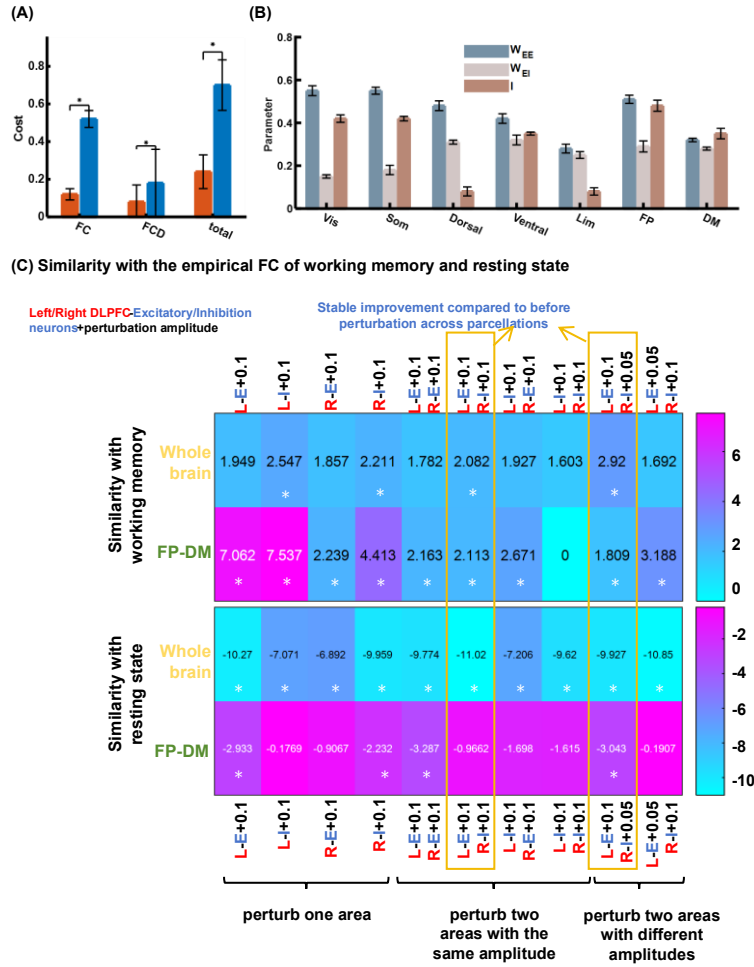


Supplemental Information for:

**Differential Excitatory-Inhibitory Balance within
Dorsolateral Prefrontal Cortex Links Inter-Network
Interactions in Working Memory**

1. Fitting and perturbing EI-MFM on Shaefer 100 parcellation



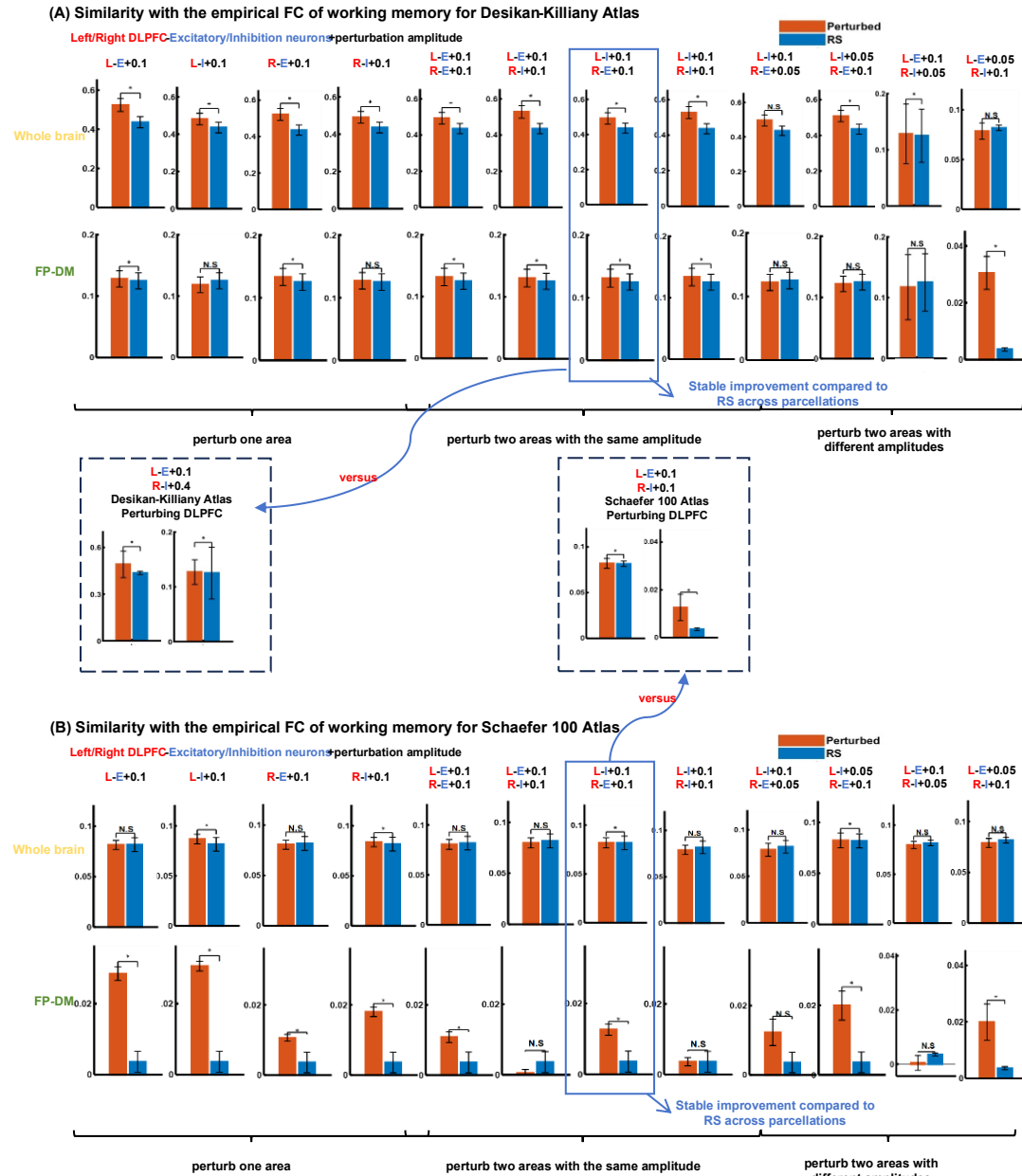
Supplementary Figure 1: Replication of Results on the EI-MFM model fitting and perturbation on Schaefer 100 parcellation. **A** The disagreement between simulation and functional measures on test set is calculated as $1-r$ for FC and KS distance for FCD and their summation for the total costs. LR and wLR denote the dynamics of inhibition neuronal populations with and without long-range excitatory inputs. * denotes a significant difference presented by permutation tests ($p < 0.05$, Bonferroni correction). **B** The estimated parameters present a gradient across brain networks, suggesting a local heterogeneity across cortex. **C** The similarity distance of simulated FC after perturbation compared to that before perturbation. The perturbations areas are applied on one side of DLPFC areas, both DLPFC areas with the same amplitude and both DLPFC areas with different amplitudes. The similarity is measured by the Pearson's correlation between the simulated and the empirical FCs of working memory (the top two rows) as well as those of resting state (the bottom two rows) among the whole brain and between the frontoparietal and the default-mode networks, respectively. A robust improvement on the similarity with empirical FC pattern of working memory across brain parcellations is found for L-E+0.1&R-I+0.1, L-E+0.1&R-E-0.05. * denotes $p < 0.05$, Bonferroni corrected.

