

Glutamate and Dysconnection in the Salience Network: Neurochemical, Effective-connectivity, and Computational Evidence in Schizophrenia

Running title: Three-level Model of Dysconnection

Roberto Limongi¹, Peter Jeon¹, Michael Mackinley¹, Tushar Das¹, Kara Dempster¹, Jean Théberge^{1,2,3}, Robert Bartha¹, Dickson Wong¹, Lena Palaniyappan¹

¹University of Western Ontario, London, Ontario

² St. Joseph's Health Care London, London, Ontario

³ Lawson Health Research Institute, London, Ontario

Correspondence concerning this article should be addressed to Roberto Limongi, Robarts Research Institute, 1151 Richmond St. N, UWO, London, Ontario, Canada, N6A 5B7. Phone: +1 226 977 6527. E-mail: rlimongi@uwo.ca

Abstract

In the dysconnection hypothesis, psychosis is caused by NMDA hypofunction resulting in aberrant network connectivity. Combining a cognitive-control task, functional magnetic resonance spectroscopy, and functional magnetic resonance imaging, we tested this hypothesis in the salience network of 20 first-episode psychosis (FEP) and 20 healthy control (HC) subjects. Across groups, glutamate concentration in the dorsal anterior cingulate cortex (dACC) was associated with higher and lower inhibitory connectivity in the dACC and in the anterior insula (AI) respectively. Crucially, glutamate concentration correlated negatively with the inhibitory influence on the excitatory neuronal population in the dACC of FEP subjects. Furthermore, aberrant computational parameters of the cognitive-control task performance were associated with aberrant inhibitory connections. Finally, the strength of connections from the dACC to the AI correlated negatively with severity of social withdrawal. These findings support a link between glutamate-mediated cortical disinhibition, deficits in effective connectivity, and computational performance in psychosis.

Keywords: dysconnection hypothesis; glutamate hypothesis; effective connectivity; dynamic causal models; predictive coding; schizophrenia

Introduction

The glutamate hypothesis (1-4) has been a central focus of interest in the study of the psychopathology of schizophrenia for more than 20 years. It states that dysfunction of glutamatergic neurotransmission is associated with signature negative symptoms of schizophrenia such as poor executive control and social withdrawal. The dysfunction of glutamatergic neurotransmission is likely caused by NMDA-receptor hypofunction (i.e., glutamate hypofunction) in inhibitory GABAergic interneurons which would lead to an increase in the synaptic gain of excitatory neurons (the disinhibition hypothesis, 5, 6). The abnormal increase of synaptic gain is also conceptualized as a disruption of the excitation-inhibition balance which, as described below, rests at the core of the dysconnection hypothesis (7) within the theoretical framework of “the Bayesian brain” (8).

The dysconnection hypothesis (c.f., disconnection hypothesis, 9)¹ states that the psychopathology of schizophrenia should be studied at three levels of analysis: neurochemical, effective-connectivity (network-connectivity), and computational levels. At the computational level, the dysconnection hypothesis states that a *suboptimal* (e.g., schizophrenia) *Bayesian brain* (10) would overly afford confidence or precision (i.e., inverse variance) to its predictions about the external stimuli and would overestimate the reliability of the prediction errors (PE), leading to false inferences (e.g., hallucinations) and cognitive failures (e.g., cognitive control, 11). At the effective-connectivity level, a predictive coding algorithm (12), namely, hierarchical message passing between lower and higher cortical levels would be altered in terms of aberrant backward

¹ Here, “dysconnection” means abnormal functional integration or effective connectivity. This meaning differs from the meaning of “disconnection” which refers to collapsed or disintegrated cognitive functions.

and forward interregional connectivity strength (13). Crucially, at the neurochemical level the aberrant connectivity strength would depend on increased synaptic gain in deep and superficial pyramidal cells, reflecting a decrease in the strength of intrinsic inhibitory connections, driven by NMDA hypofunction (7, 14). Therefore, the dysconnection hypothesis takes the glutamate hypothesis one step further by postulating that the disruption of excitation-inhibition balance makes itself evident at effective-connectivity level of analysis. In this work, we provide evidence in support of this hypothesis.

We provide evidence of the association between glutamate and dysconnection within the context of neurochemical, effective-connectivity, and computational levels of analysis. We will do this by studying the relationship between ^1H -MRS glutamate in the dorsal anterior cingulate cortex ($[\text{Glu}]_{\text{dACC}}$), the effective connectivity within the salience network, and cognitive-control dysfunction and negative symptoms in schizophrenia. Specifically, at the computational level we compare the performance of first-episode psychosis (FEP) and healthy control (HC) subjects in the Stroop task (15), which reliably engages the two nodes of the salience network (16, 17).

Suboptimal Stroop computations in FEP are reflected in long reaction times and low response accuracy (18, 19). We show that these suboptimal computations are explained in terms of a drift-diffusion model as a specific case of a Bayesian decision making (20). At the effective-connectivity level, we show that the computational parameter associated with aberrant predictions in FEP maps onto forward and backward connections. Ultimately, we demonstrate at the neurochemical level that the dysconnection within the salience network is driven by $[\text{Glu}]_{\text{dACC}}$. As we detail below, the salience network is an appropriate anatomical and functional motif (c.f., 21) to evaluate the relationship between all three levels of analysis.

The dACC is functionally specialized in conflict monitoring during the Stroop task (22). In the salience network, the right anterior cingulate cortex (dACC) is anatomically connected to the right anterior cingulate insula (AI) (23-25). Crucially, psychosis is associated with consistent structural deficits (26) as well as resting-state functional dysconnectivity in the salience network (27, 28). Given that the right AI is particularly sensitive to descending afferents from the right dACC (29), we expected the subtle variations of excitation-inhibition balance resulting from the putative glutamatergic abnormalities within the dACC to induce dysconnectivity in the salience network, affecting both intrinsic inhibitory connections of the dACC and the extrinsic connection with the AI. We anticipated the aberrant effective connectivity to account for aberrant prior beliefs and overly precise PE during cognitive conflict resolution and to the burden of negative symptoms of schizophrenia.

Methods and Materials

Participants

Twenty FEP subjects (8 female; age male M = 22.5, SD = 3.45; age female M = 22.37, SD = 3.07; age range [17 - 28]) and 20 HC subjects (9 female, age male M = 21.64, SD = 4.27; female M = 21.67, SD = 3.04; age range [16 - 29]) participated in the study. No difference in age was detected across groups ($p > |t| = 0.47$). Subjects were recruited from the “Prevention and Early Intervention Program for Psychosis” in London, Ontario. Criteria for inclusion in the FEP group included (i) first clinical presentation with psychotic symptoms and (ii) Diagnostic and Statistical Manual of Mental Disorders (5th Edition) criteria A for schizophrenia satisfied. HC subjects did not report a personal history of mental illness or family history of psychotic disorders. The subjects’ consent was obtained according to the Declaration of Helsinki, and approval for the study was obtained from the University Human Ethics Committee for Health Sciences at the University of Western Ontario. Subjects were assessed for positive and negative symptoms of schizophrenia using the eight items of the Positive and Negative Syndrome Scale (PANSS-8, 30). All relevant demographic data are provided in the Supplemental Information. As detailed below, inside an MRI scanner, subjects underwent resting-state ^1H -MRS, resting-state fMRI, and performed a color version of the Stroop task.

^1H -MRS

We measured resting-state ^1H -MRS during a four-minute block. The measurement was made in a Siemens MAGNETOM 7-Tesla MRI scanner using an 8-channel transmit/ 32-channel receive, head-only, radiofrequency coil at the Centre for Functional and Metabolic Mapping at the University of Western Ontario. A 2.0 x 2.0 x 2.0 cm (8cm³) ^1H -MRS voxel was placed on the

dACC where activation was expected based on our previous work (19) using the Stroop task (MNI: X = 1, Y = 16, Z = 38, Fig. 1). To locate the voxel, we used a two-dimensional anatomical imaging sequence in the sagittal direction (37 slices, TR = 8000 ms, TE = 70 ms, flip-angle (α) = 120°, thickness = 3.5 mm, field of view = 240×191 mm). Conventional ^1H -MRS sequences use a short or “short-as-possible” echo time to minimize the T_2 relaxation and J-coupling effects. However, a recent study by Wong, Schranz (31) proposed that a longer echo time for the semi-LASER sequence would improve glutamate measurement in the human brain. Therefore, one long echo time (100 ms) dataset was acquired using a semi-LASER ^1H -MRS pulse sequence (repetition time = 7500 ms). A 32 channel-combined, water-suppressed spectral average was acquired using the semi-LASER ^1H -MRS pulse sequence (repetition time = 7500 ms, echo time = 100 ms, N = 32). Water suppression was achieved using the VAPOR preparation sequence and a water-unsuppressed spectrum was also acquired for spectral post-processing. Spectral post-processing is described in the Supplemental Information.

Resting-state fMRI

Using the same scanner as in the ^1H -MRS measurement, we acquired 360 whole-brain functional images. A gradient echo-planar-imaging sequence was used with phase-encoding direction = A > > P, repetition time = 1000 ms, echo time = 20 ms, flip angle = 30 deg, field of view = 208 mm, field of view phase = 100 %, voxel dimension = 2 mm isotropic, slice thickness = 2 mm, multi-band acceleration factor = 3, acquisition time = 6 min. 26 s, and number of slices = 63 (interleaved slice order). Subjects were instructed to lie with their eyes open.

Stroop task

We implemented the same Stroop task we used in Taylor, Neufeld (19). By pressing a key in a four-key response pad, subjects identified as quickly and as accurate as possible the color of the ink of a string of letters (on a gray background) representing either the name of a color (yellow, green, blue, or red) or a series of ‘Xs’ (‘XXXX’) projected on screen visible via a mirror located in the scanner’s coil. The task comprised four conditions: word-only (a word in white ink saying a color name, e.g., the word ‘yellow’ written in white ink), color-only (an ‘XXXX’ string with the color ink), congruent (the ink color matching the meaning of the word, e.g., the word ‘red’ in red ink), and incongruent (the ink color differing from the meaning of the word, e.g., the word ‘red’ written in blue ink). Subjects performed 20 trials per condition. They were allotted 2 s to respond with an interstimulus trial interval of 1 s during which a fixation cross was visible. Outside the scanner, each participant rehearsed the task until reaching 80 % accuracy — collapsed across all four conditions.

Computational model of the Stroop performance

Based on our previous works (19), we were particularly interested in the incongruent condition in which we expected lower accuracy and longer reaction time in FEP than in HC. To verify this, we regressed accuracy and reaction time (separately) against condition, group, and the Condition \times Group interaction via a mixed-effects linear model (32). Subjects were included as random effects.

We fit a hierarchical drift-diffusion model to the reaction time and accuracy data. In the hierarchical drift-diffusion model, subjects accumulate information and trigger a response after reaching an accumulation threshold. Formally, the hierarchical drift-diffusion model comprises

four basic parameters representing the accumulation threshold, the starting point of the accumulation process, the accumulation (or drift) rate, and the non-decision processes such as stimulus encoding and motor execution (i.e., sensorimotor delay).

