

ORIGINAL ARTICLE

Perturbing whole-brain models of brain hierarchy: An application for depression following pharmacological treatment

Marcel Socoró-Garrigosa¹  | Yonatan Sanz Perl^{1,2} | Morten L. Kringelbach^{1,3,4} | David Erritzoe⁵ | David J. Nutt⁵ | Robin Carhart-Harris^{5,6} | Jakub Vohryzek^{1,3}  | Gustavo Deco¹

¹Center for Brain and Cognition, Computational Neuroscience Group, Department of Information and Communication Technologies, Universitat Pompeu Fabra, Barcelona, Spain

²Department of Physics, University of Buenos Aires, Buenos Aires, Argentina

³Centre for Eudaimonia and Human Flourishing, Linacre College, University of Oxford, Oxford, UK

⁴Department of Psychiatry, University of Oxford, Oxford, UK

⁵Centre for Psychedelic Research, Department of Brain Sciences, Imperial College London, London, UK

⁶Departments of Neurology, Psychiatry and Behavioral Sciences, Weill Institute for Neurosciences, University of California San Francisco, San Francisco, California, USA

Correspondence

Marcel Socoró-Garrigosa and Jakub Vohryzek, Center for Brain and Cognition, Computational Neuroscience Group, Department of Information and Communication Technologies, Universitat Pompeu Fabra, Roc Boronat 138, Barcelona 08018, Spain. Email: msocoro@student.ethz.ch and jakub.vohryzek@upf.edu

Jakub Vohryzek and Gustavo Deco share senior authorship.

Funding information

NEMESIS; ref. 101071900; EU H2020 FET Proactive project Neurotwin grant agreement no. 101017716; Jane Bradley Pettit Foundation; European Research Council Consolidator Grant: CAREGIVING, Grant/Award Number: 615539; Catalan Agency for Management of University and Research Grants Programme, Grant/Award Number: 2017SGR1545; the European Union's Horizon 2020 Research and Innovation Programme, Grant/Award Number: 720270(HBPSGA1) and 785907(HBPSGA2); Danmarks Grundforskningsfond, Grant/Award

Abstract

Determining the scale of neural representations is a central challenge in neuroscience. While localized representations have traditionally dominated, evidence suggests information is also encoded in distributed, hierarchical networks. Recent research indicates that the hierarchy of causal influences shaping functional patterns serves as a signature of distinct brain states, with implications for neuropsychiatric disorders. Here, we first explore how whole-brain models, guided by the thermodynamics of mind framework, estimate brain hierarchy and how perturbing such models enables the study of in-silico transitions represented by static functional connectivity. We then apply this to major depressive disorder, where different brain hierarchical reconfigurations emerge following psilocybin and escitalopram treatments. We build resting-state whole-brain models of depressed patients before and after interventions and conduct a dynamic sensitivity analysis to explore brain states' susceptibility—measuring their capacity to change—and their drivability to healthier states. We show that susceptibility is on average reduced by escitalopram and increased by psilocybin, and that both treatments promote healthier transitions. These results align with the post-treatment window of plasticity opened by serotonergic psychedelics and the similar clinical efficacy of both drugs in trials. Overall, this work demonstrates how whole-brain models

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2025 The Author(s). *Annals of the New York Academy of Sciences* published by Wiley Periodicals LLC on behalf of The New York Academy of Sciences.

Number: DNRF117; Carlsbergfondet; Spanish Research Project, Grant/Award Number: PSI2016-75688-P

of brain hierarchy can inform in-silico neurostimulation protocols for neuropsychiatric disorders.

KEYWORDS

depression, hierarchy, neurostimulation, psychedelics, whole-brain modeling

INTRODUCTION

Understanding how the brain represents information is one of the central challenges in neuroscience. Although not entirely clear, an accepted definition for *neural representation* is a pattern of neural activity that stands for some environmental feature in the internal workings of the brain.¹ However, a fundamental question to the notion of neural representation is at which scale does this representation occur. Historically, neural representation has been closely tied to the activity of single neurons or small clusters of neurons that encode highly specific information. This local approach stems from a wide body of research, including the discovery of place cells in the hippocampus, found to encode an organism's spatial location, or the discovery of neurons in the visual cortex which selectively respond to orientations, edges, or motion.^{2,3}

More recently, modern neuroimaging techniques like functional magnetic resonance imaging (fMRI) or magnetoencephalography (MEG) have enabled researchers to move beyond the local scale and study distributed patterns of activity across the entire brain. Moreover, techniques like diffusion tractography imaging have allowed to map the anatomical connections of the brain, giving rise to the concept of the *connectome*, defined as a comprehensive map of neural connections in the brain on many spatial scales.⁴ Furthermore, advances in graph theory have provided mathematical tools to analyze the properties of these brain networks, giving rise to the field of *network neuroscience*.⁵ For instance, we now know that the brain exhibits *small-worldness*—a balance between local specialization and global integration—and that its organization is modular.^{6,7}

Given these advances, it is reasonable to question the localized nature of neural representations and consider the possibility that the brain also encodes information within highly distributed and hierarchical networks. Indeed, measures that capture the relationships between different brain areas, such as functional connectivity (FC) and structural connectivity (SC), have been shown to identify impaired brain states in various neuropsychiatric disorders.⁸ Some of the examples include dementia,⁹ Alzheimer's disease,¹⁰ posttraumatic stress disorder,¹¹ schizophrenia,¹² and major depression.^{13,14}

Considering brain activity at a higher, network-level scale is, therefore, relevant for diagnosing brain diseases and detecting impaired states of consciousness. Below, we will review how this concept has led to increasingly sophisticated methods for analyzing and modeling personalized whole-brain activity, ultimately capturing the hierarchical nature of brain interactions. We will then examine the advantages of perturbing the constructed models to assess recovery in neuropsychiatric

disorders, which will motivate our computational study on depression recovery following distinct pharmacological treatments.

Despite identifying abnormal brain states in various disorders, the descriptive nature of FC and SC analyses falls short of explaining how brain structure gives rise to functional patterns, an essential link for understanding brain function and dysfunction. Over a decade ago, the promising field of large-scale, whole-brain models emerged to bridge this gap.¹⁵ Whole-brain models leverage SC data, derived from diffusion MRI, to inform the emergent large-scale brain dynamics. Based on a brain atlas, these models simulate interactions between regions by modeling their neural activity, often using mean-field approximations, such as the Wilson–Cowan model,¹⁶ or phenomenological models of coupled oscillators, like the Kuramoto and Hopf models.^{17,18} In other words, they explicitly express whole-brain activity by modeling local dynamics interacting on the structural scaffold, thereby bridging the explanatory gap between the structure and function of brain organization.

By fitting empirical data, these models can provide mechanistic insights into different brain states at a patient-specific level. For instance, we can optimize the anatomical connections to maximize the fit between simulated and empirical FC patterns. However, while FC connectivity is an important observable describing brain dynamics, it only captures the symmetrical relationship between different brain regions and thus fails to represent the hierarchical processing of the brain states under study. This can be overcome by using the *thermodynamics of mind* framework which considers asymmetric connectivity and thus captures the hierarchy of the causal interactions between brain regions.^{19–21}

Borrowing concepts from thermodynamics, this framework acknowledges the relationship between hierarchy and nonreversibility in neural time series. Hierarchical systems reflect a structured flow of information that operates out of thermodynamic equilibrium.^{21,22} In these systems, the asymmetries in information flow break what in physics is known as the *detailed balance*, where flux transitions between the system states disappear.^{23,24} The breaking of the detailed balance establishes an arrow of time, which in the brain manifests as neural signals becoming temporally irreversible. In other words, in highly hierarchical interactions, pairwise neural time series are causally dependent on forward time. This irreversibility or arrow of time can be estimated using temporal asymmetry by taking the difference between forward and reversal time-shifted correlations.¹⁹ (see Materials and Methods below for details). Importantly, the irreversibility estimate is, due to its asymmetry, a robust estimate of the *functional hierarchy*. By incorporating the irreversibility estimates in

the simulated-empirical fitting process, whole-brain models identify the causal, mechanistic generators of the functional hierarchy through the optimized anatomical connectivity.²¹ This is known as generative effective connectivity (GEC), extending on the concept of effective connectivity.²⁵

By using generative whole-brain modeling under the thermodynamics of mind framework, we can create personalized models of specific brain states and uncover the underlying hierarchical causal influences between brain regions. Note that personalization refers to fitting the models to each subject's individual empirical data. However, the interest is often to model the transitions between brain states, for instance, from impaired to neurotypical states in neuropsychiatric disorders.

A framework that allows studying plausible transitions between various brain states using whole-brain models is dynamic sensitivity analysis,²⁶ whose paradigm was introduced in a computational study on transitions between wakefulness and sleep.²⁷ Dynamic sensitivity analysis consists of introducing parameter perturbations in the underlying models of local dynamics. We can interpret such perturbations as in-silico stimulations, according to the underlying model. For instance, different properties can be adjusted such as the type of stimulation (i.e., increasing excitation, modulating frequency, or adding noise, among others), the stimulation intensity, the location (i.e., local or global), and the duration (i.e., long-term, pulse-based or periodic).²⁶ Indeed, driving in-silico brain transitions has been shown beneficial in clinical contexts, such as in the cases of Alzheimer's disease, psychosis, and depression.^{28–30}

Major depressive disorder (MDD) is one of the most debilitating and problematic neuropsychiatric disorders, acutely limiting psychosocial functioning across complex and heterogeneous phenotypes.^{31,32} Epidemiologically, it is estimated to affect 322 million people worldwide, and it is now the largest contributor to global disability and suicide deaths.³³ Moreover, it is expected to become the leading disease burden worldwide by 2030.³⁴

We know that MDD disrupts brain activity at a network and global level. For instance, the default-mode network (DMN) is known to be overactivated to the detriment of other areas like the executive network or the salience network, enhancing rumination and aberrant thinking.³⁵ According to the predictive coding model of MDD, these activity changes could be abnormally reducing the uncertainty of high-order priors about lower-order interoceptive and exteroceptive states.^{36,37} Furthermore, recent studies show that there is an abnormal topography shift in brain activity from unimodal lower-order sensory and motor regions to more transmodal higher-order regions like DMN and the prefrontal cortex.^{36,38,39} The hierarchy of inter-regional interactions, therefore, plays a critical role in MDD, and considering that fact might be relevant for better understanding and exploiting antidepressant medication.

The current gold-standard treatment is combined therapy of psychotherapy and antidepressants, which typically fall into the class of selective serotonin reuptake inhibitors (SSRIs).⁴⁰ However, common antidepressants present serious adverse effects and a modest efficacy, with approximately 30% of patients showing resistance to standard treatments.⁴¹ A promising avenue in the treatment of depres-

sion lies now in serotonergic psychedelics.^{42,43} Evidence suggests that they work mainly but not only through 5-HT_{2A} serotonin receptor (5-HT_{2A}R) agonism, which increases serotonin release and opens a window of plasticity from which supportive psychotherapy could benefit.⁴⁴ This key therapeutic feature is captured in the relaxed beliefs under psychedelics (REBUS) model, whereby psychedelics would alleviate depressive states by relaxing the aberrantly overweighted high-level priors and facilitating the compromised bottom-up information flow, that is, by decreasing the hierarchical constraints.⁴⁵

Using generative whole-brain modeling, a recent computational study suggests that psychedelics reconfigure the hierarchy of brain causal interactions very differently compared to standard SSRIs, namely, by flattening, instead of increasing, the overall global directedness of the hierarchy of brain dynamics.²⁰ These findings align with the REBUS model for psychedelics, and the observed difference in hierarchical reconfiguration suggests that psychedelic- and SSRI-treated brains may respond differently to stimulation protocols, which could be key to the success of combined therapies.

Given the plasticity events enhanced by psychedelics, we expected brain states following psilocybin intervention to be more susceptible compared to the SSRI treatment. However, the similar efficacy in clinical trials encouraged us to hypothesize that both psilocybin and escitalopram can unlock transitions to healthier states. Hence, our aim in this study was to investigate in silico the susceptibility of brain states—understood as their capacity to change—and their drivability to optimal states following psilocybin and escitalopram treatments. We grounded our study on the current understanding that the present drugs induce different hierarchical reconfigurations²⁰ and centered our focus on discovering the perturbational and state-transition implications of such opposing hierarchical reconfigurations.

To answer our research questions, we leveraged a double-blind phase II randomized controlled trial, where resting-state fMRI scans from all patients, following psilocybin and escitalopram treatments, were obtained at baseline and after treatment.^{46,47} As in Deco et al.,²⁰ whole-brain models of each condition were built, where the thermodynamics of mind framework was crucial to capture the hierarchy of brain causal influences of the empirical data in the GEC. Once the subject-specific models were fitted, we established a perturbation protocol following the dynamic sensitivity analysis framework,²⁶ and studied how the perturbations affect the brain hierarchy of each drug differently. Specifically, systematic perturbations were introduced by increasing the level of noise in the regional models, and, motivated by our hypotheses, we defined two measures to assess their effect on static FC: Susceptibility (S), which captures the capacity of the functional pattern to change under a perturbation; and perturbation effectivity for recovery (PER), which captures the similarity of the perturbed pattern and an optimal state extracted from a group of healthy participants. Notably, the current study extends the work of Deco et al.,²⁰ where the same fitted (unperturbed) models were obtained to study the hierarchy of brain causal influences in patients following psilocybin and escitalopram treatments, and Vohryzek et al.,³⁰ where the dynamic sensitivity analysis framework was used to study effects of psilocybin on treatment-resistant depression, by investigating how the

hierarchy of brain causal influences in the different drugs is influenced by external perturbations, a research topic that has not been explored in previous computational studies.

