



Article

<https://doi.org/10.1038/s44220-024-00298-y>

Different hierarchical reconfigurations in the brain by psilocybin and escitalopram for depression

Received: 22 November 2023

Gustavo Deco ^{1,2}✉, Yonatan Sanz Perl ^{1,3}, Samuel Johnson ^{4,5}, Niamh Bourke ^{6,7}, Robin L. Carhart-Harris ^{8,9} & Morten L. Kringelbach ^{6,7,10}✉

Accepted: 12 July 2024

Published online: 5 August 2024

Check for updates

Effective interventions for neuropsychiatric disorders may work by rebalancing the brain's functional hierarchical organization. Here we directly investigated the effects of two different serotonergic pharmacological interventions on functional brain hierarchy in major depressive disorder in a two-arm double-blind phase II randomized controlled trial comparing psilocybin therapy (22 patients) with escitalopram (20 patients). Patients with major depressive disorder received either 2×25 mg of oral psilocybin, three weeks apart, plus six weeks of daily placebo ('psilocybin arm') or 2×1 mg of oral psilocybin, three weeks apart, plus six weeks of daily escitalopram (10–20 mg; 'escitalopram arm'). Resting-state functional magnetic resonance imaging scans were acquired at baseline and three weeks after the second psilocybin dose ([NCT03429075](#)). The brain mechanisms were captured by generative effective connectivity, estimated from whole-brain modeling of resting state for each session and patient. Hierarchy was determined for each of these sessions using measures of directedness and trophic levels on the effective connectivity, which captures cycle structure, stability and percolation. The results showed that the two pharmacological interventions created significantly different hierarchical reconfigurations of whole-brain dynamics with differential, opposite statistical effect responses. Furthermore, the use of machine learning revealed significant differential reorganization of brain hierarchy before and after the two treatments. Machine learning was also able to predict treatment response with an accuracy of 0.85 ± 0.04 . Overall, the results demonstrate that psilocybin and escitalopram work in different ways for rebalancing brain dynamics in depression. This suggests the hypothesis that neuropsychiatric disorders could be closely linked to the breakdown in regions orchestrating brain dynamics from the top of the hierarchy.

A full list of affiliations appears at the end of the paper. ✉ e-mail: gustavo.deco@upf.edu; morten.kringelbach@psych.ox.ac.uk

Neuropsychiatric disorders are devastating, incurring significant disability and stigma for individuals and a serious and growing burden for society^{1,2}. Major depressive disorder (MDD) has become pervasive. A major contributor to the years lost to disability worldwide, in terms of the burden of disease, by 2030 MDD is predicted to become the largest contributor³. Adding to this problem, antidepressant drugs show modest efficacy^{4,5}, have unpleasant side effects that impact treatment adherence⁶ and are associated with high relapse rates⁷. Problems with current treatments and the scarcity of reliable animal models mean that new research strategies are needed to treat neuropsychiatric disorders^{8,9}. Yet, at the present time, it is not clear how brain dynamics change when patients get better following pharmacological intervention.

Neuroimaging has identified abnormal brain dynamics in depression^{10,11}, with changes linked primarily to higher-order brain networks including the default mode network (DMN), executive network (EN) and salience network (SN)^{11,12}. Research in healthy individuals has associated these networks with higher-order functioning, including executive control and attentional switching¹³, which are often impaired in depressed patients¹⁴. There have also been neuroimaging studies investigating the effective connectivity in depression associated with treatment response^{15–17}. Successful pharmacological intervention for depression must therefore be linked to a rebalancing of these dynamics. One candidate mechanism is the role of the serotonin 2A (5-HT2A) receptor subtype, which is closely associated with serotonergic psychedelic drugs such as psilocybin^{18,19}. The spatial distribution of 5-HT2A receptors overlaps with the DMN, EN and SN²⁰. It has been hypothesized in the ‘relaxed beliefs under psychedelics’ (REBUS) theory²¹ that psychedelics can bring about a relaxation of the precision of high-level priors or ‘beliefs’, allowing bottom-up rather than hierarchical top-down information flow, which is consistent with more hierarchy-free system dynamics, suggesting a hierarchical reconfiguration of brain dynamics²². However, a full whole-brain quantification of hierarchical processing has not yet been attempted.