Predictive coding maps elegantly onto the hierarchical drift-diffusion model (11, 20, 33). Evidence accumulation corresponds to the accumulation of presynaptic afferent activity from neuronal populations encoding PE (i.e., superficial pyramidal cells), the drift rate represents the precision of the ascending PE, and the starting point represents the prior beliefs. Formally, corrected prior beliefs (i.e., a change in the starting point parameter) correlates with a steeper slope of the accumulation process (i.e., a larger absolute value of the drift rate indicating more precise PE).

Using Bayesian and Markov chain Monte Carlo methods (34), we estimated the parameters of the hierarchical drift diffusion model from subjects of each group separately. We assumed that subjects of a particular group were alike. Therefore, their parameters were constrained by their group parameters. This also allowed us to estimate between-groups difference without over or underestimating the true group-level variance (35). Prior distributions were informative as per the default option used in the HDDM package (34). The chain length was 200,000. The number of burn-in iterations was 2000, and the chains were generated with thinning = 20. We report the proportion of the estimate's posterior distribution (PP) that differs from zero. To show differences between groups, we report the proportion of the posteriors in which the parameter estimate for one group differs from the other. Specifically, we expected larger drift rate (i.e., aberrant precision of PE) and larger starting point (aberrant prior beliefs) in the FEP group than in the HC. We did not have a-priori expectations about differences in the decision threshold and

in non-decision processes. However, we report the differences in these parameters for completeness and post-hoc interpretation.

Effective connectivity

Preprocessing and general linear model

Functional images were realigned, normalized to the Montreal Neurological Institute space, and spatially smoothed using a 4 mm (full width at half maximum) Gaussian kernel. We fit a general linear model to the images and included the six head movement parameters and the time series corresponding to the white matter and cerebrospinal fluid as regressors. In addition, we included a cosine basis set with frequencies ranging from 0.0078 to 0.1 Hz (36). Images were also high-pass filtered to remove slow frequency drifts (< 0.0078 Hz). By specifying an F-contrast across the basis set, we identified regions with blood oxygen level fluctuations within the range of frequencies specified above. On the *F*-contrast, a sphere (8-mm radius) was centered on the Montreal Neurological Institute coordinates corresponding to the dACC (the same coordinates that formed the centroid of the ¹H-MRS voxel for the [Glu]_{dACC} measurements) and to the AI (X = 38, Y = 20, Z = -4, Fig. 1) —we selected these coordinates based on our previous results (37). At the peak voxel on each region of interest, we defined a sphere (8-mm radius) and extracted the timeseries (principal eigenvariate) summarizing the activity within the sphere.

Spectral dynamic causal model (DCM)

We estimated the resting-state effective connectivity within the salience network by fitting a two-state spectral DCM (38) to the resting-state fMRI data (39). Two-state DCM assumes two populations of neurons within a region: excitatory and inhibitory. For stability of the model, each sub-population comprises self-inhibition connections (which are fixed parameters). Crucially,

two free parameters are fit to the fMRI data: interregional excitatory-to-excitatory connections and within-region inhibitory-to-excitatory connections (Fig. 1). Since we aimed at demonstrating the relationship between the effect of $[\text{Glu}]_{\text{dACC}}$ on inhibitory-to-excitatory connections and the ensuing consequences in the entire network, this two-state model was sufficient to test our hypothesis despite not capturing all the circuitry constraints of cortical columns (c.f., a four-neuronal-population model, 40). For each participant, a fully connected DCM, with no exogenous inputs was specified and inverted using spectral DCM of SPM12 (<https://www.fil.ion.ucl.ac.uk/spm/software/spm12/>).

After inverting subjects' DCMs, we estimated a series of parametric empirical Bayes models (41, 42) aiming to test, at a group level, our hypothesis against five alternative hypotheses. We fit six (general linear) models to the posterior estimates of effective connectivity from each subject. Our main hypothesis stated that $[\text{Glu}]_{\text{dACC}}$ hypofunction would cause aberrant effective connectivity in the salience network which would account for aberrant prior beliefs and aberrant precision of PE. We tested this hypothesis by comparing the evidence supporting a “three-level” model with the evidence supporting five alternative models. The three-level (neurochemical, effective-connectivity, and computational) model comprised the following four covariates: (i) group, (ii) $[\text{Glu}]_{\text{dACC}}$, (iii) precision of PE, and (iv) prior beliefs. Crucially, the model included the interactions between $[\text{Glu}]_{\text{dACC}}$, prior beliefs, and precision of PE with Group. Each covariate was mean centered, and we used effect coding for groups (HC = -1, FEP = 1). All but the “behavioral model” (see below) were reduced versions of the three-level model.

A “no-glutamate” model represented the hypothesis that glutamate did not affect the salience network, the model did not include the main effect of $[\text{Glu}]_{\text{dACC}}$ and its interaction with group. If $[\text{Glu}]_{\text{dACC}}$ did not exert an effect on the effective connectivity of the salience network, the no-

glutamate model would perform better than the three-level model. Similarly, if the computational parameters representing prior beliefs and precision of PE were not affected by the aberrant connectivity within the salience network, a model without these effects would perform better than the three-level model. We fit this alternative model and referred to it as the “no-computational” model. Alternatively, it could be that a model including the observed behavioral responses would perform better than a model comprising the computational parameters. We fit this model by substituting (in the three-level model) the behavioral responses and their interactions with group for the hierarchical drift-diffusion model parameters. We referred to this model as the “behavioral model”. For completeness, we fit a “group model” and a “null model”. The group model represented the hypothesis that changes in the salience network would be caused only by differences between groups. This model comprised only the main effect of group. The null model included only a single column of “ones” in the design matrix representing the hypothesis that neither groups nor the covariates were associated with the effective connectivity of the network.

To adjudicate between models, we performed Bayesian model comparison. At a group level, we first report the mean posterior estimate (collapsed across groups) of the strength of each connection along with the PP of models with each parameter, relative to models without. To evaluate our ‘three level’ hypothesis, we report the effect sizes; i.e., parameters β of the between-subject parametric empirical Bayesian model associated with each main effect and their interactions with group.

Goodness of fit, convergence, and acquisition quality assessments

In the hierarchical drift-diffusion model, goodness of fit was assessed via posterior predictive checks (43). Convergence was assessed by computing the R-hat statistic (44). Proportion of

variance accounted for by the dynamic causal models are reported in the Supplemental Information (Table S1). Finally, to assess the quality of our acquired ^1H -MRS data, signal-to-noise ratio (SNR_{NAA}) was calculated by dividing the amplitude of the NAA CH₃ peak in the frequency domain by the standard deviation of the noise in the last 32 most upfield points of the spectrum. The linewidth was also calculated using water-unsuppressed spectra and measuring the FWHM of the water peak in Hertz (Hz).

Data availability

Data supporting the findings of this study are available at https://uwoca-my.sharepoint.com/:f/g/personal/rlimongi_uwo_ca/EmDE6zsduKpJinRzEhWoyz8B2OX83970WxL8KnG0dUj6IA?e=0ZxdO6.

Results

[Glu]_{dACC} does not differ across groups

Glutamate MR Spectroscopy assessment could not be performed in one FEP subject. Data from this subject were not analyzed. A frequentist two-sample t-test did not reveal statistically significant difference between the mean [Glu]_{dACC} in the FEP group ($M = 6.90$ mM, $SD = 1.35$) relative to the HC group ($M = 6.38$ mM, $SD = 1.3$); $t_{(36.72)} = 1.21$, $p = 0.88$ (Fig. 1).

[Glu]_{dACC} does not account for between-groups difference in (Stroop) computational performance

Both FEP and HC subjects performed the task as instructed. Table 1 shows the summary statistics. The mixed-effects models confirmed the expected behavioral results (Table 2). Across conditions, FEP subjects performed worse than HC subjects. Planned t-tests confirmed that they were less accurate ($t_{(502)} = 3.71$, $p < 0.001$, Cohen's $d = 0.28$) and took longer ($t_{(674)} = 6.99$, $p < 0.001$, Cohen's $d = 0.52$) when resolving cognitive conflicts in the incongruent condition.

The hierarchical drift-diffusion model showed that all parameter estimates in both groups differed from 0 (PP = 1.0). Subject-wise estimates are provided in the Supplemental Information (Table S2). Fig. 2 shows that the drift-rate parameter in the HC group ($M = 2.83$, $SD = 0.18$) was larger than in the FEP group ($M = 1.83$, $SD = 0.2$), proportion of the group's posteriors in which the parameters differ (ΔPP) = 0.99. Furthermore, the starting point parameter in the FEP group was larger (i.e., closer to the decision boundary, $M = 0.44$, $SD = 0.03$) than in the HC group ($M = 0.32$, $SD = 0.03$), $\Delta PP = 0.99$. The decision threshold was lower in the FEP group ($M = 2.06$, $SD = 0.32$) than in the HC group ($M = 2.52$, $SD = 0.09$), $\Delta PP = 0.96$. Finally, the non-decision

processes took longer in the FEP group ($M = 0.64$, $SD = 0.11$) than in the HC group ($M = 0.46$, $SD = 0.09$), $\Delta PP = 0.99$.

To investigate whether $[Glu]_{dACC}$ could explain between-groups differences in precision of PE and prior beliefs, we regressed the drift-diffusion model parameters on group, $[Glu]_{dACC}$, and the $[Glu]_{dACC} \times$ Group interaction. Regarding prior beliefs, we found main effect of group, $\beta = -0.06$, $SE = 0.0002$, $t(35) = -350$, $p < .0001$, 95% CI [-0.059, -0.058] (congruent with the hierarchical drift-diffusion model results). However, we found no evidence for an effect of $[Glu]_{dACC}$, $\beta = 0.00006$, $SE = 0.0001$, $t(35) = 0.51$, $p = .61$, 95% CI [-0.0002, 0.0003], nor $[Glu]_{dACC} \times$ Group interaction, $\beta = -0.00009$, $SE = 0.0001$, $t(35) = -0.71$, $p = .48$, 95% CI [-0.0003, 0.0002].

Similarly, regarding the precision of PE we found main effect of group, $\beta = 0.5$, $SE = 0.05$, $t(35) = 9.11$, $p < .0001$, 95% CI [0.39, 0.62] —also congruent with the hierarchical drift-diffusion model results— and we found no evidence for an effect of $[Glu]_{dACC}$, $\beta = -0.009$, $SE = 0.04$, $t(35) = -0.22$, $p = .83$, 95% CI [-0.09, 0.08] nor $[Glu]_{dACC} \times$ Group interaction, $\beta = 0.01$, $SE = 0.04$, $t(35) = 0.31$, $p = .76$, 95% CI [-0.07, 0.098].

[Glu]_{dACC} affects intrinsic inhibitory connections in the salience network

Bayesian model comparison revealed that the three-level model outperformed all the alternative models (Fig. 3). The mean estimates across groups revealed that in resting state, there was intrinsic inhibition in both regions (dACC, $M = 2.25$, $PP = 1.0$; AI, $M = 2.6$, $PP = 1.0$) and an ensuing attenuation in the strength of excitatory connections between regions (dACC→AI, $M = -0.53$, $PP = 1$; AI→dACC, $M = -0.43$, $PP = 1.0$).

