Dysfunctional insular connectivity during reward prediction in patients with first-episode psychosis

André Schmidt,[1,2,3] Lena Palaniyappan,[4] Renata Smieskova,[1,3] Andor Simon,[1] Anita-Riecher Rössler,[1] Undine E. Lang,[1] Paolo Fusar-Poli,[2] Philip McGuire,[2] Stefan J. Borgwardt,[1,2,3]

[1] Department of Psychiatry (UPK), University of Basel, Switzerland

[2] Institute of Psychiatry, Department of Psychosis Studies, King's College London, UK

[3] Medical Image Analysis Center (MIAC), University Hospital Basel, Switzerland

[4] Division of Psychiatry, University of Nottingham, UK

Corresponding author: André Schmidt, Ph.D., University of Basel, Department of Psychiatry (UPK), Wilhelm Klein Strasse 27, 4012 Basel, Switzerland, Phone: +41 61 265 78 79, fax: +41 61 265 45 88, email: andre.schmidt@unibas.ch

Word count text: 3999

Word count abstract: 250

Abstract

Background: Increasing evidence indicates that psychosis is associated with abnormal reward processing. Imaging studies in patients with first-episode psychosis (FEP) revealed reduced activity in diverse brain regions including the ventral striatum, insula and anterior cingulate cortex (ACC) during reward prediction. However, whether these reductions in local brain activity are due to altered connectivity has barely been explored.

Methods: We applied dynamic causal modelling and Bayesian model selection to functional magnetic resonance imaging (fMRI) data during the Salience Attribution Task to investigate whether patients with first-episode psychosis (FEP) showed abnormal modulation of connectivity between the ventral striatum, insula and ACC induced by rewarding cues and whether these changes were related to positive psychotic symptoms and atypical antipsychotic medication.

Results: The model including reward-induced modulation of insula to ACC connectivity was the best fitting model in each group. Compared to healthy controls (n=19), FEP patients (n=29) revealed reduced connectivity from the right insula to the ACC. After subdividing patients according to current antipsychotic medication, we found that the reduced insula to ACC connectivity relative to healthy controls was only observed in untreated (n=17) but not antipsychotic-treated patients (n=12) and correlated negatively with unusual thought content in untreated FEP patients.

Limitations: Modest sample size of untreated FEP patients

Conclusion: This study indicates that insula to ACC connectivity during reward prediction is reduced in untreated FEP and related to the formation of positive psychotic symptoms. It further suggests that atypical antipsychotics may reverse connectivity between the insula and the ACC during reward prediction.

Introduction

Our brain is constantly exposed to a wide variety of stimuli, which compete for limited cognitive resources. External stimuli are processed depending on their salience so as to ignore predictable, state and task-irrelevant events while enhancing resource allocation to process unexpected or state and task-relevant events. Efficient prediction of salient stimuli such as those of rewards is thus essential for adapting ongoing behaviour. This process requires the ability to learn that a neutral stimulus becomes emotionally endowed due to its association with primary reinforcement.[1] Behavioural and fMRI studies have demonstrated impairments in patients with psychosis when anticipating reward.[2] Relative to controls, behavioural evidence indicated that FEP patients exhibited less reactivity to rewarding-predicting cues.[3] fMRI studies during reward prediction have reported reduced activity in diverse brain regions of unmedicated FEP patients including the ventral striatum (VS),[4, 5] ACC, midbrain, thalamus, and cerebellum compared with controls.[5] It has further been shown that VS activation during reward prediction was negatively related to positive psychotic symptoms in FEP patients.[4, 5]

Reward processing is critically mediated by dopamine[6, 7] and the VS response to reward-predicting cues is likely triggered by dopamine activity.[8, 9] A previous fMRI study in chronic schizophrenia patients showed that the VS response during reward prediction was only reduced in patients treated with typical antipsychotics, whereas in contrast no difference to healthy controls was observed in patients treated with atypical medication.[10] In line with this finding in chronic patients, the reduced baseline VS activation during reward prediction seen in FEP relative to healthy controls has been normalized after 6 weeks monotherapy with atypical antipsychotics.[11] The largest improvement in positive symptoms was seen in those patients with the highest VS signal increase.[11] Although not specifically during reward processing, a recent resting state fMRI study in FEP patients could also show that atypical antipsychotics increased functional connectivity between striatal regions, the ACC and right anterior insula,[12] which correlated positively with symptom improvement. Using the Salience

Attribution Task (SAT),[13] Smieskova and colleagues recently reported that FEP patients revealed a reduced right insula activity in response to high versus low-probability rewarding cues compared with controls.[14] Furthermore, the right insula and ACC activity was negatively correlated with the severity of hallucinations in unmedicated patients.[14] These three fMRI studies together show local activity changes mainly in the VS, insula and ACC in FEP patients during reward prediction[4, 5, 14] and alterations in these regions induced by antipsychotic medication.[11, 14] One previous fMRI study in unmedicated schizophrenia patients showed reduced connectivity between the prefrontal cortex and the VS during reward processing.[15] However, it remains still unclear whether the local brain activity changes in FEP patients during reward prediction may result from alterations in the underlying connectivity.

