been underpowered due to the small sample and relatively large number of variables, particularly for the follow-up behavioural analyses (correlations, behavioural CPCAs) on estimated hemodynamic response. Although the behavioural CPCAs generally supported the correlations, some findings were contrary to expectations and the correlations were not corrected for multiple comparisons. These findings should, therefore, be interpreted with caution. Although there were no significant changes in medication from time 1 to time 2, the effect of medication was not explicitly tested in the current study. Potential confounding effects of illness stage or duration were similarly not investigated. Given that both antipsychotic medications (Andreou et al., 2015) and illness duration (Eisenacher & Zink, 2017) can influence BADE behaviourally, these variables should be carefully considered in future research.

4.3.7. Conclusion

In the current study, we examined associations between changes in behavioural BADE, symptoms, neurocognition, and functional brain networks underlying evidence integration in schizophrenia patients. We replicated the functional brain networks underlying disconfirmatory evidence integration, namely, a VsAN involved in detection of the evidence,

148

attentional processes and responding, as well as a CEN involved in integrating evidence. Hyperactivity in the VsAN and hypoactivity in the CEN (especially from baseline-to-peak) was exacerbated with poorer behavioural BADE evidence integration from time 1 to time 2, indicating that activity in these networks fluctuates as a function of changes in BADE. Improvement in symptoms (including delusions) led to normalization of activity in the CEN, but further hyperactivity in the VsAN. Therefore, the effects of changes in delusions on functional brain activity (particularly in the VsAN) should be examined more thoroughly, including possible moderating variables, such as behavioural BADE. Future research should consider selecting patients for high levels of BADE evidence integration and/or greater delusion severity, and investigate potential roles of illness stage and medication use, in order to determine the degree to which these might influence changes in functional brain activity in these networks.

5. Conclusion

The goals of the current research were to (1) determine the cognitive and brain mechanisms underlying evidence integration; (2) identify which of these mechanisms underlies the bias against disconfirmatory evidence in schizophrenia and delusions; and to (3) examine associations between changes in BADE, delusions, and functional brain activity underlying disconfirmatory evidence integration in schizophrenia. In study 1, we identified sequentially-active visual attention and cognitive evaluation networks related to the detection and integration of disconfirmatory evidence in healthy individuals. Study 2 replicated these results in a novel sample of healthy controls and schizophrenia patients, and further showed that these networks were differentially related to both behavioural BADE evidence integration and delusions, associations which were separable from neurocognitive processes, such as memory and attention. In study 3, we found that poorer behavioural BADE evidence integration in patients over time was associated with exacerbation of the aberrant functional brain activity in these networks identified in study 2 (increased visual attention network activity and decreased cognitive evaluation network activity with poorer evidence integration), and that improvement in delusions led to normalization of activity in the cognitive evaluation network. Taken together, these findings highlight two functional brain networks (visual attention, cognitive evaluation) that underlie two cognitive biases (hypersalience of EVH matches, BADE) central to the maintenance of delusions in schizophrenia, and show that activity in these networks is distinct from neurocognitive functioning, and fluctuates as a function of BADE evidence integration and, to a lesser extent, symptoms. This work represents the most comprehensive research to date on functional brain networks underlying evidence integration and its relation to delusions,

cognitive biases, and neurocognitive functioning, and provides an empirical and neurobiological foundation for theoretical accounts of delusion maintenance in schizophrenia.

5.1. Study 1: Functional Brain Networks Underlying Detection and Integration of Disconfirmatory Evidence

Aim 1 (to identify cognitive and brain mechanisms underlying evidence integration) was assessed in healthy individuals by examining functional brain networks recruited during an evidence integration task (study 1, Chapter 2). We found a visual attention network and a cognitive evaluation network that were preferentially and sequentially activated during the integration of disconfirmatory evidence, and peaked following evidence presentation. The visual attention network, which included dACC and bilateral insula, peaked immediately following evidence presentation and was most active for disconfirmatory evidence, suggesting it reflects the detection of evidence, and the conflict between the evidence and belief in the case of disconfirmatory evidence. Many of the regions included in this network have been implicated in previous ROI-based research, particularly during uncertainty during belief revision (Behrens et al., 2007; Stern, Gonzalez, Welsh, & Taylor, 2010) and in the recognition of a stimulus during evidence accumulation (Krueger et al., 2017; Liu & Pleskac, 2011; Ploran et al., 2007). Moreover, the visual attention network identified here had nodes in common with the ventral attention, salience, dorsal attention and somatosensory networks (the latter primarily for studies 2 and 3) of resting-state studies. The ventral attention (Corbetta & Shulman, 2002; Yeo et al., 2011) and salience (Seeley et al., 2007) networks, which include the dACC and bilateral insula, are involved in orienting attention towards environmentally-salient stimuli (Corbetta & Shulman, 2002; Menon & Uddin, 2010), and the

salience network in particular has been shown to be important for modulating other large-scale networks, such as the dorsal attention and frontoparietal networks (Goulden et al., 2014; Menon & Uddin, 2010). The combination of these past findings with the timing of the hemodynamic response observed in the current research supports the notion that this network is recruited during evidence detection processes, which should be more involved during disconfirmatory relative to confirmatory evidence, as was observed in study 1.

The cognitive evaluation network (rostrolateral prefrontal/orbitofrontal cortex, inferior frontal gyrus, inferior parietal lobule) peaked much later in the trial than the visual attention network and was interpreted as being involved in evaluating and integrating the evidence into one's belief system. In line with this interpretation, several ROI-based studies have implicated nodes of this network in hypothesis evaluation and evidence integration (Christoff et al., 2003; Corlett et al., 2004; Sharot et al., 2012). In addition, the cognitive evaluation network has regions in common with the resting-state derived frontoparietal network (Yeo et al., 2011), which is involved in control processes and integrating information from other brain networks (Dosenbach et al., 2007; Vincent et al., 2008). Taken together, these findings suggest this network is involved in evaluating the evidence and integrating it into the belief.

Thus, our aim to identify the cognitive and brain mechanisms underlying disconfirmatory evidence integration led to the identification of two functional brain networks, a visual attention network and a cognitive evaluation network, which showed sequential activation following the presentation of evidence and were preferentially activated during disconfirmatory evidence integration. These networks can be related to two important cognitive processes underlying evidence integration and belief revision, namely, detection of the evidence for the visual attention network, and evaluation/integration of that evidence for

the cognitive evaluation network. This work lays the foundation for future research investigating evidence integration in clinical samples (for example, schizophrenia patients with delusions) by highlighting distinct cognitive processes and brain networks involved, one or both of which may underlie impairments in evidence integration.

5.2. Study 2: Functional Brain Networks Underlying Evidence Integration and Delusions in Schizophrenia

The second aim of this research was to identify the mechanism(s) underlying impaired evidence integration in schizophrenia, and specifically, the bias against disconfirmatory evidence observed in schizophrenia patients with delusions. In study 2 (Chapter 3), we partially replicated the networks identified in study 1 in healthy individuals, finding sequentially-active visual attention and cognitive evaluation networks that were preferentially active during disconfirmatory evidence integration. However, the visual attention network included more extensive recruitment of somatosensory and dorsal attention network (Yeo et al., 2011) regions, and, in conjunction with its earlier and sustained peak during the trial, likely represents additional cognitive processes occurring alongside evidence detection (e.g., attention, task demand, responding) that are grouped together into a single network due to limited temporal resolution. This complex network was interpreted as a combination of networks resulting from a similar timing of cognitive operations, which may be teased apart in future research by carefully modifying the experimental design to ensure separation of visual attention, responding, and evidence detection processes, and/or by including additional neuroimaging techniques with higher temporal resolution (e.g., electroencephalography) that could allow for separation of functional brain activity occurring in close temporal proximity.

Despite the complexity of this configuration of the visual attention network, it showed significant associations with both delusions and behavioural BADE. Specifically, schizophrenia patients with delusions showed hyperactivity in this network, and activity during confirmatory evidence integration was associated with behavioural BADE. These findings may relate to the hypersalience of EVH matches proposed to underlie delusions in schizophrenia, a cognitive bias referring to a tendency to accept a hypothesis when supporting evidence is provided (Balzan, Delfabbro, Galletly, & Woodward, 2012; Speechley et al., 2010). Hypersalience of EVH matches has been hypothesized to underlie BADE in behavioural studies, in the sense that earlier lure interpretations which confirm a belief will be viewed as sufficient, leading subsequent disconfirmatory evidence to carry less weight (Broyd et al., 2017; Sanford et al., 2014; Speechley et al., 2012). Hypersalience of EVH matches is a fairly novel finding in schizophrenia, and its neurobiological underpinnings have yet to be tested thoroughly. However, Whitman and colleagues (2013) observed hyperactivity in a dACC-related network in healthy controls during EVH matches, a network which had striking similarities to the visual attention network identified in the current study although it used a different cognitive task. Together, these findings suggest that hyperactivity in the visual attention network in schizophrenia patients with delusions may contribute to the tendency for schizophrenia patients with delusions to focus on evidence that confirms their beliefs, which may affect subsequent integration of disconfirmatory evidence.

In contrast to the visual attention network, activity in the cognitive evaluation network was attenuated in schizophrenia patients with delusions, particularly during the rise from baseline-to-peak, and was associated with behavioural BADE during disconfirmatory evidence integration only. This network was interpreted as being involved in evaluating and

integrating evidence, and hypoactivity in delusional patients might reflect another important aspect of delusion maintenance (Broyd et al., 2017; Freeman et al., 2002), the avoidance of disconfirmatory evidence. One of the regions in this network, rostrolateral prefrontal cortex, has been specifically proposed to underlie hypothesis evaluation impairments in delusions (Coltheart, 2010). This region, which shows aberrant prediction error signaling in schizophrenia patients with delusions (Corlett et al., 2007), was an important node of the cognitive evaluation network identified in the current study, and was also observed in the visual attention network, suggesting it may be part of multiple sub-systems related to belief evaluation and revision. The cognitive evaluation network also shows important similarities to the frontoparietal control network described in previous research (Dosenbach et al., 2007; Vincent et al., 2008), which supports its involvement in evaluating information. The hypoactivity in patients with delusions seen in this network, in conjunction with its association to behavioural BADE during disconfirmatory evidence integration specifically, suggests that it may contribute to delusional patients' tendency to avoid evidence that disconfirms their beliefs, and may serve as a neurobiological marker for BADE.

This work highlights distinct, sequentially active functional brain networks underlying two aspects of delusion maintenance in schizophrenia patients with delusions: the focus on obtaining confirmatory evidence, and the avoidance of disconfirmatory evidence. Theories of delusions in schizophrenia emphasize the importance of these two cognitive processes in delusion maintenance (e.g., Broyd et al., 2017; Freeman et al., 2002), and the current research shows, for the first time, alterations in functional brain networks that may underlie these cognitive biases and evidence integration deficits in schizophrenia patients with delusions.

5.3. Study 3: Changes in BADE, Delusions, and Functional Brain Activity Underlying Evidence Integration in Schizophrenia

The third aim of this research was to examine associations between changes in behavioural BADE, symptoms, and functional brain activity underlying disconfirmatory evidence integration in schizophrenia. Patients were assessed on symptoms, behavioural BADE, neurocognition, and functional brain activity during an evidence integration task at two time points at least eight weeks apart. We identified similar visual attention and cognitive evaluation networks underlying disconfirmatory evidence integration as in the previous two studies, supporting our interpretation that the visual attention and cognitive evaluation networks represent distinct aspects of evidence integration, and are associated with the focus on confirmatory, and avoidance of disconfirmatory, evidence, respectively. We found that poorer behavioural BADE evidence integration from time 1 to time 2 was associated with greater hyperactivity in the visual attention network, and greater hypoactivity in the baseline-to-peak portion of the cognitive evaluation network, both of which drove the aberrant functional brain activity observed in delusional schizophrenia patients in study 2. This is the first evidence that functional brain activity underlying evidence integration fluctuates as a function of BADE. With regard to symptoms, the hypothesis was supported for the cognitive evaluation network, but showed the opposite relation within the visual attention network. That is, decreased symptoms (including delusions) were associated with normalization of the cognitive evaluation network (particularly the baseline-to-peak aspect of the hemodynamic response), but with further hyperactivation in the VsAN. Paradoxically, patients' symptoms decreased while behavioural BADE evidence integration increased from time 1 to time 2, suggesting there may be complex associations between changes in BADE,

156

symptoms, and functional brain activity underlying evidence integration, although changes in BADE and delusions were not directly correlated. Future research should investigate this further by selecting patients for a greater bias in disconfirmatory evidence (as BADE was mild in study 3), and examining potential moderating roles of illness stage and duration, as well as medications.

5.4. Neurocognition and BADE

In studies 2 and 3, we addressed a secondary aim: whether behavioural BADE evidence integration and functional brain activity underlying evidence integration were distinct from neurocognitive functioning. We consistently found that although neurocognition was associated with these measures, it could not fully explain either behavioural BADE or functional brain activity underlying an evidence integration task. Behaviourally, we replicated previous findings in healthy individuals (Woodward et al., 2007) showing that BADE makes up a distinct component from neurocognition, extending these findings to a patient sample and additional neurocognitive measures, as well as using the more recently-developed BADE components rather than items. We also found no significant associations between changes in behavioural BADE and changes in neurocognition, suggesting that increases/decreases in each of these measures over time are independent from one another. In terms of functional brain activity, neurocognitive measures often predicted different configurations of functional brain networks (components were often clustered by networks rather than conditions or baseline-peak values). However, the similarities that emerged indicated that neurocognition does play a role in functional brain activity underlying evidence integration. This partial overlap is not unexpected, given that evidence integration requires typical neurocognitive processes such as memory and attention, and that these

processes are also implicated in our interpretations of the visual attention and cognitive evaluation networks. Therefore, these findings show that, similar to behavioural studies on BADE and neurocognition (Eisenacher & Zink, 2017), functional brain activity underlying BADE and evidence integration may be influenced, but are not entirely predicted, by neurocognition.

5.5. Conclusion

The purpose of this research was to develop a comprehensive understanding of (disconfirmatory) evidence integration in schizophrenia, how it might relate to delusions, and its underlying brain mechanisms. We first identified distinct, sequentially-active functional brain networks (visual attention, cognitive evaluation) associated with disconfirmatory evidence integration in healthy individuals, showing they could be related to detection and integration stages of evidence integration. Then, we observed that these networks were differentially affected in schizophrenia patients with delusions, and distinctly related to the bias against disconfirmatory evidence, suggesting that hyperactivity in the visual attention network reflected the tendency to focus on confirmatory evidence (i.e., hypersalience of EVH matches), and that hypoactivity in the cognitive evaluation network reflected the tendency to avoid disconfirmatory evidence (i.e., BADE), theorized to underlie delusion maintenance in schizophrenia. Finally, we found that functional brain activity in these networks fluctuates as a function of BADE (poorer evidence integration associated with more hyperactivation in the visual attention network and more hypoactivation in the cognitive evaluation network) and, to a lesser extent, symptoms (decreased symptoms associated with normalization of the cognitive evaluation, but not the visual attention, network), and that evidence integration was (partially) separable from neurocognition both behaviourally and

neurobiologically. This work provides the first evidence of differential associations between functional brain networks and two cognitive biases underlying the maintenance of delusions in schizophrenia (hypersalience of EVH matches and BADE), and highlights the complex interplay between delusions, BADE, and functional brain activity underlying evidence integration, and how these might interact to produce evidence integration deficits in schizophrenia patients with delusions.

Tables

Table 1. Study 1: Anatomical descriptions for the most extreme 10% of cognitive evaluation network loadings (Component 1), with cluster volumes, Montreal Neurological Institute (MNI) coordinates, and Brodmann's area (BA) for the peaks within each cluster.

Brain Regions	Cluster Volume (voxels)	BA for Peak Locations	MNI Coordinate for Peak Locations		
			x	y	z
Positive Loadings					
Cluster 1: Bilateral	14271				
Cerebellum Crus I		n/a	40	-76	-24
Lateral occipital cortex, inferior division		19	38	-88	-16
Occipital pole		17/18	26	-100	2
Cerebellum Crus I		n/a	-18	-82	-30
Lateral occipital cortex, inferior division		19	-40	-80	-22
Occipital pole		17	-22	-102	0
Occipital pole		18	-34	-94	-14
Cerebellum Crus II		n/a	-38	-64	-50
Cerebellum V		n/a	16	-54	-22
Cerebellum VI		n/a	-32	-44	-40
Cerebellum VIIIa		n/a	-30	-40	-42
Cluster 2: Bilateral	11626				
Supramarginal gyrus, posterior division		40	46	-44	58
Lateral occipital cortex, superior division		7	14	-70	64
Frontal orbital cortex		10	40	62	-2
Superior parietal lobule		40	-40	-46	62
Middle frontal gyrus		8	30	14	60
Frontal pole		46	44	52	-12
Frontal pole		45	42	42	28
Superior frontal gyrus		6	-6	-2	78
Postcentral gyrus		2	-56	-26	50
Lateral occipital cortex, superior division		7	-20	-68	60
Supramarginal gyrus, anterior division		2	-54	-30	52
Precentral gyrus		6	-30	-22	72
Superior frontal gyrus		6	30	4	66
Middle frontal gyrus		9/46	42	30	44
Middle frontal gyrus		9	44	28	46
Superior frontal gyrus		8	16	20	66
Postcentral gyrus		3	-32	-36	70
Precentral gyrus		4	-2	-24	82

(Table 1 continued)

Brain Regions	Cluster Volume (voxels)	BA for Peak Locations	MNI Coordinate for Peak Locations		
			x	y	z
Cluster 3: Left Hemisphere	380				
Frontal pole		10	-32	64	2
Frontal orbital cortex		11	-30	64	-8
Frontal orbital cortex		47	-40	50	-14
Cluster 4: Right Hemisphere	355				
Inferior frontal gyrus/pars opercularis		38	54	20	-2
Inferior frontal gyrus/pars opercularis		6	56	12	10
Cluster 5: Left Hemisphere	343				
Inferior frontal gyrus/pars opercularis		38	-52	18	-4
Cluster 6: Left Hemisphere	5				
Middle frontal gyrus		9	-42	24	46

Note. No negative loadings passed threshold; only clusters with volumes greater than or equal to 5 voxels are displayed.

Table 2. Study 1: Anatomical descriptions for the most extreme 10% of visual/default-mode network loadings (Component 2), with cluster volumes, Montreal Neurological Institute (MNI) coordinates, and Brodmann's area (BA) for the peaks within each cluster.

Brain Regions	Cluster Volume (voxels)	BA for Peak Locations	MNI Coordinate for Peak Locations		
			x	y	z
Positive Loadings					
Cluster 1: Bilateral	19622				
Occipital fusiform gyrus		18	26	-78	-14
Occipital pole		17	16	-100	14
Lingual gyrus		17	2	-86	-10
Occipital pole		17	-12	-100	2
Occipital fusiform gyrus		18	-22	-80	-16
Occipital pole		18	24	-94	12
Lateral occipital cortex, superior division		7	28	-64	48
Lateral occipital cortex, superior division		19	28	-72	32
Superior parietal lobule		7	-26	-62	46
Cluster 2: Right Hemisphere	208				
Middle frontal gyrus		44	48	12	34
Cluster 3: Left Hemisphere	61				
Middle frontal gyrus		44	-44	6	34
Cluster 4: Right Hemisphere	39				
Frontal pole/middle frontal gyrus		45	48	36	32
Negative Loadings					
Cluster 1: Bilateral	2479				
Ventromedial prefrontal cortex		10	-4	52	-2
Ventromedial prefrontal cortex		10	0	62	-2
Cluster 2: Left Hemisphere	1036				
Lateral occipital cortex, superior division		39	-50	-74	26
Lateral occipital cortex, superior division		37	-60	-64	14
Cluster 3: Right Hemisphere	924				
Parietal operculum cortex		48	54	-30	22
Central operculum cortex		48	58	-2	6
Cluster 4: Bilateral	851				
Precuneus cortex		23	-2	-62	26
Cluster 5: Left Hemisphere	516				
Parietal operculum cortex		42	-60	-32	20
Supramarginal gyrus, posterior division		40	-64	-44	36
Cluster 6: Left Hemisphere	428				
Middle temporal gyrus, anterior division		21	-54	-6	-16

(Table 2 continued)

Brain Regions	Cluster Volume (voxels)	BA for Peak Locations	MNI Coordinate for Peak Locations		
			x	y	z
Cluster 7: Right Hemisphere	268				
Middle temporal gyrus, anterior division		21	56	-6	-20
Cluster 8: Left Hemisphere	237				
Superior frontal gyrus		9	-24	32	46
Cluster 9: Bilateral	189				
Cingulate gyrus, posterior division		23	2	-24	46
Cluster 10: Right Hemisphere	91				
Postcentral gyrus		3	28	-38	66
Cluster 11: Right Hemisphere	16				
Angular gyrus		39	58	-66	18

Note. Only clusters with volumes greater than or equal to 5 voxels are displayed.

Table 3. Study 1: Anatomical descriptions for the most extreme 10% of response network loadings (Component 4), with cluster volumes, Montreal Neurological Institute (MNI) coordinates, and Brodmann's area (BA) for the peaks within each cluster.

Brain Regions	Cluster Volume (voxels)	BA for Peak Locations	MNI Coordinate for Peak Locations		
			x	y	z
Positive Loadings					
Cluster 1: Bilateral	20765				
Precentral gyrus		4	-40	-20	58
Postcentral gyrus		3	-54	-22	50
Lateral occipital cortex, superior division		7	-26	-60	50
Lateral occipital cortex, inferior division		18	-30	-88	6
Lateral occipital cortex, inferior division		19	-32	-88	0
Occipital fusiform gyrus		18	22	-86	-12
Cerebellum VI		n/a	32	-56	-22
Temporal occipital fusiform cortex		19	-40	-64	-16
Occipital fusiform gyrus		18	-20	-90	-12
Temporal occipital fusiform cortex		37	-38	-54	-20
Lateral occipital cortex, inferior division		18	32	-88	6
Lateral occipital cortex, superior division		19	-26	-74	28
Precentral gyrus		6	-52	4	38
Lateral occipital cortex, superior division		19	28	-74	28
Lateral occipital cortex, superior division		7	24	-64	52
Cerebellum VI		18	10	-74	-22
Cluster 2: Bilateral	1428				
Supplementary motor area		6	-4	8	54
Cluster 3: Right Hemisphere	491				
Middle frontal gyrus		6	28	-2	54
Cluster 4: Right Hemisphere	222				
Precentral gyrus		44	52	8	32
Cluster 5: Left Hemisphere	78				
Thalamus		n/a	-10	-18	8
Cluster 6: Left Hemisphere	46				
Central opercular cortex		48	-48	-22	20
Cluster 7: Right Hemisphere	12				
Supramarginal gyrus, anterior division		2	46	-32	44

(Table 3 continued)

Brain Regions	Cluster Volume (voxels)	BA for Peak Locations	MNI Coordinate for Peak Locations		
			x	y	z
Negative Loadings					
Cluster 1: Bilateral	1869				
Frontal pole		10	-4	60	22
Cluster 2: Left Hemisphere	1356				
Lateral occipital cortex, superior division		39	-50	-74	34
Cluster 3: Bilateral	201				
Cuneal cortex		18	4	-86	24
Cluster 4: Left Hemisphere	167				
Middle temporal gyrus, posterior division		21	-60	-12	-16
Cluster 5: Right Hemisphere	145				
Lateral occipital cortex, superior division		39	54	-70	32
Cluster 6: Left Hemisphere	91				
Precuneus cortex		23	-8	-52	34
Cluster 7: Left Hemisphere	64				
Middle temporal gyrus, temporooccipital part		37	-64	-56	-4
Cluster 8: Left Hemisphere	47				
Frontal pole		9	-16	42	52

Note. Only clusters with volumes greater than or equal to 5 voxels are displayed.

Table 4. Study 1: Anatomical descriptions for the most extreme 10% of visual attention network loadings (Component 5), with cluster volumes, Montreal Neurological Institute (MNI) coordinates, and Brodmann's area (BA) for the peaks within each cluster.

Brain Regions	Cluster Volume (voxels)	BA for Peak Locations	MNI Coordinate for Peak Locations		
			x	y	z
Positive Loadings					
Cluster 1: Bilateral	16843				
Occipital pole		18	18	-92	-8
Occipital fusiform gyrus		19	34	-78	-14
Occipital pole		18	-28	-96	4
Occipital fusiform gyrus		18	-20	-90	-12
Occipital fusiform gyrus		19	-36	-78	-14
Occipital pole		18	32	-90	8
Temporal occipital fusiform cortex		37	40	-60	-18
Occipital pole		18	-32	-92	-6
Lateral occipital cortex, superior division		19	30	-72	32
Superior parietal lobule		7	34	-56	48
Lateral occipital cortex, superior division		7	-30	-62	44
Supramarginal gyrus, posterior division		40	48	-42	46
Lateral occipital cortex, superior division		19	-28	-74	30
Supramarginal gyrus, posterior division		40	-46	-40	48
Cluster 2: Right Hemisphere	4694				
Middle frontal gyrus		44	48	12	34
Middle frontal gyrus		6	38	8	62
Insular cortex		47	32	26	-2
Cluster 3: Left Hemisphere	2329				
Middle frontal gyrus		44	-46	10	32
Middle frontal gyrus		45	-50	32	24
Middle frontal gyrus		6	-42	6	60
Cluster 4: Bilateral	1676				
Superior frontal gyrus		8	2	28	48
Cluster 5: Left Hemisphere	160				
Insular cortex		47	-30	24	-4
Cluster 6: Right Hemisphere	75				
Frontal pole		10	30	58	10
Cluster 7: Left Hemisphere	31				
Frontal pole		46	-44	48	-2

(Table 4 continued)

Brain Regions	Cluster Volume (voxels)	BA for Peak Locations	MNI Coordinate for Peak Locations		
			x	y	z
Negative Loadings					
Cluster 1: Bilateral	684				
Cuneal cortex		18	6	-84	26
Cluster 2: Bilateral	425				
Frontal medial cortex		11	-2	50	-8
Cluster 3: Right Hemisphere	36				
Lingual gyrus		18	10	-66	-4
Cluster 4: Right Hemisphere	28				
Lingual gyrus		37	32	-52	2

Note. Only clusters with volumes greater than or equal to 5 voxels are displayed.

Table 5. Study 2: Descriptive information per group. (standard deviations in parentheses)

Variable	Control	Non-Delusional	Delusional
N	40	31	27
Gender (male:female)	19:21	20:11	15:12
Handedness (R:M:L)	38:0:2	29:1:1	26:0:1
Age (mean (SD))	35.18 (12.64)	32.23 (9.63)	38.78 (10.07)
Education (years)[ab]*	16.83 (2.76)	14.21 (1.94)	14.78 (2.55)
ToPF	112.45 (11.80)	108.74 (15.22)	108.07 (13.67)
WASI[ab]*	111.28 (13.33)	100.10 (9.39)	96.93 (9.31)
BDI[cde]*	3.74 (3.52)	10.38 (6.54)	16.58 (11.01)
Chlorpromazine equivalents	-	748.07 (2008.72)	509.58 (888.71)

Note. BDI = Beck Depression Inventory (maximum = 63); L = left; M = mixed; R = right; SD = standard deviation; ToPF = Test of Premorbid Functioning; WASI = Wechsler Abbreviated Scale of Intelligence; [a] = controls > non-delusional; [b] controls > delusional; [c] = controls < non-delusional; [d] = controls < delusional; [e] = delusional > non-delusional; * = $p < .001$.

Table 6. Study 2: Descriptive information for symptom measures (patients only) and Schizotypal Personality Questionnaire (controls only).

Symptoms (Patients)	Non-Delusional		Delusional	
	Mean (SD)	Range	Mean (SD)	Range
PSYRATS hallucinations*	5.81 (10.35)	0-30	12.96 (15.15)	0-39
PSYRATS delusions**	4.33 (4.29)	0-11	15.89 (2.82)	12-22
SAPS total**	7.10 (7.99)	0-32	21.15 (11.93)	0-43
Hallucinations*	2.17 (3.49)	0-10	6.15 (7.23)	0-24
Delusions**	3.00 (4.00)	0-16	10.89 (5.71)	0-23
Bizarre behaviour*	0.48 (1.29)	0-5	1.33 (1.82)	0-6
Formal thought disorder	1.52 (2.93)	0-10	2.78 (3.63)	0-14
SANS total	19.03 (14.42)	0-52	15.74 (11.69)	0-46
Affective flattening	5.84 (7.36)	0-24	3.67 (4.45)	0-18
Alogia	1.94 (2.95)	0-11	0.96 (1.97)	0-6
Avolition-apathy	3.68 (2.84)	0-10	2.78 (2.79)	0-8
Anhedonia-asociality	4.90 (4.65)	0-14	5.58 (4.78)	0-16
Attention	1.71 (1.75)	0-5	1.81 (1.75)	0-6
Inappropriate affect	0.52 (1.09)	0-4	0.37 (0.97)	0-4
SPQ (Controls)	**Mean (SD)**	**Range**	**Maximum**	
Total	152.70 (34.70)	78-242	370	
Factors				
Social anxiety	22.70 (6.96)	12-39	60	
Social anhedonia	26.73 (6.81)	13-41	50	
Eccentricity/oddity	28.10 (8.74)	12-48	60	
Mistrust	22.33 (7.28)	12-42	60	
Unusual beliefs and experiences	18.05 (5.38)	12-31	60	
Subscales				
Ideas of reference	16.35 (5.01)	9-29	45	
Excessive social anxiety	21.23 (6.09)	8-32	40	
Magical ideation	10.20 (3.44)	7-20	35	
Unusual perceptual experiences	14.80 (4.83)	9-31	45	
Odd behaviour	15.03 (5.75)	7-27	35	
No close friends	21.30 (5.27)	12-33	45	
Odd speech	20.83 (6.77)	9-37	45	
Constricted affect	17.45 (4.98)	8-27	40	
Suspiciousness	15.53 (5.35)	8-28	40	

Note. PSYRATS = Psychotic Symptoms Rating Scales; SANS = Scales for the Assessment of Negative Symptoms; SAPS = Scales for the Assessment of Positive Symptoms; SD = standard deviation; SPQ = Schizotypal Personality Questionnaire; * = $p < .05$; ** = $p < .001$.