Fig. 4 shows no main effect of group on network parameters. However, main effect of $[Glu]_{dACC}$ on self-connections in both the dACC and the AI was detected. The strength of intrinsic

inhibition increased in the dACC ($\beta = 0.09$, PP = 0.98) and decreased in the AI ($\beta = -0.07$, PP = 0.96). Across groups, we did not find effect of prior beliefs on extrinsic connections. However, we found a main effect of the precision of PE on inhibitory connections. Across groups, the strength of inhibitory connections increased with the precision of PE (dACC, $\beta = 0.9$, PP = 1; AI, $\beta = 0.85$, PP = 1).

[Glu]_{dACC} correlates negatively with intrinsic inhibitory influence in the dACC of FEP subjects, accounting for aberrant prior beliefs and aberrant precision of PE

Crucially, Fig. 5 shows that there was a Group \times [Glu]_{dACC} interaction in inhibitory connections of the dACC. Connections were weaker in FEP than in HC ($\beta = -0.11$, PP = 1.0). Furthermore, there was a Group \times Prior-beliefs interaction. The effect of inhibitory connections on prior beliefs was stronger in FEP than in HC (dACC, $\beta = 39.52$, PP = 1) (AI, $\beta = 44.44$, PP = 1). Surprisingly, the effect of excitatory connections on prior beliefs was weaker in FEP than in HC (dACC \rightarrow AI, $\beta = -9.13$, PP = 1; AI \rightarrow dACC, $\beta = -7.49$, PP = 1). Finally, we found Group \times PE precision interaction in inhibitory connections. The effect of these connections on the precision of PE was weaker in FEP (dACC, $\beta = -1.56$, PP = 1; AI, $\beta = -1.08$, PP = 1). In summary, [Glu]_{dACC} differentially affected the effective connectivity of the salience network, and the connectivity of the network was differentially associated with the parameter estimates of the hierarchical drift diffusion model in the presence of FEP.

Connectivity strength correlates negatively with severity of social withdrawal

We assessed the association between the severity of negative symptoms (blunted affect, social withdrawal, and lack of spontaneity) and the strength of connections independently (i.e., three comparisons per connection) in the FEP group. Only the effect of backward connections on

social withdrawal survived Bonferroni correction for the number of negative symptoms (three symptoms, $P = 0.017$). Severity of social withdrawal increased as the connectivity strength of dACC→AI connections decreased, $\beta = -15.59$, $SE = 5.61$, $t_{(16)} = 2.78$, $p = 0.013$, 95% CI [27.43, 3.75] (Fig. 6).

Discussion

By using resting-state ^1H -MRS, DCM of resting-state fMRI, and hierarchical drift-diffusion modeling, we have demonstrated that low computational performance during cognitive-conflict resolution in schizophrenia is explained by aberrant resting-state effective connectivity (dysconnection) in the salience network. This dysconnection was associated with glutamate hypofunction in the dACC, as proven by the fact that a model without the effect of $[\text{Glu}]_{\text{dACC}}$ had low probability of having produced the fMRI data. $[\text{Glu}]_{\text{dACC}}$ was associated with opposite intrinsic connectivity strength in both the dACC and the AI, indicating that $[\text{Glu}]_{\text{dACC}}$ affected the entire network and confirming the sensitivity of the AI to subtle changes in the dACC excitation-inhibition balance. Crucially, in the FEP group the inhibitory influence on the excitatory population of the dACC decreased as a function of glutamate concentration. This finding is the first imaging evidence directly linking the glutamate hypofunction to the cortical disinhibition hypothesis in schizophrenia.

Using stochastic DCM, Bastos-Leite, Ridgway (45) found decreased intrinsic inhibition and decreased extrinsic connections in the default network of schizophrenia subjects. Similarly, using parametric empirical Bayes and DCM in a sample of schizophrenia subjects, Zhou, Zeidman (46) recently found decreased intrinsic inhibition and decreased extrinsic connectivity within the dorsal attention network. They speculated that both glutamate hypofunction and aberrant precision of prediction errors would be associated with this sort of dysconnectivity in large-scale networks. In the context of the salience network, we confirmed their hypothesis. Therefore, the fact that three independent groups have shown the same relationship between decreased intrinsic and extrinsic connections in the default, attention, and salience networks provide strong support to the dysconnection hypothesis of schizophrenia. Crucially, our results point to glutamate

hypofunction and aberrant computations of sensory and prior precision as critical causes of dysconnection and are in line with recently reported independent preliminary data showing that the effect of glutamate hypofunction should be observed at the network level (47).

These results also demonstrate that establishing the relationship between glutamate hypofunction, the disinhibition hypothesis, and the pathophysiology of schizophrenia requires the integration of the neurochemical, effective-connectivity, and computational levels of analysis. This is supported by the fact that Bayesian model comparison afforded low probability to any model not comprising all three levels ($PP \approx 0.12$, summed over all but the three-level model). Specifically, parameter estimates of resting-state aberrant connectivity (driven by resting-state $[Glu]_{dACC}$) in the salience network accounted for computational parameters of cognitive dysfunction in schizophrenia. Furthermore, an increase in the effect of the inhibitory population on the excitatory population in both regions explains the precision subjects afforded to ascending information, in line with the known role of GABA interneurons on pyramidal circuits (48).

The three-level model also explains the compensatory effect of aberrant prior beliefs to aberrant precision of ascending information in FEP. At the computational level, FEP subjects needed to accumulate less information than HC (i.e., lower decision threshold), relied more on their prior beliefs (i.e., at the beginning of each trial they were closer to the decision boundary) than on sensory information, and were less cautious when resolving cognitive conflicts —they tended to jump to conclusions (49). At a group level, we did not find evidence for a larger precision of PE in the FEP group than in the HC group. On the contrary, the HC group showed a steeper drift rate. However, within the FEP group, over affordance of precision to PE was associated with a decrease in the strength of inhibitory connections —in both regions (Fig. 5-B). Crucially, an

increase in the inhibitory activity was associated with an increase in prior beliefs (Fig. 5-C), suggesting (at the effective-connectivity level) that more precise prior beliefs compensate for the aberrant precision of PE —as predicted by the dysconnection hypothesis (7).

A three-level model could provide alternative explanations for previously reported findings. Specifically, Taylor, Neufeld (19) found no relationship between $[\text{Glu}]_{\text{dACC}}$ and deficits in Stroop performance of FEP subjects. One possible reason for this finding might be the lack of both a computational model (e.g., hierarchical drift-diffusion model) and a network model (e.g., DCM). This is supported by the fact that we found no direct association between $[\text{Glu}]_{\text{dACC}}$ on either behavioral performance (i.e., reaction time and accuracy) or hierarchical drift-diffusion model parameters. However, we demonstrated that glutamate hypofunction affected the whole salience network and the aberrant effective connectivity of this network did affect the computational performance.

The advantage of a “three-level approach” is also supported by recent findings of Shaw, Knight (14). They used a four-neuronal-population DCM to model inhibitory (GABA) connections to superficial pyramidal cells. Their DCM parameters were associated with the behavioral performance. Our Bayesian model comparison showed that a model including behavioral measurements rather than the computational parameters underperformed the three-level model —which included the computational parameters. The findings of Shaw, Knight (14), however, are in concert with our idea that establishing a relationship between synaptic dysfunction and behavior (i.e., symptoms) requires modeling effective connectivity (at least at the regional scale) that would be disrupted by the synaptic dysfunction. More in general, along with this recent work, the current results suggest that symptoms of schizophrenia emerge from the interaction between the three levels of analysis.

Our results and the idea of interactive emergence allow us to provide an alternative interpretation to recent findings of Valton, Karvelis (50), who reported a computational model of Bayesian learning in schizophrenia and HC. Unlike our work, their study did not find differences in the precision of prior beliefs or in the precision of sensory evidence across groups. However, the fact that they did not include the effective-connectivity level of analysis in their experimental design left open the possibility of a Group \times Precision interaction on the effective connectivity. Our results regarding $[\text{Glu}]_{\text{dACC}}$ support this alternative explanation. We did not find main effect of group on $[\text{Glu}]_{\text{dACC}}$ (c.f., no main effect of group on precision in 50). However, the effect of $[\text{Glu}]_{\text{dACC}}$ on inhibitory connections in the dACC varied with group.

Valton, Karvelis (50) also proposed that prior beliefs in schizophrenia subjects might not influence behavioral performance in the absence of sensory stimulation. During resting state fMRI, our FEP subjects did not expect sensory stimulation. Interestingly, the weaker the connectivity strength in descending connections (representing prior beliefs) were in resting state the more biased were their responses towards the decision threshold. Heuristically, this could be interpreted as an increase in the confidence FEP subjects afforded to their prior beliefs during task performance, caused by a decreased in the strength of backward connections. Therefore, it is possible that the effect that prior beliefs might exert on behavioral performance might depend not only on the sensory stimulation (as suggested by 50) but also on the resting-state connectivity strength of descending connections which in hierarchical message passing represent prior beliefs.

It is intriguing that the relationship between the strength of extrinsic connections and the computational parameters of precision-weighted PE (and prior beliefs) does not conform to the assumption of an asymmetric flow of information under predictive coding. Canonical formulations of predictive coding would predict an increase in the strength of ascending

connections associated with aberrant precision of PE and an increase in the strength of descending connections associated with compensatory aberrant prior beliefs. However, we found that the extrinsic connectivity strength did not reflect the precision of PE, whereas weak bidirectional connections correlated with prior beliefs. These data provide initial evidence of “bidirectional correlates of effective connectivity strength with prior beliefs” in hierarchical message passing with, as we elaborate upon below, implications for the underlying predictive coding theory of brain function and its translational corollary, the dysconnection hypothesis.

Initial works on predictive coding proposed an asymmetrical functional architecture for hierarchical message passing whereby prior beliefs would be conveyed by descending connections and PE would be conveyed by ascending connections. Not by coincidence, the connectivity asymmetry was motivated on the anatomical asymmetry widely described in early sensory areas (13). However, this scheme is not congruent with the effect of weak bidirectional connections on prior beliefs that our findings suggest.

A possible explanation of our findings is that the parameter estimates of the current DCM could be reflecting the activity of another type of excitatory inputs —for example, Von Economo neurons. DCMs do not differentiate this type of excitatory influence from others. This is a limitation of the current study that could be resolved in the future by incorporating additional excitatory neuronal populations and parameters of interlaminar connectivity (40).

We posit that the salience network might defy the canonical asymmetry of predictive coding due to the presence of Von Economo neurons, emerging as a network functionally specialized in the bidirectional processing of prior beliefs. Indeed, the “unique” role of Von Economo neurons in predictive coding has been recently suggested (51). However, asymmetric message passing does not contradict the principles of functional specialization and functional integration. For example,

classic theories of working memory have been recently updated by integrating the principle of hierarchical message passing with the widely accepted functional specialization of the prefrontal cortex (52). Working memory, in the prefrontal cortex, has been recently regarded as a locus for the generation and maintenance of prior beliefs (53) which would be passed along to lower areas in the cortical hierarchy (54). Therefore, it is possible that the prefrontal cortex generates prior beliefs and deploys such beliefs downstream to the salience network. As a hub, the salience network would then functionally specialize in broadcasting prior beliefs (i.e., predictions) to lower sensory areas and interoceptive channels. The salience network would draw upon *bidirectional* circulation of prior beliefs to execute this functional integration.