In this Article, the main aim is to study the effects of pharmacological intervention on brain dynamics and in particular to determine how this affects the hierarchy of brain processing associated with two serotonergic pharmacological interventions given to patients with treatment-resistant depression. To answer this research question, we leveraged a double-blind phase II randomized controlled trial²³, where resting-state functional magnetic resonance imaging (fMRI) scans at baseline and after treatment were obtained from all patients²². To best quantify the changes in brain dynamics, we implemented a sensitive measure of hierarchy, inspired by work in ecology capturing the hierarchical relationships between plants, herbivores and carnivores in a food web. The method uses the concept of trophic levels for each node in a directed graph, which can be thought of as the ‘height’ of a node in the hierarchy²⁴. A flat hierarchy is characterized by equal trophic levels and low directedness, which reflects low asymmetry in a network. In contrast, a strong hierarchy is associated with high directedness and strong asymmetric connections in a many-layered network. This measure of directedness (also called trophic coherence) has been related to other network properties, such as cycle structure, stability, normality and percolation²⁵.

This hierarchy measure provides the global level of directedness and regional trophic levels, which we applied to the graph of the generative effective connectivity (GEC) matrix²⁶. This was estimated from whole-brain modeling of the resting state for each session of each patient. The GEC is an extension of the classic concept of effective connectivity²⁷, but generative in the sense that the GEC is created from a whole-brain model of empirical resting-state data, where the strengths of existing anatomical connectivity are adapted iteratively until best fit, thus providing a mechanistic explanation of the resting-state data. This asymmetric measure of functional brain connectivity is fundamental for assessing hierarchy, while other simpler measures

such as functional connectivity (FC) are not suitable because they are, by definition, symmetric.

The overall result was that the two pharmacological interventions gave rise to significantly different hierarchical reconfigurations following treatment, with differential, opposite statistical effect responses. In other words, we found significant differences between measures of both directedness and the trophic level, before and after intervention. These differences were found both from direct statistical tests and when using machine learning to measure the degree of pattern separation. Furthermore, we also found significant differences in trophic levels when comparing responders with non-responders for each pharmacological intervention. Importantly, pharmacological interventions were found to work best when select cortical and subcortical regions in the so-called ‘global workspace’ were re-established as being near the top of the hierarchy²⁸. These results strongly suggest that the psilocybin and escitalopram work in very different ways when used for rebalancing the hierarchical organization of brain dynamics in depression. In general, the research presented here offers a principled route to evaluate the effects of pharmacological intervention based on empirical brain dynamics data. Many obstacles remain, including replication of the results in a much larger population and for other interventions, but the results show potential for fulfilling the great expectations of using neuroimaging for understanding the underlying mechanisms of neuropsychiatric disorders.

Results

We addressed the main research question of how pharmacological interventions affect brain dynamics by assessing hierarchical reconfiguration in the brain state in MMD patients enrolled in a double-blind phase II randomized controlled trial comparing two drugs (psilocybin and escitalopram)²³. In this study, as shown in Fig. 1a, neuroimaging resting-state data were acquired before and after the pharmacological intervention (details of the trial are provided in Methods). We used the data for these two sessions six weeks apart to quantify the effects on global directedness and regional trophic levels on the brain dynamics following either intervention.

Quantifying hierarchy in brain states

To reveal the spatiotemporal hierarchical organization of the brain dynamics for the two pharmacological interventions, we implemented a method measuring the directedness and trophic levels of a network based on previous work in ecology²⁹ and extended to general directed networks²⁴. This measure of hierarchy is very robust and has been used in many other fields; it has been described as ‘upstreamness’ in economics³⁰ and is a key ingredient in the construction of SinkRank, a measure of contribution to systemic risk³¹.

