A Systems Neuroscience Perspective of Schizophrenia and Bipolar Disorder

Sophia Frangou*

1Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY

*To whom correspondence should be addressed; Department of Psychiatry, Icahn School of Medicine at Mount Sinai, 1425 Madison Avenue, New York, NY 10029, US; tel: 212-659-1668, fax: 212-659-8576, e-mail: sophia.frangou@mssm.edu

Neuroimaging studies have generated a large body of knowledge regarding the neural correlates of schizophrenia (SZ) and bipolar disorder (BD). However, the initial goal of identifying disease-specific topographical mappings to localized brain regions or to distinct neural networks has not materialized and may be untenable. This contribution will argue that a systems neuroscience approach may prove more fruitful. The supporting evidence presented covers (a) brain structural, functional, and connectivity alterations and their implication for the clinical and cognitive manifestations of SZ and BD, (b) the prevailing system neuroscience models of the 2 disorders, and (c) key hypotheses likely to produce new insights into the mechanisms of underlying psychotic disorders.

Key words: neuroimaging/schizophrenia/bipolar disorder

Introduction

The relationship between the main psychotic disorders, schizophrenia (SZ) and bipolar disorder (BD), has been the focus of much research and debate. Both disorders are genetically related1–3 and have overlapping clinical phenomenology.4–6 They continue to rank among the leading causes of disability worldwide largely because current clinical syndromal definitions are inadequately aligned with underlying pathophysiology and hence of limited therapeutic and prognostic value.7 Accordingly, substantial research efforts are being directed toward developing biologically informed constructs of psychosis.8,9 The difficulty reaching this objective stems from the relative paucity of data to link clinical phenomenology with underlying mechanisms.

Neuroimaging has arguably had a transformative influence on the field as it has firmly established SZ and BD as brain disorders involving multiple, spatially distributed structural and functional brain abnormalities. The neuroimaging literature that either describes or contrasts the 2 disorders is expansive and has been subjected to equally extensive quantitative and narrative reviews (notable recent examples, references10–18). However, the goal of identifying disease-specific mappings to localized brain regions or to distinct neural networks has not materialized and may prove untenable. What could then be the way forward?

This contribution will argue that a systems neuroscience approach may prove particularly fruitful. First, key structural, functional, and connectivity alterations in SZ and BD are presented based on a synthesis of the evidence from the relevant neuroimaging literature. Then, proposed system neuroscience models of SZ and BD are described and discussed. Finally, key hypotheses likely to produce new insights into the mechanisms of underlying psychotic disorders are highlighted for future research.

Clinical Phenomenology as an Emergent Property of Neural Network Disruption in SZ and BD

Studies of psychopathology in SZ and BD have identified separable, transdiagnostic symptom dimensions which include hallucinations and delusions (referred to jointly as reality distortion), disorganization, amotivation/negativity symptoms (otherwise known as psychomotor poverty), depression, and mania.5,6 This factor structure is present at the first psychotic episode and remains stable at least over the ensuing 5–10 years.6

Neuroimaging has provided firm evidence linking symptom dimensions to dysfunction in multiple aspects of brain function. The mapping of 3 symptom factors, reality distortion, disorganization, and psychomotor poverty, to different patterns of cerebral blood flow in patients with SZ is a classic early example.19 Recent studies have relied mostly on magnetic resonance imaging (MRI) techniques (eg, Strakowski et al15, Bora et al15, Koutsouleris et al20, Nenadic et al21, Goghari22). Disorganization in SZ has been associated with bilateral gray matter alterations in temporal, insular, cerebellar,
and medial prefrontal cortices (MPFC)\textsuperscript{20,21} and functional deficits in the dorsolateral prefrontal cortex (DLPFC).\textsuperscript{22} In BD, disorganization has been hypothesized in the ventrolateral prefrontal cortex (VLPFC)\textsuperscript{10,12,13} and MPFC regions centered in the anterior cingulate cortex (ACC).\textsuperscript{12} Reality distortion in SZ has been associated with gray matter loss in perisylvian\textsuperscript{20,21} and thalamic regions\textsuperscript{31} and with functional abnormalities in the MPFC, amygdala (AMY), and hippocampus/parahippocampal region.\textsuperscript{22} Reality distortion in patients with BD is also associated with functional disruption in prefrontal and thalamic regions.\textsuperscript{23} Psychomotor poverty has been associated with the most extensively distributed gray matter changes in SZ encompassing the MPFC\textsuperscript{15,20} the DLPFC and VLPFC and lateral temporal cortices,\textsuperscript{20} and subcortical structures (striatum, AMY, and thalamus).\textsuperscript{20,21} Functional deficits associated with psychomotor poverty in BD have been most consistently reported in the VLPFC and ventral striatum.\textsuperscript{22} Collectively this evidence confirms that clinically recognizable symptoms in psychotic disorders emerge from the loss of structural and functional integrity within large-scale neural networks. At the same time they argue against the notion of disease-specific networks. Instead they highlight the significance of key brain regions, known to support a variety of processes as expanded below, whose internal function and spatial and temporal interaction generate brain states and corresponding behaviors that we categorize as clinical symptom dimensions.