MATERIALS AND METHODS

Empirical data

The trial design and primary clinical outcomes (ClinicalTrials.gov Identifier: NCT03429075) of this study have been previously reported and registered.^{46,47} The research was conducted at the National Institute for Health Research Imperial Clinical Research Facility under the sponsorship of Imperial College London. Ethical approval was obtained from the NHS Research and Imperial College Joint Research and Compliance Office (ID 17/LO/0389), along with approvals from the Health Research Authority and Medicines and Healthcare Products Regulatory Agency. The study was conducted in compliance with a Schedule 1 Drug License issued by the UK Home Office. All participants provided written informed consent and did not receive financial compensation for their involvement. This rigorous regulatory and ethical framework ensured the study's adherence to high standards of clinical research practice and participant protection.

Participants

To participate in the study, individuals were required to have a confirmed diagnosis of unipolar MDD from a general practitioner and a score of 16 or higher on the 21-item Hamilton Depression Rating Scale. Participants were also queried about prior use of psychedelic substances, with 31% of patients in the psilocybin group and 24% in the escitalopram group reporting previous psychedelic experience. Exclusion criteria included a personal or immediate family history of psychosis, any physician-assessed high-risk physical health condition, a history of severe suicide attempts, a positive pregnancy test, contraindications for MRI scanning, contraindications to SSRIs, or any prior use of escitalopram. Notably, treatment resistance was not considered an inclusion or exclusion criterion. All participants underwent telephone screening interviews, provided written informed consent, and completed comprehensive evaluations of their mental and physical medical histories.

Interventions

Of the 59 patients recruited with a confirmed diagnosis of MDD, 30 were randomly assigned to the psilocybin arm and 29 to the escitalopram arm using a random number generator. Consistent with Deco and colleagues²⁰ and similar to Daws and colleagues,⁴⁷ we excluded some of the patients in each arm. Specifically, for the psilocybin arm, one patient was excluded for choosing not to take the daily (placebo) capsules, and due to the COVID-19 UK lockdown, two patients did not

attend the post-treatment session; finally, five patients were excluded due to excessive fMRI head motion. For the escitalopram arm, four patients discontinued due to adverse reactions to escitalopram, one patient reported cannabis use, one patient was lost due to the COVID-19 UK lockdown, and a further three patients were excluded due to excessive fMRI head motion. Hence, the final neuroimaging sample comprised 22 patients in the psilocybin arm (mean age: 41.9 years, SD: 11.0, 8 females) and 20 patients in the escitalopram arm (mean age: 38.7 years, SD: 11.0, 6 females). Prior to treatment, all participants underwent a baseline resting-state fMRI scan with eyes closed. On the first dosing day (DD1), participants in the psilocybin arm received 25 mg of psilocybin, while those in the escitalopram arm received 1 mg of psilocybin, a presumed negligible dose. To maintain blinding, all participants were informed they would receive psilocybin but were unaware of the specific dosage. A second dosing day (DD2) occurred 3 weeks later, during which patients received the same dosages as on DD1. There was no crossover of dosages between the treatment arms. From the day after DD1, patients took daily capsules for a total of 6 weeks and 1 day. In both conditions, patients ingested one capsule per day during the initial 3 weeks and increased the dosage to two capsules per day afterward. The capsule content was either an inert placebo (microcrystalline cellulose) in the psilocybin arm or escitalopram in the escitalopram arm. In the escitalopram arm, patients received 10 mg of escitalopram for the first 3 weeks and a total of 2×10 mg (20 mg) thereafter.

MRI acquisition

Brain imaging was performed on a 3T Siemens Tim Trio at Invicro. Anatomical images were acquired using the Alzheimer's Disease Neuroimaging Initiative, Grand Opportunity (ADNI-GO56) recommended MPRAGE parameters (1-mm isotropic voxels; TR, 2300 ms; TE, 2.98 ms; 160 sagittal slices; 256×256 in-plane field of view; flip angle, 9 degrees; bandwidth, 240 Hz per pixel; GRAPPA acceleration, 2). In this study, eyes-closed resting-state fMRI data were collected with T2*-weighted echo-planar images with 3-mm isotropic voxels. A 32-channel head coil was used to acquire 480 volumes in ~10 min: TR, 1250 ms; TE, 30 ms; 44 axial slices; flip angle, 70 degrees; bandwidth, 2232 Hz per pixel; and GRAPPA acceleration, 2).

Healthy target

Due to the absence of healthy controls in the depression dataset, we accessed Human Connectome Project (HCP) fMRI data from healthy subjects to construct a healthy target. We focused on 100 unrelated subjects from the project Young Adult, as employed in a prior study.⁴⁸ Participants underwent four rs-fMRI sessions of 14 min 30 s with a repetition time of 0.72 s on a 3-T Skyra scanner (Siemens). We focused on DK80 regional fMRI signals corresponding to data from the first scanning session. Acquisition and preprocessing details can be found at the HCP website (<http://www.humanconnectome.org/>).

SC template

Given that the psilocybin and escitalopram trial did not include diffusion MRI scans, we accessed a standard SC template. This connectivity was estimated using probabilistic tractography in a prior study from two dMRI projects of the HCP.⁴⁹

fMRI data processing

The fMRI data underwent preprocessing using a custom pipeline that integrated tools from multiple neuroimaging software packages, including the FMRIB Software Library (FSL)13, Analysis of Functional NeuroImages (AFNI), FreeSurfer, and Advanced Normalization Tools (ANTs). This preprocessing protocol, previously described in detail,⁴⁷ comprised de-spiking, slice-time correction, motion correction, brain extraction, rigid-body registration of functional images to anatomical scans, nonlinear registration to a standard template, scrubbing, bandpass filtering, and nuisance regression. The fMRI DK80 parcellated time series used in the present study were extracted from this preprocessed data.

Empirical measures

Using the DK80 parcellated time series, empirical covariances $C_v^{empirical}$ and FC matrices $FC^{empirical}$ were constructed by computing for each ij -th pair of regional signals $(x_i(t), x_j(t))$

$$C_{vij}^{empirical} = \sum_t (x_i(t) - \langle x_i(t) \rangle) (x_j(t) - \langle x_j(t) \rangle), \quad (1)$$

$$FC_{ij}^{empirical} = \frac{C_{vij}^{empirical}}{\sqrt{C_{vii}^{empirical} C_{vjj}^{empirical}}}, \quad (2)$$

where $\langle \rangle$ indicates averaging across the time series. Similarly, we defined forward τ time-shifted empirical covariances $C_v(\tau)$ and forward τ time-shifted correlation matrices $FC(\tau)$ as

$$C_{vforwardij}^{empirical}(\tau) = \sum_t (x_i(t + \tau) - \langle x_i(t + \tau) \rangle) (x_j(t) - \langle x_j(t) \rangle), \quad (3)$$

$$FC_{forwardij}^{empirical}(\tau) = \frac{C_{vforwardij}^{empirical}(\tau)}{\sqrt{C_{vii}^{empirical} C_{vjj}^{empirical}}}, \quad (4)$$

where τ is the shifting time, which was set to two times the repetition time. Note that shifting makes $C_{vforward}^{empirical}(\tau)$ and $FC_{forward}^{empirical}(\tau)$ asymmetric matrices.

Whole-brain model

We modeled the local dynamics of each brain region with Stuart–Landau oscillators. This model corresponds to the normal form of a supercritical Hopf bifurcation and is widely used to investigate the transition from noisy to oscillatory dynamics.⁵⁰ It has been used to replicate important features of brain dynamics observed in electrophysiology,^{51,52} MEG,⁵³ and fMRI.^{27,54} Mathematically, let us consider a network of N parcellated regions. Each region j is represented as a node in the network, and the connectivity from node k to node j is captured by C_{jk} in the connectivity matrix C . Whole-brain dynamics can then be modeled by N Stuart–Landau oscillators coupled through the connectivity matrix C :

$$\frac{dz_j}{dt} = (a_j + i\omega_j) z_j - |z_j|^2 z_j + \sum_{k=1}^N C_{jk} (z_k - z_j) + \beta \eta_j, \quad (5)$$

where for each region j , the complex variable z_j denotes the state ($z_j = x_j + iy_j$), $\frac{\omega_j}{2\pi}$ is the intrinsic frequency (in the 0.008–0.08 Hz band), $\beta \eta_j$ is additive uncorrelated Gaussian noise with standard deviation β (for all j), and a_j is the bifurcation parameter. Intrinsic frequencies were estimated from regional narrowband blood-oxygen-level-dependent (BOLD) signals by taking their patient-averaged peak frequency. The bifurcation parameter a_j places the system in a noisy regime if $a_j < 0$ and in a stable limit cycle with self-sustained oscillations if $a_j > 0$. The point $a_j = 0$ hence represents the bifurcation point.⁵⁵ Here, we used a slightly negative a_j value very close to the bifurcation point ($a_j = -0.02$), since it has been shown to represent an optimal working point for fitting whole-brain neuroimaging data.⁵⁶ Finally, the fMRI signals were modeled by the real part of the state variables, that is, $x_j = \text{Real}(z_j)$.

Linearization of the model

The proximity to criticality given by the bifurcation parameter allows the linearization of the dynamics around the fixed point $z = 0$, which for $a < 0$ is stable.⁵⁵ This is essential because it allows us to find an analytical solution for the FC matrix. The functional correlations between all pairs of brain regions can be estimated by employing a linear noise approximation. Consequently, the dynamical system of N nodes (Equation 5) can be expressed in a vector format as follows:

$$\frac{dz}{dt} = (a - s + i\omega) \odot z - (z \odot \bar{z}) z + Cz + \eta, \quad (6)$$

where $z = [z_1, \dots, z_N]^T$, $a = [a_1, \dots, a_N]^T$, $\omega = [\omega_1, \dots, \omega_N]^T$, $\eta = [\eta_1, \dots, \eta_N]^T$, $s = [s_1, \dots, s_N]^T$ is the vector containing the strength of each node, that is, $s_j = \sum_i C_{ij}$, $[]^T$ represents the transpose and $\odot$ is the Hadamard element-wise product, that is, $u \odot v = [u_1 v_1, \dots, u_N v_N]$, and $\bar{z}$ is the complex conjugate of z .

Linearization allows dropping high-order terms in Equation (6) and the dynamics of linear fluctuations can be modeled using a Langevin stochastic differential equation

$$\frac{d}{dt}\delta\mathbf{u} = \mathbf{J}\delta\mathbf{u} + \boldsymbol{\eta}, \quad (7)$$

where fluctuations of $\mathbf{x}$ and $\mathbf{y}$ are described by a $2N$ -dimensional vector $\delta\mathbf{u} = [\delta x, \delta y]^T = [\delta x_1, \dots, \delta x_N, \delta y_1, \dots, \delta y_N]^T$, $\boldsymbol{\eta}$ is a white noise $2N$ -dimensional vector, and $\mathbf{J}$ is the $2N \times 2N$ Jacobian of the system evaluated at the fixed point $(\mathbf{x}, \mathbf{y}) = (\mathbf{0}, \mathbf{0})$, given by

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_{xx} & \mathbf{J}_{xy} \\ \mathbf{J}_{yx} & \mathbf{J}_{yy} \end{bmatrix}, \quad (8)$$

where $\mathbf{J}_{xx}, \mathbf{J}_{xy}, \mathbf{J}_{yx}, \mathbf{J}_{yy}$ are $N \times N$ matrices given by $\mathbf{J}_{xx} = \mathbf{J}_{yy} = \mathbf{C} + \text{diag}(\mathbf{a} - \mathbf{s})$ and $\mathbf{J}_{yx} = -\mathbf{J}_{xy} = \text{diag}(\boldsymbol{\omega})$, being $\text{diag}(\mathbf{v})$ a diagonal matrix with a vector $\mathbf{v}$ on the diagonal.^{20,55}

This equation helps to find the temporal evolution of $\delta\mathbf{u}$, from which network statistics can be extracted. For instance, the covariance of fluctuations $\mathbf{C}_v^{\text{simulated}} = \langle \delta\mathbf{u}\delta\mathbf{u}^T \rangle$ is here relevant, since it will allow to define simulated FC. From Equation (7), the differential equation governing $\mathbf{C}_v$ can be deduced

$$\frac{d\mathbf{C}_v}{dt} = \mathbf{J}\mathbf{C}_v + \mathbf{C}_v\mathbf{J}^T + \mathbf{Q}_n, \quad (9)$$

where $\mathbf{Q}_n$ is the noise covariance matrix $\mathbf{Q}_n = \langle \boldsymbol{\eta}\boldsymbol{\eta}^T \rangle$. If $\frac{d\mathbf{C}_v}{dt}$ is set to zero, the stationary covariance can be obtained by solving:

$$\mathbf{J}\mathbf{C}_v + \mathbf{C}_v\mathbf{J}^T + \mathbf{Q}_n = \mathbf{0}. \quad (10)$$

This is a Lyapunov equation that has a unique solution provided that $\mathbf{J}$ is asymptotically stable. It can be solved using the eigen-decomposition of the Jacobian matrix $\mathbf{J}$.^{55,57} Here, it was solved using MATLAB solver *lyap()*.