For a given directed network, this method provides both the hierarchical, node-level information (trophic level) and the global information (the directedness, or trophic coherence). In other words, the trophic level provides a measure of where a node sits in the hierarchy of a directed network. In ecology, low trophic levels would be assigned to plants, and high trophic level nodes would be assigned to carnivores, given that energy flows up the food web from low to high trophic levels. Our implementation of the hierarchy methods uses the recent extension from ref. 24 to the standard definitions of trophic level, which allows the method to overcome the limitations of requiring a basal node (that is, a node with no incoming edges) and taking into account reverse flow. As a result, this optimized method is even more robust, as it captures aspects of related network measures of stability, cycles and normality.

We thus use a well-established method to measure the spatiotemporal hierarchical organization of brain dynamics. Figure 1b presents a cartoon illustrating the trophic levels and directedness measures for flat and hierarchical networks. As can be seen in the top panel, a flat hierarchy has approximately equal trophic levels, resulting in

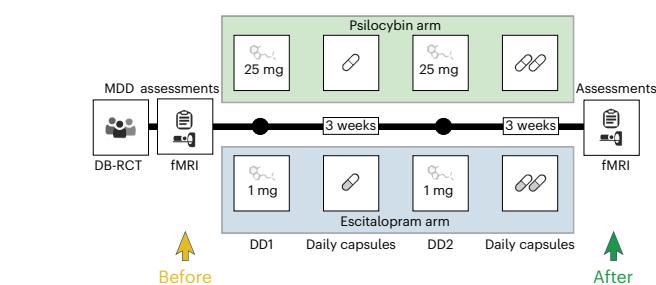
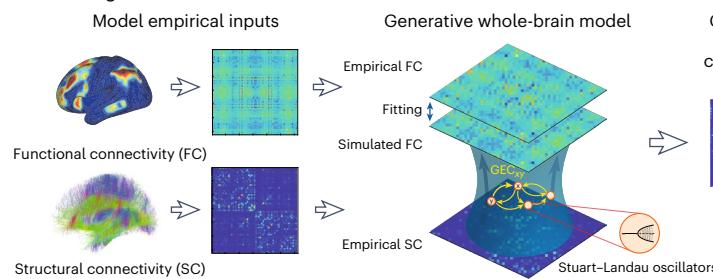
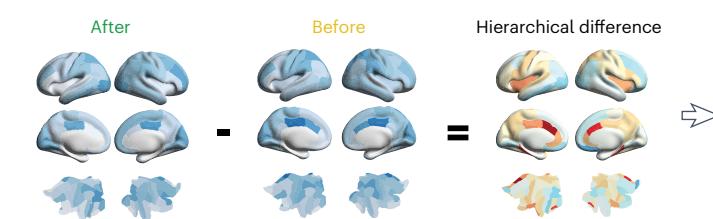
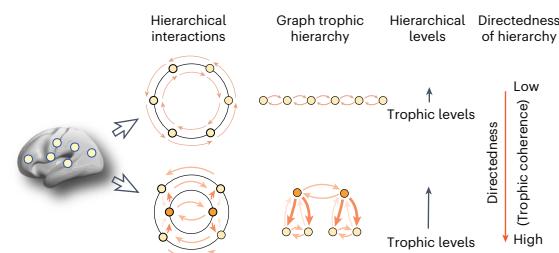
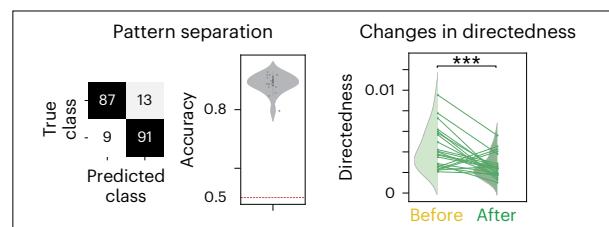
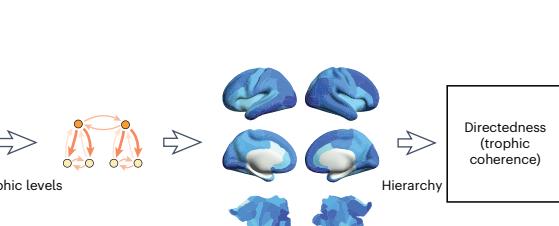
a Experimental design for pharmacological treatment of depression**c** Building whole-brain model of each brain state**e** Differences in hierarchical directedness after pharmacological treatment across patients**b** Hierarchical quantification using trophic levels**d** Directedness of brain state