**Cognitive Dysfunction as an Emergent Property of Neural Network Disruption in SZ and BD**

Cognitive dysfunction is another phenotypic dimension of SZ and BD documented in a wide array of experimental tasks in multitudinous studies.\textsuperscript{24–30} The upshot is that patients with SZ present with severe and generalized cognitive deficits.\textsuperscript{28–30} The largest and most recent meta-analysis to date, which included 9,048 patients with SZ and 8,814 healthy individuals, reported global deficits of an average weighted effect size of about 1 across tasks and across domains, that appear remarkably unaffected by geographic or temporal variation.\textsuperscript{29} In addition, SZ is associated with significant premorbid cognitive impairment and further decline prior to disease onset.\textsuperscript{31,32} Premorbid cognitive impairment is not a key feature of BD; in contrast several studies suggest better than average premorbid cognitive ability.\textsuperscript{27,31} Postonset, however, patients with BD show cognitive deficits that are qualitatively similar but quantitatively milder than those seen in patients with SZ\textsuperscript{24–27}; cognitive dysfunction is further exacerbated in BD patients with a history of psychosis.\textsuperscript{33–37}

The neural circuitry underlying cognitive dysfunction in SZ and BD has been subjected to intense scrutiny using a wide array of cognitive tasks in functional MRI (t-fMRI) studies.\textsuperscript{16,13,16,38–41} Different tasks are often discussed in terms of their relevance to specific cognitive domains although there is no general agreement for either the task groupings or the conceptual boundaries of the domains themselves.\textsuperscript{28,29} Here I focus on executive function and affect processing which are considered core domains both in SZ and BD. Neuroimaging studies of executive function (encompassing initiation, inhibition, switching, working memory, performance monitoring, planning, and sustained attention) reliably implicate the DLPFC, VLPFC, and the dorsal ACC in both disorders.\textsuperscript{13,38–40} During affect processing, most commonly examined in response to facial expressions, patients with either SZ or BD, show abnormalities in the visual association cortices, the AMY and parahippocampal gyrus, the VLPFC and the MPFC.\textsuperscript{10,13,41–43} The general pattern emerging from this line of research indicates that neural abnormalities associated with cognitive dysfunction in SZ and BD show significant topographical overlap between domains and across the 2 disorders. Moreover, the evidence reinforces the involvement of the same distributed set of brain regions implicated by studies of symptom dimensions. These observations raise the possibility that psychotic disorders involve a key set of brain regions whose internal function and spatial and temporal interaction generate brain states that we recognize both as clinical and cognitive symptoms.

**Functional Dysconnectivity in SZ and BD**

Since, symptomatic expression and cognitive dysfunction involve spatiotemporal interactions among brain regions, there has been a major shift in neuroimaging research toward defining and quantifying parameters relating to brain functional connectivity.

fMRI can be used to detect localized blood oxygen level–dependent signal change either in response to a particular experimental task or during task-free (resting state) periods. In the first instance, connectivity measures reflect interactions among the regions engaged by the experimental task. In contrast, resting state networks are defined based on correlated spontaneous fluctuations in regional brain activity.\textsuperscript{44}

**Task-Dependent Connectivity**

During executive function tasks, patients with SZ show abnormally increased connectivity between the DLPFC and temporal regions (hippocampus or superior temporal gyrus)\textsuperscript{44–46} and between the right and left DLPFC\textsuperscript{47} with additional recruitment of the VLPFC.\textsuperscript{48} Conversely, connectivity within the PFC and striatal regions is decreased.\textsuperscript{49} A similar pattern of decreased frontostriatal connectivity and increased ventral-dorsal PFC connectivity during executive function tasks has also been reported in patients with BD.\textsuperscript{49} Both in SZ and BD, connectivity abnormalities in task-dependent networks for affect
processing seem to converge on the interactions between the PFC and the AMY (and related limbic regions); frontolimbic connectivity, however, appears abnormally reduced in SZ and abnormally increased in BD.52–55

