Original Article (Schizophrenia in Translation)

Title: Schizophrenia, Subjectivity and Mindreading

Running title: Schizophrenia, Subjectivity and Mindreading

Authors: Matthew M Nour 1,2,3 & Alvaro Barrera 4,5

Author Affiliation: 1South London and Maudsley NHS Foundation Trust, London; 2Institute of Psychiatry Psychology and Neuroscience, Kings College London, London; 3Nuffield Department of Clinical Neuroscience, John Radcliffe Hospital, University of Oxford, Oxford; 4Warneford Hospital, Oxford Health NHS Foundation Trust, Oxford; 5Department of Psychiatry, University of Oxford, Oxford.

Corresponding Author:
Dr Matthew M Nour
CT1, Postgraduate Centre, Maudsley Hospital, 111 Denmark Hill, London, SE5 8AZ
T: +44 20 3228 3840, F: +44 20 3228 5201
email: matthewnour@gmail.com

Manuscript length, including abstract and acknowledgments = 2496 words
Abstract length = 161 words (excluding key words)
Text body, including acknowledgements = 2335 words
This article contains 70 references and 0 tables or figures.
Abstract

A number of recent articles, many appearing in *Schizophrenia Bulletin*, signal a renewed interest in phenomenological approaches to our understanding of schizophrenia. These approaches conceptualize schizophrenia as a disorder of altered self-awareness and decreased pre-reflective social attunement, which may manifest as an impaired understanding of self, others and the physical world. Phenomenological approaches to psychopathology are sometimes construed as being incompatible with the reductionistic methodology of contemporary neuroscience. In this article we re-examine findings from the phenomenological investigation of schizophrenia in light of an influential neurocomputational account of mindreading, which postulates that understanding-of-others is subserved by coherent internal self-models. We argue that the phenomenological approach to schizophrenia is not incompatible with a neurocomputational account of mindreading, and that the two approaches should instead be viewed as existing in a relationship of mutual constraint and enlightenment. Our hypothesis, while speculative, is an attempt to marry the phenomenological and neuronal realities of schizophrenia. Furthermore, it has implications for psychotherapeutic interventions and future research.

**Key words:** Phenomenology/ Bayesian inference/ predictive coding/self disorder
Introduction

In recent years there has been a renewed interest in phenomenological exploration of mental illness in general,1,2 and schizophrenia in particular.3-7 In parallel, researchers have sought to identify points of contact between phenomenological and neuroscientific approaches to psychopathology.1,3,8-11 This paper continues in that spirit.

First, we present work in the phenomenological tradition that identifies anomalies in self-experience and understanding-of-others as the two core experiential deficits in schizophrenia.4-6,12-14 We then present an influential neurocomputational account of mindreading15,16 as a potential link between these two experiential abnormalities. Our proposal is motivated by the recognized requirement to unify phenomenological and biological accounts in psychiatry.1,5,17-19

Schizophrenia as a self-disorder

A phenomenological approach to psychiatry seeks to understand the pervasive structural alterations in patients’ subjective experience in an attempt to unify otherwise disparate psychopathological features.1,3,20,21 Schizophrenia has been the focus of much of this work, with a particular focus on patients’ experience of ‘self’ and ‘others’.

The sense of self can be experienced on multiple inter-dependent levels,22-25 from a pre-reflective minimal-self,4,5 to a sense of self emerging from ongoing intrapersonal
and interpersonal dialogues\textsuperscript{13} that occupies the center of a personal narrative grounded in the social world.\textsuperscript{26} Abnormalities at each of these levels have been proposed in schizophrenia.

Sass and Parnas\textsuperscript{4} propose that the core psychopathological alteration in schizophrenia is an ‘instability of pre-reflective self-awareness’,\textsuperscript{5} which has been termed an ‘ipseity-disorder’ or ‘self-disorder’. They argue that saturating all subjective experience is a pre-reflective awareness of self as the unified subject of perceptual experience.\textsuperscript{4} This notion is closely related to the feeling of inhabiting a living body embedded in the world.\textsuperscript{27} Sass and Parnas suggest that the weakening of minimal self in schizophrenia has two facets: hyper-reflexivity (an automatic tendency to direct objectifying attention towards normally tacit self-experiences) and lack of self-affection (a reduced automatic sense of being the unified and vital subject of awareness).\textsuperscript{4,5} These facets are postulated to be the core generative features of schizophrenia. A reduced sense of self-affection, for example, may manifest in passivity phenomena and a feeling of depersonalization, whilst hyper-reflexivity may result in bodily sensations taking on a pathologically-salient quality. Interestingly, ‘disturbance of the perception of self’ emerged as a core dimension of the schizophrenic prodrome in an early qualitative study by Møller and Husby.\textsuperscript{28}