Supplemental Tab.1: The details of perturbation conditions

(S_E(t)) and (S_I(t)) denote the excitatory and inhibitory synaptic gating variables, respectively. Each perturbation was implemented by adding a constant offset of 0.05 or 0.1 to the selected gating variable in the left or right DLPFC after numerically solving the corresponding model equation at each time step.

Perturbation	Perturbing manner
L-E+0.1	Left DLPFC: $s_E(t) = s_E(t) + 0.1$ after numerically solving Eq. (5) within this region for each time step
L-I+0.1	Left DLPFC: $s_I(t) = s_I(t) + 0.1$ after numerically solving Eq. (6) within this region for each time step
R-E+0.1	Right DLPFC: $s_E(t) = s_E(t) + 0.1$ after numerically solving Eq. (5) within this region for each time step
R-I+0.1	Right DLPFC: $s_I(t) = s_I(t) + 0.1$ after numerically solving Eq. (6) within this region for each time step
L-E+0.1 &R-E+0.1	Left DLPFC and right DLPFC: $s_E(t) = s_E(t) + 0.1$ after numerically solving Eq. (5) within these two regions for each time step
L-E+0.1 &R-I+0.1	Left DLPFC: $s_E(t) = s_E(t) + 0.1$ after numerically solving Eq. (5) within this region for each time step Right DLPFC: $s_I(t) = s_I(t) + 0.1$ after numerically solving Eq. (6) within this region for each time step
L-I+0.1 &R-E+0.1	Left DLPFC: $s_I(t) = s_I(t) + 0.1$ after numerically solving Eq. (6) within this region for each time step Right DLPFC: $s_E(t) = s_E(t) + 0.1$ after numerically solving Eq. (5) within this region for each time step
L-I+0.1 &R-I+0.1	Left DLPFC and right DLPFC: $s_I(t) = s_I(t) + 0.1$ after numerically solving Eq. (6) within these two regions for each time step
L-E+0.1 &R-I+0.05	Left DLPFC: $s_E(t) = s_E(t) + 0.1$ after numerically solving Eq. (5) within this region for each time step Right DLPFC: $s_I(t) = s_I(t) + 0.05$ after numerically solving Eq. (6) within this region for each time step
L-E+0.05 &R-I+0.1	Left DLPFC: $s_E(t) = s_E(t) + 0.05$ after numerically solving Eq. (5) within this region for each time step Right DLPFC: $s_I(t) = s_I(t) + 0.1$ after numerically solving Eq. (6) within this region for each time step

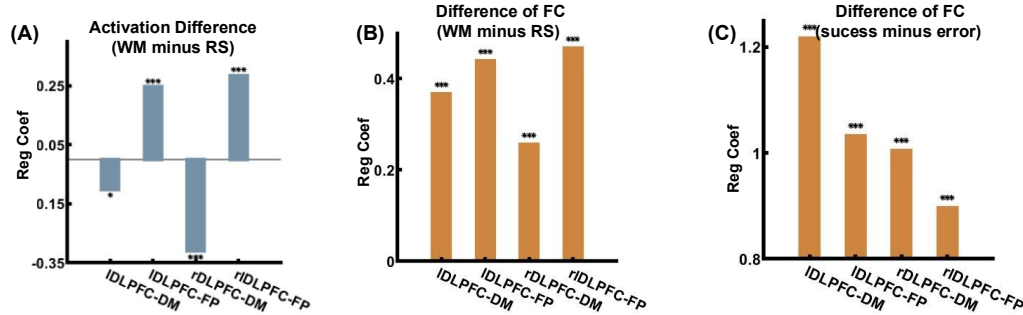
2. The control analyses of the perturbation study on the Anterior Insula



Supplementary Figure 2: The perturbation-induced brain pattern changes of the Anterior Insula.