Table 7. Study 2: Descriptive information for neurocognitive variables.

Variable	Non-Delusional		Delusional	
	Mean (SD)	Range	Mean (SD)	Range
RBANS				
Immediate memory	83.26 (20.71)	44-140	88.37 (13.24)	57-114
Visuospatial	103.68 (16.52)	69-126	105.15 (14.65)	78-121
Language	84.65 (17.54)	44-120	85.81 (20.99)	40-116
Attention	84.87 (15.08)	49-118	83.89 (15.71)	53-112
Delayed memory	84.16 (17.52)	44-112	88.78 (14.82)	48-114
TMT A	38.16 (15.17)	7-79	36.33 (12.79)	9-66
TMT B	41.68 (12.39)	19-76	40.89 (10.37)	25-62
LNS	9.45 (2.54)	6-19	8.44 (1.83)	4-14
COWAT	44.03 (10.89)	23-64	45.63 (11.70)	20-64

Note. No significant differences observed between groups. COWAT = Controlled Oral Word Association Test; LNS = Letter-Number Sequencing; RBANS = Repeatable Battery for the Assessment of Neuropsychological Status; SD = standard deviation; TMT = Trail-Making Test.

Table 8. Study 2: Behavioural BADE rotated component loadings.

Item	Component	
	Conservatism	**Evidence Integration**
Absurd 1	0.31	**0.85**
Absurd 2	0.13	**0.92**
Absurd 3	0.03	**0.94**
Neutral lure 1	**0.94**	0.17
Neutral lure 2	**0.85**	**0.37**
Neutral lure 3	0.41	**0.78**
Emotional lure 1	**0.90**	0.28
Emotional lure 2	**0.81**	**0.45**
Emotional lure 3	0.21	**0.87**
True 1	**0.88**	0.35
True 2	**0.92**	0.13
True 3	**0.58**	**-0.34**

Note. Component loadings in bold font correspond to those identified in Sanford et al. (2014) and Speechley et al. (2012).

Table 9. Study 2: Means and standard deviations (in parentheses) per group for behavioural BADE and fMRI evidence integration tasks.

Variable	Control Mean (SD)	Non-Delusional Mean (SD)	Delusional Mean (SD)
Behavioural BADE			
Evidence integration[ab]**	13.88 (55.58)	22.93 (71.89)	81.87 (108.01)
Conservatism[b]*	318.52 (109.87)	279.26 (135.91)	378.84 (122.15)
fMRI Performance			
Accuracy (%)[c]*	78.59 (10.14)	72.22 (11.50)	71.34 (12.73)
Accuracy – confirm	88.69 (15.68)	83.87 (13.54)	80.93 (17.67)
Accuracy – disconfirm[d]*	68.50 (12.48)	60.56 (12.95)	61.76 (14.32)
Reaction time (ms)[ae]**	1294.45 (227.39)	1505.73 (272.45)	1589.53 (247.52)
Reaction time – confirm[ae]**	1220.97 (239.65)	1436.54 (278.20)	1539.85 (272.56)
Reaction time – disconfirm[ae]**	1367.92 (233.50)	1436.54 (278.20)	1639.20 (237.54)
Missed responses	1.66 (4.13)	2.16 (4.08)	3.85 (7.95)

Note. BADE = bias against disconfirmatory evidence; ms = milliseconds; SD = standard deviation; [a] = delusional > controls; [b] = delusional > non-delusional; [c] = delusional < controls; [d] = non-delusional < controls; [e] = non-delusional > controls; * = $p < .05$; ** = $p < .005$.

Table 10. Study 2: Correlations between behavioural BADE components and behaviour.

Variable	BADE Evidence Integration	BADE Conservatism
Demographics/IQ (n = 96)		
Age	0.10	0.01
Education	-0.18	-0.17
Gender	-0.11	**-0.23***
Handedness	-0.02	0.00
ToPF	-0.15	0.00
WASI	**-0.32****	-0.16
Clinical		
BDI (n = 96)	**0.26***	**0.22***
PSYRATS (patients only; n = 55)		
Hallucinations	0.22	0.14
Delusions	**0.34***	**0.29***
SAPS (patients only; n = 56)		
Total	**0.32***	0.22
Hallucinations	0.22	0.23
Delusions	0.19	0.17
Bizarre behaviour	0.16	0.14
Formal thought disorder	**0.34***	0.01
SANS (patients only; n = 48-56)		
Total	0.04	-0.05
Affective flattening	-0.06	-0.10
Alogia	0.05	-0.11
Avolition-apathy	0.02	0.05
Anhedonia-asociality	0.08	0.10
Attention	0.12	0.04
Inappropriate affect	0.18	0.03
SPQ (controls only; n = 40)		
Total	0.07	-0.05
Factors		
Social anxiety	0.17	-0.28
Social anhedonia	-0.10	-0.23
Eccentricity/oddity	0.14	0.08
Mistrust	0.01	-0.07
Unusual beliefs and experiences	**0.35***	0.13
Subscales		
Ideas of reference	0.26	0.19
Excessive social anxiety	-0.08	-0.17
Magical ideation	**0.39***	0.24
Unusual perceptual experiences	0.21	0.06
Odd behaviour	0.08	0.01

(Table 10 continued)

Variable	BADE Evidence Integration	BADE Conservatism
SPQ subscales (continued)		
No close friends	-0.17	-0.28
Odd speech	0.17	0.12
Constricted affect	-0.11	-0.21
Suspiciousness	-0.14	-0.22
Neurocognition (patients only; n = 58)		
RBANS immediate memory	0.06	0.06
RBANS visuospatial	-0.01	0.03
RBANS language	0.05	0.04
RBANS attention	-0.09	-0.09
RBANS delayed memory	0.12	0.20
TMT A	0.06	0.14
TMT B	-0.06	0.01
LNS	-0.24	-0.10
COWAT	-0.04	0.09
fMRI Performance (n = 96)		
Accuracy	**-0.29****	**-0.21***
Accuracy – confirm	**-0.24***	**-0.21***
Accuracy – disconfirm	**-0.21***	-0.13
Reaction time	**0.29****	-0.02
Reaction time – confirm	**0.28****	0.00
Reaction time – disconfirm	**0.29****	-0.04
Missing responses	0.17	**0.24***

Note. Significant correlations are set in bold text. BADE = bias against disconfirmatory evidence; BDI = Beck Depression Inventory; COWAT = Controlled Oral Word Association Test; IQ = intelligence quotient; LNS = Letter-Number Sequencing; PSYRATS = Psychotic Symptoms Rating Scales; RBANS = Repeatable Battery for the Assessment of Neuropsychological Status; SANS/SAPS = Scales for the Assessment of Negative/Positive Symptoms; SPQ = Schizotypal Personality Questionnaire; TMT = Trail-Making Test; ToPF = Test of Premorbid Functioning; WASI = Wechsler Abbreviated Scale of Intelligence; $* = p < .05$; $** = p < .01$.

Table 11. Study 2: Rotated component loadings for principal component analysis on behavioural BADE and neurocognitive measures.

Variable	Attention	Memory	BADE	Visuospatial
Behavioural BADE				
Evidence integration	-0.08	0.05	**0.89**	0.01
Conservatism	0.04	0.06	**0.89**	-0.02
Neurocognition				
RBANS				
Immediate memory	0.23	**0.88**	-0.01	-0.06
Visuospatial	0.13	0.11	-0.02	**0.90**
Language	**0.51**	0.49	0.03	0.02
Attention	**0.78**	0.02	-0.09	0.34
Delayed memory	0.09	**0.84**	0.14	0.24
TMT A	**0.74**	0.24	0.15	0.26
TMT B	**0.77**	0.18	-0.01	0.24
LNS	**0.71**	0.06	-0.25	-0.23
COWAT	**0.70**	0.17	0.06	-0.10

Note. All loadings greater than 0.50 are set in bold type. BADE = bias against disconfirmatory evidence; COWAT = Controlled Oral Word Association Test; LNS = Letter-Number Sequencing; RBANS = Repeatable Battery for the Assessment of Neuropsychological Status; TMT = Trail-Making Test.

Table 12. Study 2: Anatomical descriptions for the most extreme 10% of visual attention network loadings (Component 1), with cluster volumes, Montreal Neurological Institute (MNI) coordinates, and Brodmann's area (BA) for the peaks within each cluster.

Brain Region	Cluster Volume (Voxels)	BA (Peak Locations)	MNI Coordinates		
			x	y	z
Positive Loadings					
Cluster 1: Bilateral	24246				
Lateral occipital cortex, superior division		18	30	-86	12
Occipital pole		18	-26	-90	6
Lateral occipital cortex, inferior division		18	-28	-88	8
Occipital fusiform gyrus		19	-40	-70	-12
Occipital fusiform gyrus		18	22	-82	-10
Occipital fusiform gyrus		19	38	-70	-14
Occipital pole		18	16	-94	4
Lateral occipital cortex, superior division		7	26	-64	48
Lateral occipital cortex, superior division		19	30	-72	30
Superior parietal lobule		7	30	-56	50
Temporal occipital fusiform cortex		37	-34	-52	-20
Lateral occipital cortex, superior division		19	-26	-72	28
Temporal occipital fusiform cortex		37	34	-46	-22
Lateral occipital cortex, superior division		7	-24	-66	48
Precentral gyrus		44	-48	6	34
Supramarginal gyrus, anterior division		40	-40	-38	42
Cerebellum VI		n/a	-8	-74	-24
Supramarginal gyrus, anterior division		2	46	-34	46
Cerebellum VI		n/a	8	-74	-22
Middle frontal gyrus		6	-28	-4	50
Precentral gyrus		4	-40	-16	62
Cerebellum VIIb		n/a	12	-74	-44
Cerebellum Crus II		n/a	10	-74	-40
Precentral gyrus		6	-34	-20	66
Cerebellum VIIb		n/a	-22	-70	-48
Inferior frontal gyrus, pars triangularis		48	-42	26	22
Cluster 2: Right Hemisphere	1345				
Precentral gyrus		44	48	8	32
Middle frontal gyrus		45	46	30	20
Cluster 3: Bilateral	439				
Paracingulate gyrus		32	2	16	50
Cluster 4: Right hemisphere	435				
Middle frontal gyrus		6	36	2	56

(Table 12 continued)

Brain Region	Cluster Volume (Voxels)	BA (Peak Locations)	MNI Coordinates		
			x	y	z
Cluster 5: Left Hemisphere	40				
Insular cortex		47	-30	24	0
Cluster 6: Right Hemisphere	38				
Insular cortex		47	32	26	0
Cluster 7: Left Hemisphere	8				
Thalamus		n/a	-10	-18	8

Note. No negative loadings passed threshold; only clusters with volumes greater than or equal to 5 voxels are displayed.

Table 13. Study 2: Anatomical descriptions for the most extreme 10% of visual/default-mode network loadings (Component 2), with cluster volumes, Montreal Neurological Institute (MNI) coordinates, and Brodmann's area (BA) for the peaks within each cluster.

Brain Region	Cluster Volume (Voxels)	BA (Peak Locations)	MNI Coordinates x	y	z
Positive Loadings					
Cluster 1: Left Hemisphere	2625				
Occipital pole		18	-28	-90	8
Occipital fusiform gyrus		18	-20	-86	-12
Occipital fusiform gyrus		19	-38	-76	-10
Temporal occipital fusiform cortex		37	-34	-50	-18
Lateral occipital cortex, superior division		19	-26	-70	28
Cluster 2: Right Hemisphere	2127				
Lateral occipital cortex, inferior division		18	30	-86	8
Occipital fusiform gyrus		18	22	-82	-10
Occipital fusiform gyrus		19	36	-72	-12
Occipital pole		18	16	-96	6
Temporal occipital fusiform cortex		37	34	-44	-20
Cluster 3: Right Hemisphere	291				
Lateral occipital cortex, superior division		19	28	-70	32
Lateral occipital cortex, superior division		7	26	-64	50
Superior parietal lobule		7	30	-56	50
Cluster 4: Right Hemisphere	63				
Precentral gyrus		44	46	8	28
Cluster 5: Left Hemisphere	58				
Lateral occipital cortex, superior division		7	-22	-68	46
Superior parietal lobule		7	-26	-58	48
Cluster 6: Left Hemisphere	8				
Precentral gyrus		44	-44	4	30
Negative Loadings					
Cluster 1: Bilateral	11073				
Cuneal cortex		18	-2	-76	28
Cuneal cortex		18	6	-82	28
Precuneus cortex		23	-2	-54	34
Lingual gyrus		18	10	-72	-2
Lingual gyrus		18	-10	-70	-6
Intracalcarine cortex		17	-8	-80	2
Cingulate gyrus, posterior division		23	2	-22	42
Lingual gyrus		19	16	-46	-6
Lingual gyrus		19	-20	-54	2

(Table 13 continued)

Brain Region	Cluster Volume (Voxels)	BA (Peak Locations)	MNI Coordinates x	y	z
Cluster 2: Right Hemisphere	3219				
Lateral occipital cortex, superior division		39	48	-68	32
Planum temporale		22	64	-26	12
Middle temporal gyrus, anterior division		21	58	-6	-16
Middle temporal gyrus, posterior division		21	62	-10	-10
Planum temporale		42	58	-28	14
Parietal operculum cortex		48	52	-30	20
Superior temporal gyrus, posterior division		21	66	-32	4
Planum polare		48	58	0	2
Heschl's gyrus		48	52	-18	6
Cluster 3: Bilateral	3048				
Frontal pole		10	0	56	-2
Cingulate gyrus, anterior division		32	-2	40	8
Frontal pole		10	-18	60	16
Cluster 4: Left Hemisphere	2343				
Lateral occipital cortex, superior division		19	-42	-74	36
Lateral occipital cortex, superior division		39	-52	-66	22
Planum temporale		22	-62	-30	12
Superior temporal gyrus, posterior division		22	-66	-40	6
Middle temporal gyrus, temporooccipital part		21	-64	-48	4
Middle temporal gyrus, temporooccipital part		22	-64	-44	4
Cluster 5: Left Hemisphere	785				
Superior frontal gyrus		9	-24	32	46
Cluster 6: Right Hemisphere	422				
Superior frontal gyrus		9	24	34	46
Cluster 7: Left Hemisphere	311				
Middle temporal gyrus, anterior division		21	-58	-8	-14
Cluster 8: Right Hemisphere	133				
Superior parietal lobule		2	24	-42	66
Cluster 9: Right Hemisphere	16				
Precentral gyrus		6	46	-14	50
Cluster 10: Left Hemisphere	12				
Planum temporale		41	-42	-36	14
Cluster 11: Right Hemisphere	8				
Postcentral gyrus		4	24	-30	64
Cluster 12: Right Hemisphere	6				
Precentral gyrus		4	38	-16	42

Note. Only clusters with volumes greater than or equal to 5 voxels are displayed.

Table 14. Study 2: Anatomical descriptions for the most extreme 10% of cognitive evaluation network loadings (Component 3), with cluster volumes, Montreal Neurological Institute (MNI) coordinates, and Brodmann's area (BA) for the peaks within each cluster.

Brain Region	Cluster Volume (Voxels)	BA (Peak Locations)	MNI Coordinates		
			x	y	z
Positive Loadings					
Cluster 1: Right Hemisphere	6964				
Occipital pole		18	30	-94	4
Lateral occipital cortex, inferior division		19	38	-88	-4
Occipital fusiform cortex		19	30	-78	-20
Middle temporal gyrus, posterior division		21	64	-36	-6
Cerebellum Crus I		n/a	14	-96	22
Cluster 2: Left Hemisphere	6241				
Occipital pole		18	-34	-92	0
Lateral occipital cortex, inferior division		19	-48	-72	-14
Cerebellum Crus I		n/a	-20	-82	-24
Cerebellum Crus II		n/a	-30	-74	-44
Cluster 3: Right Hemisphere	2790				
Lateral occipital cortex, superior division		7	38	-64	52
Cluster 4: Left Hemisphere	1438				
Angular gyrus		39	-54	-56	36
Lateral occipital cortex, superior division		39	-38	-64	46
Lateral occipital cortex, superior division		7	-36	-66	48
Angular gyrus		40	-42	-58	54
Cluster 5: Left Hemisphere	1362				
Frontal pole		46	-46	48	-6
Frontal orbital cortex		38	-42	22	-14
Inferior frontal gyrus, pars triangularis		48	-56	22	6
Cluster 6: Right Hemisphere	1362				
Middle frontal gyrus		9	48	20	46
Middle frontal gyrus		45	50	34	28
Cluster 7: Right Hemisphere	1072				
Frontal pole		46	46	48	-6
Frontal pole		10	38	60	2
Frontal orbital cortex		38	44	24	-14
Cluster 8: Left Hemisphere	630				
Middle frontal gyrus		9	-48	18	48
Cluster 9: Left Hemisphere	375				
Middle temporal gyrus, posterior division		21	-64	-38	-4

(Table 14 continued)

Brain Region	Cluster Volume (Voxels)	BA (Peak Locations)	MNI Coordinates		
			x	y	z
Negative Loadings					
Cluster 1: Left Hemisphere	4185				
Precentral gyrus		4	-38	-20	54
Supplementary motor area		6	-2	0	54
Precentral gyrus		6	-28	-8	54
Cluster 2: Right Hemisphere	96				
Precentral gyrus		6	30	-8	54
Cluster 3: Left Hemisphere	20				
Central opercular cortex		48	-50	-22	18
Cluster 4: Left Hemisphere	7				
Frontal pole		46	-30	38	30
Cluster 5: Left Hemisphere	7				
Precentral gyrus		6	-54	2	36

Note. Only clusters with volumes greater than or equal to 5 voxels are displayed.

Table 15. Study 2: Correlation table for associations between behaviour and activity underlying dis/confirmatory evidence integration.

Variable	VsAN Conf	VsAN Disc	VDMN Conf	VDMN Disc	CEN Conf	CEN Disc
Behavioural BADE (n = 98)						
Evidence integration	0.15	0.02	-0.11	-0.05	-0.12	**-0.24***
Conservatism	0.07	-0.07	0.02	0.10	-0.10	-0.11
fMRI Performance (n = 98)						
Accuracy	**-0.22***	-0.03	0.18	0.04	0.19	**0.25***
Accuracy – confirm	-0.17	0.01	0.18	0.05	0.10	0.14
Accuracy – disconfirm	-0.18	-0.06	0.10	0.01	**0.20***	**0.27***
Reaction time	0.13	0.02	-0.13	-0.06	-0.12	**-0.25***
Reaction time – confirm	0.10	0.00	-0.13	-0.07	-0.10	**-0.27***
Reaction time - disconfirm	0.15	0.05	-0.14	-0.07	-0.12	**-0.24***
Missed responses	0.19	0.05	-0.17	-0.12	-0.11	-0.17
Neurocognition (n = 58)						
RBANS						
Immediate memory	-0.04	-0.09	0.11	0.04	-0.05	-0.14
Visuospatial	-0.09	-0.14	0.02	0.11	0.01	-0.03
Language	-0.23	**-0.29***	0.08	0.14	-0.07	0.11
Attention	0.02	-0.06	0.12	0.22	**-0.29***	-0.06
Delayed memory	-0.07	-0.05	0.13	-0.11	-0.07	0.15
TMT A	-0.22	-0.25	0.24	0.20	-0.08	0.09
TMT B	-0.08	-0.19	0.13	0.19	-0.10	0.02
LNS	0.10	-0.02	-0.06	0.11	-0.06	-0.06
COWAT	-0.24	-0.06	**0.29***	0.07	-0.05	0.10
Symptoms (n = 57-58)						
PSYRATS hallucinations	0.20	0.11	-0.07	0.13	0.01	-0.19
PSYRATS delusions	0.08	0.07	-0.17	-0.14	0.01	-0.09
SAPS						
Total	0.20	0.18	-0.13	-0.03	-0.05	-0.21
Hallucinations	0.24	0.18	-0.15	0.00	-0.02	-0.20
Delusions	0.12	0.08	-0.10	-0.02	0.02	-0.09
Bizarre behaviour	0.01	0.07	-0.05	-0.12	-0.01	-0.07
Formal thought disorder	0.05	0.15	0.04	-0.02	-0.18	-0.21

(Table 15 continued)

Variable	VsAN Conf	VsAN Disc	VDMN Conf	VDMN Disc	CEN Conf	CEN Disc
Symptoms (n = 57-58)						
SANS						
Total	0.17	0.20	-0.19	**-0.31***	0.06	0.00
Affective flattening	0.12	**0.27***	-0.08	**-0.30***	0.16	0.01
Alogia	0.09	0.13	-0.24	**-0.31***	0.16	0.01
Avolition-apathy	0.12	0.10	-0.15	-0.21	-0.03	-0.06
Anhedonia-asociality	0.18	0.16	-0.17	-0.20	-0.08	-0.11
Attention	-0.02	-0.02	0.04	-0.01	0.08	0.06
Inappropriate affect	-0.04	-0.01	-0.02	-0.03	-0.04	-0.14

Note. Significant correlations are set in bold text. BADE = bias against disconfirmatory evidence; CEN = cognitive evaluation network; Conf = confirm; COWAT = Controlled Oral Word Association Test; Disc = disconfirm; LNS = Letter-Number Sequencing; PSYRATS = Psychotic Symptoms Rating Scales; RBANS = Repeatable Battery for the Assessment of Neuropsychological Status; SANS/SAPS = Scales for the Assessment of Negative/Positive Symptoms; TMT = Trail-Making Test; VDMN = visual/default-mode network; VsAN = visual attention network; * = $p < .05$; ** = $p < .01$.

Table 16. Study 2: Variance table for behavioural CPCA of behavioural BADE and fMRI performance on fMRI hemodynamic response averaged over poststimulus time for confirm and disconfirm conditions.

	Total	Comp 1	Comp 2	All
Overall	6.00	2.48	1.98	4.46
% Overall	100.00	41.33	33.00	74.34
Predicted	0.28	0.21	0.06	0.27
% Predicted	100.00	76.21	21.63	97.84
% Overall	4.60	3.50	0.99	4.50
Residual	5.72	2.43	1.84	4.27
% Residual	100.00	42.43	32.21	74.63
% Overall	95.40	40.48	30.73	71.20

Table 17. Study 2: Rotated component loadings (overall solution, predicted solution – dependent variables, residual solution) and predictor loadings (predicted solution – independent variables) for behavioural CPCA of behavioural BADE and fMRI performance on fMRI hemodynamic response averaged over poststimulus time for confirm and disconfirm conditions.

	Comp 1	Comp 2
Overall Solution		
VsAN confirm	**0.79**	**-0.38**
VsAN disconfirm	**0.74**	**-0.37**
VDMN confirm	**-0.78**	-0.09
VDMN disconfirm	**-0.83**	-0.19
CEN confirm	-0.02	**0.92**
CEN disconfirm	-0.08	**0.90**
Predicted Solution – Dependent Variables		
VsAN confirm	**0.16**	**-0.16**
VsAN disconfirm	0.02	-0.01
VDMN confirm	**-0.17**	0.06
VDMN disconfirm	**-0.11**	-0.03
CEN confirm	-0.14	**0.17**
CEN disconfirm	**-0.35**	0.04
Predicted Solution – Independent Variables		
fMRI accuracy	**-0.47**	**0.50**
fMRI reaction time	**0.54**	0.10
BADE evidence integration	**0.93**	-0.10
BADE conservatism	**0.42**	**-0.81**
Residual Solution		
VsAN confirm	**0.78**	**-0.33**
VsAN disconfirm	**0.76**	**-0.40**
VDMN confirm	**-0.76**	-0.16
VDMN disconfirm	**-0.82**	-0.21
CEN confirm	-0.02	**0.90**
CEN disconfirm	-0.06	**0.84**

Note. Dominant loadings are set in bold text. Predicted solution dependent and independent variables are computed on the same components, and must be interpreted together. fMRI accuracy: correct = both responses equal to condition, incorrect = one or more responses diverges from condition; fMRI reaction time: mean response time across all images, conditions, and trials; BADE (bias against disconfirmatory evidence) evidence integration & conservatism = composites score on behavioural BADE task; CEN = cognitive evaluation network; Comp = component; VDMN = visual/default-mode network; VsAN = visual attention network.

Table 18. Study 2: Variance table for behavioural CPCA of neurocognitive measures on fMRI hemodynamic response averaged over poststimulus time for confirm and disconfirm conditions.

	Total	Comp 1	Comp 2	All
Overall	6.00	2.38	1.98	4.35
% Overall	100.00	39.65	32.92	72.56
Predicted	0.82	0.40	0.21	0.61
% Predicted	100.00	48.89	25.33	74.22
% Overall	13.59	6.64	3.44	10.09
Residual	5.18	2.04	1.79	3.84
% Residual	100.00	39.43	34.56	73.98
% Overall	86.41	34.07	29.86	63.93

Table 19. Study 2: Rotated component loadings (overall solution, predicted solution – dependent variables, residual solution) and predictor loadings (predicted solution – independent variables) for behavioural CPCA of neurocognitive measures on fMRI hemodynamic response averaged over poststimulus time for confirm and disconfirm conditions.

	Comp 1	Comp 2
Overall Solution		
VsAN confirm	**0.78**	**-0.39**
VsAN disconfirm	**0.73**	**-0.37**
VDMN confirm	**-0.76**	-0.09
VDMN disconfirm	**-0.81**	-0.21
CEN confirm	-0.03	**0.92**
CEN disconfirm	-0.08	**0.89**
Predicted Solution – Dependent Variables		
VsAN confirm	**0.43**	0.10
VsAN disconfirm	0.21	**0.31**
VDMN confirm	**-0.31**	0.07
VDMN disconfirm	0.02	**-0.30**
CEN confirm	**-0.13**	-0.04
CEN disconfirm	**-0.24**	0.09
Predicted Solution – Independent Variables		
RBANS immediate memory	**-0.12**	0.02
RBANS visuospatial	-0.05	**-0.20**
RBANS language	**-0.38**	**-0.42**
RBANS attention	0.14	**-0.22**
RBANS delayed memory	-0.27	**0.41**
TMT A	**-0.47**	**-0.41**
TMT B	-0.15	**-0.37**
LNS	**0.27**	-0.14
COWAT	**-0.53**	0.11
Residual Solution		
VsAN confirm	**0.69**	**-0.31**
VsAN disconfirm	**0.66**	**-0.35**
VDMN confirm	**-0.71**	-0.10
VDMN disconfirm	**-0.79**	-0.14
CEN confirm	-0.02	**0.88**
CEN disconfirm	-0.04	**0.88**

Note. Dominant loadings are set in bold text. Predicted solution dependent and independent variables are computed on the same components, and must be interpreted together. CEN = cognitive evaluation network; Comp = component; COWAT = Controlled Oral Word Association Test (standard scores for number correct controlling for age, gender, education level); LNS = Letter-Number Sequencing (standard scores for number correct controlling for age); RBANS = Repeatable Battery for the Assessment of Neuropsychological Status (scaled scores for summed test scores controlling for age); TMT = Trail-Making Test (standard scores for time to completion controlling for age, gender, education level); VDMN = visual/default-mode network; VsAN = visual attention network.

Table 20. Study 2: Variance table for behavioural CPCA of symptoms on fMRI hemodynamic response averaged over poststimulus time for confirm and disconfirm conditions.

	Total	Comp 1	Comp 2	All
Overall	6.00	2.43	2.00	4.43
% Overall	100.00	40.57	33.32	73.88
Predicted	0.82	0.38	0.24	0.62
% Predicted	100.00	46.56	29.09	75.65
% Overall	13.70	6.38	3.99	10.36
Residual	5.18	2.08	1.76	3.84
% Residual	100.00	40.12	34.04	74.16
% Overall	86.30	34.62	29.37	64.00

Table 21. Study 2: Rotated component loadings (overall solution, predicted solution – dependent variables, residual solution) and predictor loadings (predicted solution – independent variables) for behavioural CPCA of symptoms on fMRI hemodynamic response averaged over poststimulus time for confirm and disconfirm conditions.