Note that the stationary covariance gives a measure of how fluctuations from different regions covary at stationarity, that is, when the stable origin $\mathbf{z} = \mathbf{0}$ is reached. The stationary covariance can be used to calculate the forward τ time-shifted covariance, defined as $\mathbf{C}_{v\text{forward}}(\tau) = \langle \delta\mathbf{u}(t + \tau)\delta\mathbf{u}(t)^T \rangle$, using:

$$\mathbf{C}_{v\text{forward}}(\tau) = e^{\tau\mathbf{J}} \mathbf{C}_{v\text{forward}}(0), \quad (11)$$

where $\mathbf{C}_{v\text{forward}}(0) = \mathbf{C}_v$ is the static covariance (i.e., with zero shift). For deeper mathematical insight, see Ponce-Alvarez and Deco.⁵⁵ Since we are only interested in the real part of network statistics, simulated covariance $\mathbf{C}_v^{\text{simulated}}$ and simulated forward time-shifted covariance $\mathbf{C}_{v\text{forward}}^{\text{simulated}}(\tau)$ matrices will refer to the matrices arising from taking the first N rows and N columns of $\mathbf{C}_v$ and $\mathbf{C}_{v\text{forward}}(\tau)$, respectively. Analogously to empirical measures, simulated static $\mathbf{FC}^{\text{simulated}}$ and forward τ time-shifted $\mathbf{FC}_{\text{forward}}^{\text{simulated}}(\tau)$ can be obtained with:

$$\mathbf{FC}_{ij}^{\text{simulated}} = \frac{\mathbf{C}_{vij}^{\text{simulated}}}{\sqrt{\mathbf{C}_{vii}^{\text{simulated}} \mathbf{C}_{vjj}^{\text{simulated}}}}, \quad (12)$$

$$\mathbf{FC}_{\text{forward},ij}^{\text{simulated}}(\tau) = \frac{\mathbf{C}_{v\text{forward},ij}^{\text{simulated}}(\tau)}{\sqrt{\mathbf{C}_{vii}^{\text{simulated}} \mathbf{C}_{vjj}^{\text{simulated}}}}. \quad (13)$$

Model fitting: GEC

Given a coupling connectivity matrix $\mathbf{C}$, the Hopf whole-brain model simulates the local dynamics of the parcellated brain regions, enabling the computation of both static and time-shifted simulated FC matrices through network statistic measures of real-part fluctuations. Here, patient-specific SC matrices were not available, so $\mathbf{C}$ matrices were initialized to a group SC template from the Human Connectome Project.⁴⁹ More specifically, in order to ensure consistency with the prior literature set of parameters, $\mathbf{C}$ was set as a scaled version of the template, ensuring a maximal $\mathbf{C}$ value of 0.2.⁵⁶ Models were then fitted to the empirical data by optimizing $\mathbf{C}$ following a pseudo-gradient procedure. At each iteration, $\mathbf{C}$ weights were refined according to the following updating rule:

$$\begin{aligned} \mathbf{C}_{ij} &= \mathbf{C}_{ij} + \alpha \left(\mathbf{FC}_{ij}^{\text{empirical}} - \mathbf{FC}_{ij}^{\text{simulated}} \right) \\ &+ \zeta \left(\left(\mathbf{FC}_{\text{forward},ij}^{\text{empirical}}(\tau) - \mathbf{FC}_{\text{reversal},ij}^{\text{empirical}}(\tau) \right) - \left(\mathbf{FC}_{\text{forward},ij}^{\text{simulated}}(\tau) - \mathbf{FC}_{\text{reversal},ij}^{\text{simulated}}(\tau) \right) \right), \end{aligned} \quad (14)$$

where $\mathbf{FC}_{\text{reversal},ij}^{\text{empirical}}(\tau)$ and $\mathbf{FC}_{\text{reversal},ij}^{\text{simulated}}(\tau)$ denote the empirical and simulated reversal τ time-shifted FC, analogously computed as the forward time-shifted counterparts but considering the time reversal of the empirical and simulated time series. The learning rates used were $\alpha = 0.0004$ and $\zeta = 0.0001$.

Note that the maximal resemblance between $\mathbf{FC}_{ij}^{\text{empirical}}$ and $\mathbf{FC}_{ij}^{\text{simulated}}$ was sought, but also between $\mathbf{FC}_{\text{forward},ij}^{\text{empirical}}(\tau) - \mathbf{FC}_{\text{reversal},ij}^{\text{empirical}}(\tau)$ and $\mathbf{FC}_{\text{forward},ij}^{\text{simulated}}(\tau) - \mathbf{FC}_{\text{reversal},ij}^{\text{simulated}}(\tau)$, both matrices of temporal asymmetry which estimate the pairwise levels of temporal irreversibility between time series.¹⁹ Since temporal irreversibility is linked to the hierarchy of causal interactions in brain dynamics,^{19,58} including it in model optimization is essential to encode brain hierarchy in GEC matrices. Note also that fitting the time asymmetry matrices is equivalent to fitting only the forward time-shifted correlations.

To check the quality of the fitting, every 100 iterations of the group and subject-level optimization process, a total fitting error ϵ_{total} was calculated, given by:

$$\epsilon_{total} = \frac{1}{N(N-1)} \sum_i^N \sum_{j \neq i}^N \left(FC_{ij}^{empirical} - FC_{ij}^{simulated} + FC_{forward,ij}^{empirical}(\tau) - FC_{forward,ij}^{simulated}(\tau) \right). \quad (15)$$

The optimization loop was exited whenever the error varied less than 0.1% with regard to the previous error computed. The resulting optimized $\mathbf{C}$ matrix is called GEC and contains the effective conductivity values for each existing pair of anatomical connections and not just the dMRI-based density of fibers.²⁰

To accelerate the fitting process, $\mathbf{C}$ optimization was divided into two stages. First, $\mathbf{C}$ was fitted to group average empirical FC matrices, obtaining group GECs for each of the defined conditions: psilocybin arm before treatment (PB), psilocybin arm after treatment (PA), escitalopram arm before treatment (EB), and escitalopram arm after treatment (EA). Second, group GECs were used as initial $\mathbf{C}$ matrices to fit patient-specific models, using patient-specific empirical data. The performance of model fitting was assessed by tracking the total fitting error and computing the correlation between $FC^{empirical}$ and $FC^{simulated}$ as well as between $FC_{forward}^{empirical}(\tau)$ and $FC_{forward}^{simulated}(\tau)$. Figure 1 shows a graphical representation of the pipeline for GEC generation via whole-brain modeling.

Single-site perturbation protocol

Upon computing GEC networks for both sessions in each patient, we started perturbing the corresponding—fitted—whole-brain models. Perturbed models were those arising from the optimization process, with GEC informing the asymmetric connectivity between nodes. In-silico single-site perturbations consisted of increases in the white noise standard deviation β of a particular Stuart–Landau oscillator (see Equation 5). Such disturbance can be interpreted as pushing the oscillations to a more noisy-driven regime. We also refer to the noise level β in this context as the *perturbation* or *stimulation intensity* and we call *perturbation* or *stimulation target* the region or site corresponding to the perturbed oscillator. All models underwent a systematic single-site perturbation protocol, meaning that for every model, all nodes were individually perturbed with a repertoire of intensities ranging from 0.02 to 0.5 in steps of 0.01. Each perturbed whole-brain model generated a perturbed simulated FC, capturing the functional statistical dependencies between regions under perturbation. Two measures were constructed to assess the impact of each perturbation: Susceptibility (S) and perturbation effectivity for recovery (PER).

Susceptibility to perturbation

Susceptibility aims to inform on the brain's *capacity to change* its functional patterns under a given perturbation. Consequently, we

defined susceptibility as a dissimilarity measure between perturbed and unperturbed FCs, assessed in terms of decorrelation:

$$S = 1 - \text{corr} \left(\text{vec} \left(FC_{perturbed}^{\text{simulated}, \Delta} \right), \text{vec} \left(FC_{unperturbed}^{\text{simulated}, \Delta} \right) \right), \quad (16)$$

where $\text{corr}()$ stands for Pearson's correlation and $\text{vec}(\mathbf{A}^\Delta)$ denotes the vector formed by the upper triangular elements of the matrix $\mathbf{A}$. Note that only the upper triangular elements are needed, given the symmetry of FC matrices and the trivial values of self-connections (ones) in any correlation-based FC matrix.⁵⁹

To compare the susceptibility present in psilocybin and escitalopram-induced brain states, we studied, for each patient, the average susceptibility across all possible targets. Note that for such computation, we fixed the stimulation intensity, given its influence on susceptibility.

Perturbation effectivity for recovery

Susceptibility, which captures functional malleability, has the limitation of not reporting whether the change exerted by perturbations is beneficial or not, here understood as enhancing a transition toward a healthier regime. The assessment of in-silico stimulations was, therefore, complemented with another measure, termed PER. This concept has been previously described in other computational studies,²⁸ and aims to capture the effectivity of perturbations in driving the brain toward a healthy state. We constructed a healthy functional target by averaging the resting-state FC of 100 healthy participants. As mentioned earlier, fMRI DK80-parcellated signals from 100 unrelated subjects part of the HCP Young Adult Project were accessed, and empirical FCs were computed as shown in Equation (2). In contrast with susceptibility, PER was defined as a correlation measure:

$$PER = \text{corr} \left(\text{vec} \left(FC_{perturbed}^{\text{simulated}, \Delta} \right), \text{vec} \left(FC_{healthy}^{\text{empirical}, \Delta} \right) \right). \quad (17)$$

Furthermore, to strictly measure the effect that perturbations have in driving healthy transitions, we must consider the baseline similarity of unperturbed FCs to the healthy target. We call this baseline similarity to recovery (BSR) and define it by:

$$BSR = \text{corr} \left(\text{vec} \left(FC_{unperturbed}^{\text{simulated}, \Delta} \right), \text{vec} \left(FC_{healthy}^{\text{empirical}, \Delta} \right) \right). \quad (18)$$

By removing the baseline from PER, we obtain the normalized perturbation effectivity for recovery (nPER) measure:

$$nPER = PER - BSR. \quad (19)$$

Perturbations were considered to drive a healthy transition whenever nPER was positive, that is, whenever the perturbed simulated FC was more similar to the healthy target than its unperturbed counterpart.