The perturbations areas are applied on one side of anterior insula areas, both anterior insula areas with the same amplitude and both anterior insula areas with different amplitudes using the Desikan-Killiany and the Schaefer 100 atlases. The similarity is measured by the Pearson's correlation between the simulated and the empirical FCs of working memory among the whole brain and between the frontoparietal and the default-mode networks, respectively. * denotes $p < 0.05$ after Bonferroni correction, comparing with RS. A robust improvement across brain parcellations is found for L-I+0.1&R-E-0.1.

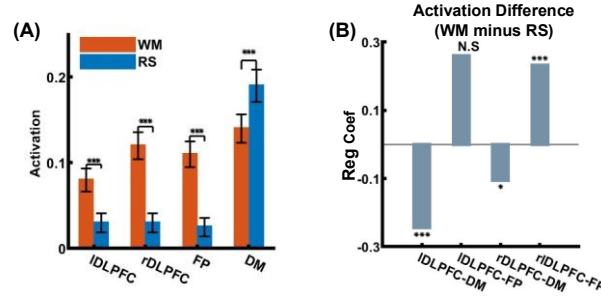
3. DLPFC accounts for the activation and interactions of the frontoparietal and default-mode networks



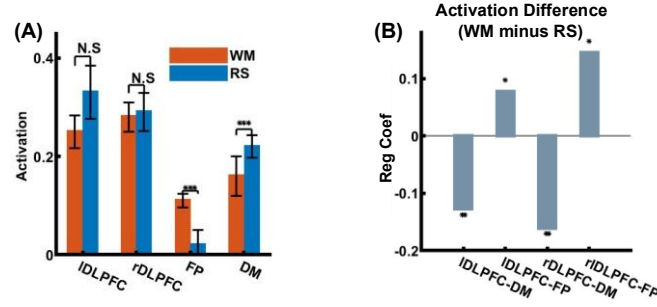
Supplementary Figure 3: Functional variations and associations of DLPFC areas, frontoparietal and default-mode networks.

A Regressing the difference of the functional activation between RS and WM for IDLPFC and rDLPFC against that for DM and FP. All the correlations, except for the correlation between rDLPFC and FP, are significant (for right DLPFC and the DM network, *linear regression*, $r = -0.3101$, $p < 0.0001$, for right DLPFC and the FP network, *linear regression*, $r = 0.2832$, $p < 0.0001$, for left DLPFC and the DM network, *linear regression*, $r = -0.1013$, $p = 0.0304$ and for left DLPFC and the FP network, *linear regression*, $r = 0.2456$, $p < 0.0001$). For X axis, each bar denotes a separate regression, i.e. FP-DM: $\Delta_{DM} \sim \Delta_{FP} + 1$, IDLPFC-DM: $\Delta_{DM} \sim \Delta_{IDLPFC} + 1$, IDLPFC-FP: $\Delta_{DM} \sim \Delta_{IDLPFC} + 1$, rDLPFC-DM: $\Delta_{DM} \sim \Delta_{rDLPFC} + 1$, where Δ_A denotes the activation difference (WM minus RS) of ROI A. **B** Regressing the difference (WM minus RS) of the functional connections between DLPFC and the two networks (X axis) against that between FP and DM (for IDLPFC-DM, *linear regression*, $r = 0.3675$, $p = 0.0003$; for IDLPFC-FP, *linear regression*, $r = 0.4385$, $p < 0.0001$; for rDLPFC-DM, *linear regression*, $r = 0.2554$, $p < 0.0001$; for rDLPFC-FP, *linear regression*, $r = 0.4666$, $p < 0.0001$). For X axis, each bar denotes a separate regression, IDLPFC-DM: $\Delta_{FP-DM} \sim \Delta_{IDLPFC-DM} + 1$, IDLPFC-FP: $\Delta_{FP-DM} \sim \Delta_{IDLPFC-FP} + 1$, rDLPFC-DM: $\Delta_{FP-DM} \sim \Delta_{rDLPFC-DM} + 1$, rDLPFC-FP: $\Delta_{FP-DM} \sim \Delta_{rDLPFC-FP} + 1$, where Δ_{A-B} denotes the difference (WM minus RS) of the function connectivity between ROI A and ROI B. **C** Regressing the difference (success minus error) of the functional connections between DLPFC and the two networks (X axis) against that between FP and DM during the 2-back tasks (for IDLPFC-DM, *linear regression*, $r = 1.2164$, $p < 0.0001$; for IDLPFC-FP, *linear regression*, $r = 1.0315$, $p < 0.0001$; for rDLPFC-DM, *linear regression*, $r = 1.0040$, $p < 0.0001$; for rDLPFC-FP, *linear regression*, $r = 0.8959$, $p < 0.0001$). For X axis, each bar denotes a separate regression with similar denotation as B, except that Δ_{A-B} denotes the difference (success minus error during 2-back tasks) of the function connectivity between ROI A and ROI B. The significance of the difference is evaluated by the permutation test. N.S denotes no significance (i.e. $p > 0.05$), * denotes $p < 0.05$, ** denotes $p < 0.01$, and *** denotes $p < 0.001$. During the regression, we excluded the DLPFC areas from the FP and DM networks to ensure the avoidance of any potential biases.