This formulation of schizophrenia as a condition of altered subjectivity and self-awareness is not new. Kraepelin wrote of a ‘disunity of consciousness’,\textsuperscript{29} whilst
Bleuler wrote of a slackening of the associative links that tie together thoughts, perceptions and affects.\cite{30,31} Jaspers and Schneider described patients experiencing depersonalization who feel ‘estranged from themselves’,\cite{23} and for whom ‘Descartes’ “cogito ergo sum” … is no longer a valid experience’.\cite{23,32} Minkowski wrote of an atrophy of Bergsonian intuition and a ‘loss of vital contact with the world’,\cite{7,33} whilst Laing’s concept of ‘primary ontological insecurity’ resembles ipseity-disorder.\cite{34}

Alternatively, Paul and John Lysaker argue that a sense of self ‘develops out of dialogues both within the individual and between the self and others’, and suggest that a ‘disruption in the dialogical self’ is often found in patients with schizophrenia.\cite{13,35} The loss of dialogical ability is likely to disrupt personal narrative formation and lead to a subjective sense of personal incoherence.\cite{36} Personal narratives have been hypothesized to occupy the highest cognitive level in an individual’s predictive hierarchy, projecting a sense of predictability, security and order onto a world that may otherwise seem chaotic and threatening.\cite{26}

**Dis-sociality, autism and loss of ‘common sense’**

As the minimal self is related to a feeling of immersion in a social world, ipseity disturbance is closely linked to a disorder of social functioning.\cite{7,27,37,38}

Bleuler postulated that schizophrenia was characterized by an autistic ‘detachment from reality’ and the social world.\cite{7,30,31} More recently, autism in patients with
schizophrenia has been recast as a qualitative disturbance of ‘attunement’ with the social environment,\textsuperscript{7} which Stanghellini & Ballerini term ‘dis-sociality’.\textsuperscript{14,39} Normally, social attunement stems from an automatic participation in a social network of shared meanings, and depends on what Blankenburg called ‘common sense’.\textsuperscript{14,38,40} For Blankenburg, a loss of common sense was prominent in certain patients with schizophrenia, who are “unable to play along with the rules of the game of interpersonal behavior.”\textsuperscript{40}

From another perspective, RD Laing postulated that patients with schizophrenia exist in a state of primary ontological insecurity, and are liable to view dialectical interpersonal interaction as threatening.\textsuperscript{34} This reasoning has recently been explored from the perspective of dialogical difficulties.\textsuperscript{12}

Dis-sociality and lack of social attunement have a number of clinical manifestations. Stanghellini talks of ‘deviant’ behaviors as the ‘epiphenomena of the disorders of primordial intersubjectivity and common sense’.\textsuperscript{14} In the case of a young woman, Blankenburg wrote that ‘the simplest of things ... seem to slip away from her: how to behave in certain situations ... how to speak with the people one meets, or what one is supposed to think about them’.\textsuperscript{40} More recently, Chris Frith has proposed that the thought disorder of schizophrenia is a manifestation of the patient’s inability to keep
track of what their interlocutor currently knows, resulting in an inability to appropriately signpost their speech.⁴¹

In the case where intersubjectivity is seen as threatening, patients may prefer social isolation.³⁴ Henriksen and Parnas have described the case of Jane, a young patient who found interpersonal contact unbearable, and recalled that even as a child her classmates made her feel uneasy as ‘they would come too close.’ Jane also struggled with a feeling of not truly existing in the world, and these complaints preceded psychosis. Minkowski describes a patient who had lost the ‘sensitivity’ to communicate with others, and viewed ‘any outside force as an attack on his personality; if he succumbed to it he would be dragged down and engulfed’.³³ Minkowski’s patient also complained of a feeling of ‘physiological decomposition’, ‘emptiness in the head’, and a feeling that his voice seemed ‘dead’.