In this study, we applied dynamical causal modelling (DCM[16]) and Bayesian model selection (BMS[17]) to the fMRI data published by Smieskova et al.[14] to address the following questions: first, among connectional models including the visual cortex, VS, insula and ACC, we investigated the regions where the high-probability rewarding cues operate and modulate connectivity strengths. We included the visual cortex as sensory input region in our models based on evidence showing that reward also modulates responses in the visual cortex.[18] Second, we investigated differences between healthy controls and FEP patients in the connectivity strengths obtained from the best fitting model and investigated possible effects of atypical antipsychotics. Finally, we explored the relation between the modulation of connectivity induced by high-probability rewarding cues and the expression of positive symptoms in FEP patients.

Methods

Patients

Participants were recruited in a specialized clinic for the early detection of psychosis at the University Hospital of Psychiatry, Basel, Switzerland. All participants provided written informed consent, and had received compensation for participating. The study was approved by the local ethics committee (Ethikkommission Nordwest- und Zentralschweiz (EKNZ)). All patients were competent to give informed consent. They were able to understand relevant study information including the reasons why they are being asked to participate and the procedures of the study and they understood the consequences of accepting or declining the invitation to participate and how to discontinue their participation.

We recruited 30 FEP patients who fulfilled criteria for acute psychotic disorder according to the ICD-10 or DSM-IV, but not yet for schizophrenia.[19] The upper limit of the duration of psychosis was 5 years and the mean duration among our included FEP patients was 7.76 months (SD=15.77 months). One patient was not able to continue the MRI examination. At study intake, we assessed subjects using the Brief Psychiatric Rating Scale (BPRS), the Scale for the Assessment of Negative Symptoms (SANS) and the Global Assessment of Functioning (GAF). Inclusion required scores of 4 or above on the hallucination item or 5 or above on the unusual thought content, suspiciousness or conceptual disorganization items of the BPRS,[19] with symptoms occurring at least several times a week and persisting for more than one week. We obtained current nicotine, cannabis and other illegal drug consumption using a semi-structured interview adapted from the Early Psychosis Prevention and Intervention Centre Drug and Alcohol Assessment Schedule (www.eppic.org.au) and applied the following exclusion criteria: history of previous psychotic disorder; psychotic symptomatology secondary to an organic disorder, recent substance abuse according to ICD-10 research criteria, psychotic symptomatology associated with an affective psychosis or a borderline personality disorder, age under 18 years; inadequate knowledge of the German language, and IQ less than 70. Furthermore, 12 FEP patients were receiving the

following atypical antipsychotics: six patients receiving quetiapine, two receiving olanzapine/aripiprazole and one receiving paliperidone/risperidone. 17 patients were without current antipsychotic medication, while 11 of them were antipsychotic-naïve and six were antipsychotic-free. Seven patients were taking antidepressants at the time of the MRI scan. We recruited 23 healthy controls (HC) from the same geographical area. Four HC had to be excluded due to brain vascular abnormalities (n=3) and arachnoid cyst (n=1). HCs had no current psychiatric disorder, no history of psychiatric illness, head trauma, neurological illness, serious medical or surgical illness, substance abuse and no family history of any psychiatric disorder as assessed by an experienced psychiatrist in a detailed clinical assessment. Table 1 presents details of subjects included in the analysis.

Insert table 1 about here

Salience attribution test (SAT)

The SAT has been previously described in more detail.[13, 20, 21] In brief, the SAT is a speeded-response game, rewarded with money, which measures responses to task-relevant and task-irrelevant cue features.[21] Participants had to respond to a briefly presented square. Before the onset of the square, participants seeing different categories of cues, indicating the likelihood of reward on a given trial. Participants receive monetary reward on 50% of trials, with more money for faster responses. The cues vary in two different visual dimensions; color (red or blue) and shape (animals or household objects), with one of these cue dimensions being task-relevant and other task-irrelevant. In the task-relevant dimension, one cue dimension is highly associated with receiving a reward, with 87.5% of these trial types rewarded (e.g. blue animals and households), whilst only 12.5% of the alternative cue dimension was rewarded (e.g. red animals and households). In the task-irrelevant dimension, 50% of both cue types were rewarded (e.g. 50% of all animals and 50% of all households). Participants were not informed about the contingencies, which remained the same over blocks, and had to learn them during the task. They were also asked to estimate reward

probabilities for each of the 4 stimulus categories after each session using visual analogue scales (VAS) ranging from 0 to 100 per cent. The SAT provides behavioural (in terms of VAS ratings and reaction times) and neuronal measures of *adaptive* (task-relevant features) and *aberrant* (task-irrelevant features) reward prediction. An exemplary trial during the SAT is shown in the supplementary material. Based on our previous findings showing neuronal differences between HC and FEP patients during adaptive reward prediction,[14] the present connectivity analysis focused on behavioural and neural effects during *adaptive* reward prediction (high-probability versus low-probability rewarding cues).