The possibility of a functionally specialized two-node network for the processing of prior beliefs is further supported both by previous works on social isolation and suicidal behavior. For example, a post-mortem study of schizophrenia subjects that died by suicide showed a larger density of Von Economo neurons in the cingulate cortex of these subjects, compared with schizophrenia subjects who died due to other causes (55). Interestingly, social isolation (i.e., social withdrawal) is strongly related with both the (aberrant) salience network (55, 56) and suicidal behavior (57). These previous works are in consonance with the relationship between dACC→AI connectivity and severity of social withdrawal we found in this work.

In conclusion, this work provides evidence to the hypothesis that the glutamate hypofunction relies on the disinhibition hypothesis and manifests itself in the effective connectivity of the salience network. The three levels of analysis provide compelling explanation to deficits in cognitive control and negative symptoms in schizophrenia. Finally, a bidirectional functional architecture of prior beliefs of the salience network could be the key to understand both its coordinating role with the central executive network (58) and its cardinal role in psychosis (26).

Funding

This study was funded by CIHR Foundation Grant (375104/2017) to LP; Schulich School of Medicine Clinical Investigator Fellowship to KD; AMOSO Opportunities fund to LP; Bucke Family Fund to LP; BrainSCAN to RL; Grad student salary support of PJ by NSERC Discovery Grant (No. RGPIN2016-05055) to JT; Canada Graduate Scholarship to KD. Data acquisition was supported by the Canada First Excellence Research Fund to BrainSCAN, Western University (Imaging Core); Innovation fund for Academic Medical Organization of Southwest Ontario; Bucke Family Fund, The Chrysalis Foundation and The Arcangelo Rea Family Foundation (London, Ontario).

Disclosures

LP reports personal fees from Otsuka Canada, SPMM Course Limited, UK, Canadian Psychiatric Association; book royalties from Oxford University Press; investigator-initiated educational grants from Janssen Canada, Sunovion and Otsuka Canada outside the submitted work. All other authors report no relevant conflicts.

References

1. Coyle JT (1996): The glutamatergic dysfunction hypothesis for schizophrenia. *Harv Rev Psychiatr.* 3:241-253.
2. Olney JW, Farber NB (1995): Glutamate receptor dysfunction and schizophrenia. *Archives of general psychiatry.* 52:998-1007.
3. Snyder MA, Gao WJ (2013): NMDA hypofunction as a convergence point for progression and symptoms of schizophrenia. *Frontiers in cellular neuroscience.* 7:31.

4. Snyder MA, Gao WJ (2019): NMDA receptor hypofunction for schizophrenia revisited: Perspectives from epigenetic mechanisms. *Schizophr Res.*
5. Homayoun H, Moghaddam B (2007): NMDA Receptor Hypofunction Produces Opposite Effects on Prefrontal Cortex Interneurons and Pyramidal Neurons. *The Journal of Neuroscience*. 27:11496.
6. Gonzalez-Burgos G, Lewis DA (2008): GABA Neurons and the Mechanisms of Network Oscillations: Implications for Understanding Cortical Dysfunction in Schizophrenia. *Schizophrenia Bulletin*. 34:944-961.
7. Friston KJ, Brown HR, Siemerkus J, Stephan KE (2016): The dysconnection hypothesis (2016). *Schizophrenia Research*. 176:83-94.
8. Friston KJ (2012): The history of the future of the Bayesian brain. *Neuroimage*. 62:1230-1233.
9. Stephan KE, Friston KJ, Frith CD (2009): Dysconnection in Schizophrenia: From Abnormal Synaptic Plasticity to Failures of Self-monitoring. *Schizophrenia Bulletin*. 35:509-527.
10. Aitchison L, Lengyel M (2017): With or without you: predictive coding and Bayesian inference in the brain. *Curr Opin Neurobiol.* 46:219-227.
11. Limongi R, Bohaterewicz B, Nowicka M, Plewka A, Friston KJ (2018): Knowing when to stop: Aberrant precision and evidence accumulation in schizophrenia. *Schizophrenia Research*, pp 386-391.

12. Spratling MW (2017): A review of predictive coding algorithms. *Brain and Cognition*. 112:92-97.
13. Bastos AM, Usrey WM, Adams RA, Mangun GR, Fries P, Friston KJ (2012): Canonical microcircuits for predictive coding. *Neuron*. 76:695-711.
14. Shaw AD, Knight L, Freeman TCA, Williams GM, Moran RJ, Friston KJ, et al. (2019): Oscillatory, Computational, and Behavioral Evidence for Impaired GABAergic Inhibition in Schizophrenia. *Schizophrenia Bulletin*.
15. Stroop JR (1935): Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*. 18:643-662.
16. Wang X, Wang T, Chen Z, Hitchman G, Liu Y, Chen A (2015): Functional connectivity patterns reflect individual differences in conflict adaptation. *Neuropsychologia*. 70:177-184.
17. Shapira-Lichter I, Strauss I, Oren N, Gazit T, Sammartino F, Giacobbe P, et al. (2018): Conflict monitoring mechanism at the single-neuron level in the human ventral anterior cingulate cortex. *Neuroimage*. 175:45-55.
18. Taylor R, Théberge J, Williamson PC, Densmore M, Neufeld RWJ (2016): ACC Neuro-over-Connectivity Is Associated with Mathematically Modeled Additional Encoding Operations of Schizophrenia Stroop-Task Performance. *Frontiers in psychology*. 7:1295-1295.
19. Taylor R, Neufeld RWJ, Schaefer B, Densmore M, Rajakumar N, Osuch EA, et al. (2015): Functional magnetic resonance spectroscopy of glutamate in schizophrenia and major

depressive disorder: anterior cingulate activity during a color-word Stroop task. *NPJ schizophrenia*. 1:15028-15028.

20. Bitzer S, Park H, Blankenburg F, Kiebel S (2014): Perceptual decision making: drift-diffusion model is equivalent to a Bayesian model. *Front Hum Neurosci*.

21. Sporns O, Kötter R (2004): Motifs in Brain Networks. *PLOS Biology*. 2:e369.

22. Pardo JV, Pardo PJ, Janer KW, Raichle ME (1990): The anterior cingulate cortex mediates processing selection in the Stroop attentional conflict paradigm. *Proceedings of the National Academy of Sciences of the United States of America*. 87:256-259.

23. Evrard HC, Forro T, Logothetis NK (2012): Von Economo neurons in the anterior insula of the macaque monkey. *Neuron*. 74:482-489.

24. Mesulam MM, Mufson EJ (1982): Insula of the old world monkey. III: Efferent cortical output and comments on function. *J Comp Neurol*. 212:38-52.

25. Mufson EJ, Mesulam MM (1982): Insula of the old world monkey. II: Afferent cortical input and comments on the claustrum. *J Comp Neurol*. 212:23-37.

26. Palaniyappan L, Liddle PF (2012): Does the salience network play a cardinal role in psychosis? An emerging hypothesis of insular dysfunction. *J Psychiatry Neurosci*, 2011/06/23 ed, pp 17-27.

27. Palaniyappan L, Simmonite M, White TP, Liddle EB, Liddle PF (2013): Neural primacy of the salience processing system in schizophrenia. *Neuron*. 79:814-828.

28. Moran LV, Tagamets MA, Sampath H, O'Donnell A, Stein EA, Kochunov P, et al. (2013): Disruption of anterior insula modulation of large-scale brain networks in schizophrenia. *Biol Psychiatry*. 74:467-474.

29. Limongi R, Sutherland SC, Zhu J, Young ME, Habib R (2013): Temporal prediction errors modulate cingulate-insular coupling. *Neuroimage*, pp 147-157.

30. Andreasen NC, Carpenter WTJ, Kane JM, Lasser RA, Marder SR, Weinberger DR (2005): Remission in schizophrenia: proposed criteria and rationale for consensus. *American Journal of Psychiatry*. 162:441 - 449.

31. Wong D, Schranz AL, Bartha R (2018): Optimized in vivo brain glutamate measurement using long-echo-time semi-LASER at 7 T. *NMR in Biomedicine*. 31:e4002.

32. Pinheiro JC, Bates DM (2000): Mixed-effects models in S and S-PLUS. New York: Springer-Verlag.

33. Fard PR, Park H, Warkentin A, Kiebel SJ, Bitzer S (2017): A Bayesian Reformulation of the Extended Drift-Diffusion Model in Perceptual Decision Making. *Frontiers in Computational Neuroscience*. 11.

34. Wiecki TV, Sofer I, Frank MJ (2013): HDDM: Hierarchical Bayesian estimation of the Drift-Diffusion Model in Python. *Front Neuroinform*. 7:14.

35. Boehm U, Marsman M, Matzke D, Wagenmakers E-J (2018): On the importance of avoiding shortcuts in applying cognitive models to hierarchical data. *Behavior Research Methods*. 50:1614-1631.

36. Kahan J, Urner M, Moran R, Flandin G, Marreiros A, Mancini L, et al. (2014): Resting state functional MRI in Parkinson's disease: the impact of deep brain stimulation on 'effective' connectivity. *Brain : a journal of neurology*. 137:1130-1144.
37. Limongi R, Perez FJ, Modrono C, Gonzalez-Mora JL (2016): Temporal Uncertainty and Temporal Estimation Errors Affect Insular Activity and the Frontostriatal Indirect Pathway during Action Update: A Predictive Coding Study. *Frontiers in Human Neuroscience*. 10:14.
38. Marreiros AC, Kiebel SJ, Friston KJ (2008): Dynamic causal modelling for fMRI: A two-state model. *Neuroimage*. 39:269-278.
39. Razi A, Kahan J, Rees G, Friston KJ (2015): Construct validation of a DCM for resting state fMRI. *Neuroimage*. 106:1-14.
40. Friston KJ, Preller KH, Mathys C, Cagnan H, Heinze J, Razi A, et al. (2017): Dynamic causal modelling revisited. *NeuroImage*.
41. Friston KJ, Litvak V, Oswal A, Razi A, Stephan KE, van Wijk BCM, et al. (2016): Bayesian model reduction and empirical Bayes for group (DCM) studies. *NeuroImage*. 128:413-431.
42. Zeidman P, Jafarian A, Seghier ML, Litvak V, Cagnan H, Price CJ, et al. (2019): A guide to group effective connectivity analysis, part 2: Second level analysis with PEB. *NeuroImage*. 200:12-25.

43. Kruschke JK (2013): Posterior predictive checks can and should be Bayesian: Comment on Gelman and Shalizi, 'Philosophy and the practice of Bayesian statistics'. *British Journal of Mathematical and Statistical Psychology*. 66:45-56.