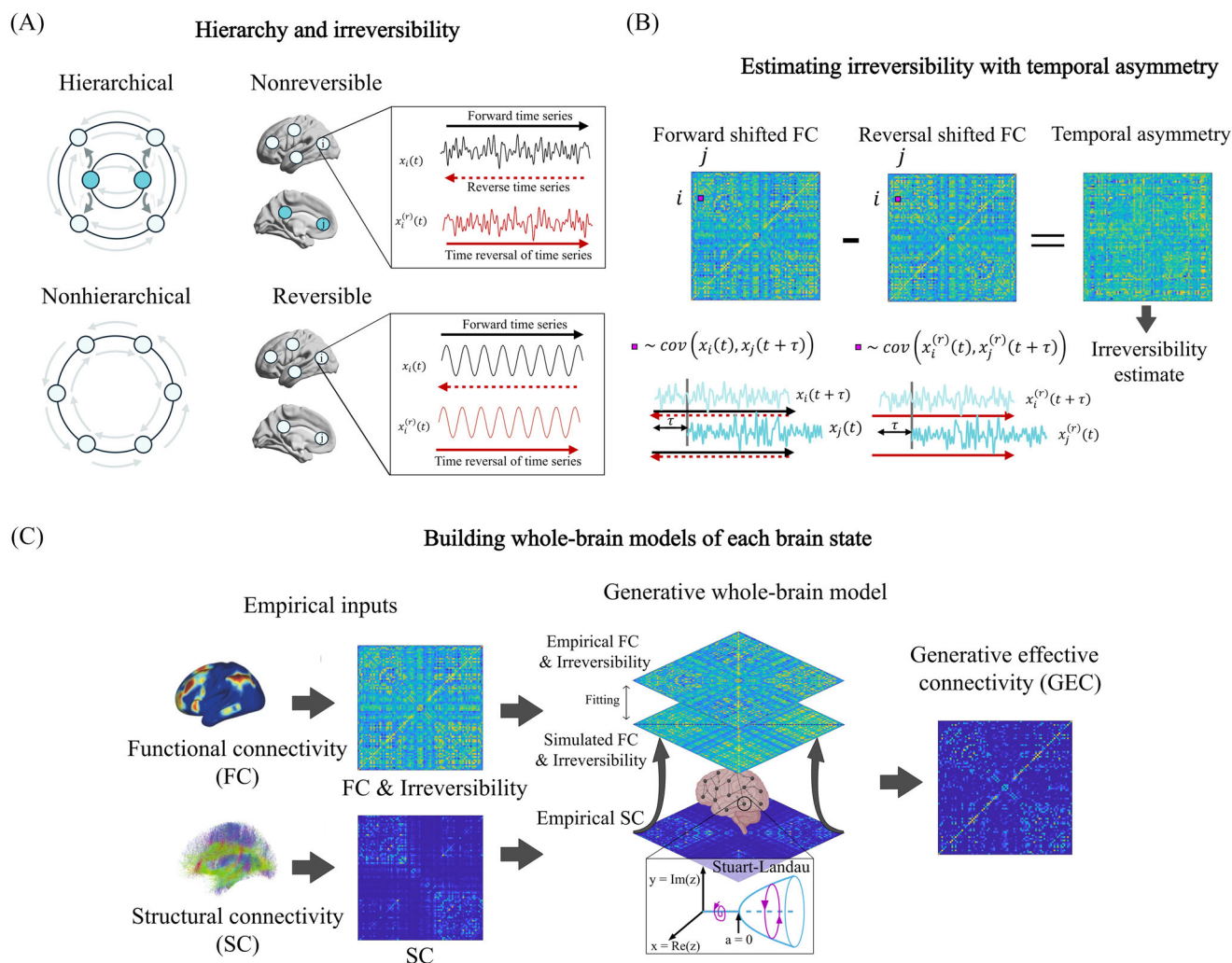


FIGURE 1 Thermodynamics of mind framework. The figure shows how the principles from thermodynamics can be used to infer the hierarchy of brain states. (A) The level of hierarchy in a brain network is deeply connected to temporal irreversibility of the produced neural time series. In hierarchical networks, information flows asymmetrically between its nodes, with intrinsically driven nodes (turquoise) driving the (extrinsic) activity of lower regions (light blue). Thermodynamically, they break the detailed balance, establishing an arrow of time. In other words, the produced time series become nonreversible. Nonhierarchical networks, instead, do not break the detail balance, making the generated signals reversible in time. (B) The irreversibility of a network can be estimated by constructing a matrix of the temporal asymmetry between regional time series. For each interaction, temporal asymmetry can be defined as the difference between forward and reversal time-shifted correlations. (C) Whole-brain modeling can be used to identify the causal, mechanistic generators of functional hierarchy. The model uses local dynamics (e.g., Stuart–Landau oscillators), initially informed by the empirical structural connectivity to fit the empirical functional connectivity and irreversibility. The optimization adjusts the weights of the anatomical connections, producing a generative effective connectivity (GEC) that estimates the generators of the functional hierarchy.

Since we want to elucidate the differences between psilocybin and escitalopram treatments in terms of how beneficial can perturbations become, we focused on identifying the healthiest transition achievable for each patient in each condition—across all regions perturbed and all stimulation intensities applied. Mathematically, this was done by determining the maximum PER in each patient (i.e., $\max_{i,\beta}(\text{PER})$). The corresponding stimulation site and stimulation intensity were collectively referred to as the patient's most *optimal* or *effective* perturbation. The general framework of the perturbation protocol and how susceptibility and PER can be used to assess brain states under perturbation is illustrated in Figure 2.

Multi-site perturbation protocol

While asking the question of how pharmacological treatments boost the therapeutic benefits of neurostimulation for depression, it is reasonable to explore the simultaneous stimulation of multiple regions rather than constraining our in-silico protocols to single-site approaches. After all, depressive states have been linked to aberrant activity in brain networks like the DMN,³⁵ suggesting that multi-site protocols might outperform local stimulations at single regions. To explore this, we employed a multi-site stimulation protocol. Given that the number of possible multi-site configurations scales exponen-

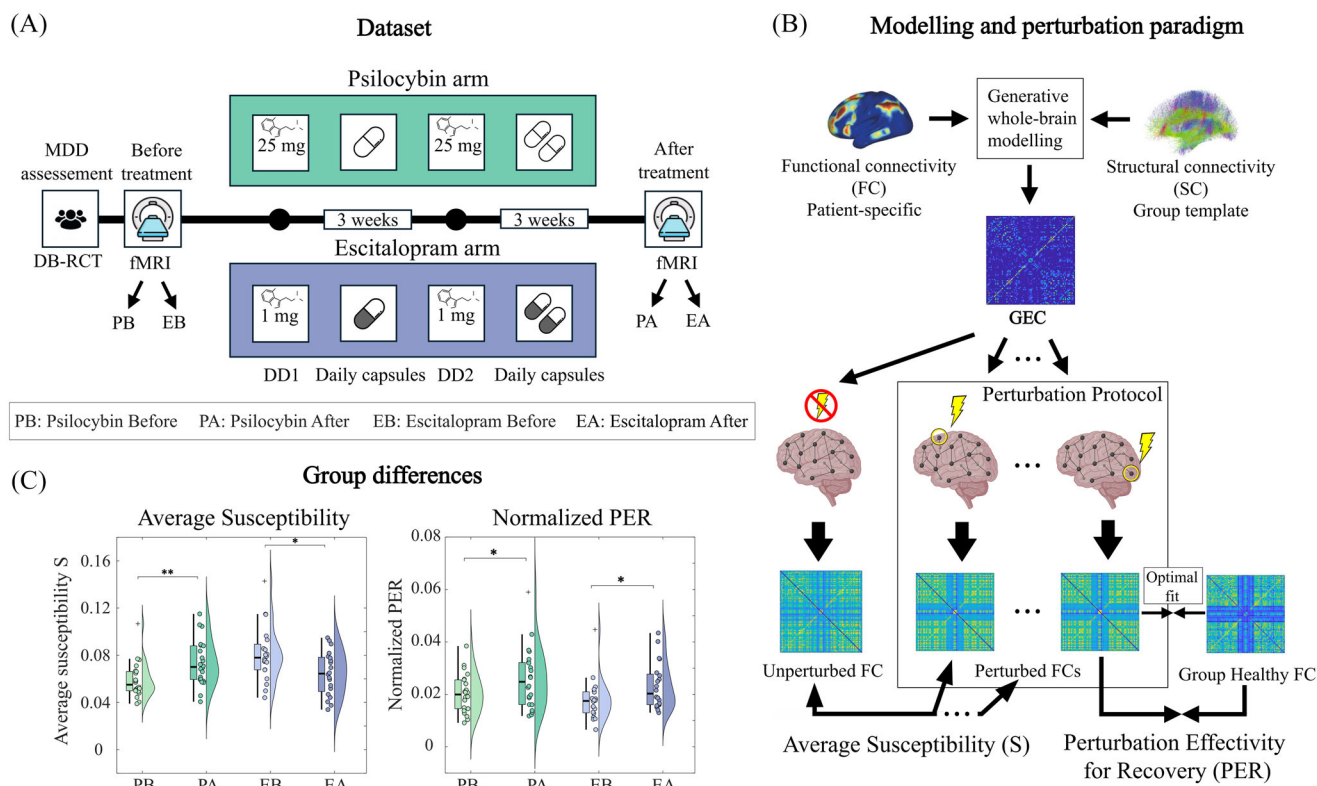


FIGURE 2 Perturbing whole-brain models of depression under different pharmacological interventions. (A) Neuroimaging resting-state data was acquired before and after a double-blind phase II randomized controlled trial (DB-RCT) comparing psilocybin therapy with escitalopram. (B) Whole-brain models were built and perturbed for each treatment arm, before and after pharmacological intervention. For each patient, a whole-brain model was first fitted to the empirical functional connectivity (FC) data via the optimization and generation of a generative effective connectivity (GEC) matrix capturing the asymmetric effective conductivity between regions. Then, patient whole-brain models informed by GECs were systematically perturbed, leading to a perturbed FC. We defined susceptibility (S) as a dissimilarity measure between unperturbed and perturbed simulated FC, and perturbation effectivity for recovery (PER) as a similarity measure between perturbed simulated FC and a healthy FC target. (C) Statistical tests of susceptibility and PER between treatment arms before and after treatment elucidate group differences.

tially with the number of regions considered, we employed a greedy perturbation algorithm, which constructed optimal configurations by iteratively adding the single-site perturbations that jointly maximize PER.

The greedy level was set to 20 areas, and the algorithm was run using PER maximization. For each greedy level, all eligible areas—all of them for the first level—were perturbed with a stimulation intensity repertoire ranging from 0.01 to 0.1 in steps of 0.01. The pair of area and noise level showing the highest PER value was selected, with that area being removed from the eligible set. Note that noise levels above 0.1 were here not considered to reduce computation time, after realizing that the highest PER values occur for rather small perturbations (Figure S5).

Statistical analysis

For a given measure of interest, statistical tests were performed to assess statistical significance between conditions. Essentially, such tests involved assessing the effects of treatment (before vs. after), as well as comparing differences between treatment arms (psilocybin

vs. escitalopram). Due to the relatively small sample size (22 patients for psilocybin and 20 for escitalopram), nonparametric Wilcoxon tests were conducted, not assuming Gaussian distributions in the studied distributions. Paired Wilcoxon tests were used to study treatment effects, while unpaired Wilcoxon tests were employed to assess differences between arms. For both cases, Wilcoxon tests were implemented using 5000 permutations to approximate the null hypothesis distribution from the data. Statistical significance was declared for p -values below 0.05 ($p < 0.05$). In the figures, three different levels of significance are denoted using the following asterisk code: *** for $p < 0.001$, ** for $p < 0.01$, and * for $p < 0.05$. Moreover, we use the abbreviations PB, PA, EB, and EA to refer to the conditions: psilocybin arm before treatment, psilocybin arm after treatment, escitalopram arm before treatment, and escitalopram arm after treatment, respectively.

RESULTS

In this work, we investigated the effects of perturbations in the depressed brain following psilocybin and escitalopram treatments. More specifically, we analyzed the susceptibility of functional brain

states to perturbations and their drivability to optimal states following both pharmacological interventions. The dataset analyzed originates from a double-blind phase II randomized controlled trial comparing psilocybin and escitalopram, where resting-state fMRI scans were taken before and after pharmacological intervention, as shown in Figure 2A (see details of the trial in Materials and Methods).

Generative whole-brain modeling and brain hierarchy

For each patient scan, whole-brain modeling, constrained by the SC, enabled the estimation of a GEC, capturing the resting-state brain dynamics. Following the DK80 parcellation, whole-brain models comprised 80 Stuart–Landau oscillators, and they were initially informed by an SC template, capturing the anatomical connections between regions. These connections were iteratively updated to make the simulated activity fit the patient-specific empirical data, which included static FC and estimates of pairwise temporal irreversibility. The latter was motivated by the thermodynamics of mind framework, which links pairwise temporal irreversibility between neural time series to functional brain hierarchy.²¹ Here, temporal irreversibility was estimated using temporal asymmetry, defined by the difference between forward and reversal time-shifted correlations and in line with the INSIDEOUT framework.¹⁹ By including estimates of pairwise irreversibility in the fitting process, the produced GEC captures the hierarchy of the causal influences between regions. Figure 1A shows the relationship between hierarchy and irreversibility. Namely, in hierarchical networks, where information flows asymmetrically between nodes, the breaking of the detailed balance establishes an arrow of time which can be observed as nonreversible time series. In contrast, networks with totally flat hierarchies produce symmetric (reversible) signals. Figure 1B illustrates how temporal irreversibility can be estimated using temporal asymmetry. Last, the process of employing generative whole-brain models to capture the hierarchy of specific brain states is shown in Figure 1C.

Whole-brain modeling under perturbation

Following the concept of dynamic sensitivity analysis, optimized whole-brain models informed by patient-specific GECs were then systematically perturbed by increasing the amount of noise in the local oscillations of perturbed nodes. The aim was, in the context of depression, to figure out under which brain states can the brain benefit more from external perturbations, treating the resulting perturbed FC as the output of interest. We first ran single-site perturbations in all nodes with a range of stimulation intensities and later explored the power of multi-site perturbation. We constructed two measures to assess the impact of perturbations on FC, susceptibility, and perturbation effectivity for recovery. The former measures the dissimilarity of the perturbed FC with respect to the unperturbed FC, reporting hence the capacity of the brain's activity to change under perturbation, and the

latter measures the proximity of the perturbed FC with respect to a healthy functional target, extracted from the average FC of 100 healthy unrelated participants.⁴⁸ Figure 2B illustrates the perturbation protocol as well as how susceptibility and PER measures were defined. By performing statistical analyses, we were able to quantify the effects of psilocybin and escitalopram treatments in these measures (Figure 2C).