Fig. 1 | Pipeline for assessing hierarchy in brain states before and after pharmacological intervention for depression. **a**, Neuroimaging resting-state data were acquired before and after a double-blind phase II randomized controlled trial (DB-RCT) comparing psilocybin therapy with escitalopram. **b**, The hierarchy was quantified using a new method measuring the directedness of a network based on trophic levels, inspired by previous work in ecology but extended to general directed networks. The top panel shows a flat hierarchy with low directedness (or trophic coherence), which reflects the low asymmetry of the network with the same trophic levels. In contrast, the bottom panel shows a strong hierarchy, associated with high directedness (and asymmetry), which emerges from the diversity of trophic levels reflecting the many layers

in the network. **c**, Whole-brain modeling is then used on individual sessions of each patient of neuroimaging resting-state data in the DK80 parcellation. This estimates the underlying mechanisms captured by the GEC. **d**, The hierarchy of the brain state in each session can be quantified using the directedness method on the individual GEC. This produces the trophic levels for each brain region and a global measure of directedness (or trophic coherence). **e**, For each patient, we then estimate the difference in hierarchical organization by subtracting the before- from the after-intervention session. This is then used as the basis of machine-learning classification and for direct comparisons of the hierarchical reconfiguration.

low directedness and reflecting the low asymmetry of the network. This is not the case for a strong hierarchy, which is associated with high directedness and consequently strong asymmetric connections. In the latter case, the diversity of trophic levels reflects that the network contains many hierarchical layers and an overall direction pointing from the lowest to the highest.

To estimate the functional hierarchy of a given brain state, we applied this method to the network produced by whole-brain modeling of a patient's individual resting state in the DK80 parcellation²⁸. This network captures the underlying mechanisms generating the resting-state dynamics³². In brief, whole-brain modeling combines the anatomical connectivity with local dynamics to fit the dynamics of empirical neuroimaging data^{32–34}. The local dynamics can be simulated using, for example, spiking, dynamical mean field and Hopf local regional models to fit many different empirical observables. Overall, the best fit is obtained with the Hopf model, so we chose to fit this model to the model-free observable of non-reversibility of the neuroimaging data^{26,35–37}. More specifically, Fig. 1c shows the procedure for fitting a whole-brain model, initially using the anatomical connectivity and then

iteratively adjusting a GEC, which is given as asymmetric weights of the existing anatomical connections²⁶. This is an extension of the concept of effective connectivity²⁷, but where GEC is generative because it uses the whole-brain model to adapt the strength of existing anatomical connectivity (that is, the effective conductive values of each fiber). In contrast to normal symmetrical measures of effective connectivity (typically using FC), here we use the asymmetrical optimization obtained from using the non-reversibility INSIDEOUT method²⁶. Recent advances in the analysis of brain-imaging data have shown that the thermodynamic concept of the arrow of time (or non-reversibility) is a very sensitive and robust measure of the underlying asymmetric information flow^{26,35–43}. This body of research has demonstrated that creating a whole-brain model of the arrow of time in empirical neuroimaging data can provide access to the generative mechanisms underlying hierarchy—which in turn can provide a direct measure of the hierarchical reconfiguration over time (Methods).

Figure 1d shows how the hierarchy of the brain state in each session can be quantified using the directedness method on the individual GEC. This produces the trophic levels for each brain region, that is, where in

the hierarchy each is placed, as well as the global measure of directedness (or trophic coherence) based on the trophic levels.

Finally, Fig. 1e illustrates the procedure for estimating the difference in hierarchical organization, which is obtained by subtracting the sessions before the pharmacological intervention from after. As shown in the following, this is used as the basis of machine learning to compare pattern separation based on regional trophic levels and for the direct comparison of the hierarchical reconfiguration of directedness.