44. Gelman A, Rubin DB (1992): Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science*. 7:457-472.

45. Bastos-Leite AJ, Ridgway GR, Silveira C, Norton A, Reis S, Friston KJ (2015): Dysconnectivity within the default mode in first-episode schizophrenia: a stochastic dynamic causal modeling study with functional magnetic resonance imaging. *Schizophr Bull*. 41:144-153.

46. Zhou Y, Zeidman P, Wu S, Razi A, Chen C, Yang L, et al. (2018): Altered intrinsic and extrinsic connectivity in schizophrenia. *NeuroImage Clinical*. 17:704-716.

47. Bryant JE, Frölich M, Tran S, Reid MA, Lahti AC, Kraguljac NV (2019): Ketamine induced changes in regional cerebral blood flow, interregional connectivity patterns, and glutamate metabolism. *Journal of Psychiatric Research*. 117:108-115.

48. Tremblay R, Lee S, Rudy B (2016): GABAergic Interneurons in the Neocortex: From Cellular Properties to Circuits. *Neuron*. 91:260-292.

49. Evans SL, Averbeck BB, Furl N (2015): Jumping to conclusions in schizophrenia. *Neuropsychiatr Dis Treat*. 11:1615-1624.

50. Valton V, Karvelis P, Richards KL, Seitz AR, Lawrie SM, Seriès P (2019): Acquisition of visual priors and induced hallucinations in chronic schizophrenia. *Brain*. 142:2523-2537.

51. Allen M, Friston KJ (2018): From cognitivism to autopoiesis: towards a computational framework for the embodied mind. *Synthese*. 195:2459-2482.
52. Miller EK, Lundqvist M, Bastos AM (2018): Working Memory 2.0. *Neuron*. 100:463-475.
53. Parr T, Friston KJ (2017): Working memory, attention, and salience in active inference. *Sci Rep*. 7:14678.
54. Pinotsis DA, Buschman TJ, Miller EK (2019): Working Memory Load Modulates Neuronal Coupling. *Cerebral Cortex*. 29:1670-1681.
55. Brüne M, Schöbel A, Karau R, Faustmann PM, Dermietzel R, Juckel G, et al. (2011): Neuroanatomical Correlates of Suicide in Psychosis: The Possible Role of von Economo Neurons. *PLOS ONE*. 6:e20936.
56. Menon V, Uddin LQ (2010): Saliency, switching, attention and control: a network model of insula function. *Brain Struct Funct*. 214:655-667.
57. Calati R, Ferrari C, Brittner M, Oasi O, Olie E, Carvalho AF, et al. (2019): Suicidal thoughts and behaviors and social isolation: A narrative review of the literature. *Journal of affective disorders*. 245:653-667.
58. Uddin LQ (2015): Salience processing and insular cortical function and dysfunction. *Nat Rev Neurosci*. 16:55-61.

Tables

Table 1. Summary statistics of accuracy and reaction time in the Stroop task

Group	Condition	Accuracy (proportion)		Reaction Time (s)	
		M	SD	M	SD
FEP	Color-only	0.99	0.09	1.06	0.27
	Congruent	0.98	0.14	0.98	0.27
	Incongruent	0.92	0.27	1.19	0.31
HC	Word-only	0.97	0.17	0.95	0.26
	Color-only	1.00	0.05	0.89	0.20
	Congruent	1.00	0.05	0.82	0.21
	Incongruent	0.98	0.14	1.04	0.25
	Word-only	1.00	0.00	0.79	0.18

Table 2. Mixed-effect models of response accuracy and reaction time in the Stroop task

Parameter	Estimate	SE	DF	t	P	Lower 95 %	Upper 95%
Accuracy	Intercept	0.979	0.006	36.69	156	<.0001	0.966 0.992
	Group [FEP]	-0.015	0.006	36.69	-2.32	0.0261	-0.027 -0.002
	Condition [Color-only]	0.015	0.004	3000.87	3.71	0.0002	0.007 0.023
	Condition [Congruent]	0.010	0.004	3000.86	2.33	0.0201	0.002 0.018
	Condition [Incongruent]	-0.030	0.004	3001.31	-7.2	<.0001	-0.038 -0.022
	Group [FEP] × Condition [Color-only]	0.011	0.004	3000.87	2.79	0.0053	0.003 0.020
	Group [FEP] × Condition [Congruent]	0.006	0.004	3000.86	1.4	0.1624	-0.002 0.014
Reaction Time	Group [FEP] × Condition [Incongruent]	-0.016	0.004	3001.32	-3.85	0.0001	-0.024 -0.008
	Intercept	0.971	0.022	36.83	45	<.0001	0.927 1.014
	Group [FEP]	0.083	0.022	36.83	3.86	0.0004	0.040 0.127
	Condition [Color-only]	0.011	0.007	2999.88	1.66	0.0968	-0.002 0.024
	Condition [Congruent]	-0.069	0.007	2999.87	-10.4	<.0001	-0.082 -0.056
	Condition [Incongruent]	0.155	0.007	2999.98	23	<.0001	0.141 0.168
	Group [FEP] × Condition [Color-only]	0.004	0.007	2999.88	0.53	0.5943	-0.010 0.017
	Group [FEP] × Condition [Congruent]	0.000	0.007	2999.87	0.06	0.9561	-0.013 0.013
	Group [FEP] × Condition [Incongruent]	-0.004	0.007	2999.98	-0.63	0.5271	-0.017 0.009

Figures Legends

Fig. 1. Voxel positioning for ^1H -MRS measurement and two-neuronal-population DCM.

Black box shows the voxel on the dACC for ^1H -MRS (top) and the same voxel for fMRI (bottom). Grey box shows the two-neuronal-population DCM of the salience network. Each region comprises one population of excitatory neurons (E) and one population of inhibitory neurons (I). Parameters of effective connectivity represent the influence of inhibitory to excitatory connections (IE, assumed to be GABAergic neurons), the influence of excitatory to inhibitory connections (EI), the influence of self-inhibitory connections within each population (SE, SI), and the influence of excitatory population of one region on the excitatory population of the other region (EE, assumed to be glutamatergic connections). Whereas EI, SE, and SI parameters are fixed in the model, IE and EE are free parameters. Small white box shows a sample spectrum obtained using ^1H -MRS semi-LASER (TE = 100 ms). Raw data are indicated in blue, fit spectrum is indicated in orange, and residual signal is shown above in green. Main glutamate signal contribution is indicated with an arrow (2.3 ppm). Since we aimed at demonstrating the relationship between the effect of $[\text{Glut}]_{\text{dACC}}$ (observed measurements on the bar chart indicating no group difference) on IE and the ensuing consequences in the whole network, this two-state model was sufficient to test our hypothesis despite not capturing all the circuitry constraints of cortical columns (c.f., a four-neuronal-population model, 40).

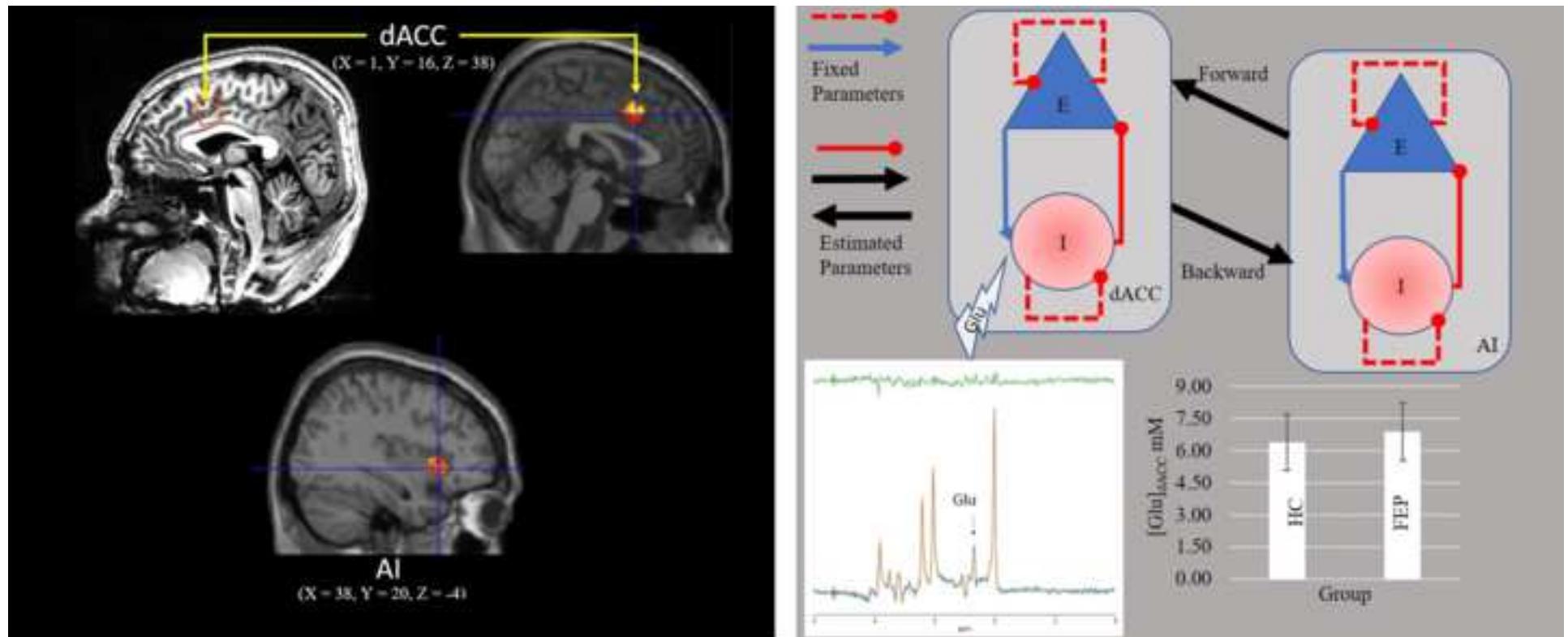
Fig. 2. Parameter estimates of the hierarchical drift diffusion model. Top right, visual depiction of the parameter estimates of the drift diffusion model (z = starting point or prior beliefs, v = drift rate or precision of PE, t = non-decision processes, a = decision threshold, RT = reaction time). Left, proportion of the posteriors in which the parameter estimate for one group differs from the other (ΔPP).