	Comp 1	Comp 2
Overall Solution		
VsAN confirm	**0.79**	**-0.38**
VsAN disconfirm	**0.72**	**-0.41**
VDMN confirm	**-0.78**	-0.11
VDMN disconfirm	**-0.82**	-0.18
CEN confirm	-0.03	**0.92**
CEN disconfirm	-0.06	**0.89**
Predicted Solution – Dependent Variables		
VsAN confirm	**0.22**	-0.18
VsAN disconfirm	**0.22**	-0.12
VDMN confirm	**-0.34**	0.05
VDMN disconfirm	**-0.35**	-0.17
CEN confirm	0.12	**0.21**
CEN disconfirm	0.05	**0.34**
Predicted Solution – Independent Variables		
PSYRATS hallucinations	-0.07	**-0.57**
PSYRATS delusions	-0.07	**-0.14**
SAPS hallucinations	0.14	**-0.62**
SAPS delusions	-0.13	**-0.31**
SAPS bizarre behaviour	0.10	-0.01
SAPS formal thought disorder	-0.15	**-0.53**
SANS affective flattening	**0.45**	0.13
SANS alogia	**0.52**	0.11
SANS avolition-apathy	**0.22**	-0.10
SANS attention	**-0.30**	0.20
SANS inappropriate affect	-0.12	**-0.20**
Residual Solution		
VsAN confirm	**0.75**	**-0.35**
VsAN disconfirm	**0.67**	**-0.39**
VDMN confirm	**-0.69**	-0.11
VDMN disconfirm	**-0.73**	-0.12
CEN confirm	-0.08	**0.89**
CEN disconfirm	-0.07	**0.83**

Note. Dominant loadings are set in bold text. Predicted solution dependent and independent variables are computed on the same components, and must be interpreted together. CEN = cognitive evaluation network; Comp = component; PSYRATS = Psychotic Symptoms Rating Scales; SANS/SAPS = Scales for the Assessment of Negative/Positive Symptoms; VDMN = visual/default-mode network; VsAN = visual attention network; PSYRATS/SAPS/SANS = total subscale scores, except for SANS inappropriate affect, which is a single item.

Table 22. Study 2: Intercorrelation table for baseline-peak fMRI hemodynamic response.

Variable	VsAN BP	VsAN PB	VDMN BP	VDMN PB	CEN BP	CEN PB
VsAN BP	1	**0.20***	**-0.75****	0.00	-0.19	0.04
VsAN PB		1	-0.15	**-0.62****	**-0.37****	0.15
VDMN BP			1	0.09	0.10	-0.03
VDMN PB				1	-0.06	0.12
CEN BP					1	0.02
CEN PB						1

Note. Significant correlations are set in bold text. BP = baseline-to-peak; CEN = cognitive evaluation network; PB = peak-to-baseline; VDMN = visual/default-mode network; VsAN = visual attention network; * = $p < .05$; ** = $p < .001$.

Table 23. Study 2: Correlation table for associations between behaviour and baseline-peak fMRI hemodynamic response.

Variable	VsAN BP	VsAN PB	VDMN BP	VDMN PB	CEN BP	CEN PB
Behavioural BADE (n = 98)						
Evidence integration	0.02	0.11	0.01	-0.15	**-0.22***	0.02
Conservatism	-0.02	0.01	0.13	-0.05	-0.14	0.06
fMRI Performance (n = 98)						
Accuracy	-0.04	-0.16	0.00	0.20	0.20	0.19
Accuracy – confirm	-0.08	-0.05	0.09	0.10	0.10	0.15
Accuracy – disconfirm	0.03	-0.20	-0.10	**0.22***	**0.22***	0.17
Reaction time	-0.04	0.13	-0.11	-0.05	**-0.21***	-0.03
Reaction time – confirm	-0.06	0.11	-0.10	-0.06	**-0.21***	-0.02
Reaction time – disconfirm	-0.01	0.16	-0.12	-0.05	**-0.21***	-0.03
Missed responses	0.14	0.07	-0.13	-0.11	-0.12	-0.12
Neurocognition (n = 58)						
RBANS						
Immediate memory	-0.08	-0.04	0.03	0.10	-0.07	0.21
Visuospatial	-0.08	-0.13	0.10	-0.01	-0.01	0.01
Language	-0.26	-0.22	-0.12	0.06	0.02	0.05
Attention	0.12	-0.13	0.07	0.24	-0.22	0.13
Delayed memory	-0.07	-0.04	0.07	-0.08	0.00	0.19
TMT A	-0.18	-0.23	0.23	0.13	-0.07	**0.26***
TMT B	-0.13	-0.11	-0.18	0.08	-0.09	0.19
LNS	0.01	0.05	-0.03	0.09	-0.10	0.12
COWAT	-0.09	-0.17	0.25	0.04	0.02	0.05
Symptoms (n = 57-58)						
PSYRATS hallucinations	0.11	0.15	-0.06	0.14	-0.13	-0.06
PSYRATS delusions	0.19	-0.03	-0.21	-0.03	-0.07	0.09
SAPS						
Total	**0.28***	0.08	-0.10	-0.02	-0.18	0.09
Hallucinations	0.23	0.15	-0.15	0.05	-0.17	0.13
Delusions	0.21	0.00	-0.06	-0.04	-0.05	0.02
Bizarre behaviour	0.01	0.06	0.05	-0.22	-0.06	0.04
Formal thought disorder	0.21	-0.01	-0.01	0.03	-0.23	0.06

(Table 23 continued)

Variable	VsAN BP	VsAN PB	VDMN BP	VDMN PB	CEN BP	CEN PB
Symptoms (n = 57-58)						
SANS						
Total	**0.34***	0.02	**-0.30***	-0.08	0.10	-0.26
Affective flattening	**0.35***	0.05	-0.24	-0.06	0.13	-0.12
Alogia	0.23	0.00	**-0.29***	-0.16	0.13	-0.15
Avolition-apathy	0.20	0.02	-0.21	-0.08	0.01	-0.23
Anhedonia-asociality	0.25	0.07	-0.17	-0.13	-0.01	**-0.34***
Attention	0.11	-0.12	-0.09	0.15	0.03	0.17
Inappropriate affect	0.00	-0.04	-0.02	-0.02	-0.10	-0.03

Note. Significant correlations are set in bold text. BADE = bias against disconfirmatory evidence; BP = baseline-to-peak; CEN = cognitive evaluation network; COWAT = Controlled Oral Word Association Test; LNS = Letter-Number Sequencing; PB = peak-to-baseline; PSYRATS = Psychotic Symptoms Rating Scales; RBANS = Repeatable Battery for the Assessment of Neuropsychological Status; SANS/SAPS = Scales for the Assessment of Negative/Positive Symptoms; TMT = Trail-Making Test; VDMN = visual/default-mode network; VsAN = visual attention network; * = $p < .05$.

Table 24. Study 2: Variance table for behavioural CPCA of behavioural BADE and fMRI performance on baseline-peak fMRI hemodynamic response.

	Total	Comp 1	Comp 2	Comp 3	Comp 4	All
Overall	6.00	1.74	1.44	1.28	1.05	5.52
% Overall	100.00	28.98	24.03	21.41	17.57	91.98
Predicted	0.27	0.16	0.10	-	-	0.26
% Predicted	100.00	59.33	35.74	-	-	95.07
% Overall	4.52	2.68	1.61	-	-	4.29
Residual	5.73	1.71	1.36	1.21	1.00	5.29
% Residual	100.00	29.80	23.74	21.21	17.54	92.29
% Overall	95.48	28.45	22.67	20.25	16.75	88.12

Table 25. Study 2: Rotated component loadings (overall solution, predicted solution – dependent variables, residual solution) and predictor loadings (predicted solution – independent variables) for behavioural CPCA of behavioural BADE and fMRI performance on baseline-peak fMRI hemodynamic response.

	Comp 1	Comp 2	Comp 3	Comp 4
Overall Solution				
VsAN BP	**0.90**	0.10	-0.23	-0.05
VsAN PB	0.16	**0.80**	**-0.47**	0.17
VDMN BP	**-0.95**	-0.02	-0.04	0.00
VDMN PB	0.00	**-0.89**	-0.29	0.20
CEN BP	-0.09	0.04	**0.96**	0.03
CEN PB	-0.03	-0.05	0.01	**0.99**
Predicted Solution – Dependent Variables				
VsAN BP	0.04	**0.10**		
VsAN PB	**0.16**	-0.07		
VDMN BP	0.00	**-0.13**		
VDMN PB	**-0.24**	0.08		
CEN BP	**-0.27**	-0.07		
CEN PB	-0.01	**-0.23**		
Predicted Solution – Independent Variables				
fMRI accuracy	**-0.48**	**-0.53**		
fMRI reaction time	0.15	**0.85**		
BADE evidence integration	**0.88**	0.17		
BADE conservatism	**0.89**	-0.29		
Residual Solution				
VsAN BP	**0.90**	0.13	-0.21	-0.05
VsAN PB	0.16	**0.86**	**-0.35**	0.13
VDMN BP	**-0.93**	-0.03	-0.05	-0.02
VDMN PB	0.00	**-0.78**	**-0.44**	0.24
CEN BP	-0.09	-0.05	**0.92**	0.02
CEN PB	-0.02	-0.02	0.00	**0.96**

Note. Dominant loadings are set in bold text. Predicted solution dependent and independent variables are computed on the same components, and must be interpreted together. fMRI accuracy: correct = both responses equal to condition, incorrect = one or more responses diverges from condition; fMRI reaction time: mean response time across all images, conditions, and trials; BADE (bias against disconfirmatory evidence) evidence integration & conservatism = composites score on behavioural BADE task; BP = baseline-to-peak; CEN = cognitive evaluation network; Comp = component; PB = peak-to-baseline; VDMN = visual/default-mode network; VsAN = visual attention network.

Table 26. Study 2: Variance table for behavioural CPCA of neurocognitive measures on baseline-peak fMRI hemodynamic response.

	Total	Comp 1	Comp 2	Comp 3	Comp 4	All
Overall	6.00	1.70	1.43	1.30	1.05	5.49
% Overall	100.00	28.37	23.81	21.69	17.57	91.44
Predicted	0.76	0.38	0.24	-	-	0.62
% Predicted	100.00	50.07	31.57	-	-	81.64
% Overall	12.66	6.34	4.00	-	-	10.34
Residual	5.24	1.43	1.28	1.15	0.95	4.81
% Residual	100.00	27.24	24.52	21.99	18.11	91.86
% Overall	87.34	23.79	21.41	19.21	15.82	80.23

Table 27. Study 2: Rotated component loadings (overall solution, predicted solution – dependent variables, residual solution) and predictor loadings (predicted solution – independent variables) for behavioural CPCA of neurocognitive measures on baseline-peak fMRI hemodynamic response.

	Comp 1	Comp 2	Comp 3	Comp 4
Overall Solution				
VsAN BP	**0.89**	0.10	-0.25	-0.01
VsAN PB	0.16	**0.78**	**-0.48**	0.18
VDMN BP	**-0.94**	-0.01	-0.05	0.02
VDMN PB	0.00	**-0.89**	-0.28	0.19
CEN BP	-0.09	0.05	**0.96**	0.03
CEN PB	-0.03	-0.05	0.01	**0.99**
Predicted Solution – Dependent Variables				
VsAN BP	**0.41**	-0.11		
VsAN PB	**0.27**	0.12		
VDMN BP	**-0.30**	-0.05		
VDMN PB	0.08	**-0.32**		
CEN BP	-0.19	**0.23**		
CEN PB	-0.04	**-0.23**		
Predicted Solution – Independent Variables				
RBANS immediate memory	-0.09	**-0.26**		
RBANS visuospatial	-0.14	0.11		
RBANS language	**-0.49**	-0.11		
RBANS attention	0.11	**-0.59**		
RBANS delayed memory	-0.12	0.06		
TMT A	**-0.57**	**-0.63**		
TMT B	**-0.33**	**-0.37**		
LNS	0.16	-0.14		
COWAT	**-0.40**	-0.15		
Residual Solution				
VsAN BP	**0.79**	0.13	-0.19	-0.02
VsAN PB	0.08	**0.74**	**-0.46**	0.20
VDMN BP	**-0.89**	0.00	-0.05	0.01
VDMN PB	-0.04	**-0.85**	-0.22	0.17
CEN BP	-0.05	0.02	**0.92**	0.05
CEN PB	-0.02	-0.03	0.03	**0.94**

Note. Dominant loadings are set in bold text. Predicted solution dependent and independent variables are computed on the same components, and must be interpreted together. BP = baseline-to-peak; CEN = cognitive evaluation network; COWAT = Controlled Oral Word Association Test (standard scores for number correct controlling for age, gender, education level); Comp = component; LNS = Letter-Number Sequencing (standard scores for number correct controlling for age); PB = peak-to-baseline; RBANS = Repeatable Battery for the Assessment of Neuropsychological Status (scaled scores for summed test scores controlling for age); TMT = Trail-Making Test (standard scores for time to completion controlling for age, gender, education level); VDMN = visual/default-mode network; VsAN = visual attention network.

Table 28. Study 2: Variance table for behavioural CPCA of symptoms on baseline-peak fMRI hemodynamic response.

	Total	Comp 1	Comp 2	Comp 3	Comp 4	All
Overall	6.00	1.71	1.43	1.30	1.05	5.50
% Overall	100.00	28.47	23.86	21.74	17.56	91.62
Predicted	0.96	0.36	0.30	-	-	0.67
% Predicted	100.00	37.94	31.54	-	-	69.48
% Overall	15.98	6.06	5.04	-	-	11.10
Residual	5.04	1.34	1.31	1.14	0.87	4.65
% Residual	100.00	26.51	25.92	22.64	17.22	92.29
% Overall	84.02	22.28	21.78	19.02	14.47	77.55

Table 29. Study 2: Rotated component loadings for (overall solution, predicted solution – dependent variables, residual solution) and predictor loadings (predicted solution – independent variables) for behavioural CPCA of symptoms on baseline-peak fMRI hemodynamic response.

	Comp 1	Comp 2	Comp 3	Comp 4
Overall Solution				
VsAN BP	**0.89**	0.09	-0.27	-0.01
VsAN PB	0.16	**0.77**	**-0.49**	0.19
VDMN BP	**-0.94**	-0.02	-0.06	0.02
VDMN PB	0.01	**-0.91**	-0.26	0.18
CEN BP	-0.09	0.05	**0.96**	0.03
CEN PB	-0.03	-0.04	0.01	**0.99**
Predicted Solution – Dependent Variables				
VsAN BP	**0.39**	0.03		
VsAN PB	0.08	-0.03		
VDMN BP	**-0.38**	0.09		
VDMN PB	0.02	**0.31**		
CEN BP	0.02	**-0.27**		
CEN PB	-0.07	**0.37**		
Predicted Solution – Independent Variables				
PSYRATS hallucinations	0.04	**0.35**		
PSYRATS delusions	-0.05	**0.29**		
SAPS hallucinations	**0.26**	**0.48**		
SAPS delusions	-0.05	**0.19**		
SAPS bizarre behaviour	**-0.27**	-0.16		
SAPS formal thought disorder	0.07	**0.37**		
SANS affective flattening	**0.44**	**-0.28**		
SANS alogia	**0.41**	**-0.36**		
SANS avolition-apathy	0.08	**-0.40**		
SANS attention	-0.14	**0.32**		
SANS inappropriate affect	-0.14	0.00		
Residual Solution				
VsAN BP	**0.75**	0.10	-0.27	-0.01
VsAN PB	0.13	**0.77**	**-0.47**	0.22
VDMN BP	**-0.82**	-0.03	-0.01	-0.01
VDMN PB	0.00	**-0.85**	-0.24	0.16
CEN BP	-0.11	0.01	**0.89**	0.06
CEN PB	0.00	0.00	0.04	**0.89**

Note. Dominant loadings are set in bold text. Predicted solution dependent and independent variables are computed on the same components, and must be interpreted together. BP = baseline-to-peak; CEN = cognitive evaluation network; Comp = component; PSYRATS = Psychotic Symptoms Rating Scales; SANS/SAPS = Scales for the Assessment of Negative/Positive Symptoms; VDMN = visual/default-mode network; VsAN = visual attention network. PB = peak-to-baseline; PSYRATS/SAPS/SANS = total subscale scores, except for SANS inappropriate affect, which is a single item.

Table 30. Study 3: Loadings and standard errors (SE) for the unconstrained and constrained generalized structured component analysis models of behavioural BADE at time 1 and time 2.

	Unconstrained				Constrained	
Variable	**Time 1**		**Time 2**		**Time 1 & Time 2**	
	Loading	**SE**	**Loading**	**SE**	**Loading**	**SE**
Evidence Integration						
Absurd 1	0.83	0.05	0.89	0.03	0.85	0.03
Absurd 2	0.84	0.03	0.82	0.02	0.87	0.02
Absurd 3	0.79	0.06	0.56	0.08	0.68	0.04
Neutral lure 2	0.82	0.06	0.88	0.02	0.85	0.03
Neutral lure 3	0.88	0.03	0.74	0.06	0.82	0.03
Emotional lure 2	0.78	0.06	0.90	0.03	0.84	0.03
Emotional lure 3	0.84	0.04	0.71	0.07	0.78	0.03
True 3	-0.39	0.15	-0.68	0.18	-0.49	0.10
Conservatism						
Neutral lure 1	0.96	0.01	0.95	0.01	0.95	0.01
Emotional lure 1	0.96	0.01	0.94	0.01	0.95	0.01
True 1	0.95	0.01	0.94	0.02	0.95	0.01
True 2	0.93	0.02	0.95	0.02	0.94	0.01
True 3	0.76	0.13	1.23	0.13	0.95	0.09

Note. SE = standard error. All loadings are significant at $p < .05$.

Table 31. Study 3: Demographic, clinical, and cognitive information for time 1 only (study 2 only) versus time 1 & time 2 (studies 2 & 3) patients (fMRI sample).

	Time 1 Only (n = 29)	Time 1 & Time 2 (n = 29)
Demographic / Intelligence		
Gender (male:female)	17:12	18:11
Handedness (R:M:L)	27:1:1	28:0:1
Age (mean(SD))	34.90 (11.50)	35.66 (9.11)
Education (years)	14.14 (2.43)	14.81 (2.02)
ToPF	106.52 (14.80)	110.35 (13.98)
WASI**	94.72 (8.88)	102.52 (8.37)
Clinical		
Chlorpromazine equivalents	837.76 (2054.15)	422.26 (813.82)
BDI	14.63 (8.88)	12.04 (9.82)
PSYRATS hallucinations	12.29 (15.10)	5.97 (10.26)
PSYRATS delusions	10.93 (7.27)	8.72 (6.39)
SAPS total	15.07 (12.95)	12.39 (11.49)
Hallucinations	5.38 (6.72)	2.68 (4.60)
Delusions	6.86 (6.25)	6.48 (6.36)
Bizarre behavior	0.72 (1.49)	1.03 (1.72)
Formal thought disorder	2.10 (3.44)	2.10 (3.23)
SANS Total	17.12 (11.72)	18.48 (15.31)
Affective flattening	4.59 (5.37)	5.07 (7.07)
Alogia	1.59 (2.54)	1.38 (2.64)
Avolition-apathy	3.21 (3.28)	3.31 (2.36)
Anhedonia-asociality	5.00 (4.89)	5.35 (4.49)
Attention*	2.21 (1.87)	1.31 (1.49)
Inappropriate affect	0.38 (0.90)	0.52 (1.15)
Neurocognitive		
RBANS immediate memory	84.10 (17.49)	87.17 (18.03)
RBANS visuospatial	101.24 (15.88)	107.48 (14.85)
RBANS language	82.28 (19.64)	88.10 (18.33)
RBANS attention	80.79 (17.89)	88.03 (11.23)
RBANS delayed memory	84.10 (16.92)	88.52 (15.73)
TMT A	37.55 (16.58)	37.07 (11.18)
TMT B	40.03 (12.79)	42.59 (9.89)
LNS	8.52 (2.08)	9.44 (2.40)
COWAT	43.66 (11.59)	45.90 (10.90)
Behavioural BADE		
Evidence integration*	85.27 (102.82)	17.43 (73.71)
Conservatism**	381.13 (135.72)	273.42 (119.26)

(Table 31 continued)

	Time 1 Only (n = 29)	Time 1 & Time 2 (n = 29)
fMRI Performance		
Accuracy (%)	70.04 (12.68)	73.58 (11.19)
Accuracy – confirm	79.05 (18.74)	85.95 (10.72)
Accuracy – disconfirm	61.03 (12.56)	61.21 (14.60)
Reaction time (ms)	1536.82 (296.10)	1552.66 (228.55)
Reaction time – confirm	1479.19 (316.30)	1490.08 (239.38)
Reaction time – disconfirm	1594.46 (289.34)	1615.24 (246.10)
Missed responses	4.12 (7.88)	1.78 (3.62)

Note. BADE = bias against disconfirmatory evidence; BDI = Beck Depression Inventory; COWAT = Controlled Oral Word Association Test; L = left; LNS = Letter-Number Sequencing; M = mixed; ms = milliseconds; PSYRATS = Psychotic Symptoms Rating Scales; R = right; RBANS = Repeatable Battery for the Assessment of Neuropsychological Status; SANS/SAPS = Scales for the Assessment of Negative/Positive Symptoms; SD = standard deviation; TMT = Trail-Making Test; ToPF = Test of Premorbid Functioning; WASI = Wechsler Abbreviated Scale of Intelligence; * = $p < .05$; ** = $p < .005$.

Table 32. Study 3: Demographic, clinical, and cognitive information by time point.

Demographic / Intelligence (n = 72)	Time 1	
Gender (male:female)	33:39	
Handedness (R:M:L)	69:1:2	
Age (mean(SD))	37.68 (9.72)	
Education (years)	14.04 (2.17)	
ToPF	106.17 (14.32)	
WASI	99.28 (10.17)	
	Time 1	Time 2
Clinical (n = 46-72)		
Chlorpromazine equivalents	321.08 (577.39)	325.57 (571.98)
BDI**	14.48 (10.43)	10.71 (9.50)
PSYRATS hallucinations	8.23 (12.28)	7.64 (11.78)
PSYRATS delusions	9.25 (6.92)	8.51 (6.37)
SAPS total*	12.40 (11.00)	10.16 (9.26)
Hallucinations	3.49 (5.48)	3.10 (4.78)
Delusions*	6.52 (6.33)	5.25 (5.51)
Bizarre behavior	0.81 (1.33)	0.86 (1.59)
Formal thought disorder*	1.73 (2.88)	1.24 (2.32)
SANS total	17.98 (12.08)	16.37 (12.06)
Affective flattening	4.94 (6.00)	4.89 (6.09)
Alogia	1.42 (2.29)	1.38 (2.32)
Avolition-apathy	3.44 (2.63)	3.28 (2.86)
Anhedonia-asociality	5.64 (4.66)	4.72 (4.30)
Attention	1.63 (1.61)	1.66 (1.69)
Inappropriate affect	0.30 (0.87)	0.17 (0.70)
Neurocognitive (n = 71-72)		
RBANS immediate memory**	85.42 (18.80)	94.50 (21.21)
RBANS visuospatial	102.89 (15.47)	103.61 (16.40)
RBANS language	87.56 (19.61)	85.04 (19.09)
RBANS attention**	83.44 (16.05)	87.65 (17.59)
RBANS delayed memory**	87.94 (16.11)	93.29 (18.22)
TMT A**	34.85 (10.95)	41.50 (11.24)
TMT B**	40.75 (11.36)	46.45 (11.62)
LNS*	8.94 (2.58)	9.67 (3.12)
COWAT	43.18 (10.67)	44.15 (10.41)
Behavioural BADE (n = 56)		
Evidence integration**	46.79 (80.23)	102.30 (88.68)
Conservatism*	322.84 (134.99)	291.74 (134.37)

(Table 32 continued)

	Time 1	Time 2
fMRI performance (n = 29)		
Accuracy (%)	73.58 (11.19)	76.55 (7.74)
Accuracy – confirm*	85.95 (10.72)	90.52 (7.30)
Accuracy – disconfirm	61.21 (14.60)	62.59 (11.64)
Reaction time (ms)	1552.66 (228.55)	1502.42 (279.34)
Reaction time – confirm	1490.08 (239.38)	1440.65 (299.33)
Reaction time – disconfirm	1615.24 (246.10)	1564.20 (282.84)
Missed Responses	1.78 (3.62)	1.19 (1.73)

Note. BADE = bias against disconfirmatory evidence; BDI = Beck Depression Inventory; COWAT = Controlled Oral Word Association Test; L = left; LNS = Letter-Number Sequencing; M = middle; ms = milliseconds; PSYRATS = Psychotic Symptoms Rating Scales; R = right; RBANS = Repeatable Battery for the Assessment of Neuropsychological Status; SANS/SAPS = Scales for the Assessment of Negative/Positive Symptoms; SD = standard deviation; TMT = Trail-Making Test; ToPF = Test of Premorbid Functioning; WASI = Wechsler Abbreviated Scale of Intelligence; * = $p < .05$; ** = $p < .005$.

Table 33. Study 3: Correlations between change (time 2-time 1) in delusions and change (time 2-time 1) in symptoms, neurocognition, behavioural BADE and fMRI performance.

Variable	PSYRATS Delusion Change	SAPS Delusion Change
Clinical		
BDI (n = 62-65)	0.00	0.12
PSYRATS (n = 61-64)		
Hallucinations	0.14	**0.31***
SAPS (n = 68-72)		
Total	**0.31****	**0.74****
Hallucinations	0.03	**0.35****
Bizarre behaviour	-0.03	-0.08
Formal thought disorder	-0.03	-0.06
SANS (n = 45-72)		
Total	-0.02	-0.01
Affective flattening	0.00	0.08
Alogia	0.09	-0.01
Avolition-apathy	-0.02	0.05
Anhedonia-asociality	-0.11	-0.18
Attention	0.09	0.03
Inappropriate affect	0.19	0.09
Neurocognitive (n = 68-72)		
RBANS immediate memory	-0.08	0.12
RBANS visuospatial	0.20	0.14
RBANS language	-0.01	0.11
RBANS attention	-0.01	-0.10
RBANS delayed memory	-0.04	0.17
TMT A	-0.17	-0.17
TMT B	-0.15	**-0.25***
LNS	-0.03	0.07
COWAT	0.01	-0.02
Behavioural BADE (n = 55-56)		
Evidence integration	-0.08	-0.13
Conservatism	-0.07	0.05

(Table 33 continued)

Variable	PSYRATS Delusion Change	SAPS Delusion Change
fMRI Performance (n = 27-29)		
Accuracy	0.07	0.13
Accuracy – confirm	0.16	-0.04
Accuracy – disconfirm	-0.01	0.21
Reaction time	0.03	0.00
Reaction time – confirm	-0.03	-0.16
Reaction time – disconfirm	0.08	0.13
Missed Responses	-0.16	-0.14

Note. Significant correlations are set in bold text. BADE = bias against disconfirmatory evidence; BDI = Beck Depression Inventory; COWAT = Controlled Oral Word Association Test; LNS = Letter-Number Sequencing; PSYRATS = Psychotic Symptoms Rating Scales; RBANS = Repeatable Battery for the Assessment of Neuropsychological Status; SANS/SAPS = Scales for the Assessment of Negative/Positive Symptoms; TMT = Trail-Making Test; $* = p < .05$; $** = p < .01$.

Table 34. Study 3: Anatomical descriptions for the most extreme 10% of visual attention network loadings (Component 1), with cluster volumes, Montreal Neurological Institute (MNI) coordinates, and Brodmann's area (BA) for the peaks within each cluster.

Brain Region	Cluster Volume (voxels)	BA for Peak Locations	MNI Coordinates		
			x	y	z
Positive Loadings					
Cluster 1: Bilateral	24222				
Lateral occipital cortex, superior division		18	-28	-90	6
Occipital fusiform gyrus		19	28	-80	-12
Occipital fusiform gyrus		18	-22	-84	-14
Occipital pole		18	28	-88	8
Occipital fusiform gyrus		19	-34	-76	-14
Occipital fusiform gyrus		37	38	-62	-12
Lateral occipital cortex, inferior division		19	-40	-76	-12
Lateral occipital cortex, superior division		18	32	-84	14
Occipital pole		18	-16	-96	0
Temporal occipital fusiform cortex		37	-40	-62	-14
Lateral occipital cortex, superior division		19	-26	-70	28
Superior parietal lobule		7	30	-52	48
Lateral occipital cortex, superior division		7	-26	-60	50
Lateral occipital cortex, superior division		7	26	-62	48
Temporal occipital fusiform cortex		37	32	-44	-22
Lateral occipital cortex, superior division		19	30	-72	26
Precentral gyrus		6	-46	4	32
Postcentral gyrus		2	-44	-32	44
Postcentral gyrus		3	-54	-24	40
Precentral gyrus		4	-40	-16	60
Lingual gyrus		17	4	-86	-6
Postcentral gyrus		3	54	-22	44
Middle frontal gyrus		6	-30	-4	50
Supramarginal gyrus, anterior division		40	40	-34	42
Cerebellum VI		n/a	-8	-74	-24
Precentral gyrus		6	-36	-4	58
Cerebellum VI		n/a	8	-72	-22
Cerebellum Crus II		n/a	10	-76	-42
Postcentral gyrus		48	-56	-18	22
Cerebellum VIIb		n/a	-24	-70	-50
Cerebellum Crus II		n/a	-8	-76	-40
Cerebellum VIIIb		n/a	16	-62	-48

(Table 34 continued)

Brain Region	Cluster Volume (voxels)	BA for Peak Locations	MNI Coordinates		
			x	y	z
Cluster 2: Right Hemisphere	853				
Precentral gyrus		44	50	10	32
Cluster 3: Bilateral	846				
Paracingulate gyrus		32	0	14	50
Supplementary motor area		6	-2	4	56
Cluster 4: Right Hemisphere	648				
Middle frontal gyrus		6	34	0	52
Cluster 5: Right Hemisphere	67				
Inferior frontal gyrus, pars triangularis		45	46	32	14
Cluster 6: Left Hemisphere	61				
Insula		47	-30	24	0
Cluster 7: Left Hemisphere	35				
Inferior frontal gyrus, pars triangularis		48	-40	26	20
Cluster 8: Left Hemisphere	27				
Thalamus		n/a	-10	-18	8
Cluster 9: Right Hemisphere	13				
Insula		47	32	24	4

Note. No negative loadings passed threshold. Only clusters with volumes greater than or equal to 5 voxels are displayed.