Susceptibility to perturbation

Susceptibility was computed for all single-site perturbations. To extract statistics across patients, we computed the average susceptibility across all perturbed targets (fixing the stimulation intensity) for the different conditions (Figure 3A). Results show that for a stimulation intensity $\beta = 0.04$, psilocybin significantly increases susceptibility ($p < 0.01$, paired Wilcoxon test with 5000 permutations), whereas escitalopram significantly decreases it ($p < 0.05$, paired Wilcoxon test with 5000 permutations). The significance of this trend was analyzed across stimulation intensities, showing that the psilocybin trend is stronger than escitalopram's one and that significance is guaranteed for intensities between 0.02 and 0.05 (Figure S3). Figure 3B shows the number of areas increasing susceptibility following psilocybin and escitalopram treatments. In line with prior results, psilocybin proves to increase susceptibility in significantly more areas than escitalopram ($p < 0.001$, unpaired Wilcoxon test with 5000 permutations), where perturbations rather tend to decrease it. Finally, the prevalence of susceptibility increases across patients is rendered region-wise for each perturbed area (Figure 3C).

Perturbation effectivity for recovery

As with susceptibility, PER was computed for every single-site perturbation. Here, however, the focus was on identifying the perturbation (intensity and location) with the highest PER, that is, driving the simulated FC the closest to the healthy target. Figure 4A shows the distributions of PER in opposition to the BSR for each group and session. For all conditions, PER values are significantly higher than BSR ($p < 0.001$, paired Wilcoxon tests with 5000 permutations), proving that if properly tailored, perturbations significantly enhance the similarity to the healthy state. Figure 3B shows the distributions of normalized PER (nPER), obtained by taking out BSR from PER. Psilocybin and escitalopram-treated brains show significantly higher nPER values for their optimal perturbations ($p < 0.05$, paired Wilcoxon tests with 5000 permutations), showing that pharmacological treatment enhances the capacity to benefit from perturbation. Finally, Figure 3C shows nPER distributions for multi-site perturbations in opposition to the single-site approach. Significantly higher nPER values are achieved by multi-site perturbations ($p < 0.001$, paired Wilcoxon tests with 5000 permutations), proving that multi-site perturbations can be more beneficial than most optimal single-site solutions if the set of stimulation targets and intensities is properly selected.

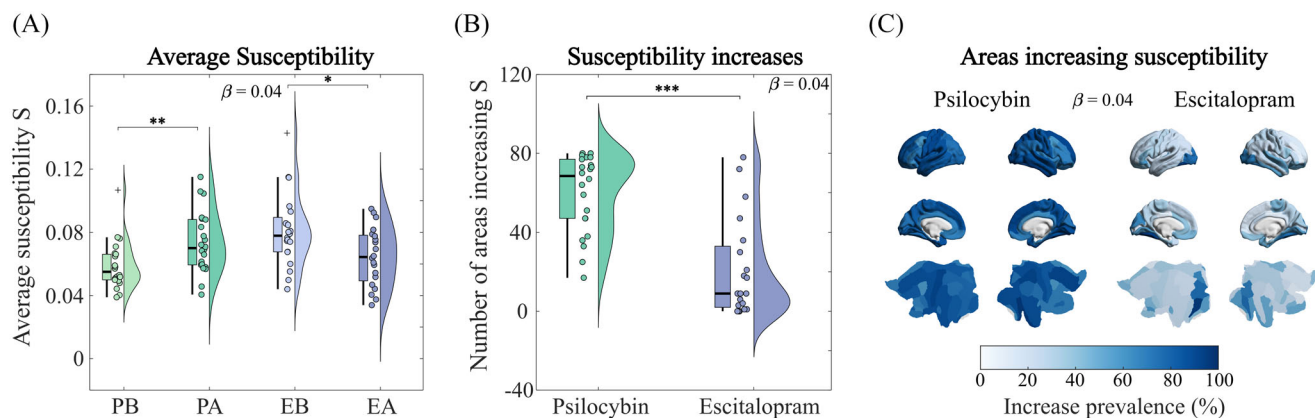


FIGURE 3 Opposite pharmacological reconfigurations in susceptibility. (A) For a fixed stimulation intensity $\beta = 0.04$, the average susceptibility (S) across perturbation sites is shown for each condition. Psilocybin significantly increases average susceptibility (** $p < 0.01$, paired Wilcoxon test with 5000 permutations), whereas escitalopram significantly decreases it (* $p < 0.05$, paired Wilcoxon test with 5000 permutations). (B) The number of areas increasing susceptibility is plotted for each treatment arm. Psilocybin significantly differs from escitalopram treatment (** $p < 0.001$, unpaired Wilcoxon test with 5000 permutations), with a tendency to increase susceptibility in more areas. (C) The renderings show the areas that increase susceptibility more consistently across patients.

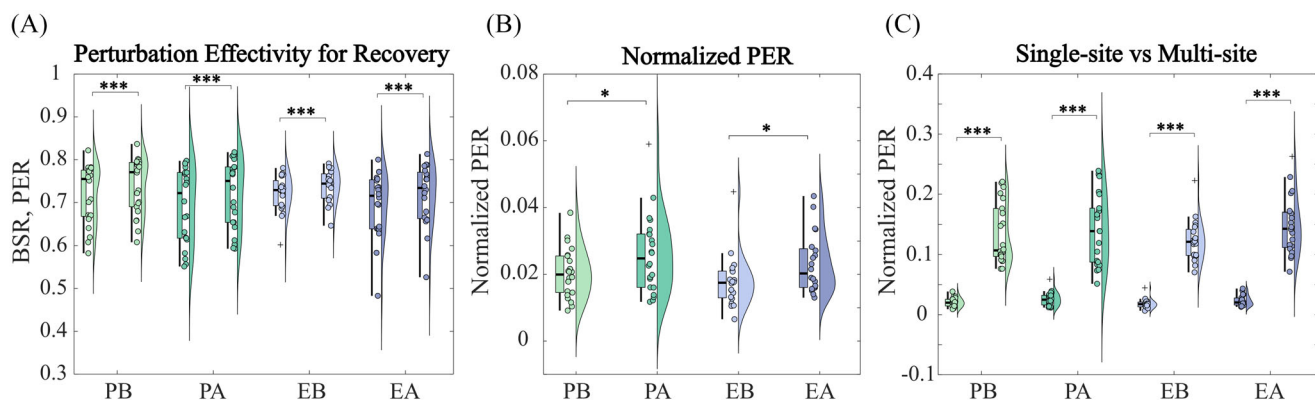


FIGURE 4 Healthier transitions unlocked by psilocybin and escitalopram. (A) The figure shows, for each condition, the baseline similarity to recovery (BSR, left) versus the highest perturbation effectivity for recovery (PER, right) achieved in each patient. These distributions remain significantly different for all conditions (** $p < 0.001$, paired Wilcoxon tests with 5000 permutations). (B) The normalized PER (nPER) is plotted for each condition. Both psilocybin and escitalopram significantly increase this quantity (* $p < 0.05$, paired Wilcoxon tests with 5000 permutations). (C) For each condition, the highest nPER value in each patient's single-site perturbation protocol is plotted (left) against the nPER value of the corresponding 20-level-greedy multi-site perturbation. Multi-site perturbation reaches significantly higher nPER values (** $p < 0.001$, paired Wilcoxon tests with 5000 permutations).

Most optimal stimulation targets

We finally examined the optimal stimulation targets following pharmacological intervention, according to the most effective single-site stimulations and the greedy multi-site solutions (Figure 5). Figure 5A shows the prevalence across patients of optimal single-site stimulation targets, colored on coronal and axial slices of the brain. The exact prevalence can be observed in Figure 5B. Overall, we found that the right amygdala is the most effective stimulation target, present in 50% of optimal single-site perturbations following both treatments. The nucleus accumbens covered the other 50%, split between right and left areas. Figure 5C shows the prevalence across patients of 20-level-greedy multi-site stimulation targets, colored on coronal and axial slices of the brain. Figure 5D shows the same data in a bar plot, focus-

ing on the first five greedy levels and displaying only areas with a prevalence greater than 30% (the equivalent analysis considering all 20 greedy levels is presented in Figure S7). The most frequently recruited areas are again the right and left nucleus accumbens as well as the right amygdala. Moreover, the right medial orbitofrontal cortex (mOFC) is the most frequently recruited cortical target.

DISCUSSION

In this work, our aim was to study in silico the susceptibility of brain states to perturbations and their drivability to optimal states following psilocybin and escitalopram treatments. Our hypothesis was that psilocybin should enhance susceptibility and improve the drivability to

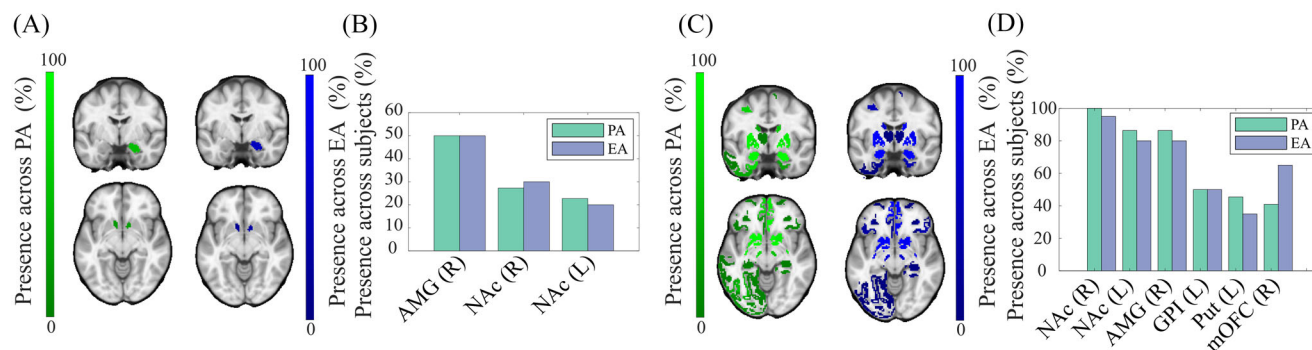


FIGURE 5 The nucleus accumbens and the right amygdala as optimal targets in single- and multi-site stimulations. (A) The presence of stimulation targets in optimal single-site perturbations is shown in coronal (top) and axial (bottom) slices of the brain following psilocybin (left) and escitalopram treatments. (B) The bar plot displays the same information as A). (C) The presence of stimulation targets in 20-level-greedy multi-site perturbations is shown in coronal (top) and axial (bottom) slices of the brain following psilocybin (left) and escitalopram treatments. (D) The bar plot shows the presence of stimulation targets for the first five greedy levels, displaying only areas with values above 30%.

healthier brain states, whereas escitalopram should on average constrain functional malleability but still enhance healthy transitions for specific stimulations. We addressed our research question by first fitting whole-brain models to the patient’s empirical data, and then systematically perturbing the fitted models by increasing the noise in the modeled local dynamics. Single-site and multi-site stimulation protocols were explored, and we assessed their impact on static FC via two distinct measures: susceptibility, which quantifies the ability of perturbations to change the FC pattern, and PER, which measures the proximity of perturbed FC to a healthy target (where an nPER was also defined strictly measuring the impact of perturbations in driving healthy transitions).

Overall, our findings suggest that psilocybin enhances the susceptibility of brain states to perturbations and improves their drivability to optimal states, whereas escitalopram generally constrains functional malleability but still facilitates healthy transitions under specific stimulations. We also demonstrate that the most optimal stimulation targets to achieve healthy transitions occur in subcortical areas like the amygdala and the nucleus accumbens, and we show that multi-site stimulation protocols can significantly outperform single-site strategies. Crucially, our analysis involved systematically perturbing whole-brain models that were fitted not only to patients’ static FC but also to an estimate of their functional hierarchy. Given that psilocybin and escitalopram have been associated with different brain hierarchical reconfigurations, the present methodology constitutes a promising avenue for the in-silico assessment of brain state transitions following pharmacological treatment for depression, with potential applications in other neuropsychiatric disorders.

Opposite overall reconfiguration in susceptibility by psilocybin and escitalopram

The analysis of susceptibility in single-site perturbations revealed an opposite pharmacological trend: psilocybin increases the average susceptibility to perturbation (for a fixed stimulation intensity), whereas

escitalopram diminishes it (Figure 3A). For this analysis, we crucially compared single-site stimulations under the same fixed stimulation intensity, since we wanted to study the exclusive dependency of susceptibility on pharmacological treatment (the dependency of susceptibility on stimulation intensity can be observed in Figure 54). In line with the first result, we also found that the number of areas where susceptibility increases with treatment is larger in the psilocybin arm than in the escitalopram arm, where most single-site perturbations diminish it (Figure 3B,C). This further confirms the opposite reconfiguration exerted by psilocybin and escitalopram in susceptibility to perturbation.