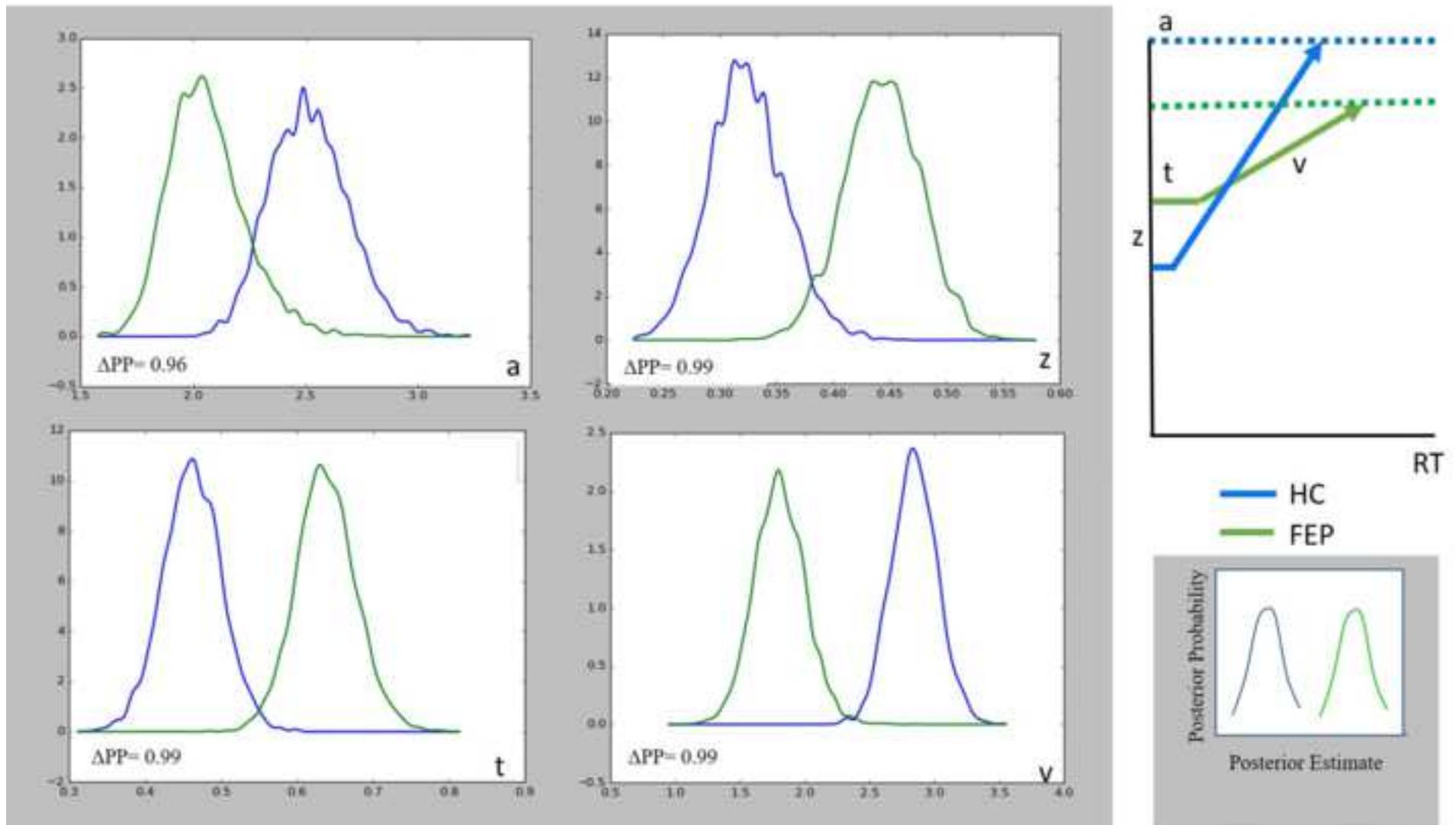
Fig. 3. Bayesian model comparison results and mean parameter estimates in the salience network across groups. * PP ≈ 0 .

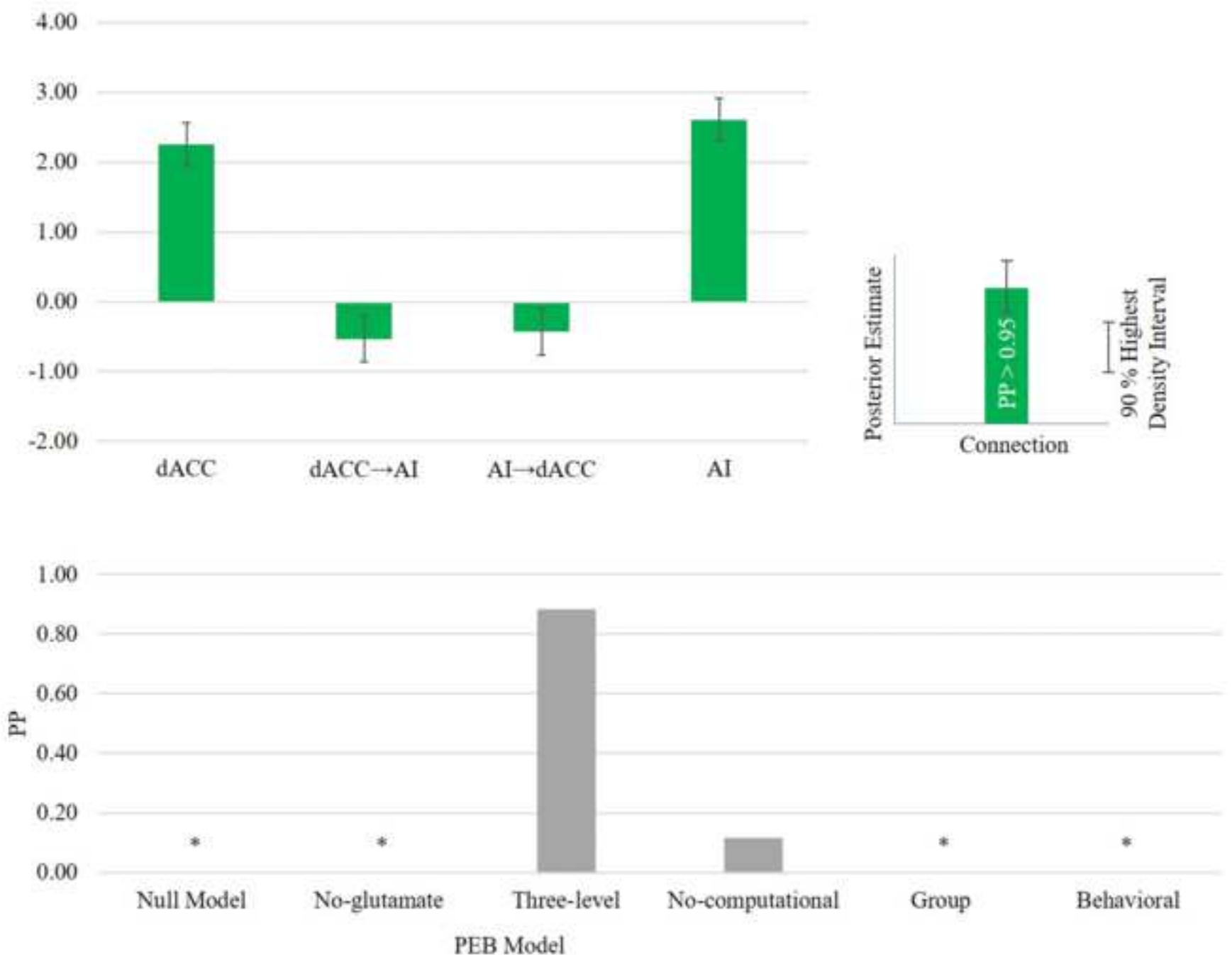
Fig. 4. Main effects of the winning three-level parametric empirical Bayes model. In A, bars represent differences between groups as defined by the effect coding (FEP = 1, HC = -1). In B, C, and D, covariates were mean centered.

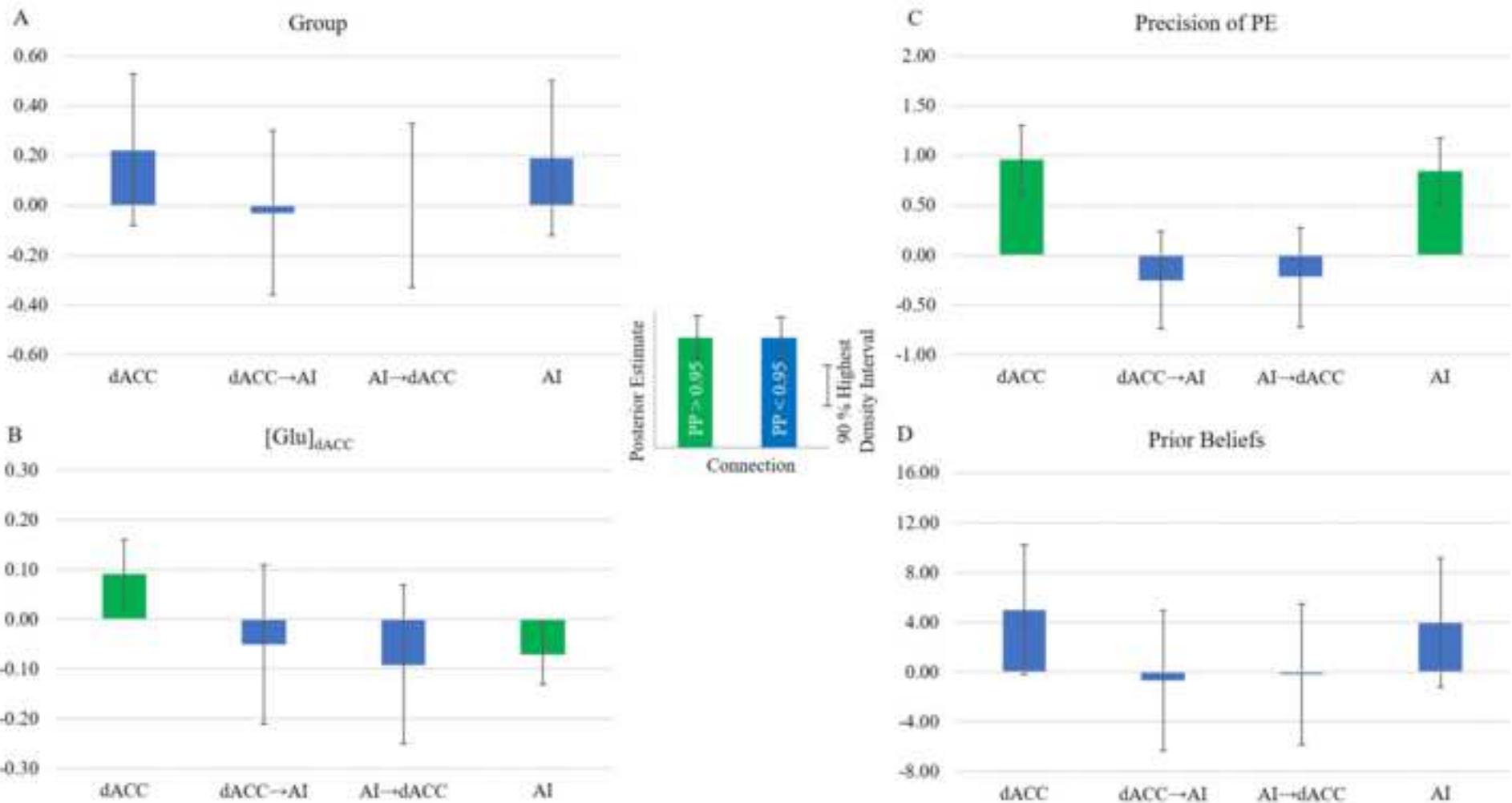
Fig. 5. Interactions between group and covariates in the winning three-level parametric empirical Bayes model. In A, B, and C, positive values represent stronger effect in the FEP group, and vice versa.