Resting State Connectivity

The first resting state network identified in healthy adults involved motor function.36 Additional resting state networks have since been described which show a close correspondence to known task-dependent networks involved in visual processing, auditory processing, memory, attention, as well as the default mode network which is task negative.57–61 These networks are highly reproducible across individuals58 and within individuals across time.57

Examination of resting state connectivity in patients with SZ or BD has, to a large extent, reaffirmed abnormalities originally reported within task-dependent networks. Despite significant between-study variability (driven by methodology and sample composition), the evidence collectively confirms that both disorders are mostly consistently associated with dysfunctional connectivity within PFC-linked networks.11,62–68 Current findings implicate a disruption in frontoparietal connectivity primarily in SZ,11,62 reduced fronto-occipital connectivity in SZ and BD,11,62,63 and increased fronto-AMY connectivity primarily in BD.65,67,68

There is also strong empirical support implicating the default mode network in psychotic disorders.11,69 The default mode network is perhaps the best researched brain network.69,70 However, it is also the only network, ie, abnormal in every neuropsychiatric condition it has been studied in, including (but not limited to) traumatic brain injury, autism, Alzheimer’s disease, SZ, major depressive disorder, and BD.69,70 These findings argue in favor of default mode network abnormalities being a rather generic marker of brain dysfunction.

Integrative System Neuroscience Models for SZ and BD

Several system neuroscience models have been proposed to provide an explanatory framework for the spectrum of phenotypic expression (clinical symptoms and cognitive dysfunction) in SZ and BD. These models rely mostly on findings from the literature on SZ. They focus on the clinical dimension of reality distortion and particularly delusional beliefs. Delusions are conceptualized as inaccurate inferences about the external environment including abnormalities in the attribution of salience (ie, meaning) to external stimuli or events.

The simplest model considers psychosis as a disorder of abnormal salience resulting from dysregulated dopamine neurotransmission within frontoparietal and frontostriatal networks.71,72 Increased dopaminergic release in the striatum has been convincingly linked to the emergence73,74 and exacerbation of reality distortion symptoms.75 Inappropriate allocation of salience is thought to underlie the formation of positive psychotic symptoms as well as the processing of reward-related information thus leading to negative symptoms.76 Consistent with this, patients with SZ appear minimally impaired in terms of hedonic experience77,78 but show significant deficits in encoding reward-related associations.79,80 Cognitive dysfunction is also considered in the context of dopamine dysregulation because executive function is known to depend on optimal dopaminergic signaling.81,82 This model has been of great heuristic value in SZ research but its transdiagnostic relevance is less clear. Direct evidence for dopamine dysregulation in BD is currently lacking although aspects of the disorder, such as mania, may involve changes in dopaminergic signaling.83

Cameron and colleagues, anchoring their model on cognitive neuroscience concepts, have proposed that psychotic disorders are disorders of deficient cognitive control.84 Cognitive control refers to processes involved in integrating sensory and motor information with higher order representations of goals or rules in order to select appropriate responses.84,85 The concept of cognitive control is therefore multifaceted and encompasses a number of mental processes often referred to as executive function; these include initiation, inhibition, switching, working memory, performance monitoring, planning, and sustained attention.84,85 The deficient cognitive control model is, at least in part, informed by recent paradigm shifts in our understanding of brain-behavior relationships. It was assumed until recently that the component mental processes of cognitive control were realized within topographically distinct regions or networks. Attempts to map these networks have generated an expansive body of evidence which, however, points to a domain-general, superordinate cognitive control network. As Niendam et al86 have most recently demonstrated, the various executive function tasks activate a shared network of regions in the frontopolar cortex, in the lateral and medial PFC bilaterally, in the dorsal ACC, and in the inferior and superior parietal lobes. The key empirically supported premise of the cognitive control deficit model of psychosis is the complete overlap between the regions that are most consistently implicated in SZ (already highlighted in the preceding sections) and those of the domain-general, superordinate cognitive control network. The model then speculates that in SZ structural and functional abnormalities seen within this network result in the loss of efficient integration of mental processes which can then lead to both cognitive dysfunction and clinical symptoms, particularly disorganization and delusional beliefs. The assertion that cognitive control dysfunction is the core unitary mechanism underlying both cognitive dysfunction and psychotic symptoms is however problematic. This is because cognitive control...
deficits (and abnormalities within the associated neural network) are highly prevalent in psychiatric and neurological conditions, even those not associated with psychosis, with depression, addiction, and Parkinson’s disease being notable examples.