Such an opposite effect aligns with the known differential effects on the pharmacology of the two drugs. On the one hand, psilocybin is known to act mainly through 5-HT_{2A}R,^{54,60,61} initiating multi-level plasticity events.⁶² On the other hand, escitalopram is an SSRI. These drugs constitute the first-line pharmacological treatment for depression and are thought to act on the serotonin transporter, where psilocybin has no appreciable affinity.⁶³

From our definition of susceptibility, we interpret its increase following psilocybin treatment as consistent with an opening of a plasticity window.^{44,64} Conversely, the reduction observed in the escitalopram arm suggests that it constrains functional malleability, limiting the brain’s capacity to change under perturbations. As mentioned in the introduction, these results align with a recent computational study where different reconfigurations of the brain’s overall hierarchy were found with psilocybin and escitalopram.²⁰ Specifically, the authors used novel topological measures adapted to directed networks to show that psilocybin flattens the brain hierarchy, whereas escitalopram increases it. Bringing it together with the present study, we demonstrate that the flattening of functional hierarchy can be related to increases in susceptibility to model perturbations, and that conversely, enhanced hierarchies difficult the ability of brain functional activity to change, on average, to local perturbations. This further supports the REBUS model, where psychedelics would open a therapeutic window of plasticity by flattening the brain’s hierarchy.⁴⁵

The optimal transitions are found after both interventions

Systematic single-site perturbations were further assessed in terms of their ability to recover healthy FC patterns, quantified through PER values. We first showed that perturbations, if properly tailored in stimulation site and intensity, lead to significantly higher PER values compared to the unperturbed BSR values (Figure 4A). This was consistent across both treatment arms, before and after pharmacological intervention. This suggests that different pharmacological treatment interventions can meaningfully benefit from additional *in-silico* stimulations, highlighting the relevance of combined therapies where pharmacological intervention is accompanied by other treatments, like transcranial direct current stimulation (tDCS) or other neurostimulation solutions.^{64,65}

We then computed the nPER values for each optimal perturbation, to strictly measure the influence of perturbations in driving healthier functional states (Figure 4B). Results demonstrate that after pharmacological intervention, nPER values significantly increase. This means that pharmacological intervention, whether psilocybin or escitalopram, enhances the therapeutic effectiveness of optimal perturbations, suggesting a synergistic effect. In the case of escitalopram, this can seem counterintuitive, especially after showing that it constrains, on average across sites, the ability of perturbations to impact FC. However, it is crucial to notice that even if susceptibility is generally decreased, some areas still experience susceptibility increases (Figure 2B,C). Since we are looking for the most optimal perturbation in each patient, the stimulation targets can and do correspond to such a minority of areas. Put together, escitalopram manages to reconfigure the brain in such a way that carefully selected perturbations in key regions still lead to meaningful healthy transitions, as meaningful as those found following psilocybin treatment. This aligns with the fact that both drugs manage to reduce depressive symptoms similarly.^{20,46} Notably, neurostimulation combined therapy has largely been studied with SSRIs rather than psychedelics, with some studies finding that SSRIs like citalopram enhance the neuroplastic after-effects of anodal tDCS.^{65,66} To this respect, our results suggest that escitalopram can enhance neurostimulation to drive functional patterns to healthier states, but since the functional malleability is constrained to the average perturbation, effective stimulations must target specific key regions. Last, the synergistic effect of psilocybin on unlocking healthier functional transitions may relate to the general increase in susceptibility. However, it is key to acknowledge that increased malleability can also imply increased vulnerability toward reaching unhealthy states. We here hypothesize that personalized stimulation protocols or supportive psychotherapy may represent the needed set of perturbations to effectively exploit the enhanced functional malleability, stimuli that might be too complex, especially in the case of psychotherapy, to be captured in this computational essay. Along with this idea, it might be that psilocybin is differentially beneficial, compared to escitalopram, in a longer term after treatment completion, given that its neuroplastic-associated events can last for weeks.^{67,68}

Given that depression alters brain function at the level of networks, that is, by hyperactivating the DMN,³⁵ we further compared the therapeutic potential of multi-site perturbations with that of the single-site approach. Since the number of possible multi-site configurations scales exponentially with the number of considered nodes, we simplified the protocol by leveraging a 20-level-greedy multi-site approach based on PER maximization, after verifying that PER values saturate once approximately 20 areas are considered (Figure S6). As expected, we found that multi-site stimulations significantly reach higher nPER values (Figure 4C), indicating that healthier FC patterns can be restored. This result aligns with the body of research supporting bilateral for transcranial magnetic stimulation (TMS)^{69,70} as well as multichannel tDCS for the treatment of depression,⁷¹ as well as the evidence showing that MDD disrupts large-scale brain networks.³⁵ It is important to note, though, that our results concern also subcortical areas, which might be more challenging to access with noninvasive brain stimulation (NIBS) techniques like TMS or tDCS.

The most optimal stimulation targets

Finally, potentially more relevant for the clinical setting, we analyzed the targets where optimal perturbations occur following pharmacological intervention (i.e., where higher PER values are reached) (Figure 5). We found that following psilocybin and escitalopram treatments, these are either the right amygdala or the right and left nucleus accumbens (Figure 5A,B). The nucleus accumbens is known to be involved in the reward circuitry and to play a crucial role in MDD patients with anhedonia.^{72,73} Our results align with existing literature and support the recent interest in targeting this area through deep brain stimulation.^{74,75} Major depression has also been linked to alterations in the amygdala, which is associated with emotion regulation, among others.^{76,77} Our study suggests that stimulation in this area could notably improve depressive symptoms, a strategy that has been recently implemented using real-time fMRI neurofeedback.⁷⁸ The recruitment of only the right amygdala suggests lateralization within the amygdala. In respect to this, the right amygdala is more commonly associated with automatic emotional processing, while the left amygdala is more linked to controlled mood regulation.^{79,80}

The targets recruited in 20-level-greedy multi-site perturbations align with single-site perturbation results (Figure 5C,D), with the right amygdala, and the nucleus accumbens (left and right) showing the highest presence across patient configurations (following both psilocybin and escitalopram treatments). Other recruited areas are the left globus pallidus internus, the left putamen, and the right mOFC, which shows up as the most optimal cortical target. This is especially relevant given the advantages of NIBS. Coherently with our results, the mOFC is related to reward and its reduced FC has been previously related to depression.⁸¹

Overall, this analysis shows that similar optimal stimulation targets are found following psilocybin and escitalopram treatments, with patient-specific differences strengthening the importance of personal-

ization in neurostimulation solutions, a concept already underscored for MDD given its heterogeneity.⁸²

Limitations and future directions

The main limitation of this study is the scale of the data. fMRI data from only 42 patients was available (22 in the psilocybin arm and 20 taking escitalopram), a number that should be increased in further studies for increased validation of the results. Building and perturbing whole-brain models of larger datasets not only would better capture the heterogeneity associated with MDD,³² but could also allow studying, for example, the ability of perturbation measures like PER to predict the depression severity after treatment.

With respect to future research directions, other perturbation protocols could be explored, for instance, by perturbing the bifurcation parameter in Equation (5) instead of the noise level or by exploring more extensively the landscape of multi-site stimulations. In line with a recent study stressing the limitations of white noise stimulation, pink noise could be used instead.⁸³ Moreover, susceptibility and PER measures could be constructed using the perturbed functional hierarchy approximated by the simulated temporal asymmetry, and not just using static FC. Perturbational measures could also be defined from measures of FC dynamics and not just with a static average across time. For example, measures like metastability or probable metastable substrates could be used.^{28,30} This could provide meaningful complementary insights, especially given that psychedelics have been shown to broaden the repertoire of connectivity states across time.^{84–88} Finally, this computational study could be complemented by examining fMRI signals some months after treatment, and not immediately upon treatment completion. Eventually, such a lapse of time should allow the eventual benefits of longer-term plasticity events to show up in the different treated groups.⁶⁸

CONCLUSION

To recapitulate, we show that psilocybin and escitalopram differentially reconfigure susceptibility to perturbation, with psilocybin increasing susceptibility on average across regions, and escitalopram decreasing it. This is consistent with the promise of psychedelics to open a window of plasticity from which supportive therapy could benefit and can be related to a flattening of the hierarchy of causal brain influences, as initially proposed by the REBUS model. We also show that healthier transitions can be promoted following the administration of both drugs. This suggests that neurostimulation can be equally useful following escitalopram treatment, but we hypothesize that the general increase of susceptibility with psilocybin may constitute a plastic advantage in a longer term after the treatment is finished. Finally, we present the right amygdala and the nucleus accumbens as potential targets for stimulation, and we show that multi-stimulation significantly outperforms single-site approaches in driving the healthy recovery of FC. Overall, the present work shows how dynamic sensitivity analysis

of whole-brain models of brain hierarchy can contribute to the in-silico personalization of effective neurostimulation protocols for MDD, following or not pharmacological intervention.

AUTHOR CONTRIBUTIONS

M.S.-G., J.V., and G.D.: Methodology, formal analysis; M.S.-G., J.V., M.L.K., R.C.-H., D.E., and D.J.N.: Resources, funding acquisition, project administration; M.S.-G., J.V., Y.S.P., and G.D.: Software, data curation; M.S.-G.: Validation, visualization; M.S.-G., J.V., Y.S.P., and G.D.: Writing—original draft, writing—review and editing; G.D. and M.L.K.: Supervision.

ACKNOWLEDGMENTS

Jakub Vohryzek is supported by EU H2020 FET Proactive project Neurotwin grant agreement no. 101017716, Yonatan Sanz-Perl is supported by “ERDF A way of making Europe,” ERDF, EU, Project NEurological MEchanismS of Injury, and Sleep-like cellular dynamics (NEMESIS; ref. 101071900) funded by the EU ERC Synergy Horizon Europe, Morten L. Kringelbach is supported by the European Research Council Consolidator Grant: CAREGIVING (615539), Pettit Foundation, Carlsberg Foundation, and Center for Music in the Brain, funded by the Danish National Research Foundation (DNRF117). Gustavo Deco is supported by the Spanish Research Project PSI2016-75688-P (Agencia Estatal de Investigación/Fondo Europeo de Desarrollo Regional, European Union); by the European Union’s Horizon 2020 Research and Innovation Programme under Grant Agreements 720270 (Human Brain Project [HBP] SGA1) and 785907 (HBP SGA2); and by the Catalan Agency for Management of University and Research Grants Programme 2017 SGR 1545. The neuroimaging analysis and whole-brain modeling is based on clinical research carried out at the National Institute for Health Research/Wellcome Trust Imperial Clinical Research Facility. The open-label trial was funded by a Medical Research Council clinical development scheme grant (MR/J00460X/1). The double-blind randomized controlled trial was funded by a private donation from the Alexander Mosley Charitable Trust, supplemented by Founders of Imperial College London’s Centre for Psychedelic Research. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the paper.

COMPETING INTERESTS

R.C.-H. reports receiving consulting fees from COMPASS Pathways, Enttheon Biomedical, Mydecine, Synthesis Institute, Tryp Therapeutics, and Usona Institute. Within the past 2 years, D.E. has received fees for scientific advisory work from the following (novel psychedelic) companies: Mydecine, Field Trip Health, Enttheon, SmallPharma Ltd, Aya Biosciences, Clerkenwell Health, and Mindstate Design Lab; D.E. has also received an honorarium fee from each of COMPASS Pathways and H. Lundbeck for a talk about psychedelic science. D.J.N. has received consulting fees from Algernon, H. Lundbeck, and Beckley Psytech, advisory board fees from COMPASS Pathways and lecture fees from Takeda, Otsuka, and Janssen; D.J.N. also owns stock in Alcarelle, Awakn, and Psyched Wellness. The other authors declare no competing interests.

DATA AND CODE AVAILABILITY

Raw empirical data were generated at Imperial College London. The code for the analysis is available on GitHub: www.github.com/marcelocoro/PertWBModelsHierarchyDep.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ORCID

Marcel Socoró-Garrigosa  <https://orcid.org/0009-0004-7430-1460>

Jakub Vohryzek  <https://orcid.org/0000-0003-0994-5054>

PEER REVIEW

The peer review history for this article is available at <https://publons.com/publon/10.1111/nyas.15391>.