4. Alternative activation metrics

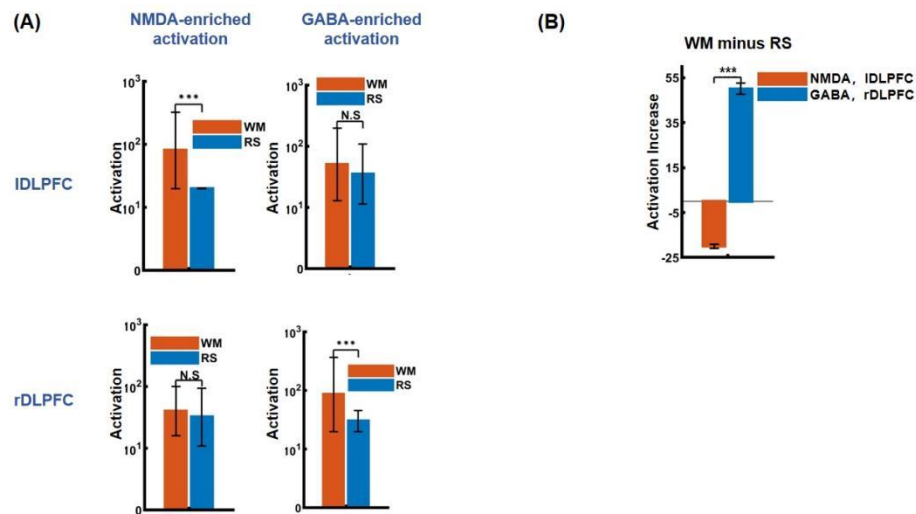


Supplementary Figure 4: Replication of the results on functional variations and associations of DLPFC areas, frontoparietal and default-mode networks (ALFF, Schaefer100). **A** The functional activation presents significant difference between resting state (RS) and working memory (WM) for left and right DLPFC, FP and DM. **B** The difference of the functional activation between RS and WM for IDLPFC and rDLPFC correlated to that for DM and FP. All the correlations, except for the correlation between rDLPFC and FP, are significant (for right DLPFC and the DM network, *linear regression*, $r = -0.1044$, $p = 0.0124$, for right DLPFC and the FP network, *linear regression*, $r = 0.2323$, $p < 0.0001$, for left DLPFC and the DM network, *linear regression*, $r = -0.2441$, $p = 0.0002$ and for left DLPFC and the FP network, *linear regression*, $r = 0.2602$, $p = 0.1429$). For X axis, each bar denotes a separate regression, i.e. IDLPFC-DM: $\Delta_{DM} \sim \Delta_{IDLPFC} + 1$, IDLPFC-FP: $\Delta_{DM} \sim \Delta_{IDLPFC} + 1$, rDLPFC-DM: $\Delta_{DM} \sim \Delta_{rDLPFC} + 1$, rDLPFC-FP: $\Delta_{FP} \sim \Delta_{rDLPFC} + 1$, where Δ_A denotes the activation difference (WM minus RS) of ROI A. The significance of the difference is evaluated by the permutation test. N.S denotes no significance (i.e. $p > 0.05$), * denotes $p < 0.05$, ** denotes $p < 0.01$, and *** denotes $p < 0.001$.