Image acquisition and analysis

Scanning was performed on a whole-body 3T MRI system (Magnetom Verio, Siemens Healthcare, Erlangen, Germany). During the SAT, we acquired T2*-weighted echo-planar images with the following parameters: 38 axial slices of 3 mm thickness, 0.5 mm interslice gap, field of view 228 x 228 cm^2 and an in-plane resolution of 3 x 3 mm^2. The repetition time was 2.5 s and the echo time 28 ms. EPIs were analyzed using SPM8 (www.fil.ion.ucl.ac.uk/spm). During pre-processing, images were realigned and unwarped, spatially normalized to the MNI space template (including reslicing to 2x2x2 mm voxels), and smoothed with a Gaussian kernel at 8 mm full half-width maximum. We first checked the realignment parameters of each individual to identify scans on which sharp movements (bigger than half of the voxel size (1.5mm) and/or more than 1.5°) had occurred and inspected those scans manually. Corrupted images were excluded and replaced with the average of the neighboring images. No subject had more than 10% corrupted images due to movement. Maximum likelihood parameter estimates were then calculated at the first level at each voxel using the general linear model. Our design matrix included an autoregressive AR(1) model of serial correlations and a high-pass filter with a cutoff of 128 s. The onsets of each event (duration 2 s for cue and 1.5 s for the outcome regressor) were convolved with the hemodynamic response function and its temporal and dispersion derivatives. The first level design matrix included four cue regressors (blue/red animals, blue/red objects, red

objects), an outcome regressor and its parametric modulation by magnitude of reward.

Volumes of interest

We selected the bilateral visual cortex (left: x=-24, y=-98, z=-8; right: x=22,y=-98, z=-6), VS (left: x=-14, y=6, z=-4; right: x=14, y=6, z=-8), insula (left: x=-34, y=14, z=0; right: x=34, y=24, z=6), as well as the dorsal ACC (x=-4, y=16, z=28) as volumes of interest (VOIs) based on following information: 1) the previously published second-level SPM analysis of these data showing reduced right insula and ACC activity in FEP patients,[14] 2) previous fMRI studies in FEP showing reduced activity in the VS[4, 5] and ACC[5] during reward prediction and 3) evidence demonstrating that reward prediction responses in the VS were normalized after atypical antipsychotic medication in FEP patients.[11] The visual cortex coordinates were based on the activation induced by all stimuli (high- and low-probability rewarding cues) collapsed across groups, while the coordinates for the VS, insula and ACC were specified from the contrast of high-probability minus low-probability rewarding cues (cluster-forming threshold of p=0.001 uncorrected, FWE-corrected at cluster-level at p<0.05). For each subject, regional time series from these VOIs were extracted within spheres of 4-mm radii centered on the peak of the contrasts of interest within the same anatomical area, as defined by the PickAtlas toolbox[22] (p<0.01 uncorrected, adjusted for effects-of-interest F contrasts).

Network analysis: Dynamic causal modelling (DCM)

DCM10 (revision No. 4290) in SPM8 was used to explore causal interaction among our VOIs. DCM[16] is a hypothesis-driven method that does not explore all possible models, but tests a specified model space based on prior knowledge about the system of interest. The bilinear DCM for fMRI infers dynamics at the neuronal level by translating modelled neuronal responses into predicted BOLD measurements. Specifically, DCM allows modelling how neural states (reflecting specific brain region) change as a function of endogenous inter-regional connections, modulatory effects on these connections, and driving inputs.[16] In this study, we particularly applied DCM to probe how the endogenous connections induced by all stimuli are modulated by high-probability rewarding cues (modulatory effect).

Model space construction

Across all models tested, we assumed the same network layout with reciprocal connections between the VS, insula and ACC. The bilateral visual cortex was further incorporated as sensory input regions, which were reciprocally connected with the insula and VS. Bilateral visual cortices and the VS exhibited inter-hemispheric connections as well. This base model was then elaborated systematically to produce alternative variants, which varied in where the effect of high-probability rewarding cues modulated connections among our VOIs (see Figure 1). These variations were guided by studies highlighting functional ACC-insula,[23, 24] ACC-VS[25] and insula-VS[18, 26] interactions during reward processing, and by studies providing evidence for an involvement of the visual cortex (and their connections to the insula and VS) in reward processing.[18, 27] In particular, we allowed high-probability rewarding cues to modulate 1) only ACC-insula connectivity, 2) ACC-insula and ACC-VS, 3) ACC-insula, ACC-VS and insula-striatum connectivity and 4) ACC-insula, ACC-VS, insula-striatum and visual cortex to insula and VS connectivity. These four options were crossed with the possibility that high-probability rewarding cues either affected i) forward, ii) backward or iii) both forward and backward connections within the hierarchical network. This additional fractioning was driven by the principle of predictive coding,[28, 29] which proposes neuronal message passing among different levels of cortical hierarchies.