REFERENCES

- Vilarroya, O. (2017). Neural representation. A survey-based analysis of the notion. *Frontiers in Psychology*, 8, 1458. <https://doi.org/10.3389/fpsyg.2017.01458>
- O'keefe, J., & Dostrovsky, J. (1971). The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Research*, 34, 171–175. [https://doi.org/10.1016/0006-8993\(71\)90358-1](https://doi.org/10.1016/0006-8993(71)90358-1)
- Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology*, 195, 215–243. <https://doi.org/10.1113/jphysiol.1968.sp008455>
- Sporns, O., Tononi, G., & Kötter, R. (2005). The human connectome: A structural description of the human brain. *PLoS Computational Biology*, 1, e42. <https://doi.org/10.1371/journal.pcbi.0010042>
- Bassett, D. S., & Sporns, O. (2017). Network neuroscience. *Nature Neuroscience*, 20, 353–364. <https://doi.org/10.1038/nn.4502>
- Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393, 440–442. <https://doi.org/10.1038/30918>
- Newman, M. E. J. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences of the United States of America*, 103, 8577–8582. <https://doi.org/10.1073/pnas.0601602103>
- Greicius, M. (2008). Resting-state functional connectivity in neuropsychiatric disorders. *Current Opinion in Neurology*, 21, 424–430. <https://doi.org/10.1097/WCO.0b013e328306f2c5>
- Rombouts, S. A. R. B., Damoiseaux, J. S., Goekoop, R., Barkhof, F., Scheltens, P., Smith, S. M., & Beckmann, C. F. (2009). Model-free group analysis shows altered BOLD fMRI networks in dementia. *Human Brain Mapping*, 30, 256–266. <https://doi.org/10.1002/hbm.20505>
- Binnewijzend, M. A. A., Schoonheim, M. M., Sanz-Arigita, E., Wink, A. M., Van Der Flier, W. M., Tolboom, N., Adriaanse, S. M., Damoiseaux, J. S., Scheltens, P., Van Berckel, B. N. M., & Barkhof, F. (2012). Resting-state fMRI changes in Alzheimer's disease and mild cognitive impairment. *Neurobiology of Aging*, 33, 2018–2028. <https://doi.org/10.1016/j.neurobiolaging.2011.07.003>
- Karl, A., Schaefer, M., Malta, L., Dorfel, D., Rohleder, N., & Werner, A. (2006). A meta-analysis of structural brain abnormalities in PTSD. *Neuroscience and Biobehavioral Reviews*, 30, 1004–1031. <https://doi.org/10.1016/j.neubiorev.2006.03.004>
- van den Heuvel, M. P., & Fornito, A. (2014). Brain networks in schizophrenia. *Neuropsychology Review*, 24, 32–48. <https://doi.org/10.1007/s11065-014-9248-7>
- Greicius, M. D., Flores, B. H., Menon, V., Glover, G. H., Solvason, H. B., Kenna, H., Reiss, A. L., & Schlaggar, A. F. (2007). Resting-state functional connectivity in major depression: Abnormally increased contributions from subgenual cingulate cortex and thalamus. *Biological Psychiatry*, 62, 429–437. <https://doi.org/10.1016/j.biopsych.2006.09.020>
- Wang, L., Hermens, D. F., Hickie, I. B., & Lagopoulos, J. (2012). A systematic review of resting-state functional-MRI studies in major depression. *Journal of Affective Disorders*, 142, 6–12. <https://doi.org/10.1016/j.jad.2012.04.013>
- Deco, G., & Kringelbach, M. L. (2014). Great expectations: Using whole-brain computational connectomics for understanding neuropsychiatric disorders. *Neuron*, 84, 892–905. <https://doi.org/10.1016/j.neuron.2014.08.034>
- Wilson, H. R., & Cowan, J. D. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical Journal*, 12, 1–24. [https://doi.org/10.1016/S0006-3495\(72\)86068-5](https://doi.org/10.1016/S0006-3495(72)86068-5)
- Cabral, J., Vidaurre, D., Marques, P., Magalhães, R., Silva Moreira, P., Miguel Soares, J., Deco, G., Sousa, N., & Kringelbach, M. L. (2017). Cognitive performance in healthy older adults relates to spontaneous switching between states of functional connectivity during rest. *Scientific Reports*, 7, 5135. <https://doi.org/10.1038/s41598-017-05425-7>
- Honey, C. J., Sporns, O., Cammoun, L., Gigandet, X., Thiran, J. P., Meuli, R., & Hagmann, P. (2009). Predicting human resting-state functional connectivity from structural connectivity. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 2035–2040. <https://doi.org/10.1073/pnas.0811168106>
- Deco, G., Sanz Perl, Y., Bocaccio, H., Tagliazucchi, E., & Kringelbach, M. L. (2022). The INSIDEOUT framework provides precise signatures of the balance of intrinsic and extrinsic dynamics in brain states. *Communications Biology*, 5, 572. <https://doi.org/10.1038/s42003-022-03505-7>
- Deco, G., Sanz Perl, Y., Johnson, S., Bourke, N., Carhart-Harris, R. L., & Kringelbach, M. L. (2024). Different hierarchical reconfigurations in the brain by psilocybin and escitalopram for depression. *Nature Mental Health*, 2, 1096–1110. <https://doi.org/10.1038/s44220-024-00298-y>
- Kringelbach, M. L., Sanz Perl, Y., & Deco, G. (2024). The thermodynamics of mind. *Trends in Cognitive Sciences*, 28, 568–581. <https://doi.org/10.1016/j.tics.2024.03.009>
- Guzmán, E., Perl, Y. S., Vohryzek, J., Eschrichs, A., Manasova, D., Türker, B., Tagliazucchi, E., Kringelbach, M., Sitt, J. D., & Deco, G. (2023). The lack of temporal brain dynamics asymmetry as a signature of impaired consciousness states. *Interface Focus*, 13, 20220086. <https://doi.org/10.1098/rsfs.2022.0086>
- Lynn, C. W., Cornblath, E. J., Papadopoulos, L., Bertolero, M. A., & Bassett, D. S. (2021). Broken detailed balance and entropy production in the human brain. *Proceedings of the National Academy of Sciences of the United States of America*, 118, e2109889118. <https://doi.org/10.1073/pnas.2109889118>
- Sanz Perl, Y., Bocaccio, H., Pallavicini, C., Pérez-Ipiña, I., Laureys, S., Laufs, H., Kringelbach, M., Deco, G., & Tagliazucchi, E. (2021). Nonequilibrium brain dynamics as a signature of consciousness. *Physical Review E*, 104, 014411. <https://doi.org/10.1103/PhysRevE.104.014411>
- Friston, K. J. (2011). Functional and effective connectivity: A review. *Brain Connect*, 1, 13–36. <https://doi.org/10.1089/brain.2011.0008>
- Vohryzek, J., Cabral, J., Castaldo, F., Sanz-Perl, Y., Lord, L. D., Fernandes, H. M., Litvak, V., Kringelbach, M. L., & Deco, G. (2023). Dynamic sensitivity analysis: Defining personalized strategies to drive brain state transitions via whole brain modelling. *Computational and Structural Biotechnology Journal*, 21, 335–345. <https://doi.org/10.1016/j.csbj.2022.11.060>
- Deco, G., Cruzat, J., Cabral, J., Tagliazucchi, E., Laufs, H., Logothetis, N. K., & Kringelbach, M. L. (2019). Awakening: Predicting external stimulation to force transitions between different brain states. *Proceedings*

- of the National Academy of Sciences, 116, 18088–18097. <https://doi.org/10.1073/pnas.1905534116>
28. Vohryzek, J., Cabral, J., Perl, Y. S., Demirtas, M., Falcon, C., Gispert, J. D., Bosch, B., Balasa, M., Kringelbach, M., Sanchez-Valle, R., Ruffini, G., & Deco, G. (2023). Design of effective personalized perturbation strategies for enhancing cognitive intervention in Alzheimer's disease. *BioRxiv*, 2023.04.20.537688. <https://doi.org/10.1101/2023.04.20.537688>
 29. Mana, L., Schwartz-Pallejà, M., Vila-Vidal, M., & Deco, G. (2024). Overview on cognitive impairment in psychotic disorders: From impaired microcircuits to dysconnectivity. *Schizophrenia Research*, 269, 132–143. <https://doi.org/10.1016/j.schres.2024.05.008>
 30. Vohryzek, J., Cabral, J., Lord, L. D., Fernandes, H. M., Roseman, L., Nutt, D. J., Carhart-Harris, R. L., Deco, G., & Kringelbach, M. L. (2024). Brain dynamics predictive of response to psilocybin for treatment-resistant depression. *Brain Communications*, 6, fcae049. <https://doi.org/10.1093/braincomms/fcae049>
 31. Fried, E. I., Flake, J. K., & Robinaugh, D. J. (2022). Revisiting the theoretical and methodological foundations of depression measurement. *Nature Reviews Psychology*, 1, 358–368. <https://doi.org/10.1038/s44159-022-00050-2>
 32. Malhi, G. S., & Mann, J. J. (2018). Depression. *Lancet*, 392, 2299–2312. [https://doi.org/10.1016/S0140-6736\(18\)31948-2](https://doi.org/10.1016/S0140-6736(18)31948-2)
 33. World Health Organization. (2017). *Depression and other common mental disorders: Global health estimates*. World Health Organization.
 34. World Health Organization. (2008). *The global burden of disease: 2004 Update*. World Health Organization.
 35. Menon, V. (2011). Large-scale brain networks and psychopathology: A unifying triple network model. *Trends in Cognitive Sciences*, 15, 483–506. <https://doi.org/10.1016/j.tics.2011.08.003>
 36. Northoff, G., & Hirjak, D. (2024). Is depression a global brain disorder with topographic dynamic reorganization? *Translational Psychiatry*, 14, 1–12. <https://doi.org/10.1038/s41398-024-02995-9>
 37. Gilbert, J. R., Wusinich, C., & Zarate, C. A. (2022). A predictive coding framework for understanding major depression. *Frontiers in Human Neuroscience*, 16, 787495. <https://doi.org/10.3389/fnhum.2022.787495>
 38. Northoff, G. (2024). Beyond mood—depression as a speed disorder: Biomarkers for abnormal slowness. *Journal of Psychiatry & Neuroscience*, 49, E357–E366. <https://doi.org/10.1503/jpn.240099>
 39. Scalabrini, A., Vai, B., Poletti, S., Damiani, S., Mucci, C., Colombo, C., Zanardi, R., Benedetti, F., & Northoff, G. (2020). All roads lead to the default-mode network—Global source of DMN abnormalities in major depressive disorder. *Neuropsychopharmacology*, 45, 2058–2069. <https://doi.org/10.1038/s41386-020-0785-x>
 40. Cuijpers, P., Sijbrandij, M., Koole, S. L., Andersson, G., Beekman, A. T., & Reynolds, C. F. (2014). Adding psychotherapy to antidepressant medication in depression and anxiety disorders: A meta-analysis. *World Psychiatry*, 13, 56–67. <https://doi.org/10.1002/wps.20089>
 41. McLachlan, G. (2018). Treatment resistant depression: What are the options? *BMJ*, 363, k5354. <https://doi.org/10.1136/bmj.k5354>
 42. Carhart-Harris, R. L., Bolstridge, M., Rucker, J., Day, C. M. J., Erritzoe, D., Kaelen, M., Bloomfield, M., Rickard, J. A., Forbes, B., Feilding, A., Taylor, D., Pilling, S., Curran, V. H., & Nutt, D. J. (2016). Psilocybin with psychological support for treatment-resistant depression: An open-label feasibility study. *Lancet Psychiatry*, 3, 619–627. [https://doi.org/10.1016/S2215-0366\(16\)30065-7](https://doi.org/10.1016/S2215-0366(16)30065-7)
 43. Palhano-Fontes, F., Barreto, D., Onias, H., Andrade, K. C., Novaes, M. M., Pessoa, J. A., Mota-Rolim, S. A., Osório, F. L., Sanches, R., Dos Santos, R. G., Tófoli, L. F., De Oliveira Silveira, G., Yonamine, M., Riba, J., Santos, F. R., Silva-Junior, A. A., Alchieri, J. C., Galvão-Coelho, N. L., Lobão-Soares, B., ... & Araújo, D. B. (2019). Rapid antidepressant effects of the psychedelic ayahuasca in treatment-resistant depression: A randomized placebo-controlled trial. *Psychological Medicine*, 49, 655–663. <https://doi.org/10.1017/S0033291718001356>
 44. Carhart-Harris, R. L., & Nutt, D. J. (2017). Serotonin and brain function: A tale of two receptors. *Journal of Psychopharmacology*, 31, 1091–1120. <https://doi.org/10.1177/0269881117725915>
 45. Carhart-Harris, R. L., & Friston, K. J. (2019). REBUS and the anarchic brain: Toward a unified model of the brain action of psychedelics. *Pharmacological Reviews*, 71, 316–344. <https://doi.org/10.1124/pr.118.017160>
 46. Carhart-Harris, R., Giribaldi, B., Watts, R., Baker-Jones, M., Murphy-Beiner, A., Murphy, R., Martell, J., Blemings, A., Erritzoe, D., & Nutt, D. J. (2021). Trial of psilocybin versus escitalopram for depression. *New England Journal of Medicine*, 384, 1402–1411. <https://doi.org/10.1056/NEJMoa2032994>
 47. Daws, R. E., Timmermann, C., Giribaldi, B., Sexton, J. D., Wall, M. B., Erritzoe, D., Roseman, L., Nutt, D., & Carhart-Harris, R. (2022). Increased global integration in the brain after psilocybin therapy for depression. *Nature Medicine*, 28, 844–851. <https://doi.org/10.1038/s41591-022-01744-z>
 48. Vohryzek, J., Deco, G., Cessac, B., Kringelbach, M. L., & Cabral, J. (2020). Ghost attractors in spontaneous brain activity: Recurrent excursions into functionally-relevant BOLD phase-locking states. *Frontiers in Systems Neuroscience*, 14, 20. <https://doi.org/10.3389/fnsys.2020.00020>
 49. Deco, G., Vidaurre, D., & Kringelbach, M. L. (2021). Revisiting the global workspace orchestrating the hierarchical organization of the human brain. *Nature Human Behaviour*, 5, 497–511. <https://doi.org/10.1038/s41562-020-01003-6>
 50. Kuznetsov, Y. A. (2023). *Elements of applied bifurcation theory*. Springer International Publishing. <https://doi.org/10.1007/978-3-031-22007-4>
 51. Freyer, F., Roberts, J. A., Becker, R., Robinson, P. A., Ritter, P., & Breakspear, M. (2011). Biophysical mechanisms of multistability in resting-state cortical rhythms. *Journal of Neuroscience*, 31, 6353–6361. <https://doi.org/10.1523/JNEUROSCI.6693-10.2011>
 52. Freyer, F., Roberts, J. A., Ritter, P., & Breakspear, M. (2012). A canonical model of multistability and scale-invariance in biological systems. *PLOS Computational Biology*, 8, e1002634. <https://doi.org/10.1371/journal.pcbi.1002634>
 53. Deco, G., Cabral, J., Woolrich, M. W., Stevner, A. B. A., Van Hartevelt, T. J., & Kringelbach, M. L. (2017). Single or multiple frequency generators in on-going brain activity: A mechanistic whole-brain model of empirical MEG data. *Neuroimage*, 152, 538–550. <https://doi.org/10.1016/j.neuroimage.2017.03.023>
 54. Kringelbach, M. L., Cruzat, J., Cabral, J., Knudsen, G. M., Carhart-Harris, R., Whybrow, P. C., Logothetis, N. K., & Deco, G. (2020). Dynamic coupling of whole-brain neuronal and neurotransmitter systems. *Proceedings of the National Academy of Sciences of the United States of America*, 117, 9566–9576. <https://doi.org/10.1073/pnas.1921475117>
 55. Ponce-Alvarez, A., & Deco, G. (2024). The Hopf whole-brain model and its linear approximation. *Scientific Reports*, 14, 2615. <https://doi.org/10.1038/s41598-024-53105-0>
 56. Deco, G., Kringelbach, M. L., Jirsa, V. K., & Ritter, P. (2017). The dynamics of resting fluctuations in the brain: Metastability and its dynamical cortical core. *Scientific Reports*, 7, 3095. <https://doi.org/10.1038/s41598-017-03073-5>
 57. Deco, G., Ponce-Alvarez, A., Hagmann, P., Romani, G. L., Mantini, D., & Corbetta, M. (2014). How local excitation-inhibition ratio impacts the whole brain dynamics. *Journal of Neuroscience*, 34, 7886–7898. <https://doi.org/10.1523/JNEUROSCI.5068-13.2014>
 58. Kringelbach, M. L., Perl, Y. S., Tagliazucchi, E., & Deco, G. (2023). Toward naturalistic neuroscience: Mechanisms underlying the flattening of brain hierarchy in movie-watching compared to rest and task. *Science Advances*, 9, eade6049. <https://doi.org/10.1126/sciadv.ade6049>
 59. Liégeois, R., Santos, A., Matta, V., Van De Ville, D., & Sayed, A. H. (2020). Revisiting correlation-based functional connectivity and its relation-