Table 35. Study 3: Anatomical descriptions for the most extreme 10% of visual/default-mode network loadings (Component 2), with cluster volumes, Montreal Neurological Institute (MNI) coordinates, and Brodmann's area (BA) for the peaks within each cluster.

Brain Region	Cluster Volume (voxels)	BA for Peak Locations	MNI Coordinates		
			x	y	z
Positive Loadings					
Cluster 1: Right Hemisphere	4106				
Occipital pole		18	28	-88	8
Occipital fusiform gyrus		18	24	-82	-10
Lateral occipital cortex, inferior division		19	38	-74	-12
Occipital fusiform gyrus		37	42	-62	-12
Lateral occipital cortex, superior division		7	26	-64	52
Superior parietal lobule		7	30	-54	48
Lateral occipital cortex, superior division		19	30	-70	28
Cluster 2: Left Hemisphere	4038				
Lateral occipital cortex, inferior division		18	-28	-90	6
Occipital fusiform gyrus		18	-22	-84	-12
Lateral occipital cortex, inferior division		19	-40	-76	-10
Lateral occipital cortex, superior division		19	-26	-70	28
Temporal occipital fusiform cortex		37	-34	-48	-20
Lateral occipital cortex, superior division		7	-22	-68	46
Superior parietal lobule		7	-26	-54	46
Cluster 3: Right Hemisphere	106				
Precentral gyrus		44	46	8	28
Cluster 4: Left Hemisphere	16				
Precentral gyrus		44	-44	4	30
Cluster 5: Left Hemisphere	7				
Supramarginal gyrus, anterior division		40	-40	-38	42
Negative Loadings					
Cluster 1: Bilateral	9319				
Precuneus		18	-2	-70	24
Cuneus		18	-4	-78	28
Cuneus		18	2	-78	30
Precuneus		23	-2	-56	34
Lingual gyrus		18	10	-68	-4
Lingual gyrus		18	-12	-70	-6
Cingulate gyrus, posterior division		23	0	-20	40
Lingual gyrus		30	14	-46	-4
Lingual gyrus		19	16	-48	-6
Lingual gyrus		19	-16	-48	-6

(Table 35 continued)

Brain Region	Cluster Volume (voxels)	BA for Peak Locations	MNI Coordinates		
			x	y	z
Cluster 2: Bilateral	2835				
Paracingulate gyrus		10	-2	54	-2
Cingulate gyrus, anterior division		24	0	36	16
Frontal pole		10	-20	58	16
Cluster 3: Left Hemisphere	1989				
Lateral occipital cortex, superior division		39	-44	-72	32
Cluster 4: Right Hemisphere	1409				
Lateral occipital cortex, superior division		39	46	-70	32
Cluster 5: Right Hemisphere	937				
Superior temporal gyrus, posterior division		22	66	-26	10
Parietal operculum cortex		48	48	-30	20
Planum temporale		42	58	-28	12
Superior temporal gyrus, posterior division		21	66	-30	2
Central opercular cortex		48	56	-4	4
Planum polare		48	40	-20	0
Heschl's gyrus		48	52	-14	6
Cluster 6: Left Hemisphere	721				
Superior frontal gyrus		9	-22	34	44
Cluster 7: Right Hemisphere	450				
Superior parietal lobule		2	24	-42	64
Postcentral gyrus		4	28	-28	58
Postcentral gyrus		3	32	-28	58
Precentral gyrus		4	38	-18	44
Cluster 8: Right Hemisphere	319				
Frontal pole		9	24	36	44
Cluster 9: Left Hemisphere	234				
Middle temporal gyrus, anterior division		21	-58	-6	-14
Superior temporal gyrus, posterior division		22	-62	-12	-2
Cluster 10: Left Hemisphere	139				
Planum temporale		22	-58	-30	12
Planum temporale		41	-44	-36	16
Cluster 11: Right Hemisphere	125				
Middle temporal gyrus, posterior division		21	60	-10	-12
Middle temporal gyrus, anterior division		21	54	-2	-22
Cluster 12: Right Hemisphere	26				
Cerebellum Crus I		n/a	28	-76	-36

Note. Only clusters with volumes greater than or equal to 5 voxels are displayed.

Table 36. Study 3: Anatomical descriptions for the most extreme 10% of cognitive evaluation network loadings (Component 3), with cluster volumes, Montreal Neurological Institute (MNI) coordinates, and Brodmann's area (BA) for the peaks within each cluster.

Brain Region	Cluster Volume (voxels)	BA for Peak Locations	MNI Coordinates x	y	z
Positive Loadings					
Cluster 1: Right Hemisphere	12090				
Occipital pole		18	24	-96	6
Lateral occipital cortex, inferior division		19	40	-86	-6
Lateral occipital cortex, inferior division		39	40	-58	50
Lateral occipital cortex, inferior division		37	46	-64	-14
Supramarginal gyrus, posterior division		40	52	-42	48
Lateral occipital cortex, superior division		7	26	-72	56
Middle temporal gyrus, temporooccipital part		21	62	-38	-4
Middle temporal gyrus, temporooccipital part		37	60	-50	-6
Lateral occipital cortex, superior division		19	28	-82	36
Cluster 2: Left Hemisphere	6716				
Occipital pole		18	-28	-96	2
Lateral occipital cortex, inferior division		19	-46	-70	-16
Occipital fusiform gyrus		18	-18	-82	-22
Occipital fusiform gyrus		19	-34	-76	-22
Cerebellum Crus II		n/a	-12	-78	-32
Cluster 3: Right Hemisphere	3420				
Middle frontal gyrus		9	44	14	50
Frontal pole		10	36	58	0
Middle frontal gyrus		45	48	32	30
Frontal pole		47	46	48	-8
Middle frontal gyrus		44	48	26	36
Frontal orbital cortex		38	48	28	-12
Inferior frontal gyrus, pars triangularis		45	56	24	22
Frontal pole		45	44	46	18
Inferior frontal gyrus, pars triangularis		48	58	22	10
Inferior frontal gyrus, pars opercularis		45	56	20	4
Cluster 4: Left Hemisphere	1394				
Frontal pole		47	-46	46	-6
Frontal orbital cortex		38	-44	22	-12
Cluster 5: Left Hemisphere	1247				
Lateral occipital cortex, superior division		7	-36	-64	54
Angular gyrus		40	-48	-52	48
Angular gyrus		39	-52	-54	42

(Table 36 continued)

Brain Region	Cluster Volume (voxels)	BA for Peak Locations	MNI Coordinates		
			x	y	z
Cluster 6: Left Hemisphere	679				
Middle frontal gyrus		9	-46	22	46
Cluster 7: Left Hemisphere	318				
Middle temporal gyrus, posterior division		21	-62	-38	0
Cluster 8: Right Hemisphere	5				
Frontal orbital cortex		38	32	24	-20
Cluster 9: Left Hemisphere	5				
Inferior frontal gyrus, pars opercularis		45	-56	20	6
Negative Loadings					
Cluster 1: Bilateral	463				
Supplementary motor area		6	-4	-2	54
Cluster 2: Left Hemisphere	440				
Precentral gyrus		4	-36	-22	50
Precentral gyrus		6	-32	-10	56

Note. Only clusters with volumes greater than or equal to 5 voxels are displayed.

Table 37. Study 3: Correlation table for associations between change (time 2-time 1) in brain activity underlying dis/confirmatory evidence integration and change (time 2-time 1) in behaviour.

Variable	VsAN Conf	VsAN Disc	VDMN Con	VDMN Disc	CEN Conf	CEN Disc
Behavioural BADE (n = 28)						
Evidence integration	**0.40***	0.25	-0.12	0.16	0.13	-0.22
Conservatism	**0.46***	0.37	-0.16	0.00	0.07	-0.11
fMRI Performance (n = 29)						
Accuracy	**-0.42***	0.09	0.30	-0.14	-0.19	-0.02
Accuracy – confirm	-0.32	-0.02	0.28	-0.13	**-0.41***	0.04
Accuracy – disconfirm	**-0.37***	0.13	0.24	-0.18	0.00	-0.05
Reaction time	0.33	0.11	-0.11	0.03	-0.20	-0.03
Reaction time – confirm	0.29	-0.02	-0.19	-0.02	-0.20	0.00
Reaction time – disconfirm	0.31	0.20	-0.03	0.06	-0.18	-0.05
Missed responses	**0.51***	0.26	**-0.50***	-0.25	**0.40***	0.01
Neurocognition (n = 29)						
RBANS immediate memory	-0.01	0.15	0.12	-0.00	0.17	0.28
RBANS visuospatial	-0.17	-0.20	0.14	0.08	-0.13	-0.05
RBANS language	-0.01	-0.05	0.17	0.31	0.06	-0.24
RBANS attention	**0.53***	0.26	-0.28	0.01	0.02	0.04
RBANS delayed memory	0.10	0.04	-0.10	-0.07	**0.41***	0.31
TMT A	-0.05	**-0.40***	-0.25	0.02	0.23	-0.10
TMT B	0.23	-0.17	-0.23	0.05	-0.12	-0.21
LNS	-0.05	-0.06	0.12	0.24	0.01	-0.15
COWAT	-0.06	-0.18	-0.23	-0.09	0.10	-0.05
Symptoms (n = 18-29)						
PSYRATS hallucinations	-0.07	-0.11	0.02	0.25	0.00	-0.33
PSYRATS delusions	-0.19	-0.05	0.24	0.06	-0.14	-0.00
SAPS						
Total	**-0.48***	0.01	**0.53***	0.28	-0.09	-0.21
Hallucinations	-0.12	-0.04	-0.05	0.06	0.06	-0.25
Delusions	-0.34	-0.02	**0.44***	0.24	0.12	-0.04
Bizarre behaviour	-0.09	0.09	0.11	-0.04	-0.12	-0.08
Formal thought disorder	-0.29	0.03	**0.41***	0.18	**-0.37***	-0.06

(Table 37 continued)

Variable	VsAN Conf	VsAN Disc	VDMN Con	VDMN Disc	CEN Conf	CEN Disc
SANS						
Total	0.07	0.12	-0.01	-0.11	0.23	0.20
Affective flattening	-0.18	0.20	0.32	-0.00	0.00	0.05
Alogia	-0.02	-0.08	0.15	0.25	-0.34	-0.19
Avolition-apathy	-0.14	-0.04	-0.09	-0.24	0.33	0.18
Anhedonia-asociality	0.23	-0.02	-0.11	0.08	0.25	0.11
Attention	-0.08	-0.10	0.10	0.04	0.12	-0.05
Inappropriate affect	0.13	0.12	-0.03	0.07	-0.09	**-0.39***

Note. Significant correlations are set in bold text. BADE = bias against disconfirmatory evidence CEN = cognitive evaluation network; Conf = confirm; COWAT = Controlled Oral Word Association Test; Disc = disconfirm; LNS = Letter-Number Sequencing; PSYRATS = Psychotic Symptoms Rating Scales; RBANS = Repeatable Battery for the Assessment of Neuropsychological Status; SANS/SAPS = Scales for the Assessment of Negative/Positive Symptoms; TMT = Trail-Making Test; VDMN = visual/default-mode network; VsAN = visual attention network; * = $p < .05$; ** = $p < .01$.

Table 38. Study 3: Variance table for behavioural CPCA of change (time 2-time 1) in behavioural BADE and fMRI performance on change (time 2-time 1) in fMRI hemodynamic response averaged over poststimulus time for confirm and disconfirm conditions.

	Total	Comp 1	Comp 2	All
Overall	6.00	2.39	1.78	4.17
% Overall	100.00	39.87	29.70	69.57
Predicted	0.85	0.38	0.35	0.73
% Predicted	100.00	44.33	41.23	85.56
% Overall	14.13	6.26	5.82	12.09
Residual	5.15	2.19	1.67	3.85
% Residual	100.00	42.45	32.37	74.81
% Overall	85.87	36.45	27.79	64.24

Table 39. Study 3: Rotated component loadings (overall solution, predicted solution – dependent variables, residual solution) and predictor loadings (predicted solution – independent variables) for behavioural CPCA of change (time 2-time 1) in behavioural BADE and fMRI performance on change (time 2-time 1) in fMRI hemodynamic response averaged over poststimulus time for confirm and disconfirm conditions.

	Comp 1	Comp 2
Overall Solution		
VsAN confirm	**0.90**	-0.14
VsAN disconfirm	**0.69**	0.17
VDMN confirm	**-0.81**	-0.13
VDMN disconfirm	**-0.67**	**-0.54**
CEN confirm	-0.03	**0.81**
CEN disconfirm	0.07	**0.88**
Predicted Solution – Dependent Variables		
VsAN confirm	**0.49**	0.12
VsAN disconfirm	0.12	**0.49**
VDMN confirm	**-0.24**	0.10
VDMN disconfirm	**0.17**	**-0.16**
CEN confirm	**0.16**	**-0.20**
CEN disconfirm	-0.11	**-0.13**
Predicted Solution – Independent Variables		
fMRI accuracy	**-0.85**	**0.44**
fMRI reaction time	**0.30**	**0.31**
BADE evidence integration	**0.85**	**0.31**
BADE conservatism	**0.80**	**0.50**
Residual Solution		
VsAN confirm	**0.76**	-0.10
VsAN disconfirm	**0.56**	0.24
VDMN confirm	**-0.85**	-0.04
VDMN disconfirm	**-0.74**	**-0.47**
CEN confirm	0.00	**0.78**
CEN disconfirm	0.12	**0.88**

Note. Dominant loadings are set in bold text. Predicted solution dependent and independent variables are computed on the same components, and must be interpreted together. fMRI accuracy: correct = both responses equal to condition, incorrect = one or more responses diverges from condition; fMRI reaction time: mean response time across all images, conditions, and trials; BADE (bias against disconfirmatory evidence) evidence integration & conservatism = composites score on behavioural BADE task; CEN = cognitive evaluation network; Comp = component; VDMN = visual/default-mode network; VsAN = visual attention network.

Table 40. Study 3: Variance table for behavioural CPCA of change (time 2-time 1) in neurocognitive measures on change (time 2-time 1) in fMRI hemodynamic response averaged over poststimulus time for confirm and disconfirm conditions.

	Total	Comp 1	Comp 2	All
Overall	6.00	2.32	1.86	4.18
% Overall	100.00	38.70	30.93	69.63
Predicted	1.75	0.68	0.58	1.26
% Predicted	100.00	38.65	33.25	71.90
% Overall	29.11	11.25	9.68	20.93
Residual	4.25	1.86	1.21	3.07
% Residual	100.00	43.81	28.40	72.21
% Overall	70.89	31.06	20.13	51.19

Table 41. Study 3: Rotated component loadings (overall solution, predicted solution – dependent variables, residual solution) and predictor loadings (predicted solution – independent variables) for behavioural CPCA of change (time 2-time 1) in neurocognitive measures on change (time 2-time 1) in fMRI hemodynamic response averaged over poststimulus time for confirm and disconfirm conditions.

	Comp 1	Comp 2
Overall Solution		
VsAN confirm	**0.90**	-0.14
VsAN disconfirm	**0.67**	0.23
VDMN confirm	**-0.81**	-0.13
VDMN disconfirm	**-0.63**	**-0.60**
CEN confirm	-0.05	**0.78**
CEN disconfirm	0.04	**0.89**
Predicted Solution – Dependent Variables		
VsAN confirm	**0.60**	-0.07
VsAN disconfirm	**0.52**	0.18
VDMN confirm	**-0.17**	-0.05
VDMN disconfirm	-0.06	**-0.28**
CEN confirm	-0.14	**0.40**
CEN disconfirm	0.12	**0.50**
Predicted Solution – Independent Variables		
RBANS immediate memory	0.08	**0.18**
RBANS visuospatial	**-0.30**	-0.22
RBANS language	-0.13	**-0.31**
RBANS attention	**0.75**	-0.09
RBANS delayed memory	0.06	**0.54**
TMT A	**-0.32**	-0.16
TMT B	0.18	**-0.49**
LNS	-0.12	**-0.34**
COWAT	**-0.16**	0.07
Residual Solution		
VsAN confirm	**0.69**	-0.16
VsAN disconfirm	**0.54**	0.16
VDMN confirm	**-0.79**	-0.02
VDMN disconfirm	**-0.69**	**-0.45**
CEN confirm	0.08	**0.61**
CEN disconfirm	0.02	**0.73**

Note. Dominant loadings are set in bold text. Predicted solution dependent and independent variables are computed on the same components, and must be interpreted together. CEN = cognitive evaluation network; Comp = component; COWAT = Controlled Oral Word Association Test (standard scores for number correct controlling for age, gender, education level); LNS = Letter-Number Sequencing (standard scores for number correct controlling for age); RBANS = Repeatable Battery for the Assessment of Neuropsychological Status (scaled scores for summed test scores controlling for age); TMT = Trail-Making Test (standard scores for time to completion controlling for age, gender, education level); VDMN = visual/default-mode network; VsAN = visual attention network.

Table 42. Study 3: Variance table for behavioural CPCA of change (time 2-time 1) in symptoms on change (time 2-time 1) in fMRI hemodynamic response averaged over poststimulus time for confirm and disconfirm conditions.

	Total	Comp 1	Comp 2	All
Overall	6.00	2.30	1.82	4.12
% Overall	100.00	38.25	30.38	68.63
Predicted	2.71	1.12	1.04	2.16
% Predicted	100.00	41.46	38.47	79.93
% Overall	45.10	18.70	17.35	36.05
Residual	3.29	1.60	0.72	2.32
% Residual	100.00	48.49	21.98	70.47
% Overall	54.90	26.62	12.07	38.69

Table 43. Study 3: Rotated component loadings (overall solution, predicted solution – dependent variables, residual solution) and predictor loadings (predicted solution – independent variables) for behavioural CPCA of change (time 2-time 1) in symptoms on change (time 2-time 1) in fMRI hemodynamic response averaged over poststimulus time for confirm and disconfirm conditions.

	Comp 1	Comp 2
Overall Solution		
VsAN confirm	**0.90**	-0.12
VsAN disconfirm	**0.64**	0.24
VDMN confirm	**-0.81**	-0.13
VDMN disconfirm	**-0.64**	**-0.57**
CEN confirm	-0.04	**0.80**
CEN disconfirm	0.07	**0.88**
Predicted Solution – Dependent Variables		
VsAN confirm	**0.57**	-0.13
VsAN disconfirm	-0.03	**0.15**
VDMN confirm	**-0.79**	-0.11
VDMN disconfirm	**-0.40**	**-0.46**
CEN confirm	0.11	**0.61**
CEN disconfirm	-0.01	**0.62**
Predicted Solution – Independent Variables		
PSYRATS hallucinations	-0.11	**-0.15**
PSYRATS delusions	**-0.21**	-0.13
SAPS hallucinations	0.02	**-0.11**
SAPS delusions	**-0.58**	0.09
SAPS bizarre behaviour	**-0.13**	-0.02
SAPS formal thought disorder	**-0.48**	**-0.24**
SANS affective flattening	**-0.37**	0.17
SANS alogia	-0.14	**-0.69**
SANS avolition-apathy	0.10	**0.40**
SANS attention	**-0.14**	0.05
SANS inappropriate affect	0.15	**-0.35**
Residual Solution		
VsAN confirm	**0.65**	-0.07
VsAN disconfirm	**0.83**	0.09
VDMN confirm	**-0.40**	-0.01
VDMN disconfirm	**-0.53**	-0.21
CEN confirm	-0.06	**0.54**
CEN disconfirm	0.18	**0.60**

Note. Dominant loadings are set in bold text. Predicted solution dependent and independent variables are computed on the same components, and must be interpreted together. CEN = cognitive evaluation network; Comp = component; PSYRATS = Psychotic Symptoms Rating Scales; SANS/SAPS = Scales for the Assessment of Negative/Positive Symptoms; VDMN = visual/default-mode network; VsAN = visual attention network. PSYRATS/SAPS/SANS = total subscale scores, except for SANS inappropriate affect, which is a single item.

Table 44. Study 3: Intercorrelation table for change (time 2-time 1) in baseline-peak fMRI hemodynamic response.

Variable	VsAN BP	VsAN PB	VDMN BP	VDMN PB	CEN BP	CEN PB
VsAN BP	1	**0.38***	**-0.73****	-0.07	0.08	-0.15
VsAN PB		1	-0.20	**-0.52****	-0.03	0.36
VDMN BP			1	**0.45***	-0.33	**0.41***
VDMN PB				1	**-0.41***	-0.19
CEN BP					1	0.08
CEN PB						1

Note. All values are based on change scores (time 2-time 1). Significant correlations are set in bold text. BP = baseline-to-peak; CEN = cognitive evaluation network; PB = peak-to-baseline; VDMN = visual/default-mode network; VsAN = visual attention network; * = $p < .05$; ** = $p < .01$.

Table 45. Study 3: Correlation table for associations between change (time 2-time 1) in behaviour and change (time 2-time 1) in baseline-peak fMRI hemodynamic response.

Variable	VsAN BP	VsAN PB	VDMN BP	VDMN PB	CEN BP	CEN PB
Behavioural BADE (n = 28)						
Evidence integration	0.33	0.29	0.02	-0.01	-0.10	0.20
Conservatism	**0.38***	**0.41***	-0.10	-0.06	-0.07	0.18
fMRI Performance (n = 29)						
Accuracy	-0.23	-0.13	0.10	0.06	-0.06	-0.24
Accuracy – confirm	-0.33	-0.01	0.22	-0.02	-0.19	-0.16
Accuracy – disconfirm	-0.11	-0.17	0.00	0.09	0.03	-0.22
Reaction time	-0.06	**0.52***	0.10	-0.30	-0.21	0.24
Reaction time – confirm	-0.15	**0.46***	0.06	**-0.39***	-0.16	0.14
Reaction time – disconfirm	0.02	**0.48***	0.12	-0.18	-0.21	0.27
Missed responses	**0.50***	0.24	**-0.44***	-0.23	0.21	0.13
Neurocognition (n = 29)						
RBANS immediate memory	-0.01	0.13	0.13	-0.07	0.18	0.34
RBANS visuospatial	-0.07	-0.29	0.04	0.24	-0.04	-0.24
RBANS language	0.05	-0.11	0.21	0.27	-0.12	0.05
RBANS attention	**0.47***	0.29	-0.19	-0.03	-0.01	0.14
RBANS delayed memory	0.13	-0.00	-0.08	-0.09	**0.39***	0.19
TMT A	-0.08	-0.33	-0.13	-0.07	0.17	-0.29
TMT B	0.01	0.08	-0.09	-0.10	-0.20	-0.00
LNS	-0.04	-0.07	0.15	0.22	-0.12	0.10
COWAT	0.08	-0.31	-0.24	0.00	0.14	**-0.40***
Symptoms (n = 18-29)						
PSYRATS hallucinations	0.02	-0.19	0.06	0.25	-0.17	-0.10
PSYRATS delusions	-0.12	-0.12	0.13	0.18	-0.06	-0.08
SAPS						
Total	-0.25	-0.23	**0.39***	**0.41***	-0.22	0.17
Hallucinations	0.06	-0.22	-0.06	0.14	-0.09	-0.10
Delusions	-0.20	-0.17	0.30	**0.38***	0.01	0.16
Bizarre behaviour	-0.21	0.20	0.18	-0.23	-0.17	0.18
Formal thought disorder	-0.11	-0.16	0.27	0.32	-0.31	0.16

(Table 45 continued)

Variable	VsAN BP	VsAN PB	VDMN BP	VDMN PB	CEN BP	CEN PB
SANS						
Total	0.05	0.12	-0.12	-0.14	0.17	0.26
Affective flattening	-0.05	0.04	0.16	0.15	-0.00	0.11
Alogia	-0.01	-0.09	0.16	0.25	-0.30	-0.08
Avolition-apathy	-0.17	-0.00	-0.07	-0.33	0.28	0.11
Anhedonia-asociality	0.14	0.06	-0.02	-0.00	0.17	0.12
Attention	-0.05	-0.12	0.08	0.04	-0.03	0.28
Inappropriate affect	0.12	0.12	0.03	-0.01	-0.27	-0.06

Note. Significant correlations are set in bold text. BADE = bias against disconfirmatory evidence; BP = baseline-to-peak; CEN = cognitive evaluation network; COWAT = Controlled Oral Word Association Test; LNS = Letter-Number Sequencing; PB = peak-to-baseline; PSYRATS = Psychotic Symptoms Rating Scales; RBANS = Repeatable Battery for the Assessment of Neuropsychological Status; SANS/SAPS = Scales for the Assessment of Negative/Positive Symptoms; TMT = Trail-Making Test; VDMN = visual/default-mode network; VsAN = visual attention network; * = $p < .05$; ** = $p < .01$.

Table 46. Study 3: Variance table for behavioural CPCA of change (time 2-time 1) in behavioural BADE and fMRI performance on change (time 2-time 1) in baseline-peak fMRI hemodynamic response.

	Total	Comp 1	Comp 2	Comp 3	All
Overall	6.00	1.99	1.63	1.41	5.03
% Overall	100.00	33.15	27.20	23.56	83.91
Predicted	0.88	0.58	0.24	-	0.82
% Predicted	100.00	65.86	27.69	-	93.55
% Overall	14.61	9.62	4.04	-	13.67
Residual	5.12	1.89	1.69	-	3.58
% Residual	100.00	36.93	33.02	-	69.95
% Overall	85.39	31.54	28.20	-	59.73

Table 47. Study 3: Rotated component loadings (overall solution, predicted solution – dependent variables, residual solution) and predictor loadings (predicted solution – independent variables) for behavioural CPCA of change (time 2-time 1) in behavioural BADE and fMRI performance on change (time 2-time 1) in baseline-peak fMRI hemodynamic response.

	Comp 1	Comp 2	Comp 3
Overall Solution			
VsAN BP	**0.88**	0.22	-0.11
VsAN PB	0.28	**0.90**	0.01
VDMN BP	**-0.90**	0.01	**-0.36**
VDMN PB	-0.21	**-0.52**	**-0.67**
CEN BP	0.06	-0.07	**0.90**
CEN PB	**-0.52**	**0.71**	0.04
Predicted Solution – Dependent Variables			
VsAN BP	0.07	**0.43**	
VsAN PB	**0.61**	0.03	
VDMN BP	0.11	**-0.15**	
VDMN PB	**-0.27**	**0.15**	
CEN BP	**-0.26**	0.09	
CEN PB	**0.21**	0.08	
Predicted Solution – Independent Variables			
fMRI accuracy	-0.19	**-0.47**	
fMRI reaction time	**0.85**	**-0.35**	
BADE evidence integration	**0.54**	**0.68**	
BADE conservatism	**0.57**	**0.79**	
Residual Solution			
VsAN BP	**0.73**	0.12	
VsAN PB	0.17	**0.51**	
VDMN BP	**-0.90**	-0.30	
VDMN PB	-0.28	**-0.78**	
CEN BP	0.12	**0.68**	
CEN PB	**-0.66**	**0.50**	

Note. Dominant loadings are set in bold text. Predicted solution dependent and independent variables are computed on the same components, and must be interpreted together. fMRI accuracy: correct = both responses equal to condition, incorrect = one or more responses diverges from condition; fMRI reaction time: mean response time across all images, conditions, and trials; BADE (bias against disconfirmatory evidence) evidence integration & conservatism = composites score on behavioural BADE task. BP = baseline-to-peak; CEN = cognitive evaluation network; Comp = component; PB = peak-to-baseline; VDMN = visual/default-mode network; VsAN = visual attention network.

Table 48. Study 3: Variance table for behavioural CPCA of change (time 2-time 1) in neurocognitive measures on change (time 2-time 1) in baseline-peak fMRI hemodynamic response.

	Total	Comp 1	Comp 2	Comp 3	All
Overall	6.00	1.98	1.62	1.43	5.04
% Overall	100.00	33.07	27.04	23.91	84.03
Predicted	1.89	0.91	0.47	-	1.38
% Predicted	100.00	47.94	24.91	-	72.86
% Overall	31.56	15.13	7.86	-	22.99
Residual	4.11	1.59	1.15	0.77	3.52
% Residual	100.00	38.77	28.07	18.81	85.65
% Overall	68.44	26.53	19.21	12.87	58.62

Table 49. Study 3: Rotated component loadings (overall solution, predicted solution – dependent variables, residual solution) and predictor loadings (predicted solution – independent variables) for behavioural CPCA of change (time 2-time 1) in neurocognitive measures on change (time 2-time 1) in baseline-peak fMRI hemodynamic response.