Insert Figure 1 about here

Bayesian Model selection (BMS)

We used BMS[17] to determine the most plausible model of the ones we considered. The BMS method rests on comparing the (log) evidence of a predefined set of models (see model space construction). The model evidence is the probability of observing the empirical data, given a model, and represents a principled measure of model quality derived from probability theory.[17] We used a random-effects BMS approach for group studies, which is capable of quantifying the degree of heterogeneity in a population while being extremely robust to potential outliers.[30] One common way to summarize the results of random-effects BMS is to

report the exceedance probability (EP) of each model (i.e. the probability that this model is more likely than any other of the models tested, given the group data).

Group statistics

1-way analysis of variance (ANOVA) and Chi-Quadrat tests, respectively, were used to examine between-group differences in clinical, demographical, behavioural parameters and Bonferroni post-hoc testing was applied to correct for multiple comparisons. The connectivity analysis was based on the summary statistics approach in DCM, that is, model selection followed by interrogation of posterior estimates.[31] In particular, we used the posterior means reflecting the modulatory effect from the best fitting model obtained from BMS for the ANOVA analysis. In a first step, all FEP patients were treated as one group. A second ANOVA with three groups was then applied to address the effect of antipsychotics. Finally, Pearson correlation analysis was used to assess the relationship between significant group differences in connectivity strengths and positive psychotic symptoms (indexed by BPRS items 9, 10, 11 and 15) in treated and untreated FEP patients. The statistical threshold was adjusted for the number of correlations performed for both patient groups separately (n=4; $p<0.5/4$). The influence of potential outliers for each correlation was tested with Cook's distance test (critical value: $4/(n-k-1)=0.33/0.57$). No outliers were detected.

Results

Behavioural scores on adaptive reward prediction

Compared with HCs, FEP patients showed reduced VAS ratings at trend level ($F(1,47)=2.906$, $p=0.095$). No group difference for reaction times was found ($F(1,47)=2.561$, $p=0.116$). Subsequent ANOVA analysis with three groups revealed no differences between HCs, treated and untreated FEP patients for both VAS ratings ($F(1,47)=2.165$, $p=0.127$) and reaction times ($F(1,47)=1.379$, $p=0.262$).

Network analysis (DCM results)

Bayesian Model selection

Random-effect BMS revealed model 1 as the best fitting model in HCs (EP: 56%) and all FEP patients (EP: 65%). Model 1 was also superior to all other models tested if patients were separated in treated (EP: 29%) and untreated FEP patients (EP: 41%) (Figure 2A).

Group differences in effective connectivity

In our final group-level analysis, we were able to test for differences in 2 parameters describing the modulation of connections induced by high-probability rewarding cues (cf. model 1). We found a significant reduction in the modulation of right insula to ACC ($F(1,47)=5.976$, $p=0.018$) but not in the modulation of left insula to ACC connectivity in all FEP patients relative to HCs ($F(1,47)=0.320$, $p=0.574$).

Effects of antipsychotics on effective connectivity

The subsequent three-group ANOVA analysis revealed a significant group effect on the modulation of right insula to ACC ($F(2,47)=3.823$, $p=0.029$) but not left insula to ACC connectivity ($F(2,47)=0.281$, $p=0.756$). Compared to HC, post-hoc testing showed that the modulation of right insula to ACC connectivity induced by high-probability rewarding cues was significantly reduced in untreated ($p=0.025$) but not antipsychotic-treated FEP patients ($p=0.695$) (Figure 2B, Table 2).

Insert Table 2 and Figure 2 about here

Relation between abnormal connectivity and positive symptoms

Pearson correlation analysis indicated a significant negative correlation between the modulatory effect on right insula to ACC connectivity induced by high-probability rewarding cues and the formation of unusual thought content (BPRS item 11) in untreated ($r=-0.593$, $p=0.012$, corrected for multiple testing) but not treated FEP patients ($r=0.127$, $p=0.694$) (Figure 3). No correlations between right insula to ACC connectivity and BPRS items 9,10 and 15 were found.

Insert Figure 3 about here

Discussion

This study demonstrates that right insula to ACC connectivity during reward prediction is significantly reduced in FEP patients compared to HCs. Importantly, this reduced insula to ACC connectivity is only evident in untreated but not treated FEP patients and negatively related to the formation of unusual thought content in untreated patients.

Irrespective of the diagnostic group, the BMS results revealed that rewarding cues essentially modulated insula-ACC connectivity within our network, supporting the key role of this functional coupling during salience processing.[23, 32] This finding dovetails with the concept of proximal salience.[24] This concept proposes that the processing of incoming stimuli induces a proximal salience signal in the insula depending on its predictability, which indicates whether further downstream processing is required to adjust one's predictive model. The downstream processing includes motor action, updating the prefrontal fund of knowledge or stopping an activity that is ongoing. All of these downstream activities require resource allocation to appropriate networks and are initiated by insula-ACC interactions. With respect to the SAT, high-probability rewarding cues are the ones that require further downstream processing and action. The observation that these stimuli modulate insula-ACC connectivity adds support to the notion that the role of the insula-ACC network lies in the formation of stimulus-response association (proximal salience), which precedes the learning of stimulus-reinforcement associations (motivational salience) in which hippocampal-midbrain-striatal connections may play a more crucial role.

