Bayesian inference, dysconnectivity and neuromodulation in schizophrenia

This scientific commentary refers to ‘Estimating changing contexts in schizophrenia’, by Kaplan et al. (doi:10.1093/brain/aww095).

The paper by Kaplan et al. in this issue of Brain addresses one of the most interesting questions in contemporary schizophrenia research: the role of uncertainty during perception (Kaplan et al., 2016). Uncertainty enjoys much interest in schizophrenia research as it may provide a crucial link between core clinical symptoms of schizophrenia—aberrant perceptual inference (e.g. hallucinations) and abnormal beliefs (delusions)—and longstanding neurobiological findings that patients with schizophrenia display widespread alterations in structural and functional brain connectivity (dysconnectivity).

These two cardinal features of schizophrenia have been integrated into disease theories, which have developed in three waves. A first influential proposal was that dysconnectivity in schizophrenia arises from abnormal regulation of NMDA receptor (NMDAR)-dependent transmission by neuromodulatory (dopaminergic and cholinergic) influences (Friston, 1998). Given the critical role of NMDARs for synaptic plasticity and myelination, this suggested that both neurodevelopmental aspects of schizophrenia (cf. abnormal pruning of connections by altered experience-dependent plasticity) and structural dysconnectivity might arise from a primary disturbance of NMDAR-dependent plasticity due to aberrant neuromodulatory control. Second, these putatively abnormal NMDAR-neuromodulator interactions (NNI) were proposed to cause a central computational impairment in schizophrenia: abnormal hierarchical Bayesian inference in the cortex (Stephan et al., 2006). This proposal was inspired by the notion that the brain constructs a hierarchical and probabilistic model of the world in order to infer the environmental causes of its sensory inputs (predictive coding), and by the increasingly discernible importance of NMDAR-neuromodulator interactions for implementing hierarchical Bayesian inference in the brain (Fig. 1). Under generic conditions, belief updates in Bayesian inference are driven by prediction errors (the difference between actual and predicted inputs) but, critically, weighted by how uncertain or precise both predictions and sensory inputs are. While prediction (error) signalling relies on glutamatergic transmission (NMDA and AMPA receptors), uncertainty-weighting may draw on tonic neuromodulatory signals, e.g. dopaminergic or cholinergic volume transmission. This view puts uncertainty (or its inverse, precision) at the core of hypotheses of schizophrenia in the construction of bridges from neurophysiology to clinical symptoms and led to influential conceptualizations of perceptual aberrations and delusion formation in schizophrenia (Fletcher and Frith, 2009; Corlett et al., 2010; Adams et al., 2013).

While these theories are appealing in that they integrate physiological and computational mechanisms and link them to clinical symptoms, testing their predictions has been a slow process. This is partially because the necessary tools have been under development. Recent years, however, have seen major methodological advances in computational neuroimaging. For example, trajectories of individual belief updates, and how they are driven by prediction errors and uncertainty, can be inferred from individual behaviour by hierarchical Bayesian models (Nassar et al., 2010; Iglesias et al., 2013). Furthermore, dynamic causal modeling has made it possible to characterize effective (directed) connectivity between neuronal populations based on functional MRI or electrophysiological data.

The article by Kaplan et al. provides a compelling demonstration of how these techniques can be combined to probe hypothesized abnormalities of Bayesian inference and neuromodulation in schizophrenia. Using both hierarchical Bayesian
modelling of behaviour (Nassar et al., 2010) and dynamic causal modelling of functional MRI, their work focuses on one important aspect of uncertainty in hierarchical Bayesian inference: disambiguating changes in context from noise in sensory inputs.

Kaplan et al. compared medicated patients with schizophrenia to healthy controls, using a paradigm in which numbers drawn from a Gaussian probability distribution were presented sequentially to the individual. With a certain probability (unknown to the participants), the mean of this distribution would shift on any given trial, and the participants’ task was to indicate when they believed that a shift had taken place. Mastering this task requires hierarchical inference, where estimates of the distribution’s mean are informed by estimates of higher-order statistical structure (i.e. probability of context changes).

Model-based analysis of the behavioural data indicated that, compared to healthy controls, patients with schizophrenia overestimated context change probability. Additionally, they showed a greater sum of prediction errors across the task and a higher learning rate in response to perceived change points than controls. Neurophysiologically, these behavioural differences were accompanied by reduced activity of anterior prefrontal cortex in patients at the time they believed they had detected a context change. Conversely, patients showed increased midbrain activity on those post-decision trials that featured numbers suggesting that the previous decision was likely to have been correct (implicit confirmatory feedback). These (and other) functional MRI findings suggest that interactions between the dopaminergic midbrain and a previously described hierarchical cortical network comprising parietal cortex, dorsolateral prefrontal cortex (DLPFC), and anterior prefrontal cortex (Badre and D’Esposito, 2009) are critical to master the inferential challenge posed by the task.