The hierarchical predictive coding model, proposed by Friston and colleagues and the hierarchical temporal processing deficit model, proposed by Krishnan and colleagues, will be considered together as they both emphasize aberrant prediction as the key unitary mechanism underlying psychosis. However, the predictive coding model uses concepts derived from computational neuroscience based on Bayesian principles while the hierarchical temporal processing deficit model employs linguistic conventions from the field of cognitive psychology. The conceptual similarity of these models is therefore obscured by the use of different terminology. For that reason, when describing the models below, different terms that relate to the same underlying concepts are juxtaposed. To begin with, both models emphasize that the normal micro- and macro-anatomical and functional organization of the human brain supports mental processes through hierarchical strategies. Forward connections from sensory to associative cortices depend on local neuronal adaptations to sensory stimuli. Backward connections from higher order cortical areas provide signals that disambiguate activity in sensory areas through the process of comparing the incoming information to preexisting representations. These representations incorporate future expectations (ie, predictions) either in terms of the next sensory input or the likely behavioral output. Comparisons between preexisting representations and incoming sensory information are also hierarchically organized. Highly associative cortical areas are involved in integrating information from multiple sensory domains, in maintaining representations that allow predictions over longer timescales and in modulating the function of lower brain regions through changing “post-synaptic gain” (predictive coding) or through “priming” (hierarchical temporal processing deficit model). New evidence is incorporated into the higher order representations resulting in updating the “posterior beliefs” (predictive coding) about the new evidence, or in the “resolution of the mismatch” (hierarchical temporal processing deficit model) between the new input and preexisting representations. According to the hierarchical temporal processing deficit model, reality distortion arises from impaired “memory-based prediction of perception.” The predictive coding model identifies this impairment as “aberrant encoding of precision,” whereby precision represents the likelihood of the predicted event. The model further postulates that this impairment is likely to arise because of deficient identification of prediction errors by principal or pyramidal neurons within the superficial layers of the associative cortical regions.

Common across all the models discussed is the notion that environmental insults (eg, trauma, nutritional deficits) and genetic variation (eg, DISC1, NRG1) that are known to increase the risk for psychosis, disrupt processes responsible for the orderly neuronal configuration (ie, migration, differentiation, and adhesion) and for the efficient neuronal communication (ie, synaptic integrity, neurotransmission). An important advantage of the predictive coding model is that it allows computational formulations of its tenants and predictions. This is because biologically restrained computational models of the brain may prove invaluable in allowing us to test hypotheses regarding the effect of microscale mechanisms on synaptic and global processes that may generate new insights regarding the pathogenesis of psychosis and could potentially identify and evaluate new interventions.

Also common across all the models is that disturbances in affect processing are either not discussed or implicitly assumed to arise from the same mechanisms postulated to give rise to psychotic symptoms and cognitive dysfunction. However, affect processing is highly relevant as it may prove particularly informative in defining syndromal boundaries between SZ and BD. Affect processing is a multifaceted concept that involves the generation, the behavioral, and experiential inhibition of emotional states and their recognition in oneself and in others. As is the case with cognitive control, it was assumed until recently that these component processes were realized within topographically distinct regions or networks comprising the “affective” and the “social” brain. However, such distinctions are beginning to give way to more unitary considerations of affect as the physiological, behavioral, and experiential outcome of domain-general, superordinate networks. Facial affect processing will be used here to exemplify this new approach as it is the most extensively studied aspect of affect processing in terms of its neural correlates and its relevance to psychotic disorders. The domain-general, superordinate network engaged in facial affect processing also shows evidence of hierarchical organization into perceptual and higher order associative brain regions. Perceptual regions in the visual and inferior temporal cortices (mainly the fusiform gyrus) are primarily concerned with the processing of facial features; higher order regions include the AMY and VLPFC which are primarily concerned with contextual appraisal and regulation. Similarities but, importantly, also differences between SZ and BD have been documented in the facial affect processing network based on meta-analyses of the relevant fMRI literature. Both diagnoses were associated with similar reduction in VLPFC engagement but differed in the degree of engagement of the AMY (which was greater in BD) and the visual association cortices (which was greater in SZ). These differences in activation are also reflected in altered connectivity. In both disorders, there is evidence of a reduction in regulatory signals from the PFC.
connectivity between AMY and VLPFC has been found to be increased in BD\textsuperscript{55,107} while in SZ increased connectivity has been noted within the perceptual regions of the network.\textsuperscript{108} The patterns of abnormalities in the regional engagement and connectivity of the facial affect processing network could be accommodated within each of the system neuroscience models described above. In general terms, in patients with SZ, abnormal inferences about the facial stimuli (be they attributable to abnormal salience, impaired memory-based perceptual prediction, or abnormal encoding of precision) could lead to inefficient engagement and connectivity within this domain-general network. However, it is difficult to apply these concepts to patients with BD because if a unitary mechanism applied to both disorders then this should be evident at the neural system level.