	Comp 1	Comp 2	Comp 3
Overall Solution			
VsAN BP	**0.88**	0.22	-0.10
VsAN PB	0.29	**0.89**	0.00
VDMN BP	**-0.90**	0.01	**-0.36**
VDMN PB	-0.20	**-0.52**	**-0.68**
CEN BP	0.07	-0.07	**0.91**
CEN PB	**-0.51**	**0.71**	0.08
Predicted Solution – Dependent Variables			
VsAN BP	**0.32**	0.25	
VsAN PB	**0.65**	0.02	
VDMN BP	0.09	**-0.34**	
VDMN PB	-0.13	**-0.26**	
CEN BP	-0.11	**0.43**	
CEN PB	**0.59**	-0.11	
Predicted Solution – Independent Variables			
RBANS immediate memory	**0.29**	-0.13	
RBANS visuospatial	**-0.42**	-0.25	
RBANS language	-0.03	**-0.39**	
RBANS attention	**0.45**	**0.33**	
RBANS delayed memory	0.13	**0.44**	
TMT A	**-0.49**	0.12	
TMT B	0.07	-0.15	
LNS	-0.02	**-0.46**	
COWAT	**-0.49**	**0.40**	
Residual Solution			
VsAN BP	**0.78**	0.04	-0.02
VsAN PB	0.18	**0.64**	-0.15
VDMN BP	**-0.84**	-0.23	-0.18
VDMN PB	-0.12	**-0.78**	-0.26
CEN BP	0.12	0.14	**0.74**
CEN PB	**-0.49**	0.23	0.23

Note. Dominant loadings are set in bold text. Predicted solution dependent and independent variables are computed on the same components, and must be interpreted together. BP = baseline-to-peak; CEN = cognitive evaluation network; Comp = component; COWAT = Controlled Oral Word Association Test (standard scores for number correct controlling for age, gender, education level); LNS = Letter-Number Sequencing (standard scores for number correct controlling for age); PB = peak-to-baseline; RBANS = Repeatable Battery for the Assessment of Neuropsychological Status (scaled scores for summed test scores controlling for age); TMT = Trail-Making Test (standard scores for time to completion controlling for age, gender, education level); VDMN = visual/default-mode network; VsAN = visual attention network.

Table 50. Study 3: Variance table for behavioural CPCA of change (time 2-time 1) in symptoms on change (time 2-time 1) in baseline-peak fMRI hemodynamic response.

	Total	Comp 1	Comp 2	Comp 3	All
Overall	6.00	1.97	1.65	1.39	5.02
% Overall	100.00	32.83	27.58	23.19	83.61
Predicted	2.59	1.13	0.94	-	2.07
% Predicted	100.00	43.76	36.19	-	79.96
% Overall	43.09	18.86	15.60	-	34.45
Residual	3.41	1.26	1.17	-	2.43
% Residual	100.00	37.02	34.29	-	71.31
% Overall	56.91	21.07	19.51	-	40.58

Table 51. Study 3: Rotated component loadings (overall solution, predicted solution – dependent variables, residual solution) and predictor loadings (predicted solution – independent variables) for behavioural CPCA of change (time 2-time 1) in symptoms on change (time 2-time 1) in baseline-peak fMRI hemodynamic response.

	Comp 1	Comp 2	Comp 3
Overall Solution			
VsAN BP	**0.86**	0.23	-0.09
VsAN PB	0.19	**0.92**	-0.01
VDMN BP	**-0.90**	-0.03	-0.37
VDMN PB	-0.19	**-0.58**	**-0.64**
CEN BP	0.08	-0.04	**0.91**
CEN PB	**-0.58**	**0.64**	0.06
Predicted Solution – Dependent Variables			
VsAN BP	-0.05	**0.48**	
VsAN PB	**0.29**	0.02	
VDMN BP	-0.31	**-0.62**	
VDMN PB	**-0.69**	-0.17	
CEN BP	**0.65**	0.03	
CEN PB	0.00	**-0.53**	
Predicted Solution – Independent Variables			
PSYRATS hallucinations	**-0.22**	0.00	
PSYRATS delusions	**-0.24**	0.00	
SAPS hallucinations	**-0.20**	0.19	
SAPS delusions	**-0.22**	**-0.42**	
SAPS bizarre behaviour	**0.21**	**-0.38**	
SAPS formal thought disorder	**-0.41**	**-0.31**	
SANS affective flattening	-0.05	**-0.21**	
SANS alogia	**-0.64**	-0.01	
SANS avolition-apathy	**0.45**	-0.10	
SANS attention	-0.03	**-0.24**	
SANS inappropriate affect	**-0.14**	0.11	
Residual Solution			
VsAN BP	0.16	**0.78**	
VsAN PB	**0.75**	**0.25**	
VDMN BP	0.03	**-0.64**	
VDMN PB	**-0.39**	-0.14	
CEN BP	-0.10	**0.24**	
CEN PB	**0.68**	-0.13	

Note. Dominant loadings are set in bold text. Predicted solution dependent and independent variables are computed on the same components, and must be interpreted together. BP = baseline-to-peak; CEN = cognitive evaluation network; Comp = component; PB = peak-to-baseline; PSYRATS = Psychotic Symptoms Rating Scales; SANS/SAPS = Scales for the Assessment of Negative/Positive Symptoms; VDMN = visual/default-mode network; VsAN = visual attention network. PSYRATS/SAPS/SANS = total subscale scores, except for SANS inappropriate affect, which is a single item.

Figures

Figure 1. Study 1: Timeline of the evidence integration tasks (disconfirm condition). Each trial began with the presentation of a distorted image of two animals (e.g., bird and dolphin) morphed together at a ratio of 60:40 (Version 1) or 70:30 (Version 2) for 500ms. After a 250ms mask (version 2 only), participants were presented with a 16-point rating scale and were asked to indicate the degree to which the image appeared to be of one animal or the other. After 6 seconds, or once a rating was made, a less distorted image of the same animals morphed together at a different ratio was displayed on screen for 3 seconds, and participants were asked to re-rate the image. A = Version 1; B = Version 2.
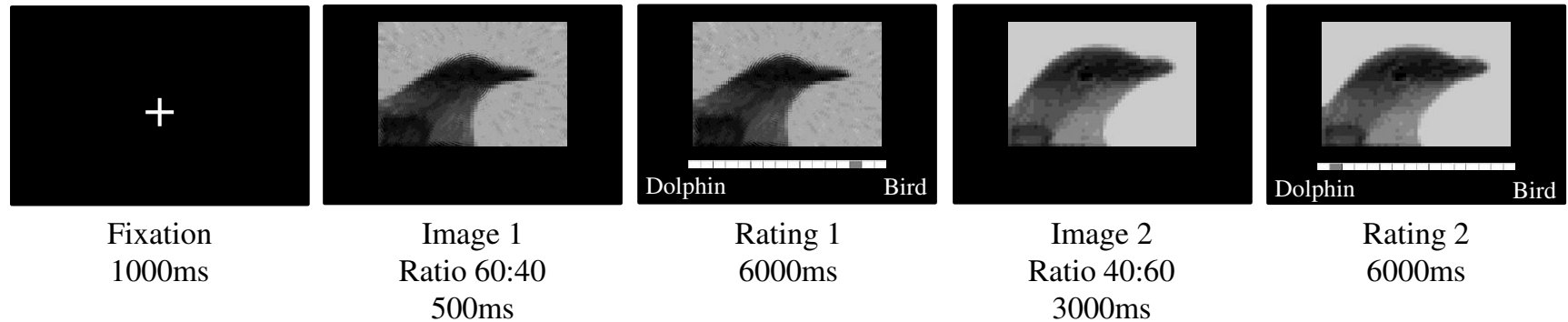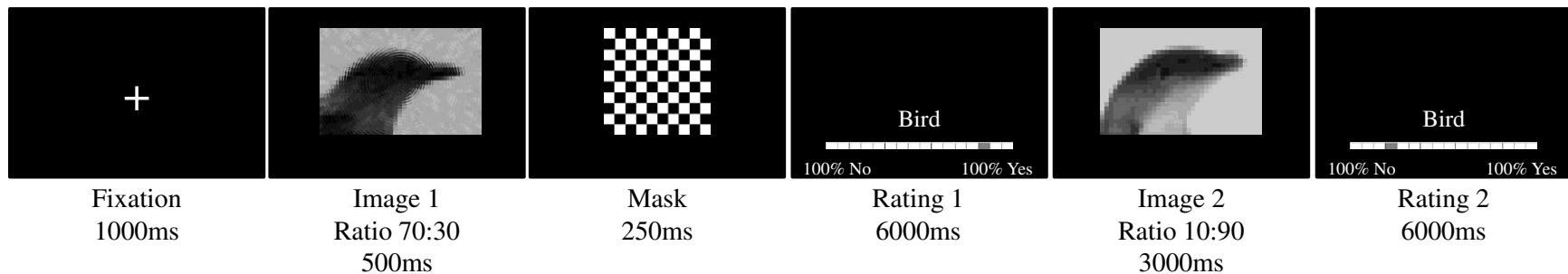
Figure 2. Study 1: A: Dominant 10% of component loadings for cognitive evaluation network (Component 1; red/yellow = positive loadings, threshold = 0.16, maximum = 0.28; no negative loadings passed threshold). Montreal Neurological Institute Z-axis coordinates are displayed; left = left. B: Mean finite impulse response (FIR)-based predictor weights averaged over conditions and plotted as a function of poststimulus time. C: Mean FIR-based predictor weights averaged over versions and plotted as a function of poststimulus time; [a] = confirm > no change; [b] = disconfirm > no change; [c] = disconfirm > confirm. Error bars are standard errors. * = $p < .005$, ** = $p < .001$.
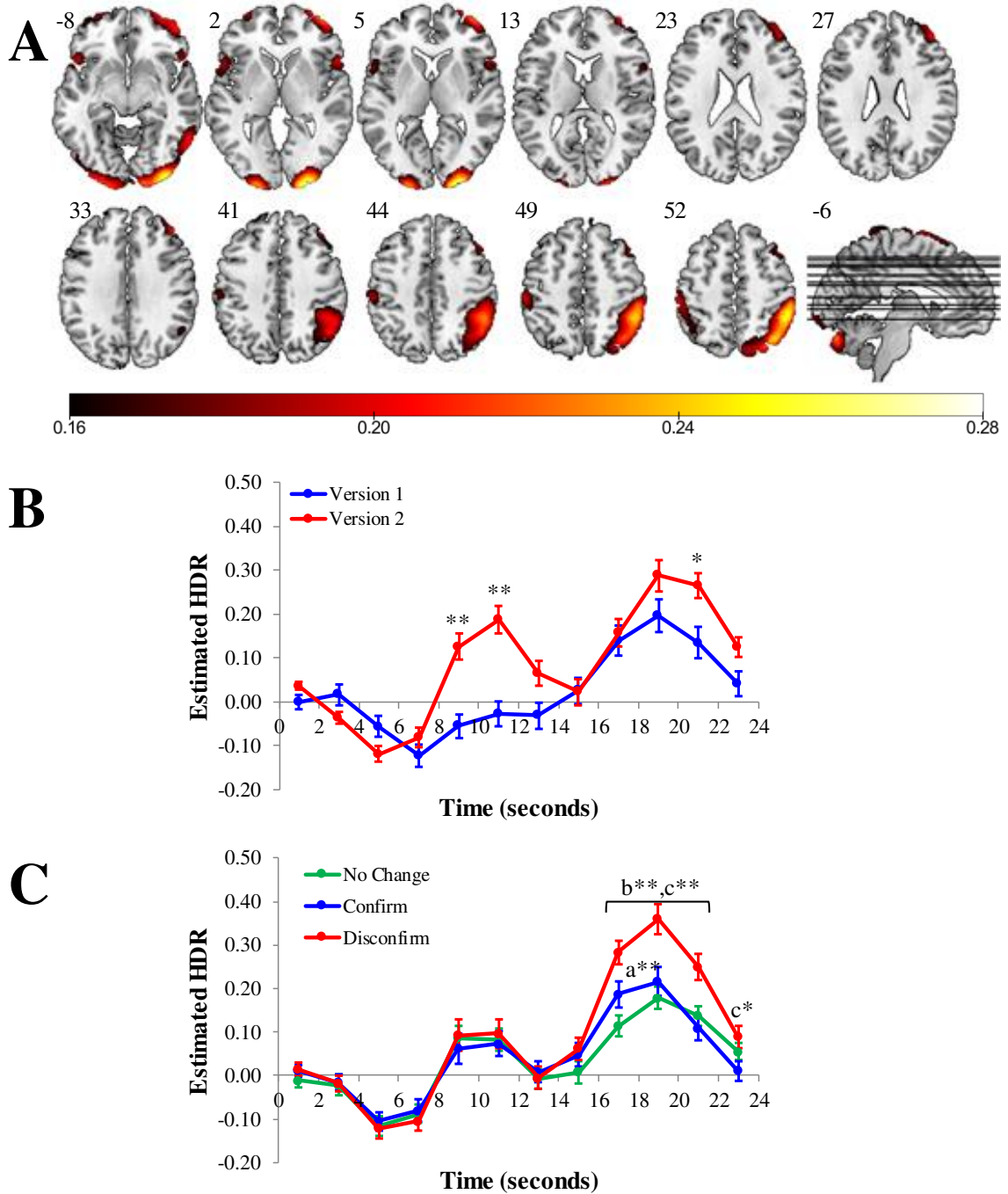
Figure 3. Study 1: A: Dominant 10% of component loadings for visual/default-mode network (Component 2; red/yellow = positive loadings, threshold = 0.16, maximum = 0.40; blue/green = negative loadings, threshold = -0.13, minimum = -0.23). Montreal Neurological Institute Z-axis coordinates are displayed; left = left. B: Mean finite impulse response (FIR)-based predictor weights averaged over conditions and plotted as a function of poststimulus time. C: Mean FIR-based predictor weights averaged over versions and plotted as a function of poststimulus time; [a] = disconfirm > no change; [b] = disconfirm > confirm; [c] = no change > disconfirm. Error bars are standard errors. * = $p < .01$, ** = $p < .001$.
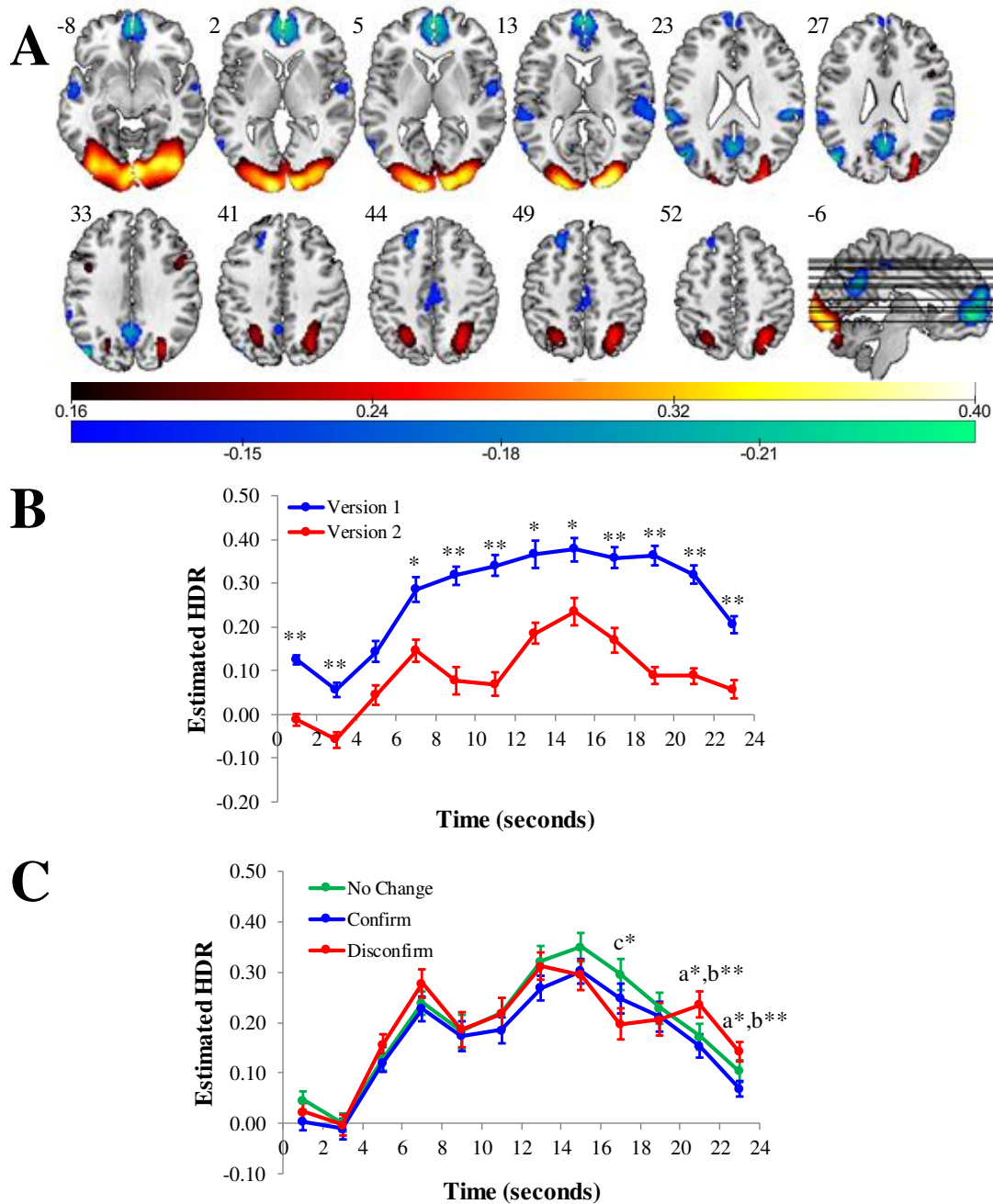
Figure 4. Study 1: A: Dominant 10% of component loadings for response network (Component 4; red/yellow = positive loadings, threshold = 0.15, maximum = 0.33; blue/green = negative loadings, threshold = -0.23, minimum = -0.15). Montreal Neurological Institute Z-axis coordinates are displayed; left = left. B: Mean finite impulse response (FIR)-based predictor weights averaged over conditions and plotted as a function of poststimulus time. C: Mean FIR-based predictor weights averaged over versions and plotted as a function of poststimulus time; [a] = confirm > no change; [b] = disconfirm > no change; [c] = disconfirm > confirm; [d] = confirm > disconfirm. Error bars are standard errors. * = $p < .01$, ** = $p < .001$.
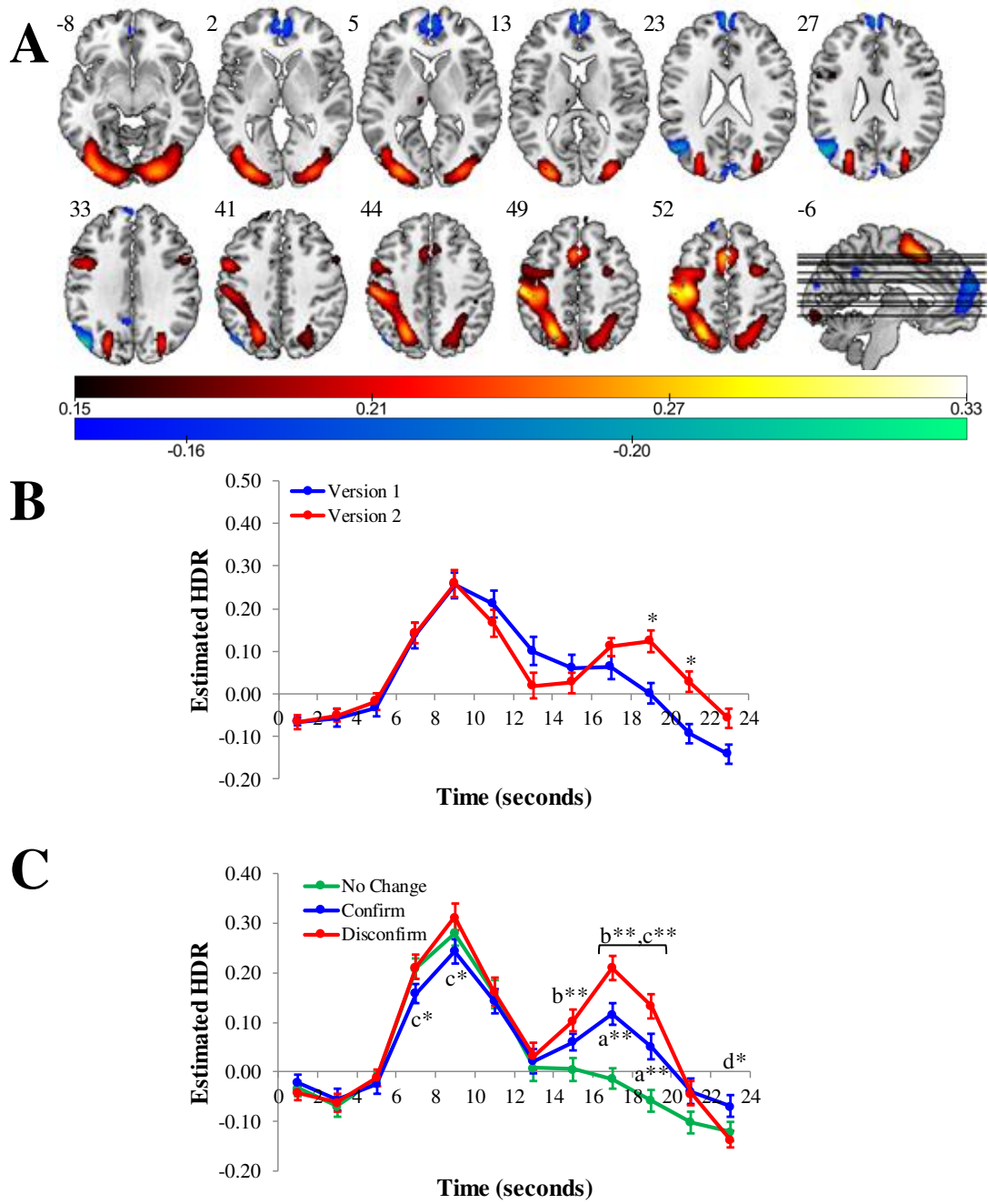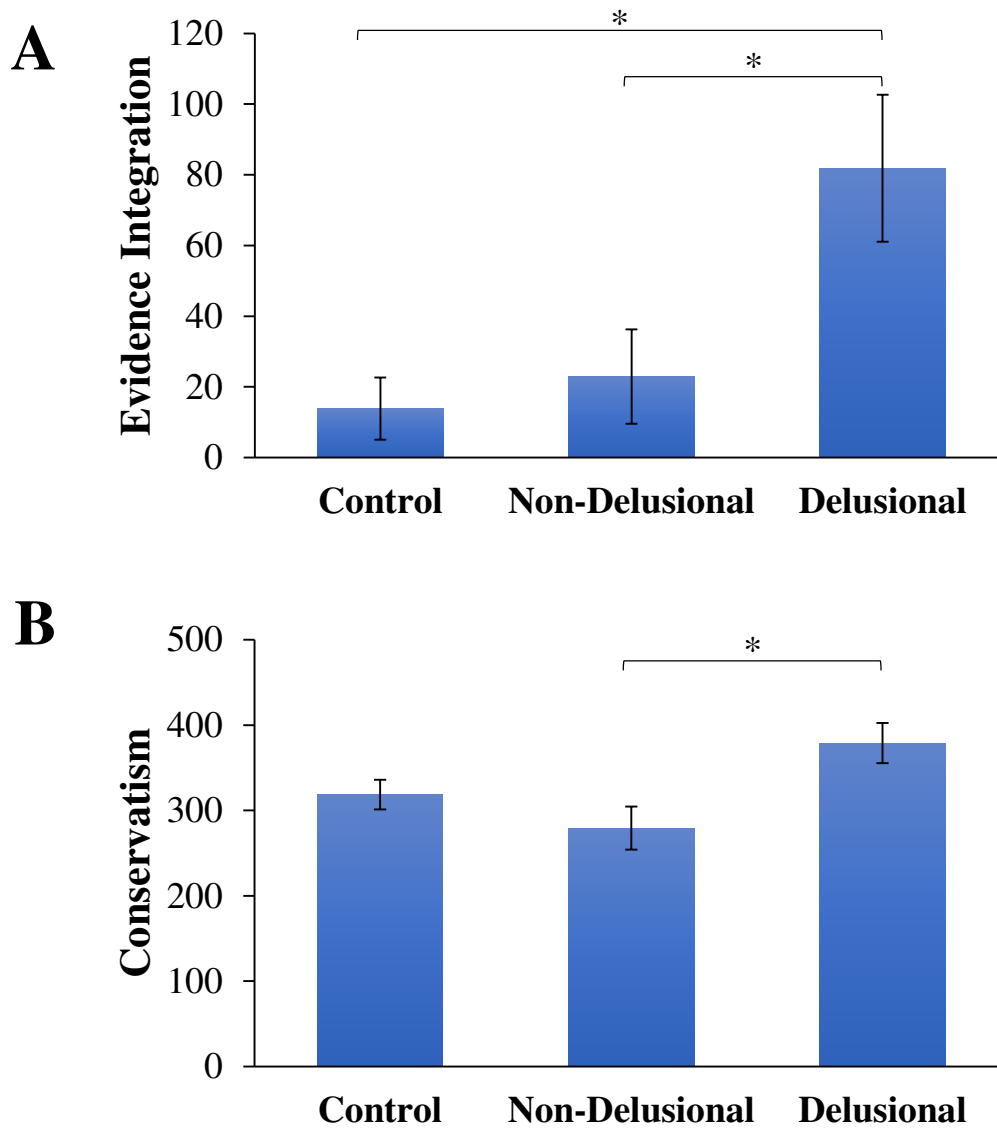
Figure 5. Study 1: A: Dominant 10% of component loadings for visual attention network (Component 5; red/yellow = positive loadings, threshold = 0.13, maximum = 0.30; blue/green = negative loadings, threshold = -0.13, minimum = -0.18). Montreal Neurological Institute Z-axis coordinates are displayed; left = left. B: Mean finite impulse response (FIR)-based predictor weights averaged over conditions and plotted as a function of poststimulus time. C: Mean FIR-based predictor weights averaged over versions and plotted as a function of poststimulus time; [a] = disconfirm > no change; [b] = disconfirm > confirm; [c] = no change > confirm; [d] = no change > disconfirm. Error bars are standard errors. * = $p < .01$, ** = $p < .001$.
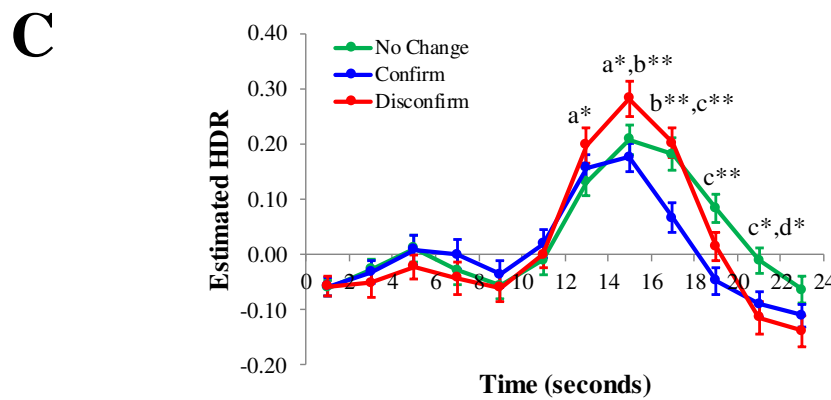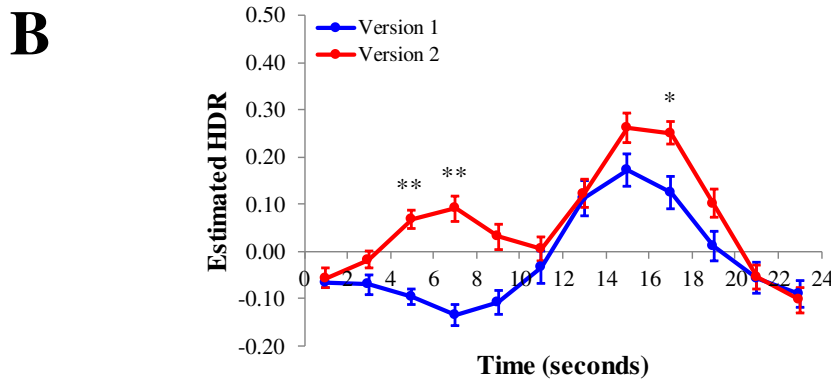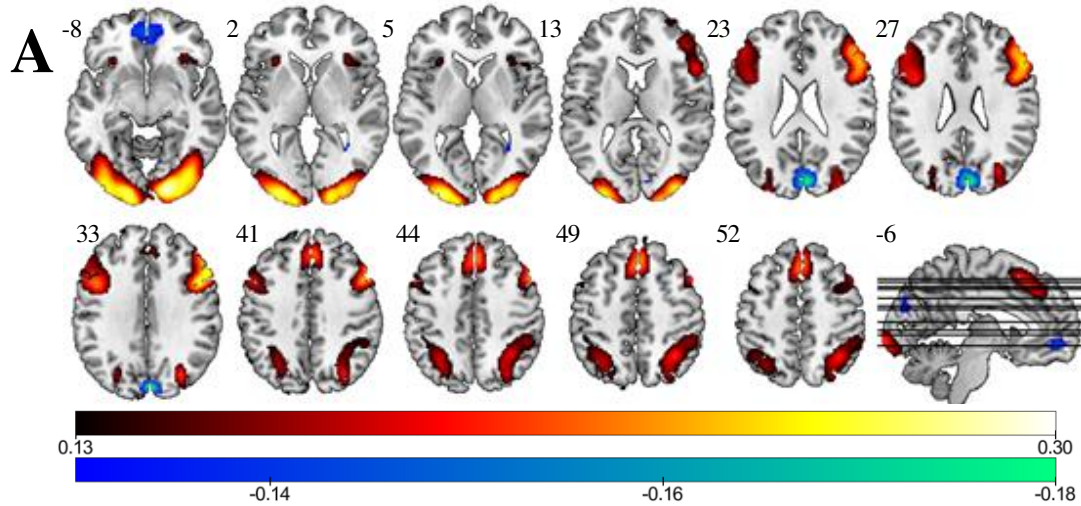
Figure 6. Study 2: Behavioural BADE task (order of interpretations: neutral lure, emotional lure, absurd, true).