Machine learning of hierarchical reconfigurations

We used machine learning to establish the significant hierarchical reorganization between before and after treatment. To do this, we first computed the GEC for both sessions in each patient receiving either psilocybin or escitalopram, independent of treatment response. This allowed us to compute the trophic levels and directedness before and after intervention in each patient. The before and after scan sessions for each of the two types of intervention were classified using machine learning (Fig. 2a). For classification we used the hierarchical trophic levels for each region (for each patient and condition) as input features. We sorted the regions according to each region's statistical significant hierarchical trophic level across patients, region by region, using Wilcoxon 10,000-permutation tests. To estimate with the highest possible accuracy, we selected sequentially the minimum numbers of regions from the sorted list of statistically significant trophic levels with the minimum number of inputs that yielded the largest accuracy. We found that the highest accuracy was obtained using 69 regions for the psilocybin treatment arm and six regions for the escitalopram treatment arm. Importantly, we trained the support vector machine (SVM) with the leave-one-out cross-validation procedure; that is, we randomly chose one patient for generalization and the rest for training, repeated, and shuffled 1,000 times. Furthermore, we made sure that the training set was balanced in terms of number of examples for each class label, and randomly selected the patients in each class for each shuffling iteration (Methods). The results show significant effects above chance levels with an accuracy of 0.89 ± 0.03 (mean \pm s.d.) for psilocybin and 0.89 ± 0.02 for escitalopram. In addition, we also carried out machine learning for all 80 regions for both treatments to show that the preselection of regions does not affect the results (Supplementary Fig. 5).

Directedness of hierarchical reconfiguration

The overall global hierarchical reconfiguration can be directly quantified using the measure of directedness, which is a global measure of hierarchy based on regional trophic levels (Methods). Figure 2b shows the significant differences between before and after treatment ($P < 0.001$, paired Wilcoxon test using 10,000 permutations) for both the psilocybin and escitalopram treatment arms. Importantly, with psilocybin treatment, the directedness significantly decreases, but it increases for the escitalopram treatment, reflecting the differential hierarchical reconfiguration for the two pharmacological interventions. This demonstrates differential, opposite statistical effect responses, whereby the directedness decreases for the psilocybin treatment but increases for the escitalopram treatment.

Fig. 2 | Hierarchical reconfiguration following the administration of psilocybin and escitalopram. Throughout, the full results are shown for psilocybin in the left column and for escitalopram in the right column for both responders and non-responders. **a**, Machine learning of the hierarchical organization reflected in the trophic levels was used to classify the scan sessions, that is, the brain dynamics before and after the intervention. The confusion matrices show the percentage results of the 22 patients treated with psilocybin and the 20 patients treated with escitalopram. This gives an accuracy of 0.89 ± 0.03 (mean \pm s.d.) for psilocybin and 0.89 ± 0.02 for escitalopram, revealing a significant level of hierarchical reconfiguration. **b**, The hierarchical reorganization is shown by comparing the directedness (or trophic coherence) as a global measure of the hierarchical reconfiguration.

Importantly, the finding of a differential treatment-dependent hierarchical reconfiguration was confirmed by additional analyses. First, we computed the correlations between changes in Beck Depression Inventory (BDI)/treatment response (post–pre) for hierarchy (changes in directedness post–pre) for the psilocybin and escitalopram groups. As shown in Supplementary Fig. 6, this yielded no significance for psilocybin (corr: -0.023 , $P = 0.91$ (not significant (NS)), red line) and for escitalopram (corr: 0.16 , $P = 0.50$ (NS), blue line). Yet, as expected when we combine the two groups, this yields a significant correlation of 0.3 ($P = 0.049$, black line). The results mean that the change in BDI score is not predictive for the change in hierarchy for either group, but only when combining them. Overall, this shows that treatment is the important variable.

Second, further validating this important result, we ran an analysis of variance (ANOVA) with three variables: (1) binary variable of whether a patient was classified as responder, (2) change in BDI score (post–pre) and (3) type of treatment (psilocybin or escitalopram). The results confirmed that only the treatment is significant: factor 1 (response): F -statistics 0.063 , $P = 0.81$ (NS); factor 2 (BDI change): F -statistics 1.22 , $P = 0.35$ (NS); factor 3 (treatment): F -statistics 16.086 , $P = 0.0013$.