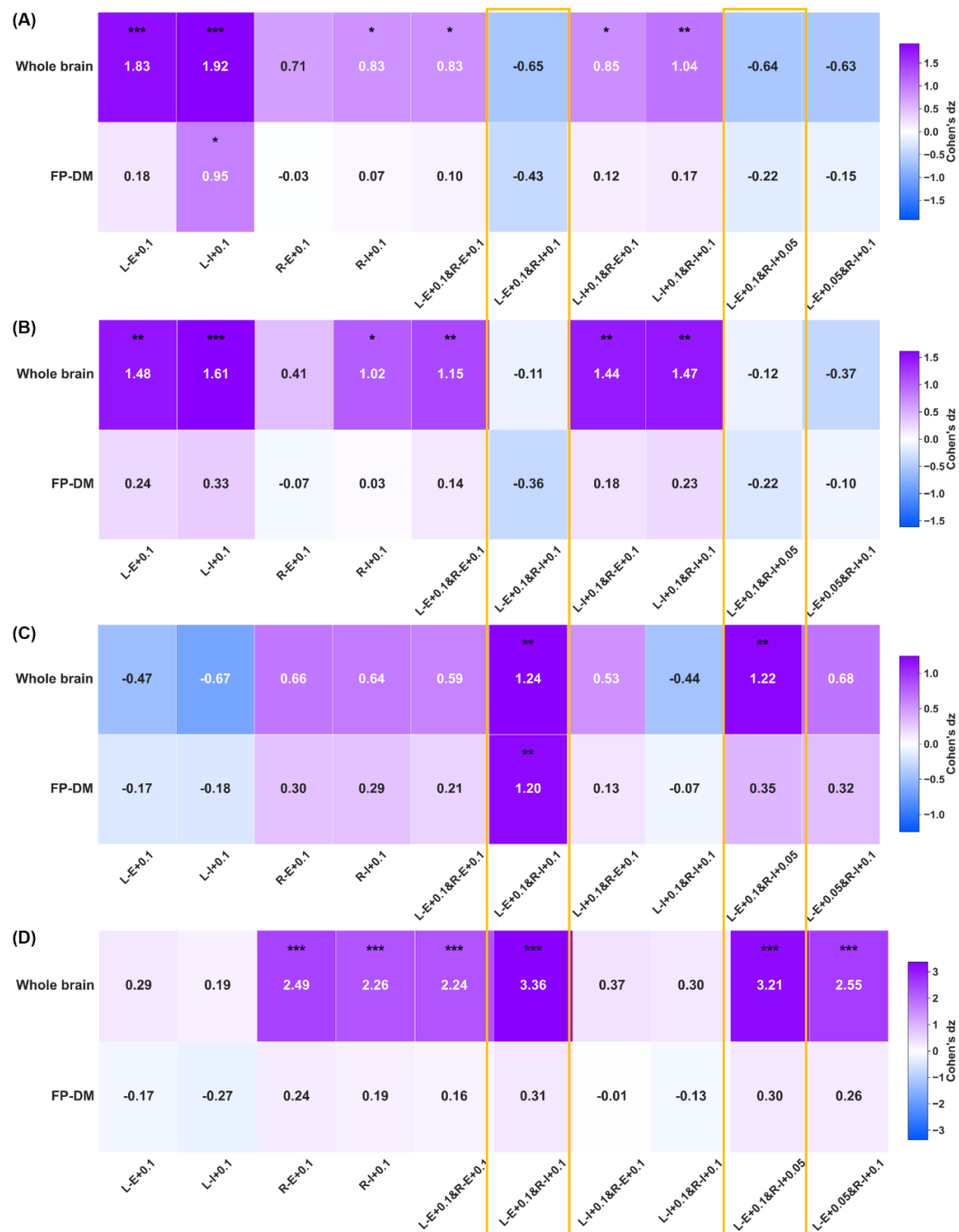
Supplementary Figure 5: Republication of the results on functional variations and associations of DLPFC areas, frontoparietal and default-mode networks (ALFF, Desikan-Killiany). **A** The functional activation presents significant difference between resting state (RS) and working memory (WM) for FP and DM. The difference is not significant for left and right DLPFC. **B** The difference of the functional activation between RS and WM for IDLPFC and rDLPFC correlated to that for DM and FP (for right DLPFC and the DM network, *linear regression*, $r = -0.1605$, $p = 0.0016$, for right DLPFC and the FP network, *linear regression*, $r = 0.21451$, $p = 0.0175$, for left DLPFC and the DM network, *linear regression*, $r = -0.1269$, $p = 0.0014$ and for left DLPFC and the FP network, *linear regression*, $r = 0.0778$, $p = 0.0404$). For X axis, each bar denotes a separate regression, i.e. IDLPFC-DM: $\Delta_{DM} \sim \Delta_{IDLPFC} + 1$, IDLPFC-FP: $\Delta_{DM} \sim \Delta_{IDLPFC} + 1$, rDLPFC-DM: $\Delta_{DM} \sim \Delta_{rDLPFC} + 1$, rDLPFC-FP: $\Delta_{FP} \sim \Delta_{rDLPFC} + 1$, where Δ_A denotes the activation difference (WM minus RS) of ROI A. The significance of the difference is evaluated by the permutation test. N.S denotes no significance (i.e. $p > 0.05$), * denotes $p < 0.05$, ** denotes $p < 0.01$, and *** denotes $p < 0.001$.

5. The GABA and NMDA receptor-enriched network analysis on Schaefer 100

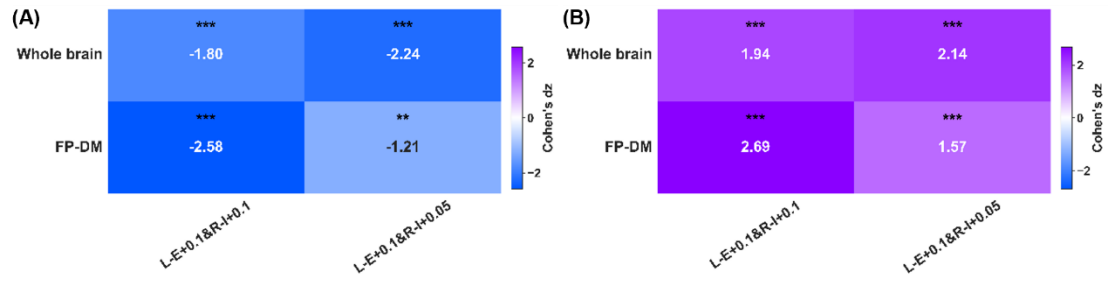


Supplementary Figure 6: Republication of results on GABA and NMDA-enriched analysis on the activation of left and right DLPFC on Schaefer 100 parcellation. We perform the molecular-enriched analysis on how the receptor density if NMDA and GABA would influence the functional activation of the left and right DLPFC areas. **A** NMDA-enriched and GABA-enriched activation during working memory and resting state for the left DLPFC under the Schaefer 100 parcellation. **B** The working memory-aroused increase of the NMDA-enriched activation of the left DLPFC versus that of the GABA-enriched activation of the right DLPFC under the Schaefer 100 parcellation.