**Conclusions and Future Directions**

We have accumulated a large body of knowledge regarding the clinical and cognitive manifestations of SZ and BD and their neural underpinnings. Substantial evidence points to the involvement of multiple large-scale neural networks and alterations in local microscale circuitry within associative and sensory cortices. Nonetheless, a fully mechanistic description of either psychotic disorder remains elusive.

The neuroimaging data available suggest 4 new hypotheses for further testing. First, the data urge us to consider the possibility that the clinical, cognitive, and affective features of SZ and BD emerge from more general processes which themselves map on to domain-general, superordinate networks. This point is illustrated in figure 1 which highlights the domain-general networks implicated in SZ and BD and the phenomena they have been associated with most frequently. Second, microscale abnormalities in neuronal function and the associated disturbances in connectivity within these domain-general networks are likely to differ between SZ and BD. Current system neuroscience models represent valiant attempts to synthesize the multitude of data into a coherent framework and to provide a unifying principle that could account for the phenotypic expression of psychotic disorders. One could argue that in actual fact these models provide interesting and heuristic accounts of mechanisms likely to produce a neurodevelopmental disorder dominated by pervasive distortion of reality and generalized cognitive dysfunction (including cognitive control deficits) leading to profound social disability. There are groups of patients across a number of diagnostic categories (not just the “psychosis continuum”) that could fit this description. This does not mean that the current models are “dimensional.” It could be argued that they are in fact “syndromal,” as they attempt to provide

![Fig. 1. System neuroscience framework for psychosis. For ease of visualization, perception is represented by the visual network only; subdivisions within cortical regions are not shown. ACC, anterior cingulate cortex; AMY, amygdala; PAR, parietal cortex; PFC, prefrontal cortex; VC, visual cortex.](http://schizophreniabulletin.oxfordjournals.org/)

527
an explanatory framework for multiple dimensions of phenomenology. This leads to the third hypothesis. A system neuroscience approach may be usefully employed to define a more biological informed nosology for psychiatric disorders. Neuroimaging research could play a crucial part in such an endeavor because disturbances within neural networks are the necessary condition for the clinical expression of any mental disorder or dysfunction. This has at least 2 important ramifications. One is the possibility that neuroimaging data can differentiate between patients and healthy individuals or among patients with different psychiatric disorders. The feasibility of this approach has been demonstrated by the successful application of multivariate pattern classification analyses in SZ and BD. In principle, this type of analyses could lead to the identification of more homogeneous groups of patients, in terms of their neural patterns, which would also be expected to share similar disturbances at the level of microscale circuits. The use of unsupervised classifiers represents another avenue of research, ie, likely to prove particular fruitful. Unsupervised classifiers attempt to distinguish groups of individuals defined solely by their neural profiles while being agnostic about their diagnostic status. This is fundamentally different from supervised multivariate decoding that relies on diagnostic categorization during the training phase of the algorithm. The presence of subgroups of patients defined by their neural architecture could lead to new and more biologically meaningful phenotypes for psychosis. The final hypothesis acknowledges that not all processes relevant to the pathogenesis of psychosis can be examined in biological systems. The hierarchical predictive coding model exemplifies a new approach to brain modeling that allows computational formulations of brain function at multiple levels, from changes to synaptic gain to connectivity alterations within large-scale networks. A sufficiently detailed model of neuronal groupings and their connections could help define the set of conditions under which phenomena relevant to psychosis might emerge. Potentially it could also help explore the effect of diverse risk factors on neuronal computations in order to understand the mechanisms that render neural networks vulnerable to psychosis. Conversely, it could be used to model the effect of treatments and even assist in the selection of new pharmacological interventions based on their predicted effect on the model.

Acknowledgment

The authors have declared that there are no conflicts of interest in relation to the subject of this study.

References