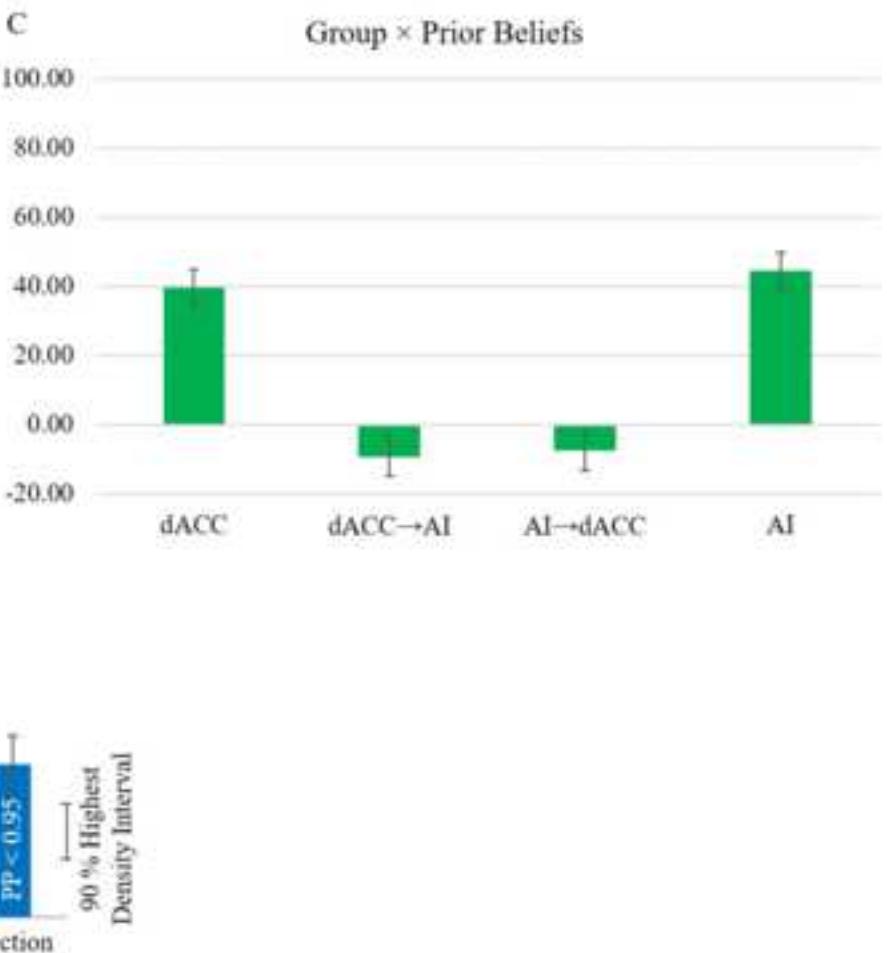
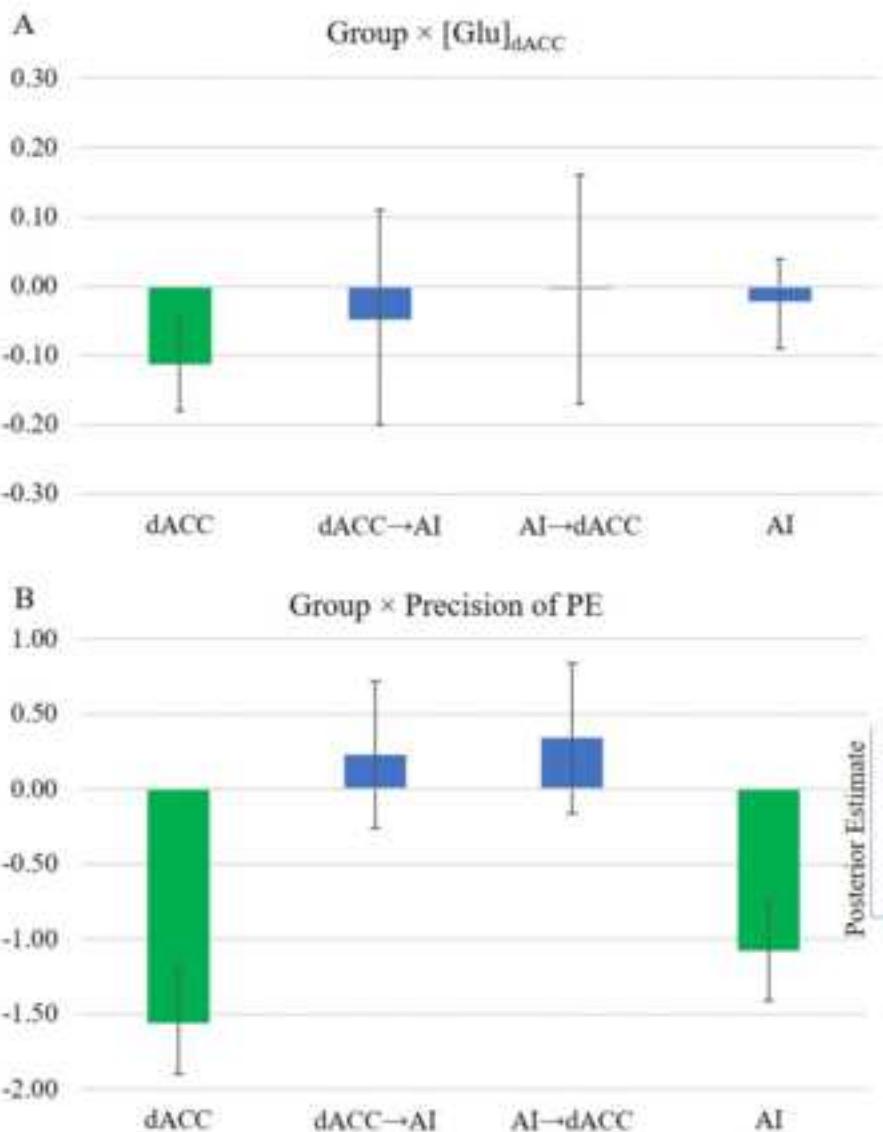
Fig. 6. Association between Extrinsic Connectivity Strength and Severity of Social Withdrawal.

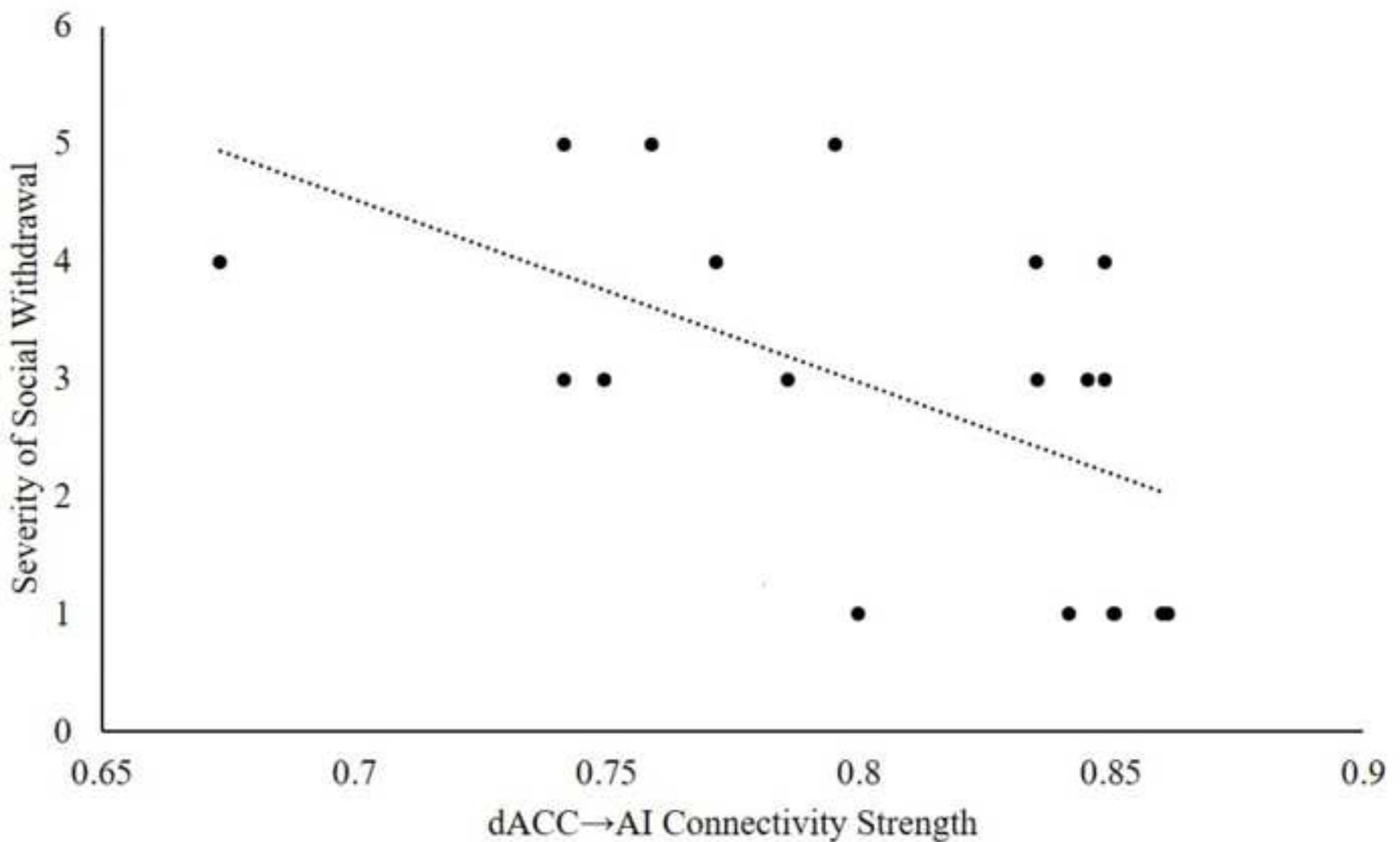












Supplementary Information

Table S1. Subject-level values of proportion of variance accounted for the three-level parametric-empirical-Bayes model

Subject	HC	FEP
1	5.1	21.44
2	6.21	9.48
3	18.23	37.41
4	32.19	20.83
5	32.55	20.13
6	11.97	18.03
7	31.64	3.6
8	15.28	8.37
9	19.77	5.39
10	12.18	31.89
11	5.76	38.74
12	21.81	46
13	24.01	27.92
14	6.48	31.96
15	30.61	36.86
16	27.6	30.61
17	11.82	9.82
18	7.61	6.09
19	15.3	26.32
20	8.51	

Note. The mean proportion of variance accounted for by the model was 0.17 (SD = 9.73, 95 % CI [12.67, 21.79]) in the HC group and 0.23 (SD = 12.93, 95% CI = [16.44, 28.91]) in the FEP group. —a proportion of at least 0.17 of variance accounted for by the model (averaged across subjects) is considered acceptable (1).