Jenny can't fall asleep.

Jenny can't wait until it is finally morning.

Jenny wonders how many presents she will find under the tree.

Jenry is nervous about her exam the next day.

0.2

implausible   somewhat plausible   very plausible

Jenry is worried about her ill mother.

0.9

implausible   somewhat plausible   very plausible

Jenry loves her bed.

0.9

implausible   somewhat plausible   very plausible

Jenry is excited about Christmas morning.

9.8

implausible   somewhat plausible   very plausible

Figure 7. Study 2: fMRI evidence integration task (example of lure image "umbrella" for picture "bat").

Full picture (bat)   Lure image (umbrella)

Figure 8. Study 2: Examples of the four conditions for the fMRI evidence integration task.
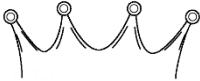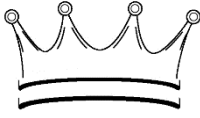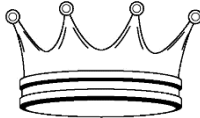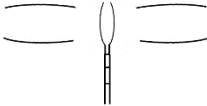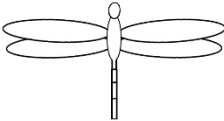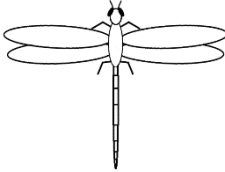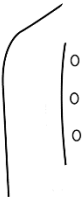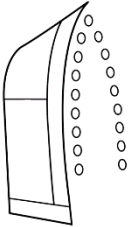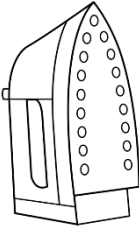
Condition	Image 1	Image 2	Image 3 (Full Picture)
Yes-Yes (YY)			
Prompt word (type)	Crown (true)	Crown (true)	Crown (true)
No-No (NN)			
Prompt word (type)	Sofa (absurd)	Sofa (absurd)	Dragonfly (true)
No-Yes (NY)			
Prompt word (type)	Ladybug (true)	Ladybug (true)	Ladybug (true)
Yes-No (YN)			
Prompt word (type)	Jacket (lure)	Jacket (lure)	Iron (true)

Figure 9. Study 2: Timing of the fMRI evidence integration task.

4000ms	4000ms	1000ms
(Image 1 & Rating)	(Image 2 & Rating)	(Image 3)
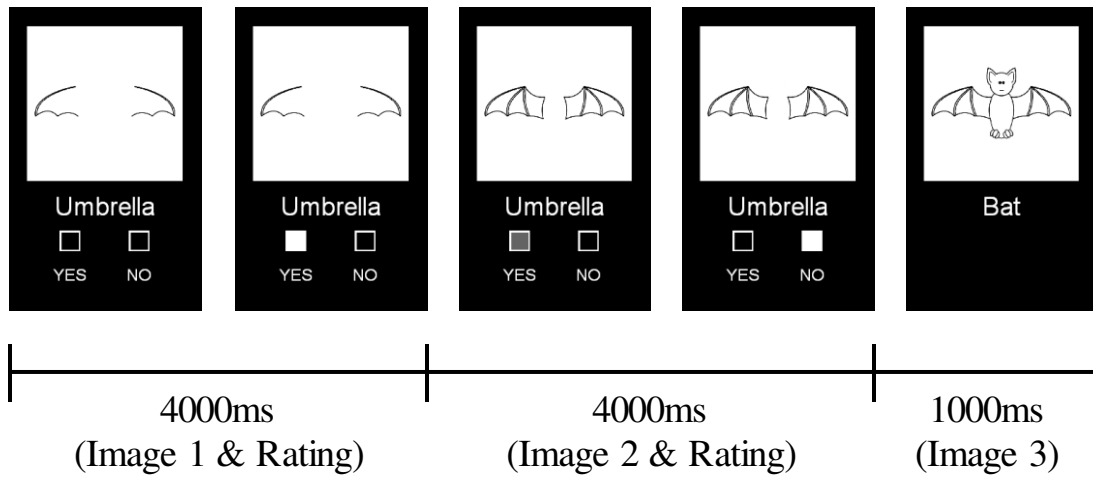
Figure 10. Study 2: Group differences on behavioural BADE evidence integration (A) and conservatism (B). Error bars are standard errors.

Figure 11. Study 2: Dominant 10% of component loadings for visual attention network (Component 1: red/yellow = positive loadings, threshold = 0.24, maximum = 0.47; no negative loadings passed threshold). Montreal Neurological Institute Z-axis coordinates are displayed; left = left.
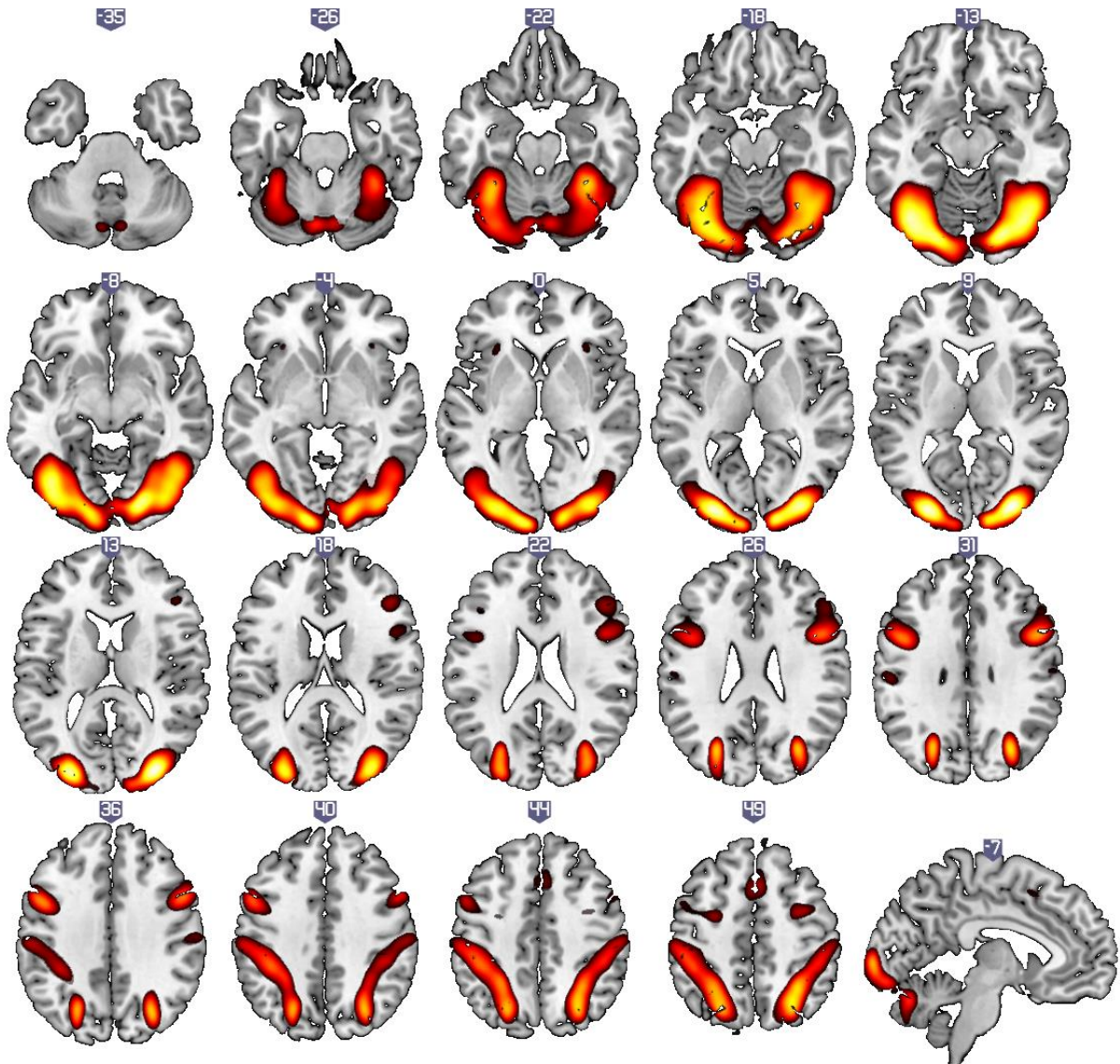
Figure 12. Study 2: Estimated hemodynamic response (HDR) for visual attention network (Component 1). A: Main effect of Condition. B: Main effect of Group. C: Interaction between Condition and Time. [a] = YY > NN; [b] = NN > YY; [c] = NY > YY; [d] = YY > YN; [e] = YN > YY; [f] = NY > NN; [g] = YN > NN; [h] = NY > YN; [i] = YN > NY; * = $p < .05$; ** = $p < .01$. Error bars are standard errors.
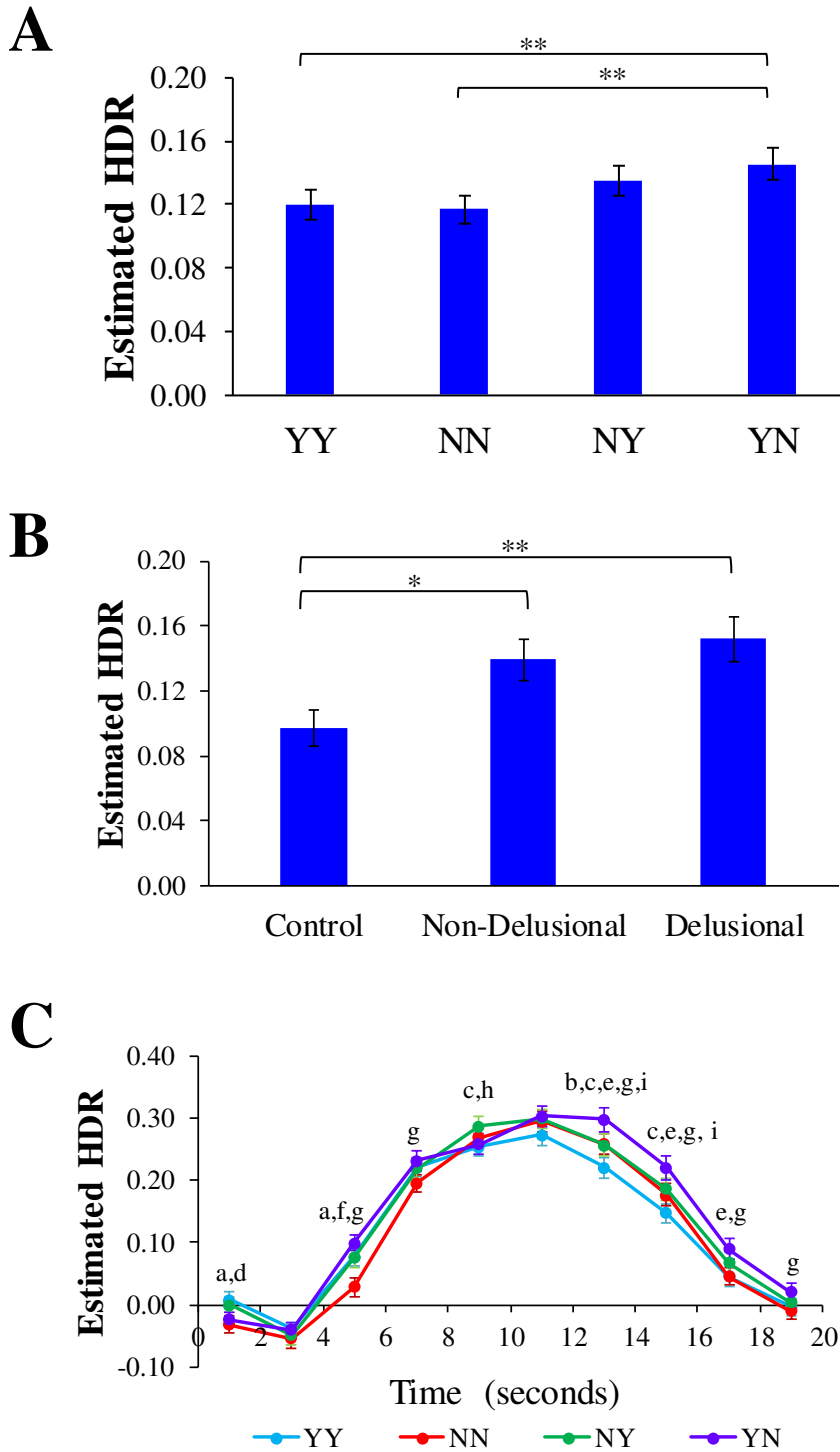
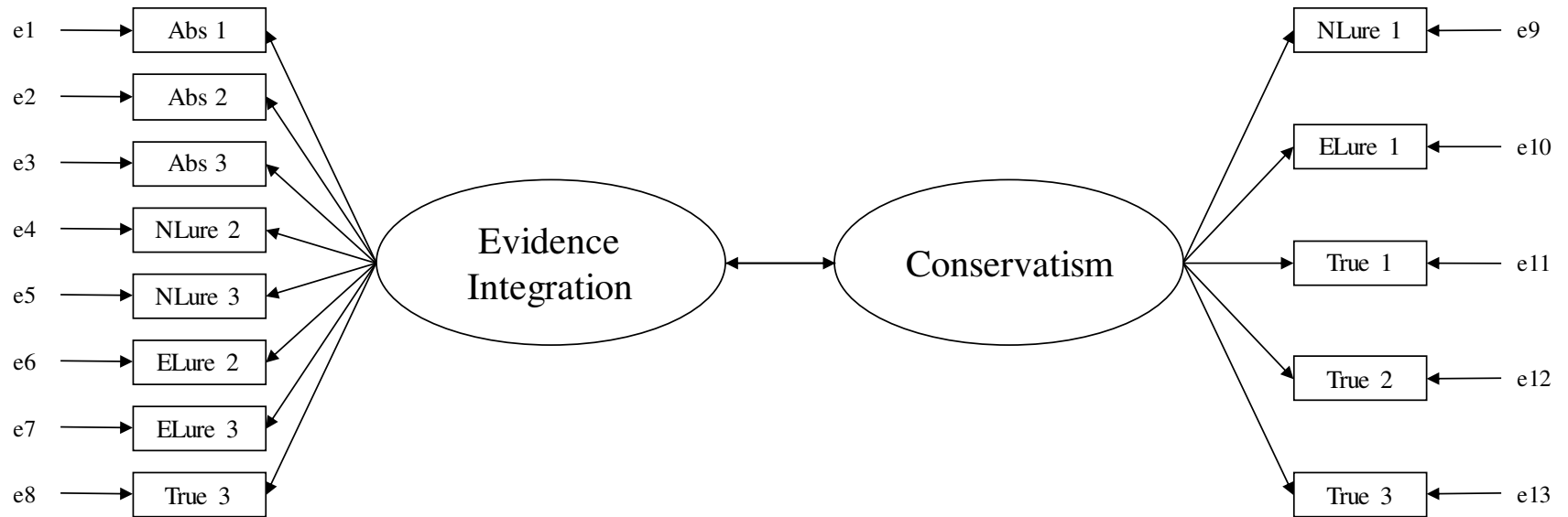Figure 13. Study 2: Dominant 10% of component loadings for visual/default-mode network (Component 2: red/yellow = positive loadings, threshold = 0.19, maximum = 0.35; blue/green = negative loadings; threshold = -0.19, minimum = -0.41). Montreal Neurological Institute Z-axis coordinates are displayed; left = left.
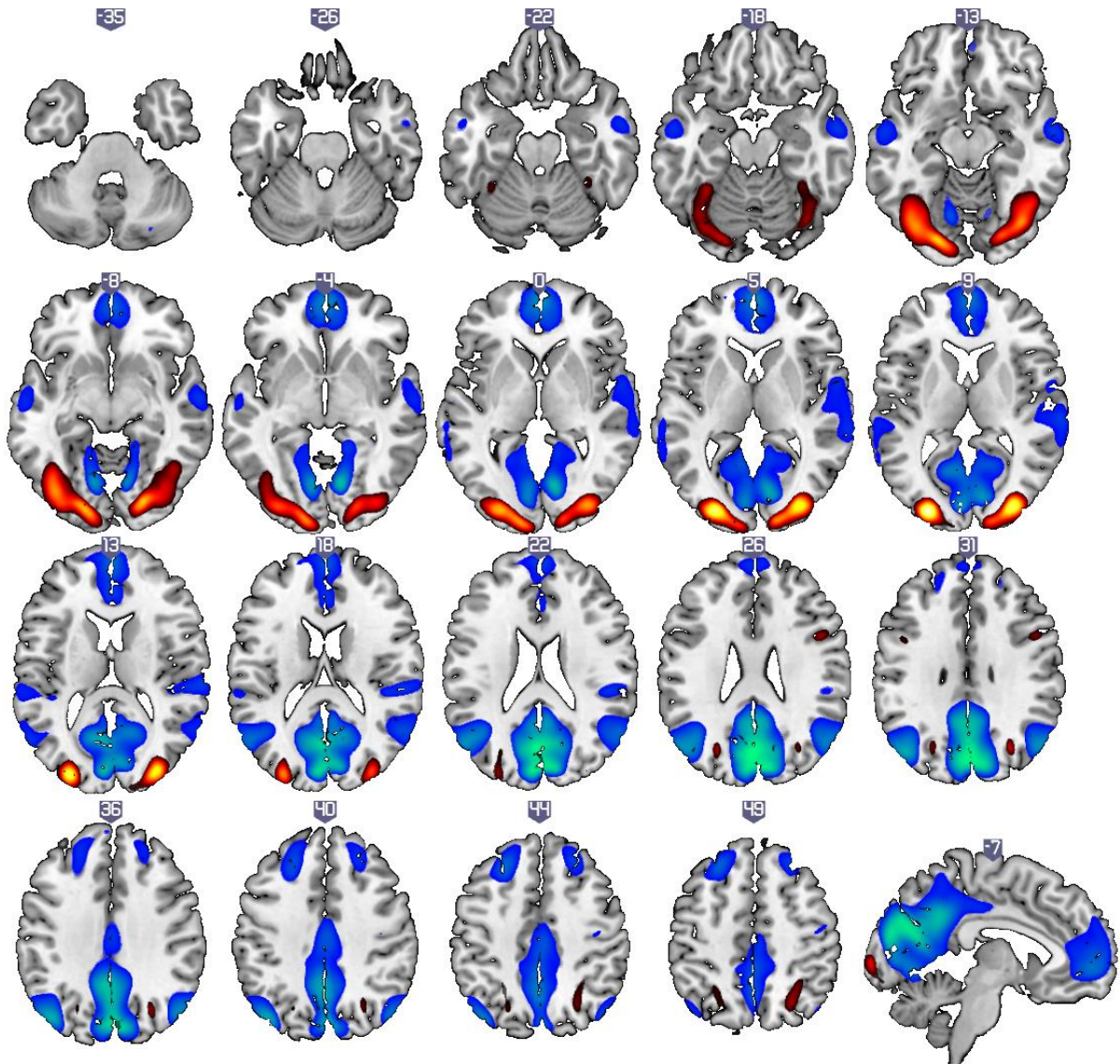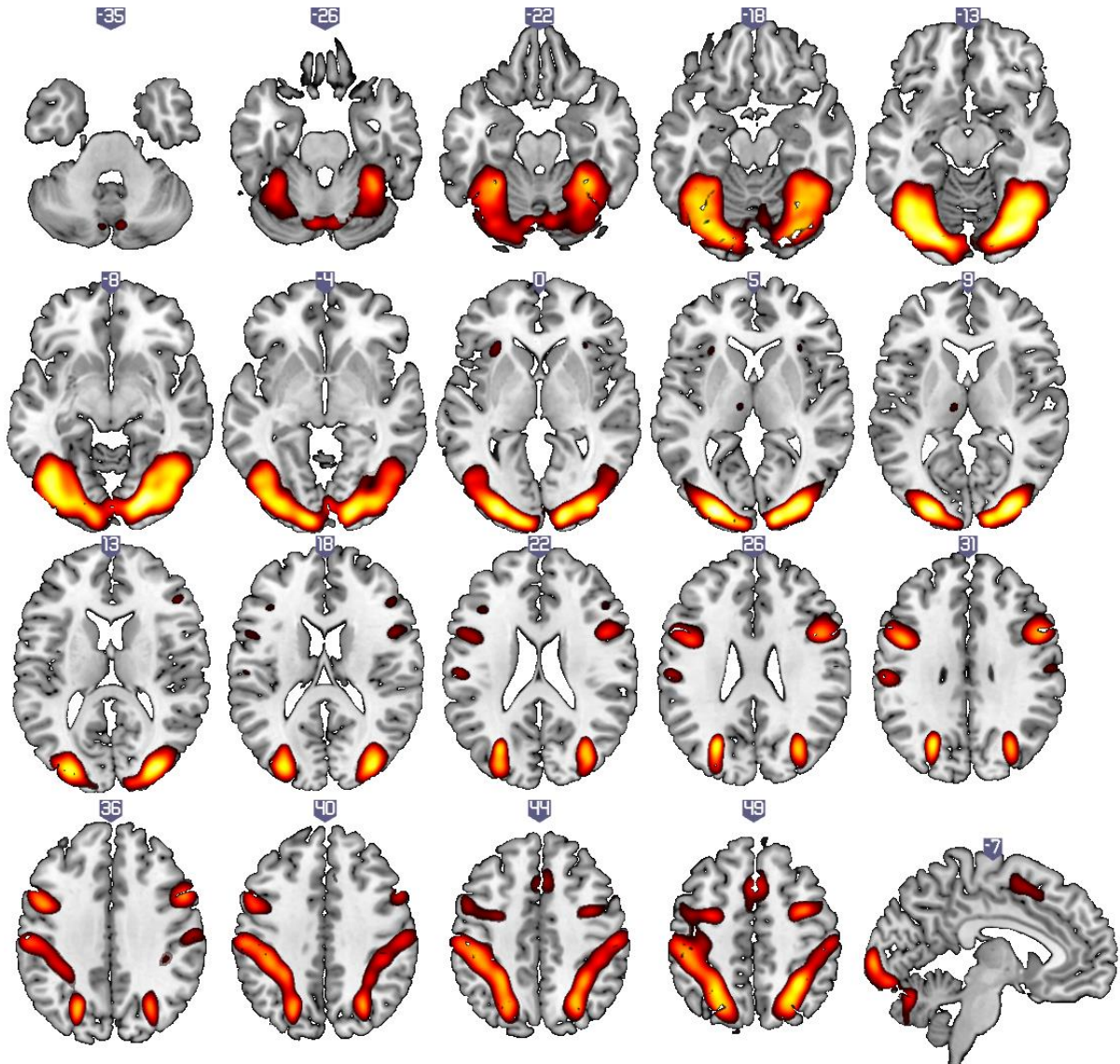
Figure 14. Study 2: Estimated hemodynamic response (HDR) for visual/default-mode network (Component 2). A: = Main effect of Group. B: Interaction between Condition and Time. [a] = YN > YY; [b] = YN > NN; [c] = YN > NY; [d] = YY > NN; [e] = NY > YY; * = $p < .005$. Error bars are standard errors.
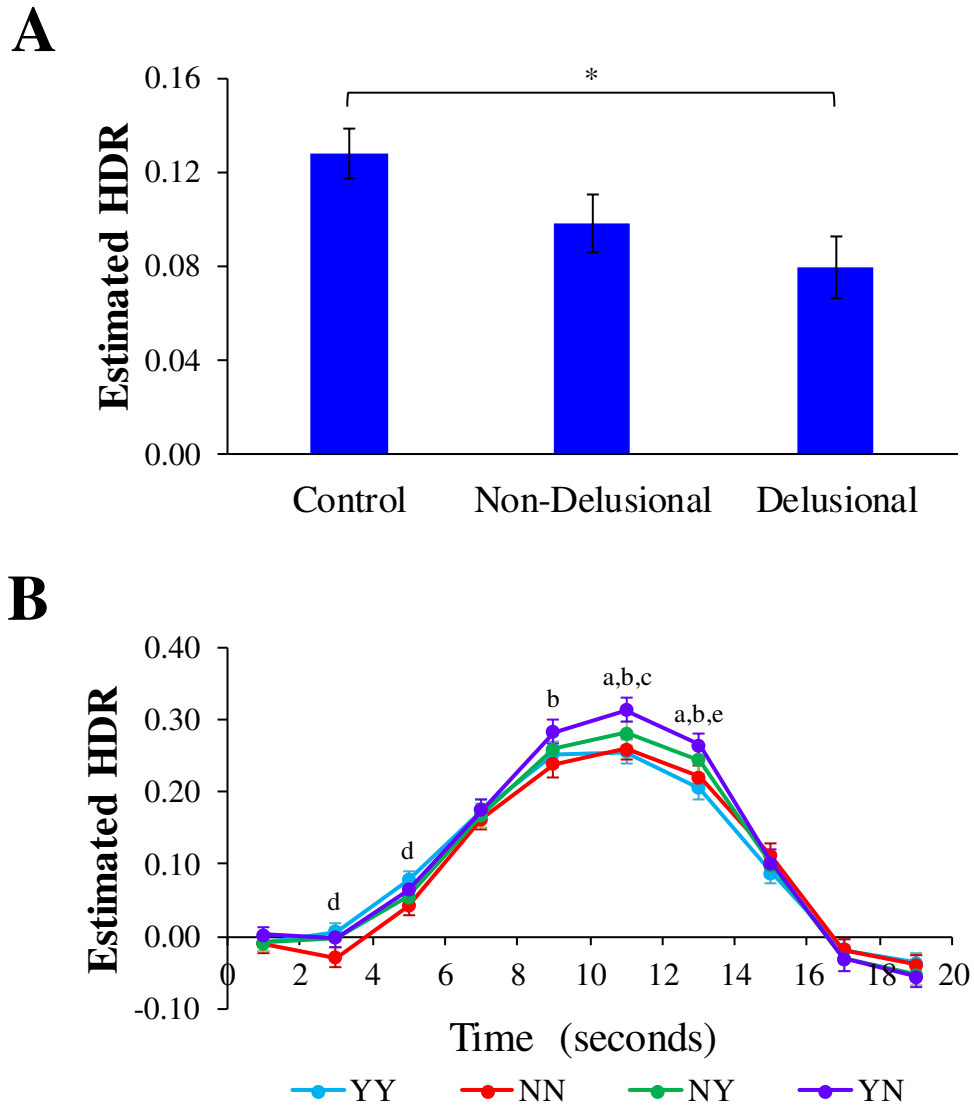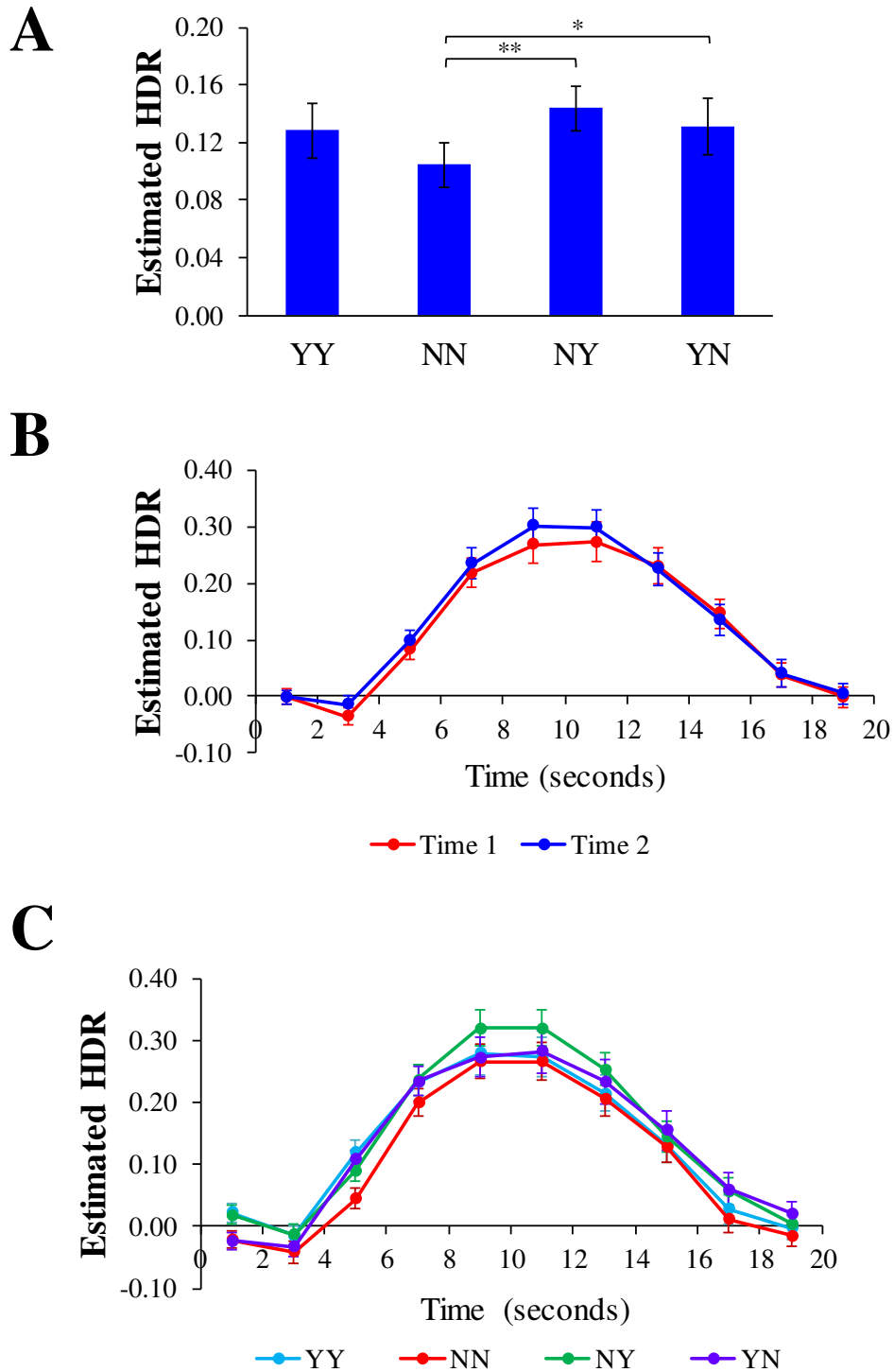
Figure 15. Study 2: Dominant 10% of component loadings for cognitive evaluation network (Component 3: red/yellow = positive loadings, threshold = 0.11, maximum = 0.25; blue/green = negative loadings, threshold = -0.11, minimum = -0.24). Montreal Neurological Institute Z-axis coordinates are displayed; left = left.
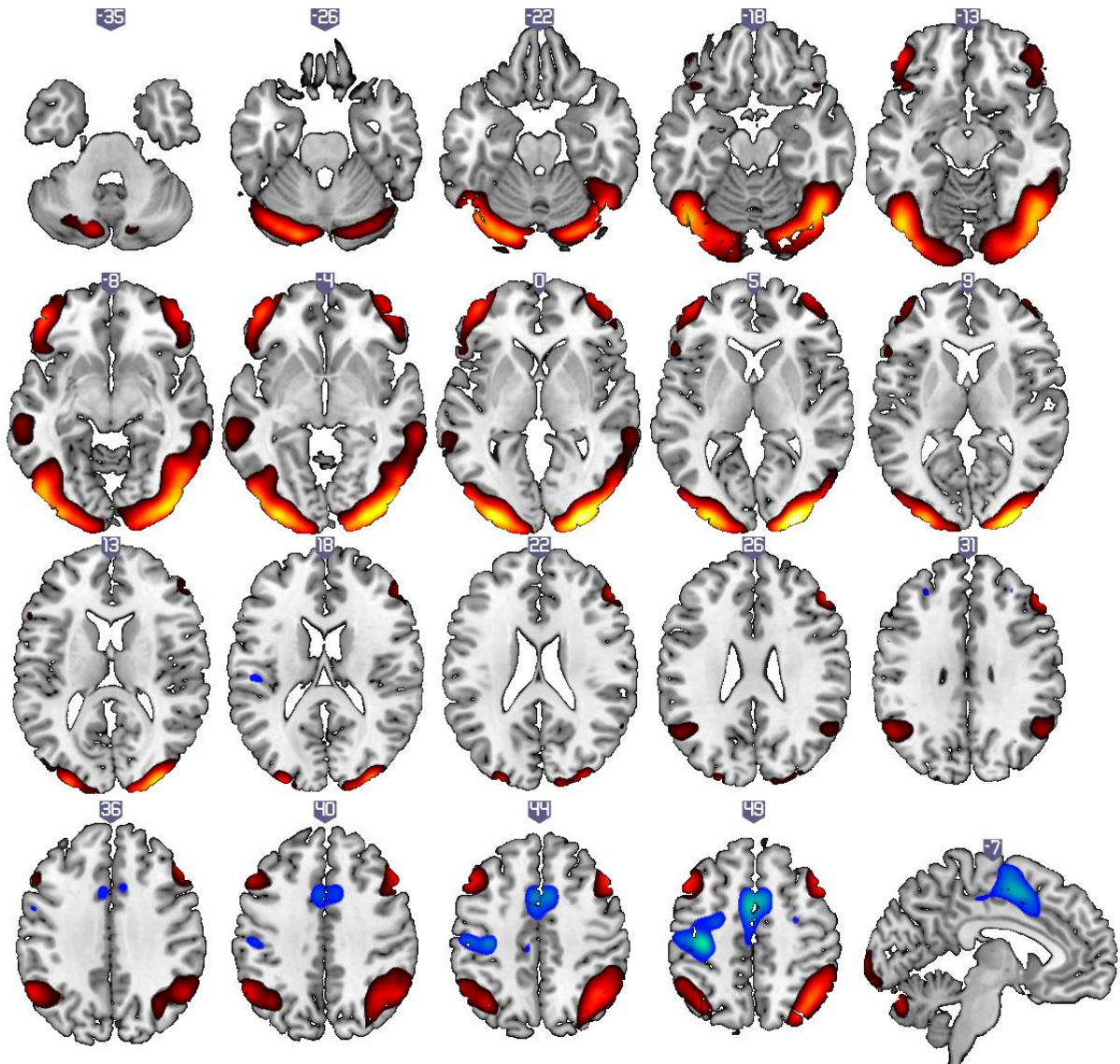
Figure 16. Study 2: Estimated hemodynamic response (HDR) for cognitive evaluation network (Component 3). A: Main effect of Condition. B: Main effect of Group (non-significant). C: Interaction between Condition and Time. [a] = YN > YY; [b] = YN > NN; [c] = YN > NY; [d] = NY > YY; [e] = NY > NN; [f] = YY > NN; * = $p < .05$; ** = $p < .001$. Error bars are standard errors.
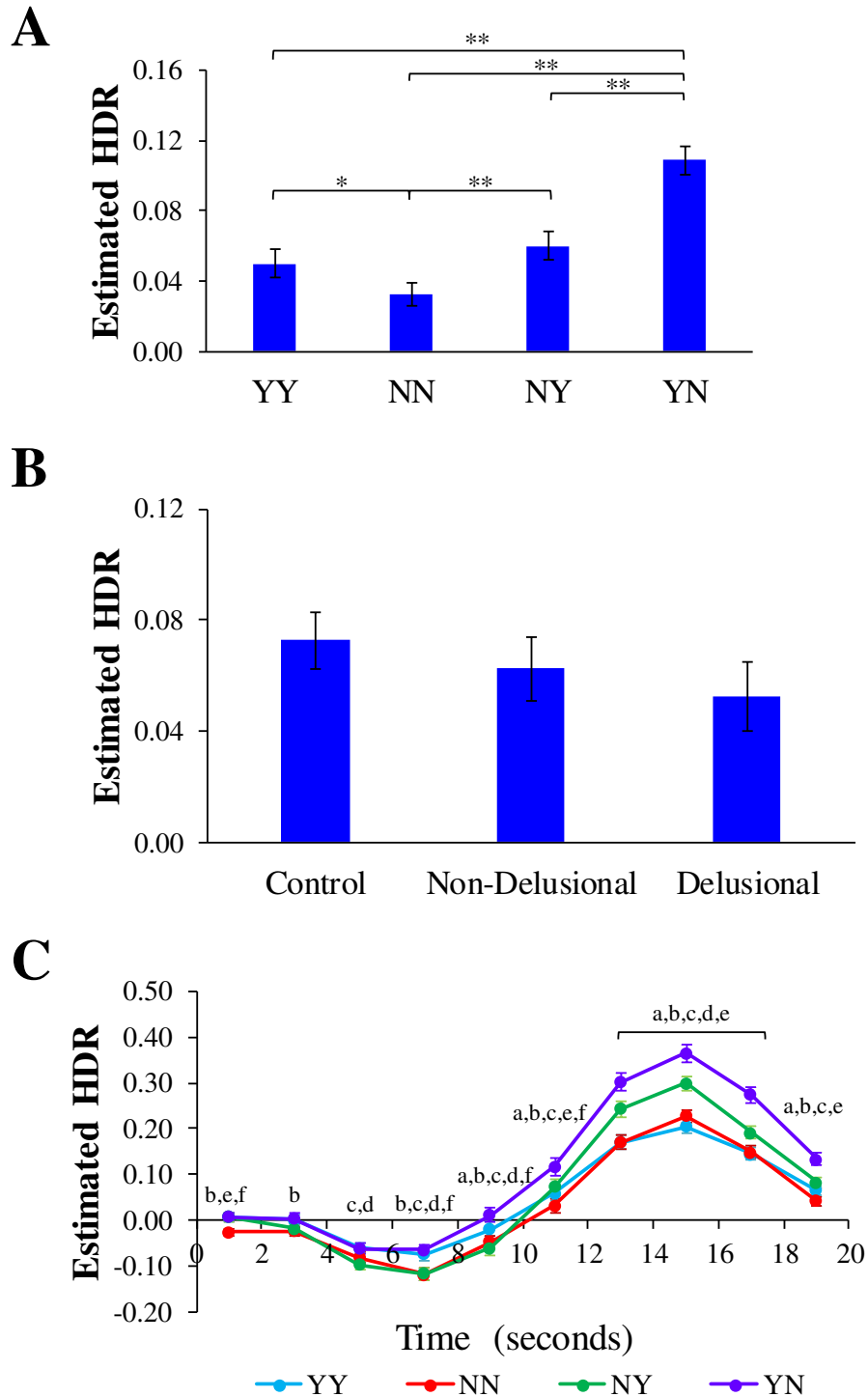
Figure 17. Study 3: Two-dimensional model of the bias against disconfirmatory evidence task used for generalized structured component analysis. Abs = absurd; e = error; ELure = emotional lure; NLure = neutral lure.