We also investigated possible self-reported sex effects using multi-way ANOVA of the differences in directedness before and after intervention. This showed non-significant effects of sex ($P > 0.05$) on the mean of the differences in directedness ($F(1,39) = 3.08$, $P = 0.087$), whereas the intervention effect was highly significant ($F(1,39) = 33.28$, $P < 0.001$).

It is also important to remark that although the interventions had different remission rates (based on the BDI: 64% for psilocybin and 30% for escitalopram), the main finding of different reconfigurations between the treatment arms is independent of remission rates, but based on the significant differential, opposite statistical effect responses in the changes in hierarchy after intervention. In the case of patients given psilocybin, the mean directedness of the hierarchy is decreasing, whereas it increases for patients given escitalopram. Furthermore, this reconfiguration in hierarchy is also found at the node level, as shown in the new analysis presented in Supplementary Fig. 1, which shows the statistical significance of the differential reconfiguration with differential, opposite statistical effect responses (one is positive and the other negative).

We also carried out a rigorous analysis matching patients with a similar baseline and change in depression score and tested for differential, opposite statistical effect responses. Supplementary Fig. 2 shows the changes in directedness for patients with similar initial BDI baseline scores (>23) and changes in BDI after and before (<-10). The results show that the mean directedness hierarchy decreases for the psilocybin arm ($P < 0.05$, paired Wilcoxon test using 10,000 permutations), but increases for the escitalopram arm ($P < 0.05$, paired Wilcoxon test using 10,000 permutations).

More detailed information was obtained from running four different linear mixed-effect models. Supplementary Table 4 provides the details of a linear mixed-effect model for paired repeated measures analysis with a random effect and matching (grouping) patients

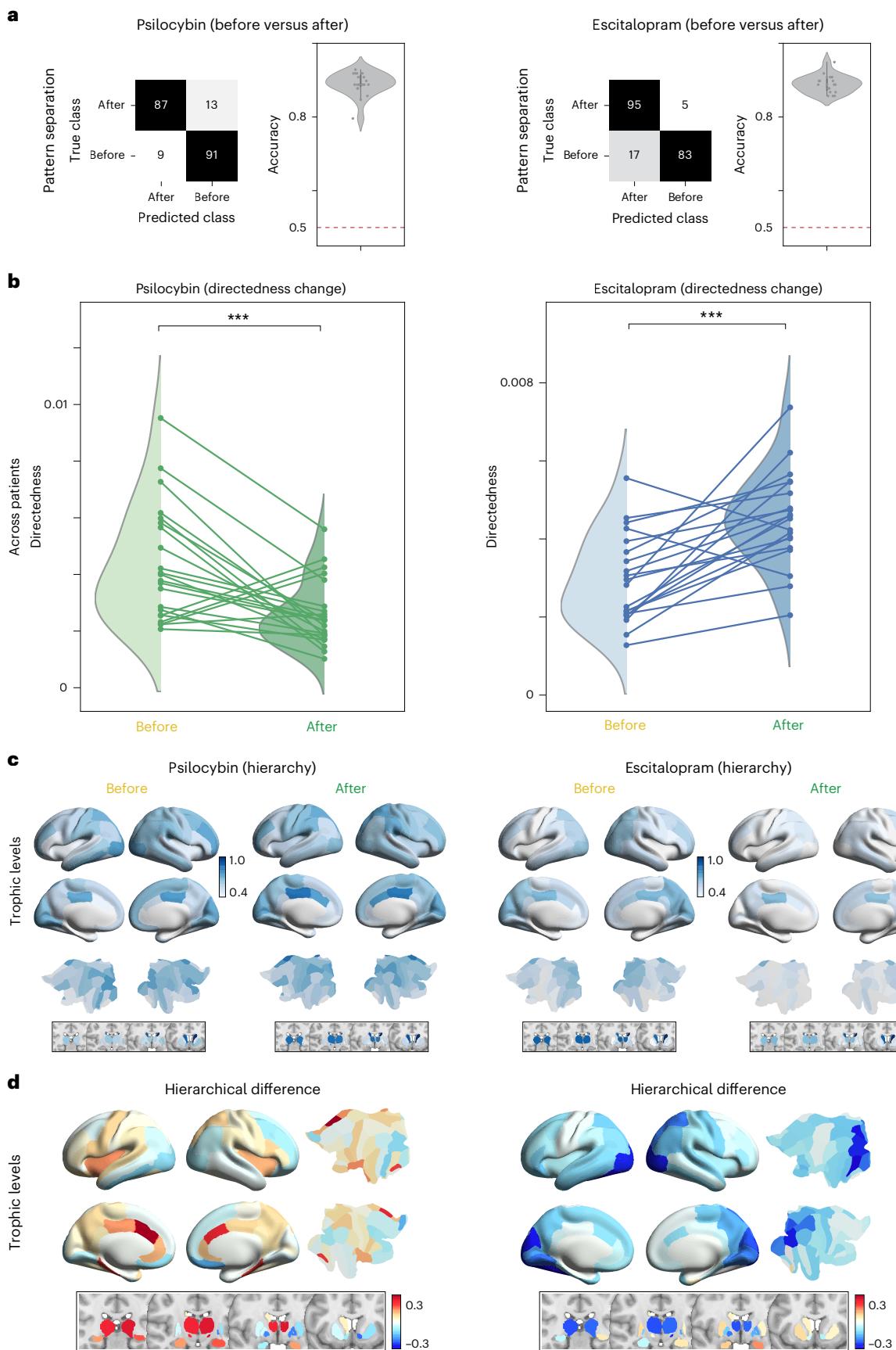
This shows significant differences between the before and after (paired Wilcoxon rank test 10,000 permutations, two-sided and Cohen's d for the effect size; psilocybin, $***P = 0.002$, $d = 1.1$; escitalopram, $***P = 0.001$, $d = 1.4$). Importantly, the change in directedness goes in opposite directions, reflecting the differential hierarchical reconfiguration for the two pharmacological interventions.