6. Sensitivity analysis

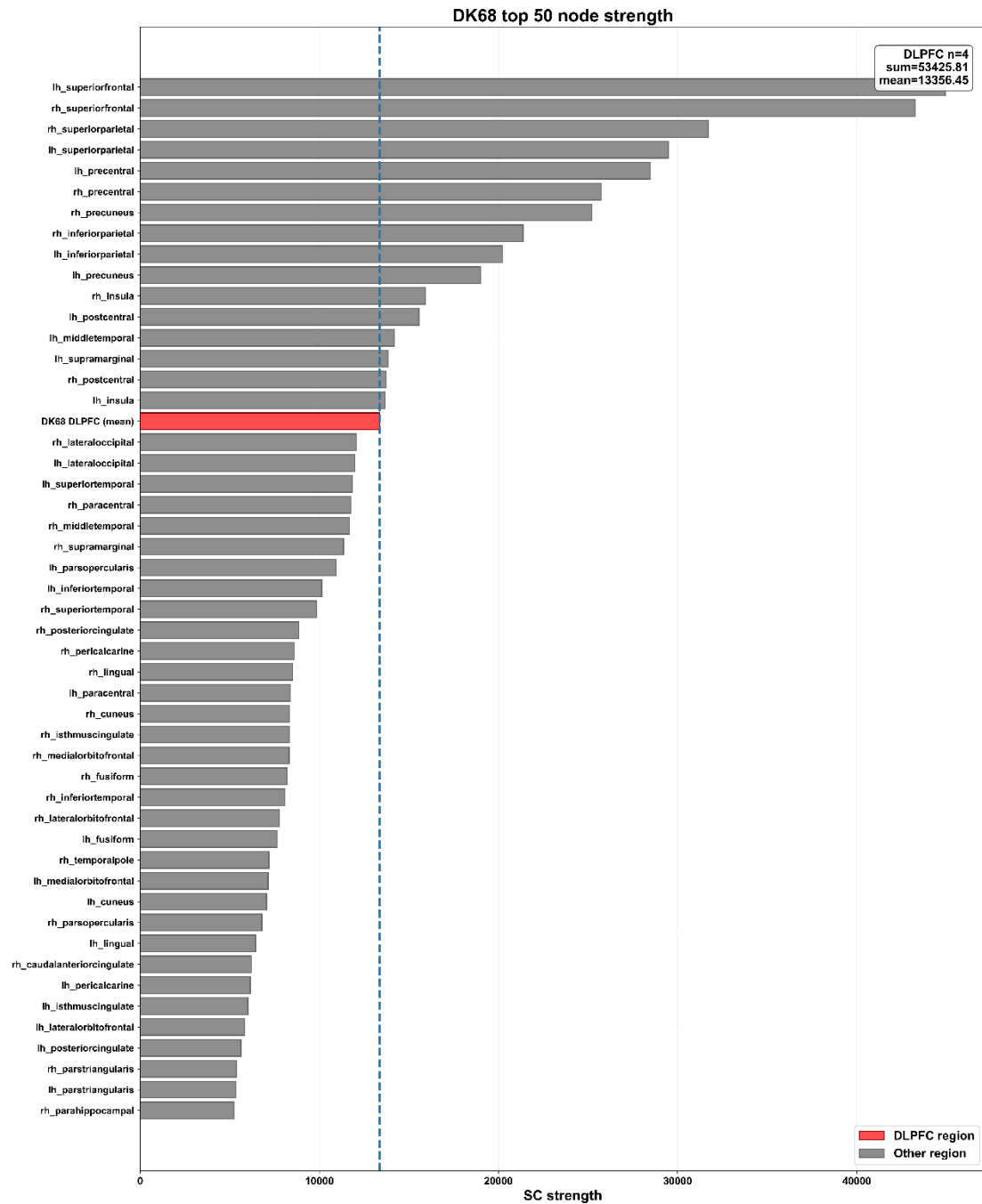


Supplementary Figure 7: Perturbation effects on simulated FC similarity under lambda=1 model configurations. Standardized pre-to-post perturbation effects on simulated FC similarity are shown as Cohen's d. Effects were evaluated for whole-brain FC and FP-DM network FC across different DLPFC perturbation settings. Yellow boxes mark the two key left-excitatory/right-inhibitory perturbations examined in the main analysis. Positive values indicate stronger perturbation-related shifts in the tested similarity metric. **A-B** Effects on similarity to empirical resting-state FC, where negative values indicate reduced rest-like similarity after perturbation. **A** Results in DK68 parcellation. **B** Results in Schaefer 100 parcellation. **C-D** Effects on similarity to empirical working-memory FC, where positive values indicate increased working-memory-like similarity after perturbation. **C** Results in DK68 parcellation. **D** Results in Schaefer 100 parcellation.