Kaplan et al. scrutinized this idea by constructing and comparing four dynamic causal models, which captured the above hierarchy and its interaction with the midbrain in different ways. They found that both the detection of a perceived context change and processing subsequent information in favour of this decision led to a reduction of most cortico-cortical connection strengths; by contrast, the latter condition enhanced cortical connections from parietal cortex and DLPFC to the midbrain. Notably, the increase in the DLPFC→midbrain connection was related to clinical symptoms: it was significantly more pronounced in patients with strong delusions, compared to patients with mild delusional symptoms.

Prefrontal-midbrain connections have been examined in detail by previous neuroanatomical and pathophysiological studies, and represent a particularly prominent case of NMDAR-neuromodulator interactions in the dysconnection theory of schizophrenia. Prefrontal connections to the midbrain utilize NMDA receptors in order to exert a potent drive on dopaminergic neurons, where the strength of these glutamatergic synapses is regulated by both cholinergic afferents from the brainstem as well as autocrine dopamine release (for review, see Stephan et al., 2009). The dynamic causal modelling results by Kaplan et al. reveal that the DLPFC→midbrain connection is enhanced in patients with schizophrenia, both when context changes are perceived and when subsequent confirmatory information is processed. This is consistent with a dysregulation of glutamatergic prefrontal influences on the midbrain, possibly leading to excessive and/or ill-timed dopamine release (cf. ‘aberrant salience’; Kapur, 2003). The resulting maladaptive plasticity induced by aberrant dopamine release via efferent connections from the midbrain might represent a key trigger for the
formation of abnormal beliefs and, eventually, delusions (Stephan et al., 2009; Adams et al., 2013). The relation between DLPFC→midbrain connection strength and delusion severity that Kaplan et al. report lends empirical support to this previous proposal. Overall, the study by Kaplan et al. makes important contributions to characterizing abnormalities of Bayesian inference in schizophrenia and their potential relation to altered neuromodulation. Having said this, the present study also has a number of limitations that deserve consideration. First, the functional MRI results derive from categorical analyses of task events, such as post-decision trials that were labelled as representing implicit confirmatory feedback. This labelling was partially based on applying binary thresholds, which, although not implausible, essentially represent an arbitrary cut-off. Future work should exploit the inferred computational quantities more directly and examine how functional MRI activity and connectivity relate to trial-by-trial estimates of prediction errors and uncertainty at different hierarchical levels (e.g. trial outcomes and change points). Second, while systematic model selection was applied to dynamic causal models of functional MRI data, the analysis of behavioural data took one specific model for granted, without comparing it to alternative possibilities of how full hierarchical Bayesian inference could be approximated in the brain, and without considering how this might vary across subjects. Finally, the patients were receiving antipsychotic medication. This might represent a particular confound for the neuroimaging results on activity and connectivity of dopaminergic and dopaminoreceptive regions.

Despite these caveats, the study of Kaplan et al. represents an important step forward in schizophrenia research. It demonstrates the potential of computational neuroimaging investigations and contributes novel evidence in support of theories positing a link between abnormalities of NMDA-receptor modulator interactions and Bayesian inference in schizophrenia. As our understanding of the physiological implementation of hierarchical Bayesian inference in cortex advances (Fig. 1), one might hope that suitably validated computational assays will become useful tools to support differential diagnosis and individual treatment predictions for patients with schizophrenia.

**Funding**

We acknowledge support by the René and Susanne Braginsky Foundation and the University of Zurich.

Klaas E. Stephan,1,2,3 Andrea O. Diaconescu,4 and Sandra Iglesias1

1 Translational Neuromodeling Unit (TNU), Institute for Biomedical Engineering, University of Zurich and ETH Zurich, 8032 Zurich, Switzerland
2 Wellcome Trust Centre for Neuroimaging, University College London, London, WC1N 3BG, UK
3 Max Planck Institute for Metabolism Research, 50931 Cologne, Germany

Correspondence to: Klaas E. Stephan. E-mail: stephan@biomed.ee.ethz.ch
doi:10.1093/brain/aww120

**References**


Friston KJ. The disconnection hypothesis. Schizophr Res 1998; 30: 115–25.