We further found a reduced right insula to ACC connectivity in untreated FEP patients compared with HC. Moreover, the degree of insula-ACC connectivity was negatively correlated with the formation of unusual thought content in these patients. These findings extend our previous result of reduced ACC activity in unmedicated FEP patients and the relationship between positive symptoms in untreated FEP patients and regional activity in the right insula and ACC in response to high-probability reward cues.[14] Given that the

psychopathological assessment was made at study intake and imaging later, dysfunctional insular connectivity could thus reflect vulnerability to positive symptom formation. Although functional connectivity studies extract a bilateral salience network pattern involving both right and left insula and ACC,[23] the right-hemispheric asymmetry is reminiscent of studies that use temporal information e.g. Granger causality or DCM.[33-35] A meta-analysis revealed that both the insula and ACC were accompanied by significant gray matter reductions in FEP patients,[36] which might provide a scaffold for the reduction of insula-ACC connectivity observed here. In accordance with this, deficits in gray matter volumes in the insula and ACC were also negatively related with delusion and hallucinations in psychotic patients.[37] However, gray matter losses in the ACC and insula have been detected across different psychiatric diagnoses and may not be specific to psychosis.[38] Within the framework of proximal salience, deficient insular detection of external salient events such as those of rewarding cues might lead to a faulty allocation of salience to internally generated thoughts and impede the attention to relevant external information.[24] The internal mental state might be further enhanced by inappropriate salience, promoting the formation of various psychotic symptoms such as hallucinations and delusions.[24] Unlike hallucinations and delusions, illogical thinking may be more pronounced when subjects are interacting with stimuli as in carrying out a task inside a scanner. The here found relationship between ACC-insula dysconnectivity when processing rewarding cues and the severity of thought content suggests that aberrant assignment of salience to task-relevant stimuli at hand may enhance the emergence of illogical and bizarre ideas in psychosis.

The putative imbalance between active inference processes about external phenomena and self-generated internal reflections may result from a failure of the insula-ACC network and in particular of the insula to switch between these two alternating systems. This interpretation is motivated by a recent model proposing that activation in the insula-ACC network is negatively correlated with the engagement of the default mode network,[23] a system that is active during the construction of self-relevant mental simulations.[39] Reduced negative

correlation between the default mode network and the task-positive network has already been observed in clinical high-risk subjects for psychosis. Notably, a negative relation was found between the correlation of default mode network and the task-positive network and the expression of cognitive impairments.[40]

Importantly, the reduced insula-ACC connectivity was only evident in untreated but not antipsychotic-treated FEP, suggesting a normalization of this functional coupling via D2 receptor antagonism together with 5-HT2A receptor antagonism.[41] This result corresponds to conclusions from a recent review that the BOLD signal in specific neural regions normalizes over the course of antipsychotic treatment[42] and to a recent resting state fMRI study showing that antipsychotic-induced improvement of psychotic symptoms was accompanied by increased functional connectivity between striatal regions, the ACC and the anterior insula.[12] The antipsychotic effect in treated patients can perhaps be explained by the underlying structure as well, given that insular and ACC volumes increase with increasing antipsychotic exposure in psychotic patients.[43, 44] However, meta-analytical evidence indicates that ACC and insula volume is particularly decreased in treated FEP patients.[45] More studies are needed to understand the structure-function relationship of the insula-ACC network in psychosis and alterations induced by antipsychotics.

Limitations

There are some limitations to be considered in the present study. We restricted our analysis to striatal-insular-ACC connectivity although there are also other regions activated in response to high-probability rewarding cues during the SAT such as the midbrain, medial dorsal thalamus and prefrontal cortex[20] and a previous study during the processing of aversive outcomes showed reduced functional connectivity between the medial prefrontal cortex and the VS in unmedicated schizophrenia patients compared with healthy controls.[15] More research is required to study (abnormal) functional connectivity during reward processing including feedback phases and the processing of aversive stimuli. We cannot completely rule out that smoking has confounded our findings given the impact of smoking

on the connectivity between the ACC and insula in schizophrenia.[46] However, there were no correlations between left (r=-0.67, p=0.652) and right insula to ACC connectivity (r=-0.65, p=0.659) and smoking behaviour across all subjects. Furthermore, abnormal insula-ACC connectivity seems to be task-specific. While insula-ACC dysconnectivity is not prominent in resting state conditions,[33] our results showed that when high-probability rewarding cues were presented, this network is not generating the neural readiness that is required for further action on the reward predicting stimuli as for example the formation of stimulus-reinforcement association. Another point of contention is that we found connectivity differences across groups in relation to antipsychotic medication, while no significant effects were found for the behavioural indices (though at statistical trend level). However, significant effect on brain activations but not behavioural performance is a common finding in fMRI studies and can be explained by the fact that functional neuroimaging techniques detect changes at the physiological level and are more sensitive than behavioural measures.[47] Finally, this study analysed a relatively modest number of treated and untreated FEP patients. Larger samples sizes are needed to replicate our findings.