c, The average regional trophic levels are rendered on the brain before and after the intervention. The subcortical regions are rendered on four coronal slices in Montreal Neurological Institute (MNI) space ($y = [-26, -14, -2, -10]$ mm).

d, The rendering shows the differences in trophic levels between before and after treatment. As can be seen, the effects of the two interventions are very different, reflecting the fact that psilocybin mostly increases the hierarchical organization, whereas escitalopram causes a general decrease in hierarchical organization.

according to (1) their BDI baseline, (2) their BDI change (post–pre) and (3) their treatment type (psilocybin or escitalopram) to predict hierarchy change (directedness change post–pre). For this procedure, we first used *k*-means clustering to find the optimal silhouette

and used $n = 5$ clusters to define the grouping. We thus matched five similar groups of participants across the whole group of participants and ran the linear mixed-effect model in MATLAB (2022b). The results showed an intercept (bias) t -statistic = -0.394 ($P = 0.695$ (NS)), BDI



baseline t -statistic = 0.431 ($P = 0.668$ (NS)), BDI change t -statistic = 0.399 ($P = 0.692$ (NS)) and treatment t -statistic = -5.152 ($P < 0.001$). This highly significant result demonstrates that only the treatment is important, without any confounding factors.

Supplementary Table 5 provides the details of a linear mixed model of directedness using BDI change as a fixed effect grouped by treatment as the random effect. The results showed an intercept (bias) t -statistic = -0.051 ($P = 0.96$ (NS)) and BDI change t -statistic = 0.409 ($P = 0.68$ (NS)). The lack of significance for intercept and BDI change when using treatment as the random effect shows that treatment is the important factor.

Supplementary Table 6 provides the details of a linear mixed model of directedness change using BDI change, BDI baseline, QIDS baseline, QIDS change, age, sex, illness duration, medication withdrawal and medication as a fixed effect grouped by treatment as the random effect. As can be seen in the table, none of the results are significant and are therefore not confounding factors.

Supplementary Table 7 provides the details of a linear mixed model of directedness change using treatment as a fixed effect and using other variables as random effects: BDI baseline, QIDS baseline, QIDS change, age, sex, illness duration, medication withdrawal and medication. The results showed an intercept (bias) t -statistic = 0.526 ($P = 0.60$ (NS)) and treatment t -statistic = -4.08 ($P < 0.001$). This highly significant result demonstrates that only the treatment is important, without any confounding factors.