Table S2. Parameter estimates of the hierarchical drift-diffusion model

Group	Parameter								
	a		v		t		z		
	mean	std	mean	std	mean	std	mean	std	
HC	1	2.535	0.259	2.872	0.263	0.455	0.057	0.326	0.034
	2	2.558	0.275	2.884	0.272	0.540	0.064	0.326	0.034
	3	2.650	0.299	2.735	0.253	0.524	0.065	0.325	0.034
	4	2.538	0.236	2.845	0.260	0.300	0.038	0.326	0.034
	5	2.535	0.241	2.881	0.267	0.363	0.044	0.326	0.034
	6	2.678	0.295	2.561	0.285	0.563	0.058	0.324	0.034
	7	2.373	0.248	2.841	0.263	0.377	0.050	0.327	0.034
	8	2.592	0.266	2.863	0.270	0.447	0.052	0.326	0.034
	9	2.541	0.258	2.871	0.265	0.454	0.055	0.326	0.034
	10	2.439	0.241	2.795	0.254	0.455	0.054	0.326	0.034
	11	2.475	0.235	2.554	0.291	0.548	0.064	0.325	0.034
	12	2.462	0.237	2.902	0.268	0.360	0.048	0.327	0.035
	13	2.431	0.263	3.051	0.331	0.458	0.059	0.327	0.035
	14	2.344	0.282	2.926	0.288	0.503	0.067	0.327	0.035
	15	2.510	0.236	2.912	0.275	0.324	0.041	0.326	0.034
	16	2.656	0.300	2.701	0.256	0.486	0.062	0.325	0.034
	17	2.509	0.245	2.767	0.262	0.590	0.064	0.325	0.034
	18	2.507	0.272	3.011	0.319	0.582	0.068	0.326	0.035
	19	2.505	0.244	2.891	0.266	0.389	0.050	0.326	0.034
	20	2.474	0.245	2.804	0.262	0.528	0.059	0.325	0.034
FEP	1	2.365	0.486	2.305	0.403	0.821	0.082	0.443	0.035
	2	2.619	0.548	1.954	0.377	0.739	0.090	0.442	0.035
	3	1.371	0.173	1.322	0.366	0.533	0.029	0.441	0.035
	4	1.916	0.368	2.524	0.434	0.743	0.063	0.444	0.035
	5	1.797	0.243	1.633	0.335	0.642	0.056	0.443	0.035
	6	1.943	0.263	1.784	0.321	0.543	0.050	0.444	0.035
	7	1.854	0.275	1.202	0.350	0.800	0.095	0.443	0.035
	8	2.173	0.390	2.422	0.421	0.652	0.056	0.443	0.035
	9	1.707	0.268	2.259	0.404	0.640	0.046	0.443	0.035
	10	2.006	0.221	0.746	0.280	0.681	0.065	0.442	0.035

Spectral post-processing

The spectral acquisition was analyzed to measure glutamate concentration using methods previously developed in our laboratory (2-5). Specifically, the 32 spectral acquisition was corrected for frequency and phase drifts as described in Near, Edden (6). The corrected spectra were then averaged into one spectrum to represent the entire four-minute resting-state block. The averaged spectrum underwent post-processing using combined QUALITY Eddy Current Correction (7), using 400 QUALITY (quantification improvement by converting lineshapes to the Lorentzian type) points, to reduce linewidth distortions as well as Hankel Singular Value Decomposition (HSVD) water removal to remove any residual water signal between 4.2 ppm and 5.7 ppm before being fit with fitMAN (2, 3), a time-domain fitting algorithm that uses a non-linear, iterative Levenberg-Marquardt minimization algorithm to echo time-specific prior knowledge templates. The metabolite fitting template included eighteen brain metabolites: alanine, aspartate, choline, creatine, γ -aminobutyric acid (GABA), glucose, glutamate, glutamine, glutathione, glycine, lactate, myo-inositol, N-acetyl aspartate, N-acetyl aspartyl glutamate, phosphorylethanolamine, scyllo-inositol, and taurine. As well, a single peak was used to fit the water unsuppressed spectrum.

Using Barstool (2), tissue-specific (gray matter, white matter, and CSF) T1 and T2 relaxations were corrected through partial volume segmentation calculations of voxels mapped onto T1-weighted images acquired using a 0.75 mm isotropic MP2RAGE sequence (repetition time = 6000 ms, TI1 = 800 ms, TI2 = 2700 ms, flip-angle 1 (α 1) = 4°, flip-angle 2 (α 2) = 5°, field of view = 350 mm \times 263 mm \times 350 mm, acquisition time = 9 min. 38 s, iPATPE = 3 and 6/8 k-space). Finally, glutamate concentration quantification was calculated using the postprocessed water suppressed and unsuppressed spectra for each participant along with voxel-appropriate, tissue-specific relaxation time adjustments (8).

References

1. Zeidman P, Jafarian A, Corbin N, Seghier ML, Razi A, Price CJ, et al. (2019): A guide to group effective connectivity analysis, part 1: First level analysis with DCM for fMRI. *NeuroImage*. 200:174-190.
2. Wong D, Schranz AL, Bartha R (2018): Optimized in vivo brain glutamate measurement using long-echo-time semi-LASER at 7 T. *NMR in Biomedicine*. 31:e4002.
3. Bartha R, Drost DJ, Williamson PC (1999): Factors affecting the quantification of short echo in-vivo ¹H MR spectra: prior knowledge, peak elimination, and filtering. *NMR Biomed*. 12:205-216.
4. Ryan K, Wawrzyn K, Gati J, Chronik BA, Wong D, Duggal N, et al. (2018): ¹H MR Spectroscopy of the Motor Cortex Immediately following Transcranial Direct Current Stimulation at 7 Tesla. *bioRxiv*.322941.
5. Penner J, Bartha R (2015): Semi-LASER ¹H MR spectroscopy at 7 Tesla in human brain: Metabolite quantification incorporating subject-specific macromolecule removal. *Magnetic Resonance in Medicine*. 74:4-12.
6. Near J, Edden R, Evans CJ, Paquin R, Harris A, Jezzard P (2015): Frequency and phase drift correction of magnetic resonance spectroscopy data by spectral registration in the time domain. *Magnetic Resonance in Medicine*. 73:44-50.
7. Bartha R, Drost DJ, Menon RS, Williamson PC (2000): Spectroscopic lineshape correction by QUECC: Combined QUALITY deconvolution and eddy current correction. *Magnetic Resonance in Medicine*. 44:641-645.
8. Gasparovic C, Song T, Devier D, Bockholt HJ, Caprihan A, Mullins PG, et al. (2006): Use of tissue water as a concentration reference for proton spectroscopic imaging. *Magnetic Resonance in Medicine*. 55:1219-1226.

ID	Group	AGESCAN	DOB	AgePsychosisOnset	DUPMonths	DaysOfAntipsychotic
1	HC	26.00	23-Jun-90		0.00	0.00
2	HC	23.00	16-Oct-93		0.00	0.00
3	HC	17.00	18-Nov-99		0.00	0.00
4	HC	23.00	22-Jul-93		0.00	0.00
5	HC	16.00	8-Nov-00		0.00	0.00
6	HC	25.00	10-Sep-91		0.00	0.00
7	HC	16.00	24-Oct-00		0.00	0.00
8	HC	16.00	4-Jan-01		0.00	0.00
9	HC	29.00	15-Apr-88		0.00	0.00
10	HC	22.00	18-May-95		0.00	0.00
11	HC	23.00	31-Jan-94		0.00	0.00
12	HC	20.00	6-Apr-97		0.00	0.00
13	HC	20.00	20-Nov-96		0.00	0.00
14	HC	20.00	9-May-97		0.00	0.00
15	HC	20.00	31-Jul-97		0.00	0.00
16	HC	20.00	13-Sep-97		0.00	0.00
17	HC	27.00	6-Sep-90		0.00	0.00
18	HC	22.00	3-Jan-95		0.00	0.00
19	HC	18.00	1-Oct-99		0.00	0.00
20	HC	22.00	27-Feb-96		0.00	0.00
1	FEP	19.00	25-Apr-97	17.00	24.00	0.00
2	FEP	20.00	10-Oct-96	20.00	4.00	0.00
3	FEP	19.00	28-Nov-97	19.00	1.00	7.00
4	FEP	17.00	3-Oct-99	17.00	2.00	0.00
5	FEP	18.00	18-Oct-98	18.00	2.00	7.00
6	FEP	17.00	10-Feb-00	16.00	9.00	0.00
7	FEP	24.00	19-Jan-93	23.00	12.00	0.00
8	FEP	21.00	24-Nov-95	21.00	1.00	0.00
9	FEP	25.00	23-Jul-91	20.00	59.00	7.00
10	FEP	28.00	6-Jul-89	28.00	3.00	0.00
11	FEP	20.00	21-Nov-96	18.00	14.00	0.00
12	FEP	23.00	29-Apr-94			5.00
13	FEP	23.00	21-Sep-93	22.00	6.00	0.00
14	FEP	24.00	22-Nov-92	24.00	0.50	7.00
15	FEP	23.00	7-Mar-94			0.00
16	FEP	20.00	2-May-97	19.00	9.00	0.00
17	FEP	27.00	7-Dec-89	21.00	72.00	0.00
18	FEP	26.00	19-Nov-91	26.00	1.00	7.00
19	FEP	19.00	5-Dec-98	19.00	0.50	10.00

Antipsychotic	Dose	duration	Gender	PANSS8P1	PANSS8P3	PANSS8G9	PANSS8P2	PANSS8G5
			1	1	1	1	1	1
			2	1	1	1	1	1
			1	1	1	1	1	1
			1	1	1	1	1	1
			2	1	1	1	1	1
			1	1	1	1	1	1
			1	1	1	1	1	1
			1	1	1	1	1	1
			1	1	1	1	1	1
			1	1	1	1	1	1
			2	1	1	1	1	1
			2	1	1	1	1	1
			1	1	1	1	1	1
			1	1	1	1	1	1
			1	1	1	1	1	1
			2	1	1	1	1	1
			2	1	1	1	1	1
			2	1	1	1	1	1
			2	1	1	1	1	1
			2	1	1	1	1	1
0	0	0	1	4	4	4	4	2
0	0	0	1	5	5	4	4	3
Invega	6mg	1wk	1	4	4	3	5	1
0	0	0	1	5	5	4	1	1
Abilify	5mg	1wk	1	5	4	3	3	1
0	0	0	2	5	5	2	1	3
0	0	0	1	6	6	5	5	4
0	0	0	1	5	4	4	4	1
Invega	3mg, 6mg	3days, 4days	1	7	5	6	3	2
Olanzapine		1day	1	6	2	3	4	1
0	0	0	2	6	2	4	3	1
Risperidone	1.5mg	5 days	2	4	5	3	2	1
0	0	0	1	5	5	5	6	3
Olanzapine	7.5mg	1wk	2	5	2	5	4	1
0	0	0	2	7	6	6	4	1
0	0	0	1	5	5	3	1	1
0	0	0	1	7	2	3	3	1
Risperidone	1mg	1wk	2	5	5	3	1	1
Abilify	5mg	10days	2	5	1	4	3	1

PANSS8N1 PANSS8N4 PANSS8N6