- ship with structural connectivity. *Network Neuroscience*, 4, 1235–1251. https://doi.org/10.1162/netn_a_00166
60. Nichols, D. E. (2016). Psychedelics. *Pharmacological Reviews*, 68, 264–355. <https://doi.org/10.1124/pr.115.011478>
 61. Vargas, M. V., Dunlap, L. E., Dong, C., Carter, S. J., Tombari, R. J., Jami, S. A., Cameron, L. P., Patel, S. D., Hennessey, J. J., Saeger, H. N., Mccorvy, J. D., Gray, J. A., Tian, L., & Olson, D. E. (2023). Psychedelics promote neuroplasticity through the activation of intracellular 5-HT_{2A} receptors. *Science*, 379, 700–706. <https://doi.org/10.1126/science.adf0435>
 62. Carhart-Harris, R. L. (2019). How do psychedelics work? *Current Opinion in Psychiatry*, 32, 16–21. <https://doi.org/10.1097/YCO.0000000000000467>
 63. Chu, A., & Wadhwa, R. (2025). Selective serotonin reuptake inhibitors. In *StatPearls*. StatPearls Publishing.
 64. Ruffini, G., Lopez-Sola, E., Vohryzek, J., & Sanchez-Todo, R. (2024). Neural geometrodynamics, complexity, and plasticity: A psychedelics perspective. *Entropy*, 26, 90. <https://doi.org/10.3390/e26010090>
 65. Ruffini, G., Castaldo, F., Lopez-Sola, E., Sanchez-Todo, R., & Vohryzek, J. (2024). The algorithmic agent perspective and computational neuropsychiatry: From etiology to advanced therapy in major depressive disorder. *Entropy*, 26, 953. <https://doi.org/10.3390/e26110953>
 66. Nitsche, M. A., Kuo, M. F., Karrasch, R., Wächter, B., Liebetanz, D., & Paulus, W. (2009). Serotonin affects transcranial direct current-induced neuroplasticity in humans. *Biological Psychiatry*, 66, 503–508. <https://doi.org/10.1016/j.biopsych.2009.03.022>
 67. Calder, A. E., & Hasler, G. (2023). Towards an understanding of psychedelic-induced neuroplasticity. *Neuropsychopharmacology*, 48, 104–112. <https://doi.org/10.1038/s41386-022-01389-z>
 68. Pasquini, L., Vohryzek, J., Escrichs, A., Sanz Perl, Y., Ponce-Alvarez, A., Idesis, S., Girn, M., Roseman, L., Mitchell, J. M., Gazzaley, A., Kringelbach, M., Nutt, D. J., Lyons, T., Carhart-Harris, R. L., & Deco, G. (2024). Long-term effects of psilocybin on dynamic and effectivity connectivity of fronto-striatal-thalamic circuits. *BioRxiv*, 2024.11.06.622302. <https://doi.org/10.1101/2024.11.06.622302>
 69. Mutz, J., Vipulanathan, V., Carter, B., Hurlmann, R., Fu, C. H. Y., & Young, A. H. (2019). Comparative efficacy and acceptability of non-surgical brain stimulation for the acute treatment of major depressive episodes in adults: Systematic review and network meta-analysis. *BMJ*, 364, l1079. <https://doi.org/10.1136/bmj.l1079>
 70. Trevisol, A. P., Goldberger, K. W., Mulsant, B. H., Rajji, T. K., Downar, J., Daskalakis, Z. J., & Blumberger, D. M. (2019). Unilateral and bilateral repetitive transcranial magnetic stimulation for treatment-resistant late-life depression. *International Journal of Geriatric Psychiatry*, 34, 822–827. <https://doi.org/10.1002/gps.5091>
 71. Ruffini, G., Salvador, R., Castaldo, F., Baleeiro, T., Camprodon, J. A., Chopra, M., Cappon, D., & Pascual-Leone, A. (2024). Multichannel tDCS with advanced targeting for major depressive disorder: A tele-supervised at-home pilot study. *Medrxiv*, 2024.03.04.24303508. <https://doi.org/10.1101/2024.03.04.24303508>
 72. Klawonn, A. M., & Malenka, R. C. (2018). Nucleus accumbens modulation in reward and aversion. *Cold Spring Harbor Symposia on Quantitative Biology*, 83, 119–129. <https://doi.org/10.1101/sqb.2018.83.037457>
 73. Hu, Y., Zhao, C., Zhao, H., & Qiao, J. (2023). Abnormal functional connectivity of the nucleus accumbens subregions mediates the association between anhedonia and major depressive disorder. *BMC Psychiatry [Electronic Resource]*, 23, 282. <https://doi.org/10.1186/s12888-023-04693-0>
 74. Bewernick, B. H., Hurlmann, R., Matusch, A., Kayser, S., Grubert, C., Hadrysiewicz, B., Axmacher, N., Lemke, M., Cooper-Mahkorn, D., Cohen, M. X., Brockmann, H., Lenartz, D., Sturm, V., & Schlaepfer, T. E. (2010). Nucleus accumbens deep brain stimulation decreases ratings of depression and anxiety in treatment-resistant depression. *Biological Psychiatry*, 67, 110–116. <https://doi.org/10.1016/j.biopsych.2009.09.013>
 75. Karrouri, R., Hammani, Z., Benjelloun, R., & Otheman, Y. (2021). Major depressive disorder: Validated treatments and future challenges. *World Journal of Clinical Cases*, 9, 9350–9367. <https://doi.org/10.12998/wjcc.v9.i31.9350>
 76. Perlman, K., Benrimoh, D., Israel, S., Rollins, C., Brown, E., Tunteng, J. F., You, R., You, E., Tanguay-Sela, M., Snook, E., Miresco, M., & Berlim, M. T. (2019). A systematic meta-review of predictors of antidepressant treatment outcome in major depressive disorder. *Journal of Affective Disorders*, 243, 503–515. <https://doi.org/10.1016/j.jad.2018.09.067>
 77. Phillips, M. L., Chase, H. W., Sheline, Y. I., Etkin, A., Almeida, J. R. C., Deckersbach, T., & Trivedi, M. H. (2015). Identifying predictors, moderators, and mediators of antidepressant response in major depressive disorder: Neuroimaging approaches. *American Journal of Psychiatry*, 172, 124–138. <https://doi.org/10.1176/appi.ajp.2014.14010076>
 78. Young, K. D., Siegle, G. J., Zotev, V., Phillips, R., Misaki, M., Yuan, H., Drevets, W. C., & Bodurka, J. (2017). Randomized clinical trial of real-time fMRI amygdala neurofeedback for major depressive disorder: Effects on symptoms and autobiographical memory recall. *American Journal of Psychiatry*, 174, 748–755. <https://doi.org/10.1176/appi.ajp.2017.16060637>
 79. Dyck, M., Loughhead, J., Kellermann, T., Boers, F., Gur, R. C., & Mathiak, K. (2011). Cognitive versus automatic mechanisms of mood induction differentially activate left and right amygdala. *Neuroimage*, 54, 2503–2513. <https://doi.org/10.1016/j.neuroimage.2010.10.013>
 80. Murphy, J. E., Yanes, J. A., Kirby, L. A. J., Reid, M. A., & Robinson, J. L. (2020). Left, right, or bilateral amygdala activation? How effects of smoothing and motion correction on ultra-high field, high-resolution functional magnetic resonance imaging (fMRI) data alter inferences. *Neuroscience Research*, 150, 51–59. <https://doi.org/10.1016/j.neures.2019.01.009>
 81. Rolls, E. T., Cheng, W., & Feng, J. (2020). The orbitofrontal cortex: Reward, emotion and depression. *Brain Communications*, 2, fcaa196. <https://doi.org/10.1093/braincomms/fcaa196>
 82. Gogulski, J., Ross, J. M., Talbot, A., Cline, C. C., Donati, F. L., Munot, S., Kim, N., Gibbs, C., Bastin, N., Yang, J., Minasi, C., Sarkar, M., Truong, J., & Keller, C. J. (2023). Personalized repetitive transcranial magnetic stimulation for depression. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 8, 351–360. <https://doi.org/10.1016/j.bpsc.2022.10.006>
 83. Klar, P., Çatal, Y., Langner, R., Huang, Z., & Northoff, G. (2023). Scale-free dynamics in the core-periphery topography and task alignment decline from conscious to unconscious states. *Communications Biology*, 6, 499. <https://doi.org/10.1038/s42003-023-04879-y>
 84. Atasoy, S., Vohryzek, J., Deco, G., Carhart-Harris, R. L., & Kringelbach, M. L. (2018). Common neural signatures of psychedelics: Frequency-specific energy changes and repertoire expansion revealed using connectome-harmonic decomposition. *Progress in Brain Research*, 242, 97–120. <https://doi.org/10.1016/bs.pbr.2018.08.009>
 85. Cruzat, J., Perl, Y. S., Escrichs, A., Vohryzek, J., Timmermann, C., Roseman, L., Luppi, A. I., Ibañez, A., Nutt, D., Carhart-Harris, R., Tagliazucchi, E., Deco, G., & Kringelbach, M. L. (2022). Effects of classic psychedelic drugs on turbulent signatures in brain dynamics. *Network Neuroscience*, 6, 1104–1124. https://doi.org/10.1162/netn_a_00250
 86. Tagliazucchi, E., Carhart-Harris, R., Leech, R., Nutt, D., & Chialvo, D. R. (2014). Enhanced repertoire of brain dynamical states during the psychedelic experience. *Human Brain Mapping*, 35, 5442–5456. <https://doi.org/10.1002/hbm.22562>
 87. Varley, T. F., Carhart-Harris, R., Roseman, L., Menon, D. K., & Stamatakis, E. A. (2020). Serotonergic psychedelics LSD & psilocybin increase the fractal dimension of cortical brain activity in spatial and temporal domains. *Neuroimage*, 220, 117049. <https://doi.org/10.1016/j.neuroimage.2020.117049>
 88. Vohryzek, J., Cabral, J., Timmermann, C., Atasoy, S., Roseman, L., Nutt, D. J., Carhart-Harris, R. L., Deco, G., & Kringelbach, M. L. (2024). The flattening of spacetime hierarchy of the N,N-dimethyltryptamine

brain state is characterized by harmonic decomposition of spacetime (HADES) framework. *National Science Review*, 11, nwae124. <https://doi.org/10.1093/nsr/nwae124>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Socoró-Garrigosa, M., Perl, Y. S., Kringelbach, M. L., Erritzoe, D., Nutt, D. J., Carhart-Harris, R., Vohryzek, J., & Deco, G. (2025). Perturbing whole-brain models of brain hierarchy: An application for depression following pharmacological treatment. *Ann NY Acad Sci.*, 1–18. <https://doi.org/10.1111/nyas.15391>