**Empirical studies**

Evidence for self-disorder as a specific feature of schizophrenia-spectrum disorders has emerged in the last fifteen years,²⁸,⁴³,⁴⁴ most recently in two heterogeneous samples of first-admission patients using the validated Examination of Anomalous Self-Experiences (EASE) tool.³⁷,⁴⁵ Self-disorder appears to show only a weak correlation with positive symptom score,³⁷ suggesting that it is a trait or vulnerability⁴⁶ phenotype of schizophrenia-spectrum disorders, akin to Bleuler’s ‘fundamental symptoms’.⁶,³¹
Social functioning and mindreading in schizophrenia has been studied using ‘Theory of Mind’ (ToM) paradigms. A large meta-analysis found consistent ToM impairments in patients with schizophrenia, both in psychotic and remission phases. In a subsequent meta-analysis, the same authors found ToM impairments not only in first-episode psychosis patients, but also in their non-psychotic relatives and in ultra-high-risk patients. The magnitude of the impairment in the relatives and ultra-high-risk group was intermediate between the psychotic group and control group. ToM impairment may therefore represent a trait feature of schizophrenia, conferring a vulnerability phenotype. ToM impairments in schizophrenia are clinically important, correlating with quality of life measures and insight. They may also be accompanied by a sense of loss of immediate social ‘attunement’ and feelings of ‘invasiveness’.

**Self-disorder underlies dis-sociality**

Self-understanding and understanding-of-others are clearly interrelated, with a number of papers proposing that dis-sociality is secondary to self-disorder in schizophrenia. In this paper we present a neurocomputational hypothesis that supports this notion. Computational neuroscience aims to understand the ability of neural circuits to extract information about the environment from sensory data, where this task is seen as a type of statistical inference. From this perspective, understanding the goals and intentions of another person is among the most...
challenging (statistical) inferential task that neuronal circuits must perform, as any number of ‘hidden’ motives could underlie the observable behavior of another person.\textsuperscript{15,16,56} At first glance this account of mindreading could not be further from that of phenomenology, in which empathic understanding of others is emphatically non-inferential, and often ‘goes beyond’\textsuperscript{57} the immediate sensory data. It must be appreciated, however, that there is no simple isomorphism between the descriptions of the ‘same’ phenomena at neuronal and experiential levels of analysis. The understanding of another person’s intentions can be non-inferential at the level of phenomenological analysis, whilst simultaneously being subserved by mechanistic computational processes at the subpersonal neuronal level.\textsuperscript{9} Indeed, this must be the case.

**Perception as unconscious inference**

In predictive coding accounts, mindreading is construed as a special case of perceptual inference.\textsuperscript{15,16} In perceptual inference the brain infers the current ‘state of the world’ by combining incoming sensory data with prior knowledge gained through experience.\textsuperscript{56,58,59}

Neuronal sensory processing circuits are arranged hierarchically.\textsuperscript{56,58,60} New sensory information propagates up neuronal hierarchies, starting at the bottom level of sensory epithelia. In parallel, high levels of the processing hierarchy make on-line predictions about the expected sensory input given the likely current ‘state of the world’ (known as ‘prior probabilities’ or ‘priors’). Prior sensory predictions
cascade down the circuit and constrain the interpretation of incoming sensory signals, which would otherwise be open to an infinite variety of interpretations. At all levels of the hierarchy the (top-down) prior predictions and (bottom-up) sensory data are compared, and any discrepancy (‘prediction error’) is used to fine-tune the current estimate of the ‘state of the world’.

This iterative process results in an automatic convergence towards an internally-coherent hierarchical representation of the state of the world that is accurate and multi-layered, and represents the circuit's best estimate of the current state of the world (the ‘posterior probability’).

Importantly, the top-down generation of sensory predictions relies on the brain’s internal ‘generative model’ of the causal structure of the world, which enables sensory signals to be predicted from representations of the likely ‘state of the world’. These generative models are sculpted and optimized in light of experience. By suppressing the generative model’s prediction errors the process described above achieves a Bayesian inversion of these hard-won generative models. This inversion essentially reverses the inputs and outputs of the generative model, resulting in a ‘recognition model’ that can identify the ‘states of the world’ given new sensory data (i.e. a model capable of perceptual inference). Mathematical treatments of this proposal are outlined elsewhere.