Conclusion

In summary, this study demonstrates that FEP patients exhibit reduced right insula to ACC connectivity during reward prediction and that abnormal insula to ACC connectivity may make patients more vulnerable to the formation of psychotic symptoms. Our findings also suggest that atypical antipsychotics reverse insula-ACC connectivity during reward prediction in FEP patients. Longitudinal studies with larger samples are needed to draw robust inferences on medication effects on insula to ACC connectivity and to validate whether the assessment of effective insula connectivity during reward prediction may reflect an important brain marker of treatment effectiveness in psychosis.

Acknowledgements

This work was supported by the Swiss National Science Foundation (A.S. grant number 155184; R.S. and S.J.B. grant number 3232BO_119382). The authors declare that there are no conflicts of interest in relation to the subject of this study. We thank Jon Roiser for providing the SAT task.

References

1. Berridge KC, Robinson TE. What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Res Brain Res Rev* 1998; 28(3): 309-369.

2. Ziauddeen H, Murray GK. The relevance of reward pathways for schizophrenia. *Curr Opin Psychiatry* 2010; 23(2): 91-96.

3. Murray GK, Clark L, Corlett PR, et al. Incentive motivation in first-episode psychosis: a behavioural study. *BMC Psychiatry* 2008; 8: 34.

4. Esslinger C, Englisch S, Inta D, et al. Ventral striatal activation during attribution of stimulus saliency and reward anticipation is correlated in unmedicated first episode schizophrenia patients. *Schizophr Res* 2012; 140(1-3): 114-121.

5. Nielsen M, Rostrup E, Wulff S, et al. Alterations of the brain reward system in antipsychotic naïve schizophrenia patients. *Biol Psychiatry* 2012; 71(10): 898-905.

6. Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. *Science* 1997; 275(5306): 1593-1599.

7. Berridge KC, Robinson TE. Parsing reward. *Trends Neurosci* 2003; 26(9): 507-513.

8. Knutson B, Bjork JM, Fong GW, et al. Amphetamine modulates human incentive processing. *Neuron* 2004; 43(2): 261-269.

9. da Silva Alves F, Bakker G, Schmitz N, et al. Dopaminergic modulation of the reward system in schizophrenia: a placebo-controlled dopamine depletion fMRI study. *Eur Neuropsychopharmacol* 2013; 23(11): 1577-1586.

10. Juckel G, Schlagenhauf F, Koslowski M, et al. Dysfunction of ventral striatal reward prediction in schizophrenic patients treated with typical, not atypical, neuroleptics. *Psychopharmacology (Berl)* 2006; 187(2): 222-228.

11. Nielsen MO, Rostrup E, Wulff S, et al. Improvement of brain reward abnormalities by antipsychotic monotherapy in schizophrenia. *Arch Gen Psychiatry* 2012; 69(12): 1195-1204.

12. Sarpal DK, Robinson DG, Lencz T, et al. Antipsychotic Treatment and Functional Connectivity of the Striatum in First-Episode Schizophrenia. *JAMA Psychiatry* 2014; 72(1): 5-13.

13. Roiser JP, Stephan KE, den Ouden HE, et al. Do patients with schizophrenia exhibit aberrant salience? *Psychol Med* 2009; 39: 199-209.

14. Smieskova R, Roiser JP, Chaddock CA, et al. Modulation of motivational salience processing during the early stages of psychosis. *Schizophr Res* 2015; 166(1-3): 17-23.

15. Schlagenhauf F, Sterzer P, Schmack K, et al. Reward feedback alterations in unmedicated schizophrenia patients: relevance for delusions. *Biol Psychiatry* 2009; 65: 1032-1039.

16. Friston KJ, Harrison L, Penny W. Dynamic causal modelling. *Neuroimage* 2003; 19(4): 1273-1302.

17. Penny WD, Stephan KE, Mechelli A, et al. Comparing dynamic causal models. *Neuroimage* 2004; 22(3): 1157-1172.

18. Rothkirch M, Schmack K, Deserno L, et al. Attentional modulation of reward processing in the human brain. *Hum Brain Mapp* 2014; 35(7): 3036-3051.

19. Yung AR, Phillips LJ, McGorry PD, et al. Prediction of psychosis. A step towards indicated prevention of schizophrenia. *Br J Psychiatry Suppl* 1998; 172(33): 14-20.

20. Roiser JP, Stephan KE, den Ouden HE, et al. Adaptive and aberrant reward prediction signals in the human brain. *Neuroimage* 2010; 50(2): 657-664.

21. Roiser JP, Howes OD, Chaddock CA, et al. Neural and behavioral correlates of aberrant salience in individuals at risk for psychosis. *Schizophr Bull* 2013; 39(6): 1328-1336.