245

Figure 18. Study 3: Dominant 10% of component loadings for visual attention network (Component 1: red/yellow = positive loadings, threshold = 0.23, maximum = 0.46; no negative loadings passed threshold). Montreal Neurological Institute Z-axis coordinates are displayed; left = left.
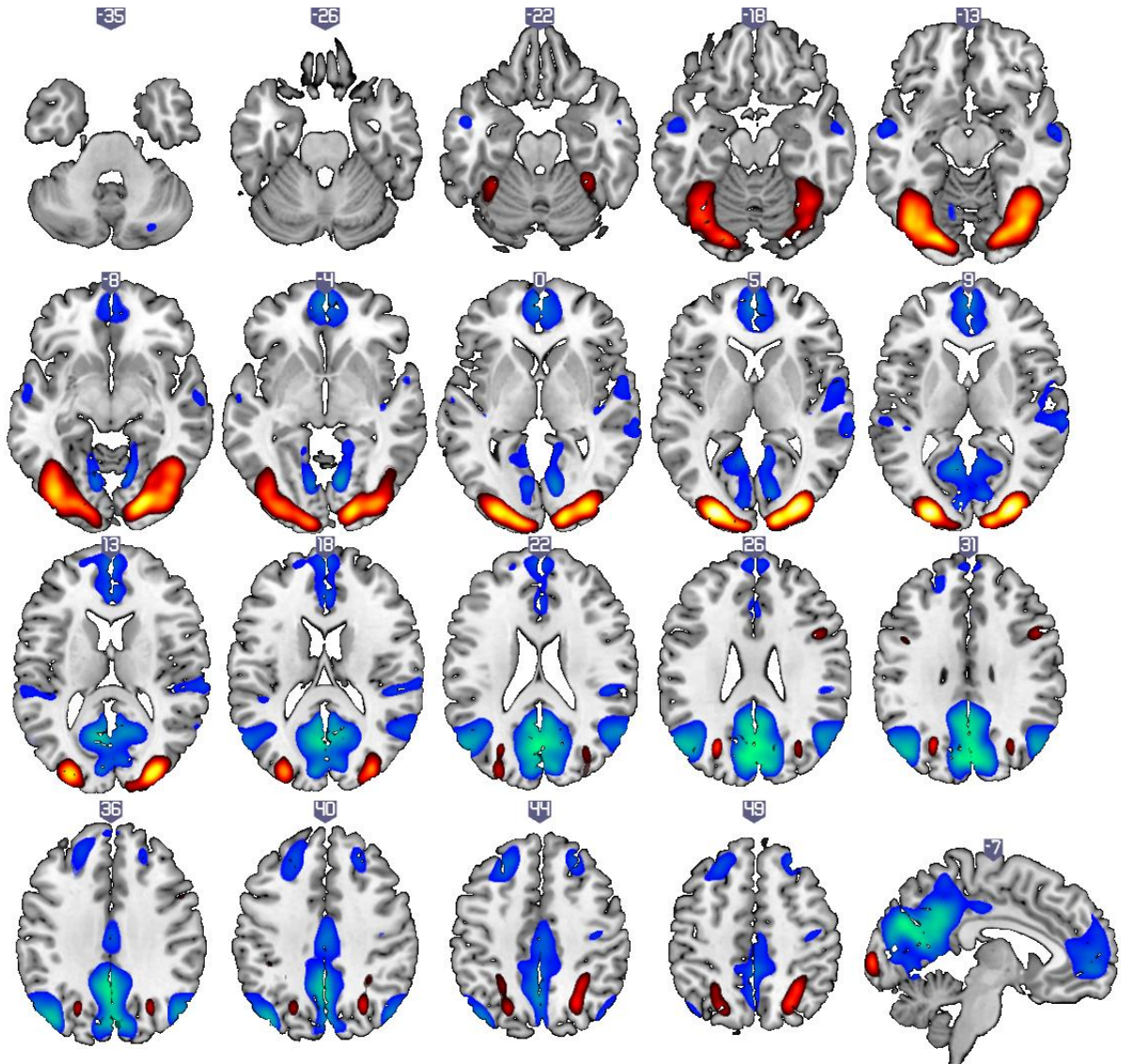
Figure 19. Study 3: Estimated hemodynamic response (HDR) for visual attention network (Component 1). A: Main effect of Condition. B: Interaction between Time Point and Time (non-significant). C: Interaction between Condition and Time (non-significant). $*$ = $p < .05$; $**$ = $p < .005$. Error bars are standard errors.
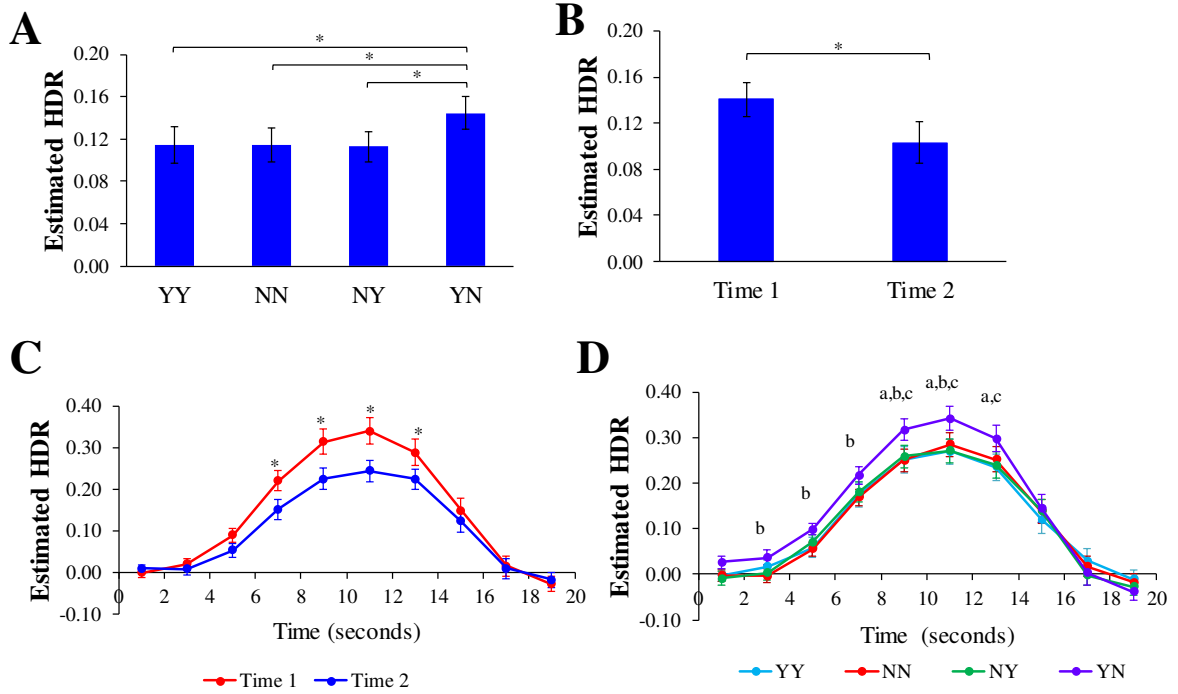

247

Figure 20. Study 3: Dominant 10% of component loadings for visual/default-mode network (Component 2: red/yellow = positive loadings, threshold = 0.18, maximum = 0.36; blue/green = negative loadings, threshold = -0.18, min = -0.37). Montreal Neurological Institute Z-axis coordinates are displayed; left = left.
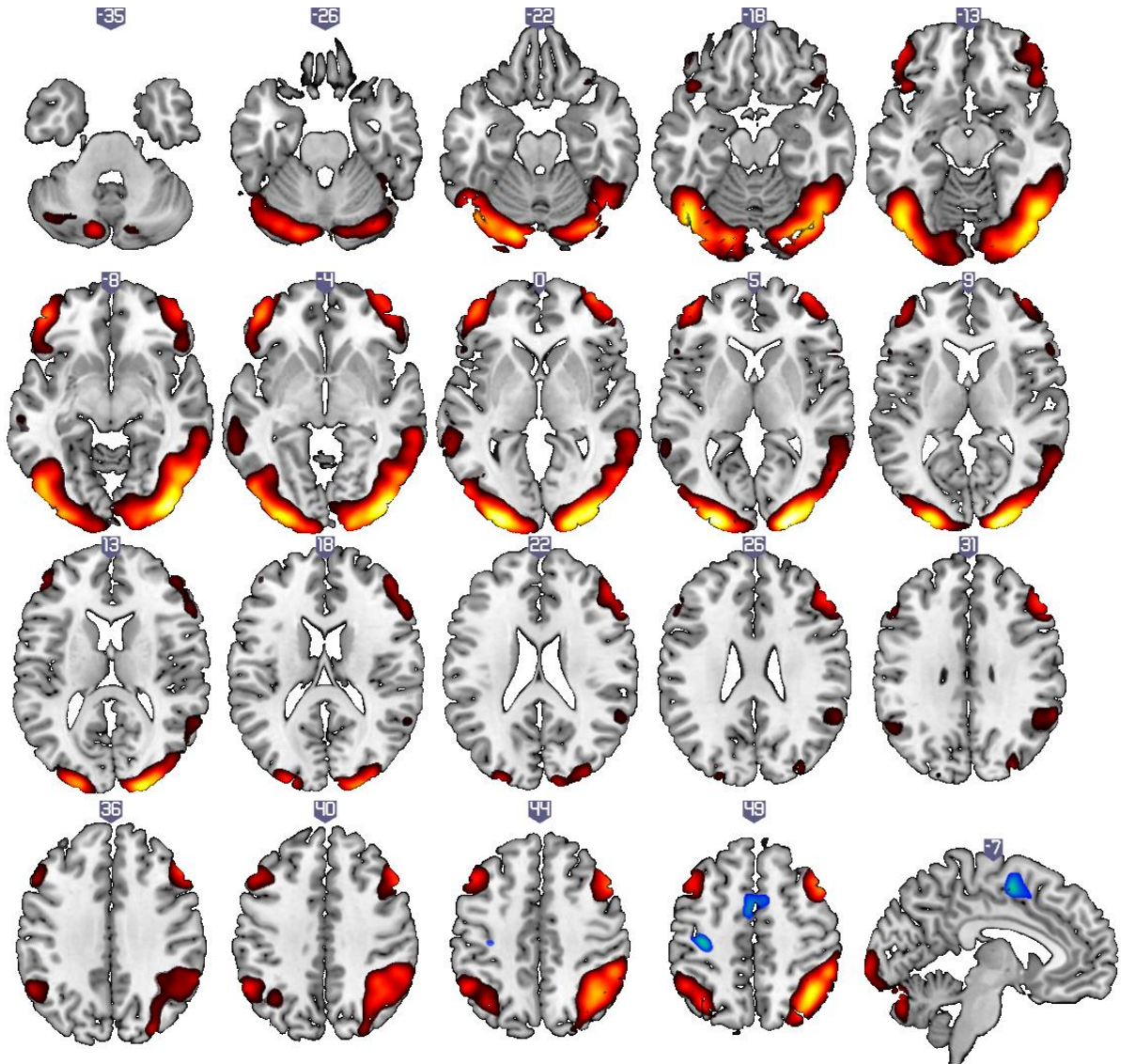
Figure 21. Study 3: Estimated hemodynamic response (HDR) for visual/default-mode network (Component 2). A: Main effect of Condition. B: Main effect of Time Point. C: Interaction between Time Point and Poststimulus Time. D: Interaction between Condition and Poststimulus Time. [a] = YN > YY; [b] = YN > NN; [c] = YN > NY; * = $p < .05$. Error bars are standard errors.

Figure 22. Study 3: Dominant 10% of component loadings for cognitive evaluation network (Component 3: red/yellow = positive loadings, threshold = 0.12, maximum = 0.26; blue/green = negative loadings, threshold = -0.12, minimum = -0.17). Montreal Neurological Institute Z-axis coordinates are displayed; left = left.
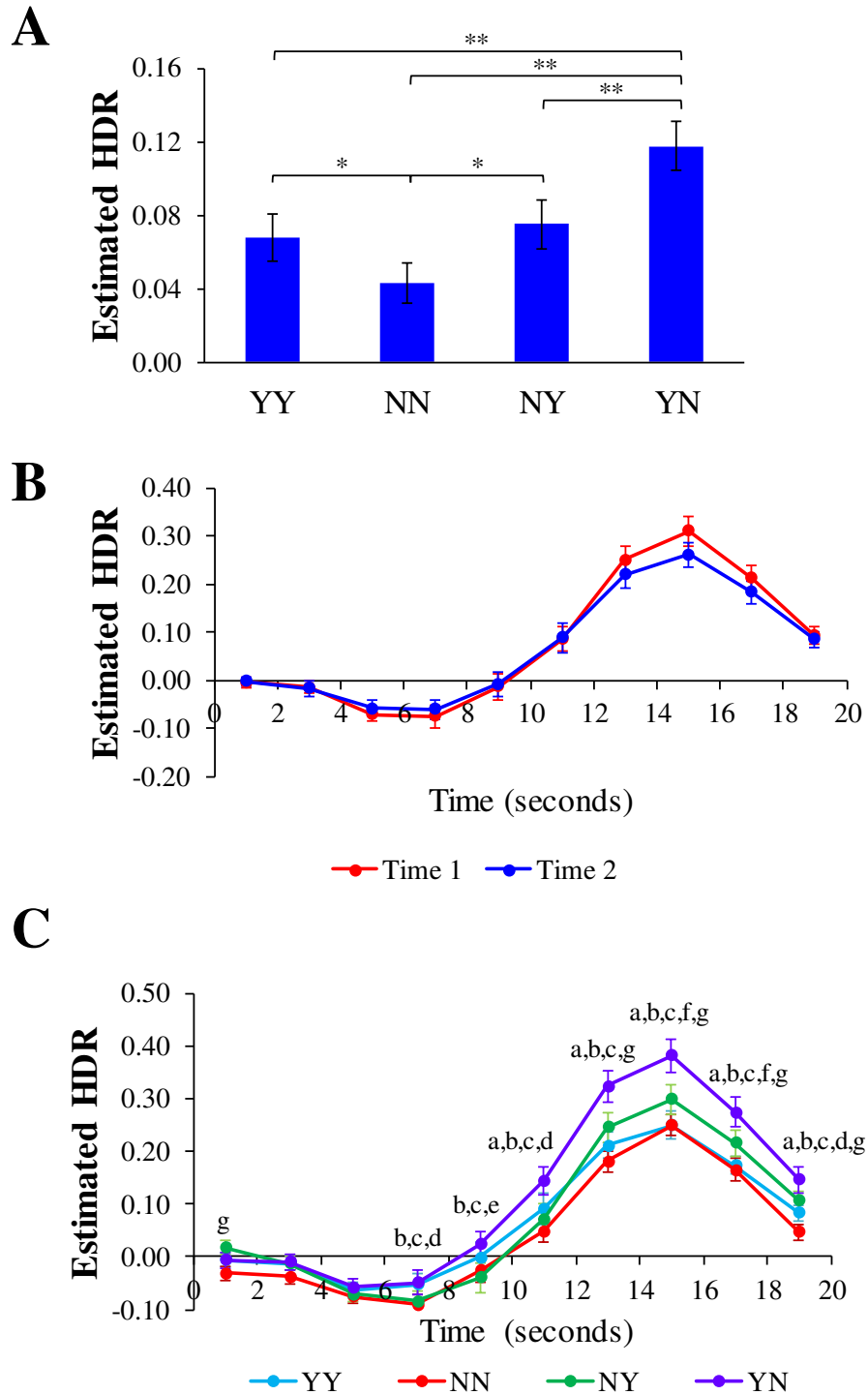
Figure 23. Study 3: Estimated hemodynamic response (HDR) for cognitive evaluation network (Component 3). A: Main effect of Condition. B: Interaction between Time Point and Time (non-significant). C: Interaction between Condition and Time. [a] = YN > YY; [b] = YN > NN; [c] = YN > NY; [d] = YY > NN; [e] = YY > NY; [f] = NY > YY; [g] = NY > NN; * = $p < .05$; ** = $p < .001$. Error bars are standard errors.

Bibliography

Addington, J., Addington, D., & Maticka-Tyndale, E. (1991). Cognitive functioning and positive and negative symptoms in schizophrenia. *Schizophrenia Research, 5*(2), 123-134.

Aleman, A., Hijman, R., de Haan, E. H., & Kahn, R. S. (1999). Memory impairment in schizophrenia: a meta-analysis. *American Journal of Psychiatry, 156*(9), 1358-1366.

American Psychiatric Association. (2000). *DSM-IV-TR: Diagnostic and statistical manual of mental disorders, text revision*. Washington, DC: American Psychiatric Association.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th Ed.). Washington, DC: American Psychiatric Association.

Andreasen, N. C. (1984a). *Scale for the assesment of negative symptoms (SANS)*. Iowa City, IA: University of Iowa.

Andreasen, N. C. (1984b). *Scale for the assessment of positive symptoms (SAPS)*. Iowa City, IA: University of Iowa.

Andreou, C., Schneider, B. C., Braun, V., Kolbeck, K., Gallinat, J., & Moritz, S. (2015). Dopamine effects on evidence gathering and integration. *J Psychiatry Neurosci, 40*(6), 422-428.

Annett, M. (1970). A classification of hand preference by association analysis. *British Journal Of Psychology (London, England: 1953), 61*(3), 303-321.

Baker, J. T., Holmes, A. J., Masters, G. A., & et al. (2014). Disruption of cortical association networks in schizophrenia and psychotic bipolar disorder. *JAMA Psychiatry, 71*(2), 109-118.

Balzan, R., Delfabbro, P., Galletly, C., & Woodward, T. S. (2012). Reasoning heuristics across the psychosis continuum: The contribution of hypersalient evidence-hypothesis matches. *Cognitive Neuropsychiatry, 17*(5), 431-450.

Basso, M. R., Nasrallah, H. A., Olson, S. C., & Bornstein, R. A. (1998). Neuropsychological correlates of negative, disorganized and psychotic symptoms in schizophrenia. *Schizophrenia Research, 31*(2), 99-111.

Beck, A. T., Steer, R. A., & Brown, G. K. (1961). Beck Depression Inventory--II. In P. A. Arbisi & R. F. Farmer (Eds.), *BDI-II*.

Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience, 10*(9), 1214-1221.

Bell, V., Halligan, P. W., & Ellis, H. D. (2008). Are anomalous perceptual experiences necessary for delusions? *Journal of Nervous and Mental Disease, 196*(1), 3-8.

Bentall, R. P., Corcoran, R., Howard, R., Blackwood, N., & Kinderman, P. (2001). Persecutory delusions: a review and theoretical integration. *Clinical Psychology Review, 21*(8), 1143-1192.

Bentall, R. P., Kinderman, P., & Kaney, S. (1994). The self, attributional processes and abnormal beliefs: towards a model of persecutory delusions. *Behaviour Research and Therapy, 32*(3), 331-341.

Benton, A. L. (1967). Problems of Test Construction in the Field of Aphasia. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior, 3*(1), 32-58.

Berman, I., Viegner, B., Merson, A., Allan, E., Pappas, D., & Green, A. I. (1997). Differential relationships between positive and negative symptoms and neuropsychological deficits in schizophrenia. *Schizophrenia Research, 25*(1), 1-10.