Mindreading and perceptual inference
An intriguing suggestion is that the task of inferring the hidden mental states of others from observations of their actions uses a special class of internal generative models that have been fine-tuned to optimize self-generated motor actions.\textsuperscript{15,16}

These specialized generative models transform self-generated motoric commands, themselves hierarchically determined by personal goals and intentions, into predicted sensory consequence of the voluntary action. The models may be thought of as ‘self-models’ composed of self-organizing hierarchies of personal priorities, goals and intentions, which relate to a given social context.\textsuperscript{16} Kilner, Friston and Frith propose that an iterative Bayesian inversion of these generative models, occurring at the subpersonal level as described above, would result in recognition models capable of grasping another person’s goals and intentions from observations of their actions.\textsuperscript{16} This inversion would allow the neuronal circuit to converge on the internally-coherent hierarchical configuration of ‘hidden causes’ that best predicts the observed actions of the other person, and is as experientially ‘inferential’ as seeing the world when you open your eyes.

What makes mindreading a special case of perceptual inference is that this final hierarchical configuration is actually taken from the observer’s own personal repertoire of generative models linking an understanding of environmental context to self-generated action, through a hierarchy of intentions, priorities and goals specific to the observer.\textsuperscript{15,16} Thus, an accurate prediction of the intentions of another person will depend on an intuitive grasp of how intentions, goals and desires are
organized within a person, and how they relate to wider environmental and social contexts. This understanding is reminiscent of Blankenburg’s ‘common sense’.40

If schizophrenia-spectrum disorders are characterized by instability in the self-model4 and a reduced ability to form stable dialogical self-representations13 then the foundation (generative models) upon which intuitive mindreading is based would be disrupted. Consequently, social interactions become less predictable, and hence more threatening. In this way mindreading is secondary to the presence of a stable and internally coherent self-model that is attuned to the world of social and environmental contexts. Interestingly, patients diagnosed with certain personality disorders, in which the sense of self is affected, have also been found to have impairments in mindreading.53

**Therapeutic and research implications**

The speculative hypothesis presented in this paper has implications for therapeutic approaches and future research. If coherent self-models of goal-directed action subserve mindreading abilities, therapies that improve self-awareness and help re-construct self-narratives in schizophrenia may enhance patients’ experiences of living in the social world. Paul Lysaker and colleagues have demonstrated that, in the course of individual psychotherapy, increased reflective self-awareness often emerges many months before the capacity to form mature relationships with other people.35,62–64 Therapeutic approaches that strengthen the patient’s sense of self and
improve the coherence of personal life narratives may aid recovery and social functioning.\textsuperscript{12,65–70}

Future studies should address how the different facets of altered self-awareness and dis-sociaility are related within an individual. Furthermore, disorders of social functioning and self-awareness occur in psychiatric disorders other than schizophrenia, and further work will be required to fully characterize the features that are specific to schizophrenia.

\textbf{Conclusions}

We have presented theoretical and empirical work suggesting that alterations in self-awareness and awareness-of-others form the core experiential abnormalities in schizophrenia. We proposed a speculative hypothesis connecting these findings to an influential neurocomputational account of mindreading, which suggests that mindreading is subserved by internal self-models. The hypothesis presented here is an example of how phenomenological and neuroscientific approaches to psychopathology can coexist in a relationship of mutual enlightenment and constraint.\textsuperscript{21} The phenomenological finding that schizophrenia is characterized by disorders of self-awareness and social functioning raises the question of how these facets may be related at the neuronal level, whilst a neurocomputational formulation positing an asymmetry in the relationship between the core experiential deficits has implications for phenomenologically-minded psychotherapeutic interventions.
The current effort to empirically demonstrate the reality of phenomenological abnormalities in schizophrenia and relate them to neuroscientific theories is an encouraging step towards a richer and more valid understanding of psychopathology.

**Acknowledgements**

We thank two anonymous reviewers for their thoughtful comments.

**Conflicts of interest**

MMN and AB have no conflicts of interest in relation to the subject of this study.

**Funding**

MMN and AB received no funding for this work.

**References**