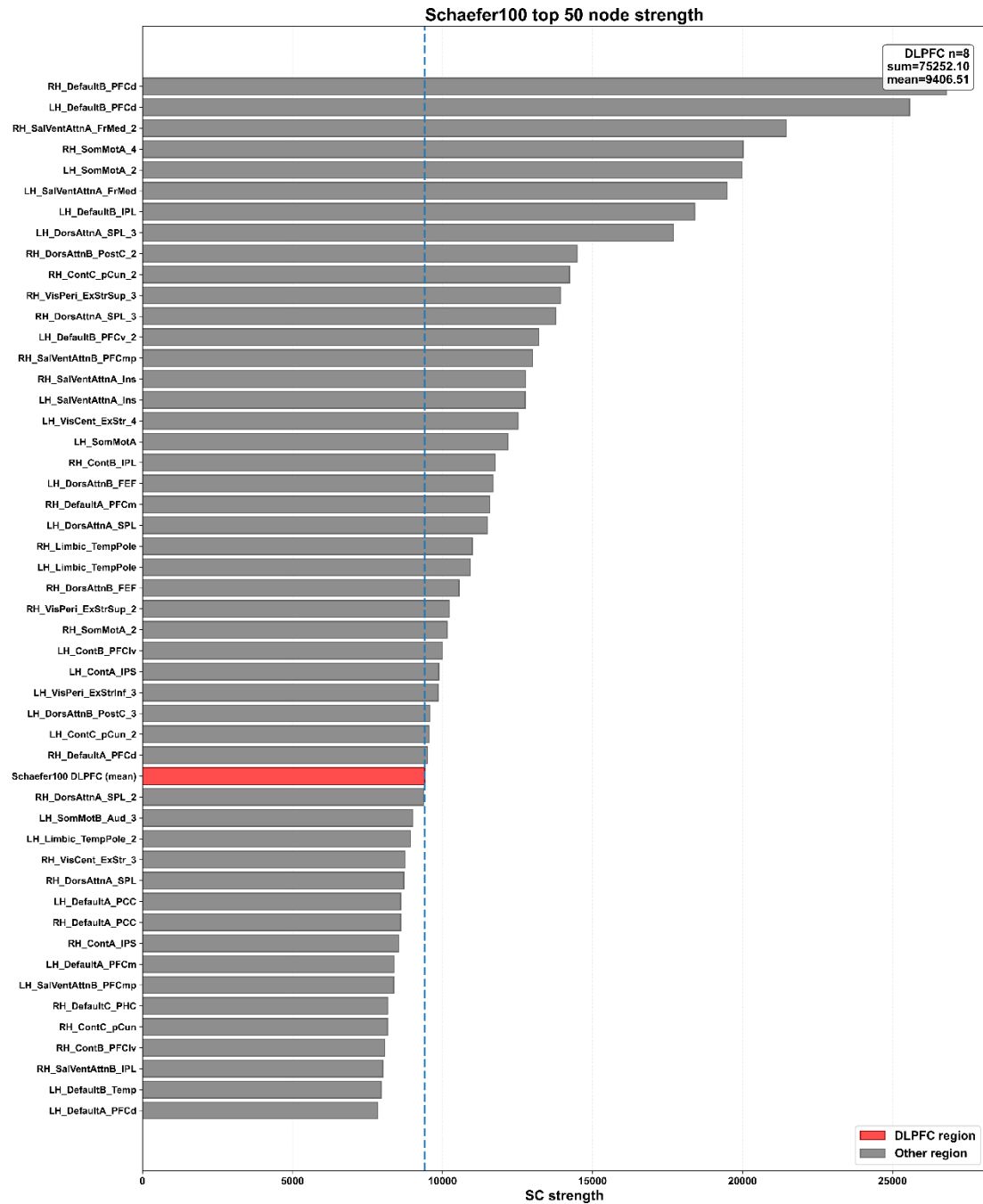
Supplementary Figure 8: Robustness of DLPFC perturbation effects under the Glasser parcellation.

A Perturbation-related decreases in similarity to empirical resting-state FC. **B** Perturbation-related increases in similarity to empirical working-memory FC. Effects are shown for whole-brain FC and FP-DM network FC under the two key perturbation settings, L-E+0.1 & R-I+0.1 and L-E+0.1 & R-I+0.05.



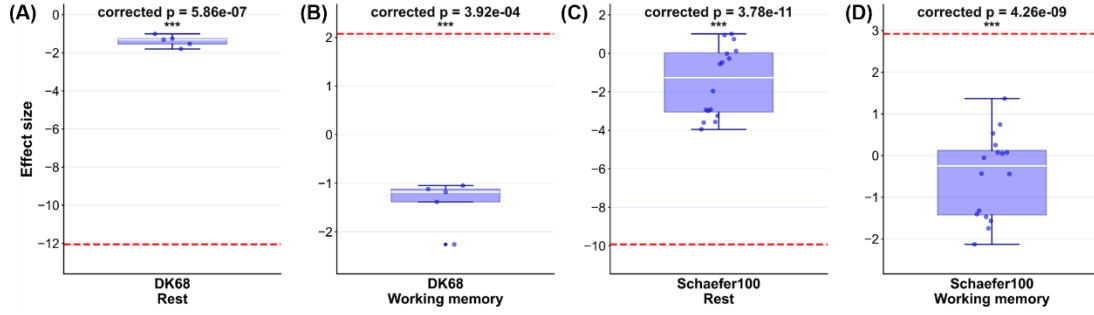
Supplementary Figure 9: Top 50 regions ranked by structural connectivity strength in the DK68 parcellation.