22. Maldjian JA, Laurienti PJ, Kraft RA, et al. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage* 2003; 19(3): 1233-1239.

23. Menon V, Uddin LQ. Saliency, switching, attention and control: a network model of insula function. *Brain Struct Funct* 2010; 214(5-6): 655-667.

24. Palaniyappan L, Liddle PF. Does the salience network play a cardinal role in psychosis? An emerging hypothesis of insular dysfunction. *J Psychiatry Neurosci* 2012; 37(1): 17-27.

25. Gluth S, Rieskamp J, Büchel C. Neural evidence for adaptive strategy selection in value-based decision-making. *Cereb Cortex* 2014; 24(8): 2009-2021.

26. Chikama M, McFarland NR, Amaral DG, et al. Insular cortical projections to functional regions of the striatum correlate with cortical cytoarchitectonic organization in the primate. *J Neurosci* 1997; 17(24): 9686-9705.

27. Pessoa L, Engelmann JB. Embedding reward signals into perception and cognition. *Front Neurosci* 2010; 4.

28. Friston K. A theory of cortical responses. *Philos Trans R Soc Lond B Biol Sci* 2005; 360: 815-836.

29. Friston K. The free-energy principle: a unified brain theory? *Nat Rev Neurosci* 2010; 11(2): 127-138.

30. Stephan KE, Penny WD, Daunizeau J, et al. Bayesian model selection for group studies. *Neuroimage* 2009; 46(4): 1004-1017.

31. Stephan KE, Penny WD, Moran RJ, et al. Ten simple rules for dynamic causal modeling. *Neuroimage* 2010; 49(4): 3099-3109.

32. Seeley WW, Menon V, Schatzberg AF, et al. Dissociable intrinsic connectivity networks for salience processing and executive control. *J Neurosci* 2007; 27(9): 2349-2356.

33. Palaniyappan L, Simmonite M, White TP, et al. Neural primacy of the salience processing system in schizophrenia. *Neuron* 2013; 79(4): 814-828.

34. Sridharan D, Levitin DJ, Menon V. A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. *Proc Natl Acad Sci U S A* 2008; 105(34): 12569-12574.

35. Goulden N, Khusnulina A, Davis NJ, et al. The salience network is responsible for switching between the default mode network and the central executive network: replication from DCM. *Neuroimage* 2014; 99: 180-190.

36. Ellison-Wright I, Glahn DC, Laird AR, et al. The anatomy of first-episode and chronic schizophrenia: an anatomical likelihood estimation meta-analysis. *Am J Psychiatry* 2008; 165(8): 1015-1023.

37. Palaniyappan L, Mallikarjun P, Joseph V, et al. Reality distortion is related to the structure of the salience network in schizophrenia. *Psychol Med* 2011; 41(8): 1701-1708.

38. Goodkind M, Eickhoff SB, Oathes DJ, et al. Identification of a Common Neurobiological Substrate for Mental Illness. *JAMA Psychiatry* 2015; 72(4): 305-15.

39. Buckner RL, Andrews-Hanna JR, Schacter DL. The brain's default network: anatomy, function, and relevance to disease. *Ann N Y Acad Sci* 2008; 1124: 1-38.

40. Wotruba D, Michels L, Buechler R, et al. Aberrant Coupling Within and Across the Default Mode, Task-Positive, and Salience Network in Subjects at Risk for Psychosis. *Schizophr Bull* 2013; 40(5): 1095-104.

41. Meltzer HY, Li Z, Kaneda Y, et al. Serotonin receptors: their key role in drugs to treat schizophrenia. *Prog Neuropsychopharmacol Biol Psychiatry* 2003; 27(7): 1159-1172.

42. Abbott CC, Jaramillo A, Wilcox CE, et al. Antipsychotic drug effects in schizophrenia: a review of longitudinal FMRI investigations and neural interpretations. *Curr Med Chem* 2013; 20(3): 428-437.

43. Pressler M, Nopoulos P, Ho BC, et al. Insular cortex abnormalities in schizophrenia: Relationship to symptoms and typical neuroleptic exposure. *Biol Psychiatry* 2005; 57(4): 394-398.

44. Tomelleri L, Jogia J, Perlini C, et al. Brain structural changes associated with chronicity and antipsychotic treatment in schizophrenia. *Eur Neuropsychopharmacol* 2009; 19(12): 835-840.

45. Radua J, Borgwardt S, Crescini A, et al. Multimodal meta-analysis of structural and functional brain changes in first episode psychosis and the effects of antipsychotic medication. *Neurosci Biobehav Rev* 2012; 36(10): 2325-33.

46. Moran LV, Sampath H, Kochunov P, et al. Brain circuits that link schizophrenia to high risk of cigarette smoking. *Schizophr Bull* 2013; 39(6): 1373-1381.

47. Wilkinson D, Halligan P. The relevance of behavioural measures for functional-imaging studies of cognition. *Nat Rev Neurosci* 2004; 5(1): 67-73.