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where Is the Semantic System? A Critical Review and Meta-Analysis of 120 Functional Neuroimaging Studies. *Cerebral Cortex, 19*(12), 2767-2796.

Blackwood, N. J., Howard, R. J., Bentall, R. P., & Murray, R. M. (2001). Cognitive neuropsychiatric models of persecutory delusions. *American Journal of Psychiatry, 158*(4), 527-539.

Broyd, A., Balzan, R. P., Woodward, T. S., & Allen, P. (2017). Dopamine, cognitive biases and assessment of certainty: A neurocognitive model of delusions. *Clinical Psychology Review, 54*, 96-106.

Buchsbaum, B. R., Olsen, R. K., Koch, P. F., Kohn, P., Kippenhan, J. S., & Berman, K. F. (2005). Reading, hearing, and the planum temporale. *NeuroImage, 24*(2), 444-454.

Buchy, L., Woodward, T. S., & Liotti, M. (2007). A cognitive bias against disconfirmatory evidence (BADE) is associated with schizotypy. *Schizophrenia Research, 90*, 334-337.

Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The Brain's Default Network. *Annals of the New York Academy of Sciences, 1124*, 1-38.

Buckner, R. L., Krienen, F. M., Castellanos, A., Diaz, J. C., & Yeo, B. T. (2011). The organization of the human cerebellum estimated by intrinsic functional connectivity. *Journal of Neurophysiology, 106*(5), 2322-2345.

Buonocore, M., Bosia, M., Riccaboni, R., Bechi, M., Spangaro, M., Piantanida, M., . . . Cavallaro, R. (2015). Combined neurocognitive and metacognitive rehabilitation in schizophrenia: Effects on bias against disconfirmatory evidence. *European Psychiatry, 30*(5), 615-621.

Burgess, P. W., Dumontheil, I., & Gilbert, S. J. (2007). The gateway hypothesis of rostral prefrontal cortex (area 10) function. *Trends in Cognitive Sciences, 11*(7), 290-298.

Burgess, P. W., Simons, J. S., Dumontheil, I., & Gilbert, S. J. (2005). The gateway hypothesis of rostral prefrontal cortex (area 10) function. In J. Duncan, P. McLeod, & L. Phillips (Eds.), *Measuring the Mind: Speed, Control, and Age* (pp. 215-246). Oxford: Oxford University Press.

Carter, C. S., & van Veen, V. (2007). Anterior cingulate cortex and conflict detection: an update of theory and data. *Cognitive, Affective & Behavioral Neuroscience, 7*(4), 367-379.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research, 1*(2), 245-276.

Cattell, R. B., & Vogelmann, S. (1977). A comprehensive trial of the scree and kg criteria for determining the number of factors. *Multivariate Behavioral Research, 12*(3), 289-325.

Chmielewski, M., & Watson, D. (2008). The heterogeneous structure of schizotypal personality disorder: Item-level factors of the schizotypal personality questionnaire and their associations with obsessive-compulsive disorder symptoms, dissociative tendencies, and normal personality. *Journal of Abnormal Psychology, 117*(2), 364-376.

Choi, E. Y., Yeo, B. T., & Buckner, R. L. (2012). The organization of the human striatum estimated by intrinsic functional connectivity. *Journal of Neurophysiology, 108*(8), 2242-2263.

Chong, H. Y., Teoh, S. L., Wu, D. B., Kotirum, S., Chiou, C. F., & Chaiyakunapruk, N. (2016). Global economic burden of schizophrenia: a systematic review. *Neuropsychiatric Disease and Treatment, 12*, 357-373.

Christoff, K., Ream, J. M., Geddes, L. P. T., & Gabrieli, J. D. E. (2003). Evaluating self-generated information: anterior prefrontal contributions to human cognition. *Behavioral Neuroscience, 117*(6), 1161-1168.

Coltheart, M. (2010). The neuropsychology of delusions. *Annals of the New York Academy of Sciences, 1191*, 16-26.

Coltheart, M., Langdon, R., & McKay, R. (2011). Delusional belief. *Annual Review of Psychology, 62*, 271-298.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews: Neuroscience, 3*(3), 201-215.

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 favor of an associative model of learning. *Neuron, 44*(5), 877-888.

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.

Cuesta, M. J., & Peralta, V. (1995). Cognitive disorders in the positive, negative, and disorganization syndromes of schizophrenia. *Psychiatry Research, 58*(3), 227-235.

d'Acremont, M., Schultz, W., & Bossaerts, P. (2013). The human brain encodes event frequencies while forming subjective beliefs. *Journal of Neuroscience, 33*(26), 10887-10897.

de Zubicaray, G. I., Hansen, S., & McMahon, K. L. (2013). Differential processing of thematic and categorical conceptual relations in spoken word production. *Journal of Experimental Psychology: General, 142*(1), 131-142.

Dealberto, M. J. (2013). Are the rates of schizophrenia unusually high in Canada? A comparison of Canadian and international data. *Psychiatry Research, 209*(3), 259-265.

Dosenbach, N. U. F., Fair, D. A., Miezin, F. M., Cohen, A. L., Wenger, K. K., Dosenbach, R. A. T., . . . Petersen, S. E. (2007). Distinct brain networks for adaptive and stable task control in humans. *Proceedings of the National Academy of Sciences of the United States of America, 104*(26), 11073-11078.

Egner, T. (2011). Surprise! A unifying model of dorsal anterior cingulate function? *Nature Neuroscience, 14*(10), 1219-1220.

Eifler, S., Rausch, F., Schirmbeck, F., Veckenstedt, R., Englisch, S., Meyer-Lindenberg, A., . . . Zink, M. (2014). Neurocognitive capabilities modulate the integration of evidence in schizophrenia. *Psychiatry Research, 219*(1), 72-78.

Eisenacher, S., Rausch, F., Mier, D., Fenske, S., Veckenstedt, R., Englisch, S., . . . Zink, M. (2016). Bias against disconfirmatory evidence in the 'at-risk mental state' and during psychosis. *Psychiatry Research, 238*, 242-250.

Eisenacher, S., & Zink, M. (2017). Holding on to false beliefs: The bias against disconfirmatory evidence over the course of psychosis. *Journal of Behavior Therapy and Experimental Psychiatry, 56*, 79-89.

Flashman, L. A., & McAllister, T. W. (2002). Lack of awareness and its impact in traumatic brain injury. *NeuroRehabilitation, 17*(4), 285-296.

Fletcher, P. C., Anderson, J. M., Shanks, D. R., Honey, R., Carpenter, T. A., Donovan, T., . . . 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.

Fox, M. D., Corbetta, M., Snyder, A. Z., Vincent, J. L., & Raichle, M. E. (2006). Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems. *Proceedings of the National Academy of Sciences of the United States of America, 103*(26), 10046-10051.

Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., & Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences of the United States of America, 102*(27), 9673-9678.

Freeman, D., Garety, P. A., Kuipers, E., Fowler, D., & Bebbington, P. E. (2002). A cognitive model of persecutory delusions. *British Journal of Clinical Psychology, 41*(4), 331-347.

Freeman, D., Garety, P. A., Kuipers, E., Fowler, D., Bebbington, P. E., & Dunn, G. (2007). Acting on persecutory delusions: The importance of safety seeking. *Behaviour Research and Therapy, 45*, 89-99.

Frith, C., Leary, J., Cahill, C., & Johnstone, E. (1991). Performance on psychological tests. Demographic and clinical correlates of the results of these tests. *The British journal of psychiatry. Supplement*(13), 26-29, 44-26.

Frith, C. D. (1992, 2014). *The cognitive neuropsychology of schizophrenia*: Psychology Press.

Frith, C. D., Friston, K. J., Liddle, P. F., & Frackowiak, R. S. J. (1991). A PET study of word finding. *Neuropsychologia, 29*(12), 1137-1148.

Fromm-auch, D., & Yeudall, L. T. (1983). Normative data for the halstead-reitan neuropsychological tests. *Journal of Clinical Neuropsychology, 5*(3), 221-238.

Garety, P. A. (1992). Making sense of delusions. *Psychiatry, 55*(3), 282-291.

Garety, P. A., & Freeman, D. (1999). Cognitive approaches to delusions: A critical review of theories and evidence. *British Journal of Clinical Psychology, 38*, 113-154.

Garety, P. A., & Freeman, D. (2013). The past and future of delusions research: from the inexplicable to the treatable. *The British Journal of Psychiatry, 203*(5), 327-333.

Garety, P. A., Kuipers, E., Fowler, D., Freeman, D., & Bebbington, P. E. (2001). A cognitive model of the positive symptoms of psychosis. *Psychological Medicine, 31*(2), 189-195.

Garrity, A. G., Pearlson, G. D., McKiernan, K., Lloyd, D., Kiehl, K. A., & Calhoun, V. D. (2007). Aberrant "default mode" functional connectivity in schizophrenia. *American Journal of Psychiatry, 164*(3), 450-457.

Gilbert, S. J., Spengler, S., Simons, J. S., Frith, C. D., & Burgess, P. W. (2006a). Differential functions of lateral and medial rostral prefrontal cortex (area 10) revealed by brain-behavior associations. *Cerebral Cortex, 16*(12), 1783-1789.

Gilbert, S. J., Spengler, S., Simons, J. S., Steele, J. D., Lawrie, S. M., Frith, C. D., & Burgess, P. W. (2006b). Functional Specialization within Rostral Prefrontal Cortex (Area 10): A Meta-analysis. *Journal of Cognitive Neuroscience, 18*(6), 932-948.

Gilleen, J., & David, A. S. (2005). The cognitive neuropsychiatry of delusions: from psychopathology to neuropsychology and back again. *Psychological Medicine, 35*(1), 5-12.

Goeree, R., Farahati, F., Burke, N., Blackhouse, G., O'Reilly, D., Pyne, J., & Tarride, J. E. (2005). The economic burden of schizophrenia in Canada in 2004. *Current Medical Research and Opinion, 21*(12), 2017-2028.

Goulden, N., Khusnulina, A., Davis, N. J., Bracewell, R. M., Bokde, A. L., McNulty, J. P., & Mullins, P. G. (2014). The salience network is responsible for switching between the default mode network and the central executive network: replication from DCM. *NeuroImage, 99*, 180-190.

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

Heinrichs, R. W. (2005). The primacy of cognition in schizophrenia. *American Psychologist, 60*(3), 229-242.

Heinrichs, R. W., & Zakzanis, K. K. (1998). Neurocognitive deficit in schizophrenia: A quantitative review of the evidence. *Neuropsychology, 12*(3), 426-445.

Howell, G. T., & Lacroix, G. L. (2012). Decomposing interactions using GLM in combination with the COMPARE, LMATRIX and MMATRIX subcommands in SPSS. *Tutorials in Quantitative Methods for Psychology, 8*(1), 1-22.

Hunter, M. A., & Takane, Y. (2002). Constrained principal component analysis: Various applications. *Journal of Educational and Behavioral Statistics, 27*(2), 105-145.

Hwang, H., & Takane, Y. (2004). Generalized structured component analysis. *Psychometrika, 69*(1), 81-99.

Jin, H., & Mosweu, I. (2017). The Societal Cost of Schizophrenia: A Systematic Review. *Pharmacoeconomics, 35*(1), 25-42.

Joseph, R. (1986). Confabulation and delusional denial: Frontal lobe and lateralized influences. *Journal of Clinical Psychology, 42*(3), 507-520.

Kahn, R. S., & Keefe, R. E. (2013). Schizophrenia is a cognitive illness: Time for a change in focus. *JAMA Psychiatry, 70*(10), 1107-1112.

Kaiser, H. F. (1991). Coefficient Alpha for a Principal Component and the Kaiser-Guttman Rule. *Psychological Reports, 68*(3), 855-858.

Kim, S., Cardwell, R., & Hwang, H. (2016). Using R Package gesca for generalized structured component analysis. *Behaviormetrika, 44*(1), 3-23.

Krueger, P. M., van Vugt, M. K., Simen, P., Nystrom, L., Holmes, P., & Cohen, J. D. (2017). Evidence accumulation detected in BOLD signal using slow perceptual decision making. *J Neurosci Methods, 281*, 21-32.

Kumar, D., Ul Haq, M. Z., Dubey, I., Dotivala, K. N., Siddiqui, S. V., Prakash, R., . . . Nizamie, S. H. (2010). Effect of meta-cognitive training in the reduction of positive symptoms in schizophrenia. *European Journal of Psychotherapy and Counselling, 12*(2), 149-158.

Laurienti, P. J., Burdette, J. H., Wallace, M. T., Yen, Y.-F., Field, A. S., & Stein, B. E. (2002). Deactivation of Sensory-Specific Cortex by Cross-Modal Stimuli. *Journal of Cognitive Neuroscience, 14*(3), 420-429.

Lavigne, K. M., Menon, M., & Woodward, T. S. (2016). Impairment in subcortical suppression in schizophrenia: Evidence from the fBIRN Oddball Task. *Human Brain Mapping, 37*(12), 4640-4653.

Lavigne, K. M., Metzak, P. D., & Woodward, T. S. (2015a). Functional brain networks underlying detection and integration of disconfirmatory evidence. *NeuroImage, 112*, 138-151.

Lavigne, K. M., Rapin, L. A., Metzak, P. D., Whitman, J. C., Jung, K., Dohen, M., . . . Woodward, T. S. (2015b). Left-dominant temporal-frontal hypercoupling in schizophrenia patients with hallucinations during speech perception. *Schizophr Bull, 41*(1), 259-267.

Liu, T., & Pleskac, T. J. (2011). Neural correlates of evidence accumulation in a perceptual decision task. *Journal of Neurophysiology, 106*(5), 2383-2398.

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

Maher, B. A. (1988). *Anomalous experience and delusional thinking: The logic of explanations*: John Wiley & Sons.

Manoliu, A., Riedl, V., Zherdin, A., Mühlau, M., Schwerthöffer, D., Scherr, M., . . . Sorg, C. (2014). Aberrant dependence of default mode/central executive network interactions on anterior insular salience network activity in schizophrenia. *Schizophrenia Bulletin, 40*(2), 428-437.

Marsh, R., Horga, G., Parashar, N., Wang, Z., Peterson, B. S., & Simpson, H. B. (2014). Altered activation in fronto-striatal circuits during sequential processing of conflict in unmedicated adults with obsessive-compulsive disorder. *Biological Psychiatry, 75*(8), 615-622.

McDermott, K. B., Petersen, S. E., Watson, J. M., & Ojemann, J. G. (2003). A procedure for identifying regions preferentially activated by attention to semantic and phonological relations using functional magnetic resonance imaging. *Neuropsychologia, 41*(3), 293-303.

McGrath, J., Saha, S., Chant, D., & Welham, J. (2008). Schizophrenia: a concise overview of incidence, prevalence, and mortality. *Epidemiologic Reviews, 30*, 67-76.

McLean, B. F., Mattiske, J. K., & Balzan, R. P. (2017). Association of the Jumping to Conclusions and Evidence Integration Biases With Delusions in Psychosis: A Detailed Meta-analysis. *Schizophrenia Bulletin, 43*(2), 344-354.

Menon, V., & Uddin, L. Q. (2010). Saliency, switching, attention and control: a network model of insula function. *Brain Struct Funct, 214*(5-6), 655-667.

Metzak, P. D., Feredoes, E., Takane, Y., Wang, L., Weinstein, S., Cairo, T., . . . Woodward, T. S. (2011). Constrained principal component analysis reveals functionally connected load-dependent networks involved in multiple stages of working memory. *Human Brain Mapping, 32*(6), 856-871.

Metzak, P. D., Lavigne, K. M., & Woodward, T. S. (2015). Functional brain networks involved in reality monitoring. *Neuropsychologia, 75*, 50-60.

Metzak, P. D., Meier, B., Graf, P., & Woodward, T. S. (2013). More than a surprise: The bivalency effect in task switching. *Journal of Cognitive Psychology, 25*(7), 833-842.

Metzak, P. D., Riley, J. D., Wang, L., Whitman, J. C., Ngan, E. T. C., & Woodward, T. S. (2012). Decreased efficiency of task-positive and task-negative networks during working memory in schizophrenia. *Schizophrenia Bulletin, 38*(4), 803-813.

Millier, A., Schmidt, U., Angermeyer, M. C., Chauhan, D., Murthy, V., Toumi, M., & Cadi-Soussi, N. (2014). Humanistic burden in schizophrenia: a literature review. *Journal of Psychiatric Research, 54*, 85-93.

Moritz, S., Heeren, D., Andresen, B., & Krausz, M. (2001). An analysis of the specificity and the syndromal correlates of verbal memory impairments in schizophrenia. *Psychiatry Research, 101*(1), 23-31.

Moritz, S., Veckenstedt, R., Hottenrott, B., Woodward, T. S., Randjbar, S., & Lincoln, T. M. (2010). Different sides of the same coin? Intercorrelations of cognitive biases in schizophrenia. *Cognitive Neuropsychiatry, 15*(4), 406-421.

Moritz, S., Veckenstedt, R., Randjbar, S., Vitzthum, F., & Woodward, T. S. (2011). Antipsychotic treatment beyond antipsychotics: metacognitive intervention for schizophrenia patients improves delusional symptoms. *Psychological Medicine, 41*(9), 1823-1832.

Moritz, S., & Woodward, T. S. (2006). A generalized bias against disconfirmatory evidence in schizophrenia. *Psychiatry Research, 142*, 157-165.

O'Leary, D. S., Flaum, M., Kesler, M. L., Flashman, L. A., Arndt, S., & Andreasen, N. C. (2000). Cognitive correlates of the negative, disorganized, and psychotic symptom dimensions of schizophrenia. *The Journal of neuropsychiatry and clinical neurosciences, 12*(1), 4-15.

Palaniyappan, L., & Liddle, P. F. (2012). Does the salience network play a cardinal role in psychosis? An emerging hypothesis of insular dysfunction. *Journal of Psychiatry and Neuroscience, 37*(1), 17-27.

Palaniyappan, L., Mallikarjun, P., Joseph, V., White, T. P., & Liddle, P. F. (2011). Reality distortion is related to the structure of the salience network in schizophrenia. *Psychological Medicine, 41*(8), 1701-1708.

Ploran, E. J., Nelson, S. M., Velanova, K., Donaldson, D. I., Petersen, S. E., & Wheeler, M. E. (2007). Evidence accumulation and the moment of recognition: dissociating perceptual recognition processes using fMRI. *Journal of Neuroscience, 27*(44), 11912-11924.

Raichle, M. E., & MacLeod, A. M. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America, 98*(2), 676-682.

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., Reynolds, C., Lencz, T., Scerbo, A., Triphon, N., & Kim, D. (1994). Cognitive-perceptual, interpersonal, and disorganized features of schizotypal personality. *Schizophrenia Bulletin, 20*(1), 191-201.

Randolph, C. C., Tierney, M. M. C., Mohr, E. E., & Chase, T. T. N. (1998). The Repeatable Battery for the Assessment of Neuropsychological Status (RBANS): Preliminary Clinical Validity. *Journal of Clinical and Experimental Neuropsychology, 20*(3), 310-319.

Reitan, R. M., & Wolfson, D. (1985). *The Halstead-Reitan neuropsychological test battery: Theory and clinical interpretation* (Vol. 4): Reitan Neuropsychology.

Reynolds, C. R. (2002). *Comprehensive Trail-Making Test (CTMT)*. Austin, TX: PRO-ED, Inc.

Ribary, U., Mackay, A. L., Rauscher, A., Tipper, C. M., Giaschi, D., Woodward, T. S., . . . Moiseey, A. (2017). Emerging neuroimaging technologies: Towards future personalized diagnostics, prognosis, targeted intervention and ethical challenges. In J. Illes & S. Hossain (Eds.), *Neuroethics: Defining the Issues in Theory, Practice and Policy - Looking to the Future*: Oxford University Press.

Riccaboni, R., Fresi, F., Bosia, M., Buonocore, M., Leiba, N., Smeraldi, E., & Cavallaro, R. (2012). Patterns of evidence integration in schizophrenia and delusion. *Psychiatry Research, 200*(2), 108-114.

Ross, K., Freeman, D., Dunn, G., & Garety, P. (2011). A randomized experimental investigation of reasoning training for people with delusions. *Schizophrenia Bulletin, 37*(2), 324-333.

Rossler, W., Salize, H. J., van Os, J., & Riecher-Rossler, A. (2005). Size of burden of schizophrenia and psychotic disorders. *European Neuropsychopharmacology, 15*(4), 399-409.

Saha, S., Chant, D., Welham, J., & McGrath, J. (2005). A systematic review of the prevalence of schizophrenia. *PLoS Medicine, 2*(5 e141), 413-433.

Sanford, N., Lecomte, T., Leclerc, C., Wykes, T., & Woodward, T. S. (2013). Change in jumping to conclusions linked to change in delusions in early psychosis. *Schizophrenia Research, 147*(1), 207-208.

Sanford, N., Veckenstedt, R., Moritz, S., Balzan, R., & Woodward, T. S. (2014). Impaired integration of disambiguating evidence in delusional schizophrenia patients. *Psychological Medicine, 44*, 2729-2738.

Seeley, W. W., Menon, V., Schatzberg, A. F., Keller, J., Glover, G. H., Kenna, H., . . . Greicius, M. D. (2007). Dissociable intrinsic connectivity networks for salience processing and executive control. *Journal of Neuroscience, 27*(9), 2349-2356.

Serences, J. T. (2004). A comparison of methods for characterizing the event-related BOLD timeseries in rapid fMRI. *NeuroImage, 21*(4), 1690-1700.

Sharot, T., Kanai, R., Marston, D., Korn, C. W., Rees, G., & Dolan, R. J. (2012). Selectively altering belief formation in the human brain. *Proceedings of the National Academy of Sciences of the United States of America, 109*(42), 17058-17062.

Sharot, T., Korn, C. W., & Dolan, R. J. (2011). How unrealistic optimism is maintained in the face of reality. *Nature Neuroscience, 14*(11), 1475-1479.

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Weiller, E., Hergueta, T., . . . Dunbar, G. C. (1998). The Mini-International Neurpsychiatric Interview (M.I.N.I):

The develoment and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *The Journal of Clinical Psychiatry, 59*(Suppl. 20), 22-33.

Shulman, G. L., Corbetta, M., Buckner, R. L., Raichle, M. E., Fiez, J. A., Miezin, F. M., & Petersen, S. E. (1997). Top-down modulation of early sensory cortex. *Cerebral Cortex, 7*(3), 193-206.

Simeone, J. C., Ward, A. J., Rotella, P., Collins, J., & Windisch, R. (2015). An evaluation of variation in published estimates of schizophrenia prevalence from 1990 horizontal line 2013: a systematic literature review. *BMC Psychiatry, 15*, 193-206.

Simpson, J., & Done, D. J. (2002). Elasticity and confabulation in schizophrenic delusions. *Psychological Medicine, 32*(3), 451-458.

So, S. H., Freeman, D., Dunn, G., Kapur, S., Kuipers, E., Bebbington, P., . . . Garety, P. A. (2012). Jumping to conclusions, a lack of belief flexibility and delusional conviction in psychosis: a longitudinal investigation of the structure, frequency, and relatedness of reasoning biases. *Journal of Abnormal Psychology, 121*(1), 129-139.

Speechley, W. J., Moritz, S., Ngan, E. T. C., & Woodward, T. S. (2012). Impaired evidence integration and delusions in schizophrenia. *Journal of Experimental Psychopathology, 3*(4), 688-701.

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

Stern, E. R., Gonzalez, R., Welsh, R. C., & Taylor, S. F. (2010). Updating beliefs for a decision: neural correlates of uncertainty and underconfidence. *The Journal Of Neuroscience: The Official Journal Of The Society For Neuroscience, 30*(23), 8032-8041.

Takane, Y., & Hunter, M. A. (2001). Constrained principal component analysis: A comprehensive theory. *Applicable Algebra in Engineering, Communication and Computing, 12*, 391-419.

Takane, Y., & Shibayama, T. (1991). Principal component analysis with external information on both subjects and variables. *Psychometrika, 56*(1), 97-120.

Turner, D. C., Aitken, M. R., Shanks, D. R., Sahakian, B. J., Robbins, T. W., Schwarzbauer, C., & Fletcher, P. C. (2004). The role of the lateral frontal cortex in causal associative learning: exploring preventative and super-learning. *Cerebral Cortex, 14*(8), 872-880.

Turner, M. E., & Pratkanis, A. R. (1998). Twenty-Five Years of Groupthink Theory and Research: Lessons from the Evaluation of a Theory. *Organizational Behavior and Human Decision Processes, 73*(2/3), 105-115.

Uddin, L. Q. (2015). Salience processing and insular cortical function and dysfunction. *Nature Reviews: Neuroscience, 16*(1), 55-61.

Vincent, J. L., Kahn, I., Snyder, A. Z., Raichle, M. E., & Buckner, R. L. (2008). Evidence for a frontoparietal control system revealed by intrinsic functional connectivity. *Journal of Neurophysiology, 100*(6), 3328-3342.

Volz, K. G., Schubotz, R. I., & von Cramon, D. Y. (2004). Why am I unsure? Internal and external attributions of uncertainty dissociated by fMRI. *NeuroImage, 21*(3), 848-857.

Walsh, B. J., Buonocore, M. H., Carter, C. S., & Mangun, G. R. (2011). Integrating Conflict Detection and Attentional Control Mechanisms. *Journal of Cognitive Neuroscience, 23*(9), 2191-2201.

Wechsler, D. (2009). *Test of Premorbid Functioning*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (2011). *Wechsler Abbreviated Scale of Intelligence–Second Edition*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (2014). Wechsler Adult Intelligence Scale–Fourth Edition (WAIS–IV).

White, T. P., Joseph, V., Francis, S. T., & Liddle, P. F. (2010). Aberrant salience network (bilateral insula and anterior cingulate cortex) connectivity during information processing in schizophrenia. *Schizophrenia Research, 123*(2/3), 105-115.

Whitfield-Gabrieli, S., & Ford, J. M. (2012). Default mode network activity and connectivity in psychopathology. *Annual Review of Clinical Psychology, 8*, 49-76.

Whitman, J. C., Menon, M., Kuo, S., & Woodward, T. S. (2012). Bias in favour of self-selected hypotheses is associated with delusion severity in schizophrenia. *Cognitive Neurpsychiatry, 18*(5), 376-389.

Whitman, J. C., Metzak, P. D., Lavigne, K. M., & Woodward, T. S. (2013). Functional connectivity in a frontoparietal network involving the dorsal anterior cingulate cortex underlies decisions to accept a hypothesis. *Neuropsychologia, 51*(6), 1132-1141.

Woodward, T. S., Buchy, L., Moritz, S., & Liotti, M. (2007). A bias against disconfirmatory evidence is associated with delusion proneness in a nonclinical sample. *Schizophrenia Bulletin, 33*(4), 1023-1028.

Woodward, T. S., Feredoes, E., Metzak, P. D., Takane, Y., & Manoach, D. S. (2013). Epoch-specific functional networks involved in working memory. *NeuroImage, 65*, 529-539.

Woodward, T. S., Jung, K., Hwang, H., Yin, J., Taylor, L., Menon, M., . . . Erickson, D. (2014). Symptom Dimensions of the Psychotic Symptom Rating Scales in Psychosis: A Multisite Study. *Schizophrenia Bulletin, 40*(Suppl_4), S265-S274.

Woodward, T. S., Metzak, P. D., Meier, B., & Holroyd, C. B. (2008). Anterior cingulate cortex signals the requirement to break inertia when switching tasks: a study of the bivalency effect. *NeuroImage, 40*(3), 1311-1318.

Woodward, T. S., Moritz, S., & Chen, E. Y. H. (2006a). The contribution of a cognitive bias against disconfirmatory evidence (BADE) to delusions: A study in an Asian sample with first episode schizophrenia spectrum disorders. *Schizophrenia Research, 83*, 297-298.

Woodward, T. S., Moritz, S., Cuttler, C., & Whitman, J. C. (2006b). The contribution of a cognitive bias against disconfirmatory evidence (BADE) to delusions in schizophrenia. *Journal of Clinical and Experimental Neuropsychology, 28*, 605-617.

Woodward, T. S., Munz, M., Leclerc, C., & Lecomte, T. (2009). Change in delusions is associated with change in "jumping to conclusions". *Psychiatry Research, 170*, 124-127.

Woodward, T. S., Tipper, C. M., Leung, A. L., Lavigne, K. M., Sanford, N., & Metzak, P. D. (2015). Reduced functional connectivity during controlled semantic integration in schizophrenia: A multivariate approach. *Human Brain Mapping, 36*(8), 2948-2964.

Wuthrich, V., & Bates, T. C. (2005). Reliability and validity of two Likert versions of the Schizotypal Personality Questionnaire (SPQ). *Personality and Individual Differences, 38*(7), 1543-1548.

Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., . . . Buckner, R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology, 106*(3), 1125-1165.

Yoshida, W., & Ishii, S. (2006). Resolution of uncertainty in prefrontal cortex. *Neuron, 50*(5), 781-789.

Zhou, L., Pu, W., Wang, J., Liu, H., Wu, G., Liu, C., . . . Liu, Z. (2016). Inefficient DMN Suppression in Schizophrenia Patients with Impaired Cognitive Function but not Patients with Preserved Cognitive Function. *Scientific Reports, 6*, 21657-21657.