Structural connectivity strength was calculated as the weighted node strength for each cortical region. Regions are ranked in descending order according to structural connectivity strength. The red bar indicates the mean structural connectivity strength of DLPFC regions in the DK68 parcellation, and the blue dashed line marks this mean value. Grey bars indicate non-DLPFC regions. This ranking was used to identify non-DLPFC control regions with structural connectivity strength matched to DLPFC.



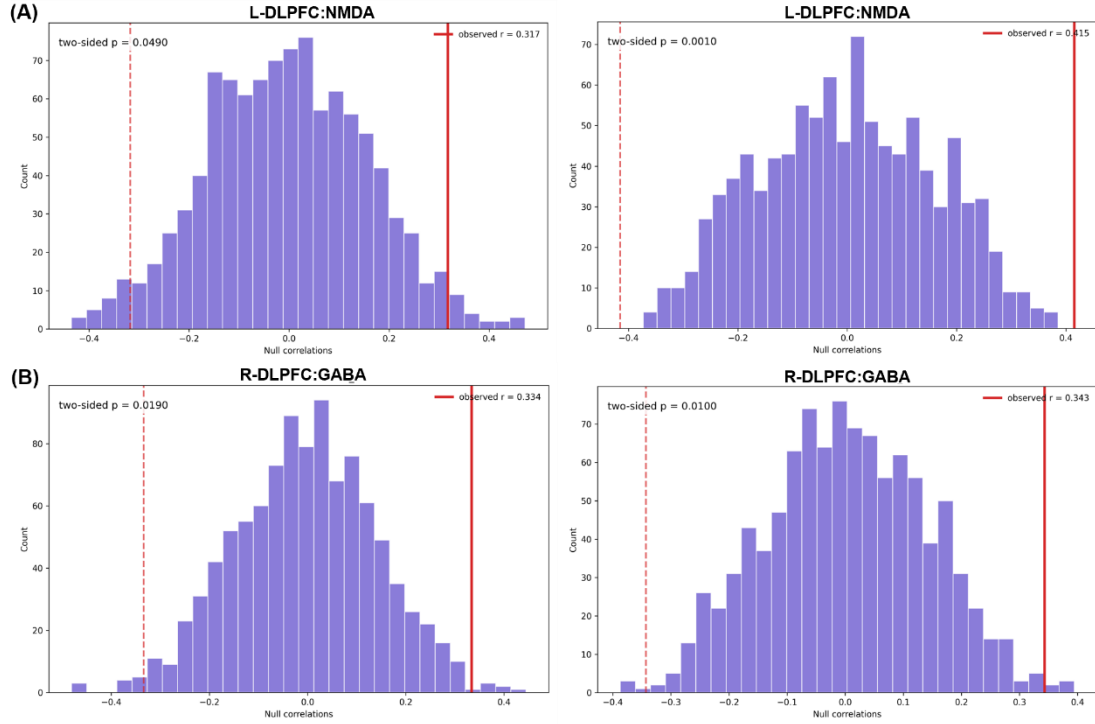
Supplementary Figure 10: Top 50 regions ranked by structural connectivity strength in the Schaefer 100 parcellation.

Structural connectivity strength was calculated as the weighted node strength for each cortical region. Regions are ranked in descending order according to structural connectivity strength. The red bar indicates the mean structural connectivity strength of DLPFC regions in the Schaefer100 parcellation, and the blue dashed line marks this mean value. Grey bars indicate non-DLPFC regions. This ranking was used to identify non-DLPFC control regions with structural connectivity strength matched to DLPFC.



Supplementary Figure 11: Structural-strength-matched control analysis of DLPFC perturbation effects.

Boxplots show the distribution of perturbation effect sizes obtained from non-DLPFC control regions with structural connectivity strength matched to the mean DLPFC strength ($\pm 10\%$). For each candidate control region, the same perturbation setting, L-E+0.1 & R-I+0.05, was applied under the $\lambda=0$ model configuration. The perturbation effect was quantified as Cohen's d of the pre-to-post change in FC similarity across ten fitted parameter sets. Red dashed lines indicate the corresponding DLPFC perturbation effect specified for each condition. (A) DK68 parcellation, similarity to resting-state FC. (B) DK68 parcellation, similarity to working-memory FC. (C) Schaefer100 parcellation, similarity to resting-state FC. (D) Schaefer100 parcellation, similarity to working-memory FC. Corrected p values were obtained using one-sample t-tests comparing the control-region effect-size distribution against the corresponding DLPFC effect, followed by Bonferroni correction across the four comparisons.



Supplementary Figure 12: Spatial permutation analysis of receptor-enriched DLPFC effects.

Spatial autocorrelation-preserving permutation tests were used to evaluate the spatial specificity of the left-DLPFC NMDA-enriched effect (A) and right-DLPFC GABA-A-enriched effect (B). The correlation refers to the Pearson correlation across cortical parcels between the PET-derived NMDA or GABA-A receptor-density map and the corresponding WM-RS difference map of receptor-enriched functional expression. Histograms show null distributions of spatial correlations obtained after randomly rotating the cortical receptor-density maps on the spherical surface. The left and right columns show DK68 and Schaefer100 results, respectively. Red solid lines indicate the observed spatial correlations, and red dashed lines indicate two-sided significance thresholds from the null distributions. Two-sided p values are shown in each panel.