Figure legends

Figure 1. Model space construction. Numbers 1 through 7 indicate left and right visual cortex, right and left striatum, left and right insula, and ACC, respectively. 12 different variations of DCMs were created depending on where the modulation of high-probability rewarding cues exerted its effect (red arrows) on the endogenous connections (black arrows).

Figure 2. A) Bayesian Model Selection results among all 12 DCMs for each group separately. Results are expressed in terms of exceedance probability, the relative probability that this model is more likely than any other of the models tested, given the group data. **B)** Significant group differences in the modulation of right insula to ACC connectivity induced by high-probability rewarding cues. In particular, the modulation of right insula->ACC connectivity was significantly reduced in untreated FEP patients compared with healthy controls, whose connectivity strengths did not differ from those of treated FEP patients.

Figure 3. Negative correlation between the modulation of right insula->ACC connectivity and unusual thought content across untreated ($r=-0.593$, $p=0.012$) but not treated FEP patients ($r=0.127$, $p=0.694$). The x-axis represents patients' unusual thought content as indexed by the BPRS item 11. The y-axis represents the posterior mean (1/s) of the modulation of right insula to ACC connectivity induced by high-probability rewarding cues.

Tables

Table 1: Study population

	HC (n=19)	FEP-treated (n=12)	FEP-untreated (n=17)	ANOVA/ Chi-quadrat	Bonferroni post-hoc
Age in y (SD)	26.42 (4.11)	27.42 (7.93)	24.82 (1.38)	F(2,47)=0.749, P=0.479	/
Gender (n/% female)	9 (47)	6 (50)	4 (24)	$\chi^2(2)$=2.858, p=0.240	/
Handedness (n/% right)	18 (95)	11 (92)	16 (94)	$\chi^2(2)$=0.124, p=0.940	/
MWT (SD)	113 (9.88)	105 (19.63)	103 (12.27)	F(2,47)=2.570, P=0.093	/
BPRS total (SD)	24.53 (1.7)	42.75 (14.75)	51.71 (15.53)	F(2,47)=24.687, P<0.0001	HC<FEP-treated, HC<FEP-untreated
Suspiciousness (BPRS 9)	1.00 (0.00)	3.00 (1.71)	3.47 (1.38)	F(2,47)=22.059, P<0.0001	HC<FEP-treated, HC<FEP-untreated
Hallucinations (BPRS 10)	1.00 (0.00)	2.42 (2.15)	3.53 (2.0)	F(2,47)=11.781, P<0.0001	HC<FEP-treated, HC<FEP-untreated
Unusual thought content (BPRS 11)	1.00 (0.00)	3.25 (1.87)	3.71 (1.9)	F(2,47)=17.431, P<0.0001	HC<FEP-treated, HC<FEP-untreated

Conceptual disorganization (BPRS 15)	1.00 (0.00)	2.08 (1.31)	2.06 (1.30)	F(2,47)=6.561, P=0.003	HC<FEP-treated, HC<FEP-untreated
SANS total (SD)	0.00 (0.0)	17.08 (16.21)	21.82 (14.88)	F(2,47)=16.396, P<0.0001	HC<FEP-treated, HC<FEP-untreated
GAF total (SD)	88.63 (4.52)	63.50 (9.65)	53.06 (17.95)	F(2,47)=41.171, P<0.0001	HC<FEP-treated, HC<FEP-untreated
Antidepressants (n/% user)	0 (0)	3 (25)	4 (24)	$\chi^2(2)$=5.381, p=0.068	/
Cannabis (n/% user)	4 (21)	1 (8)	7 (41)	$\chi^2(2)$=4.308, p=0.116	/
Cigarettes per day (n)	2.47 (5.834)	9.42 (8.207)	10.88 (11.522)	F(2,47)=4.618, P=0.015	HC<FEP-untreated

Abbreviations: HC, healthy controls, FEP-treated, antipsychotic-treated patients with first-episode psychosis; FEP-untreated, untreated patients with first-episode psychosis; MWT, “Mehrfachwahl-Wortschatz-Intelligenz-Test”, a multiple choice-vocabulary-intelligence test; BPRS, brief psychiatric rating scale; GAF, Global Assessment of Functioning; SANS; Scale for the Assessment of Negative Symptoms. Numbers in brackets represent degrees of freedom.

Table 2. DCM parameters from the best fitting model

	HC (N=19)	FEP-TREATED (N=12)	FEP-UNTREATED (N=17)
Right insula to ACC connectivity[a]	0.1867 (0.3064)[b]	0.0651 (0.1567)	-0.0642 (0.2877)
Left insula to ACC connectivity	0.0976 (0.3138)	0.1182 (0.2234)	0.1879 (0.4933)

Mean and SD reflecting the modulatory effect induced by high-probability rewarding cues. [a]$F(2,47)= 3.823$, P=.029 for analysis of variance, and P=.025 (healthy controls greater than untreated FEP patients) for Bonferroni-corrected post hoc t test. [b]Significant t tests within each group compared with zero ($p< 0.05$).