Overall, all these investigations (Fig. 2b, Supplementary Figs. 3, 4 and 6 and Supplementary Tables 4–7) converge on the fact that the differential, opposite statistical effect responses in hierarchy are consistently changed in opposite directions for the psilocybin and escitalopram arms. This strongly supports the main finding of a differential reconfiguration of hierarchy.

In addition to this hierarchy analysis, we also tested this claim using more traditional methods based on FC. We first tested the ability of FC (Supplementary Fig. 3a) to distinguish the effects of psilocybin and escitalopram before and after intervention. We compared the mean across all elements in FC matrices and did not find any significant differences between before and after intervention. Second, we tested the ability of time-shifted FC (with a shift of 1 repetition time (TR); Supplementary Fig. 3b) using a measure of asymmetry of these matrices before and after intervention. In particular, we computed asymmetry as the mean of the absolute difference between the respective time-shifted FC matrix and the transposed matrix. We found a significant difference for the psilocybin session ($P < 0.05$, paired Wilcoxon test using 10,000 permutations) but not for the escitalopram session. Still, the effect for psilocybin is less statistically significant than what we found using the hierarchy measure.

This lack of sensitivity of both alternate measurements is not surprising given that these methods are focused on grand average static, spatial correlation and lack sensitivity to temporal dynamics. Even more importantly, in the case of FC, this measure is symmetric, meaning that FC does not capture the asymmetry intrinsic to hierarchical organization. On the other hand, the time-shifted FC is a simple but effective measure to capture asymmetry. Yet, it only captures static spatial correlations across the entire time window. In contrast, the hierarchy measure captures the hierarchical organization of spatiotemporal dynamics generated from a whole-brain model. It is therefore more sensitive to changes for both interventions and, more importantly, reveals the differential, opposite statistical effect responses.

Additionally, we tested the power of using our measure of hierarchical organization to distinguish treatment effects in a relatively small sample of patients and repeated these procedures using another measure of global brain connectivity (GBC; Methods), which captures the functional coupling of each region with the rest of the brain. Supplementary Fig. 4 shows that this measure performs significantly worse using GBC than when using hierarchical levels. This poorer

performance can be seen in the differences in the scatterplots in Supplementary Fig. 4a for GBC and hierarchical levels before and after both psilocybin and escitalopram. To further quantify this difference, we directly tested the inability of FC for distinguishing treatment effects by comparing the individual correlations after versus before treatment: psilocybin ($P < 0.001$, paired Wilcoxon test with 10,000 permutations) and escitalopram ($P < 0.04$, paired Wilcoxon test with 10,000 permutations) (see the violin plots in Supplementary Fig. 4b).

Regional differences in hierarchical reconfiguration

To achieve a better understanding of the regional changes underlying the differential hierarchical reconfiguration for each treatment arm, Fig. 2c shows cortical and subcortical renderings of the average trophic regional levels (across patients) for before and after treatment with psilocybin (left column) and escitalopram (right column). Figure 2d shows renderings of the differential trophic levels, demonstrating the hierarchical reconfiguration and how the differential effects of the two interventions result in strikingly different patterns of regional trophic levels. Although the small sample size makes statistical testing at the regional level more difficult due to the multiple comparisons, Supplementary Table 1 presents quantitative information reporting on the top 20% of regions involved in the hierarchical reorganization overall in each intervention arm (irrespective of whether responders or non-responders). As can be seen, the psilocybin treatment changes the hierarchy for a large number of mainly cortical regions, many of which have been shown to be part of the global workspace²⁸. The top regions include both cortical (left posterior cingulate and left rostral anterior cingulate cortices) and subcortical (right hippocampus and right amygdala) regions. In contrast, the overall response to escitalopram treatment does not change the hierarchy of cortical regions, but only subcortical regions (of which the amygdala, left putamen and right hippocampus are found in the global workspace).

In spite of the small sample size, we also carried out rigorous statistical testing. Supplementary Table 2 shows the results for the full psilocybin treatment arm (22 patients), where 61 of 80 regions survive statistical testing ($P < 0.05$, paired Wilcoxon test using 10,000 permutations and false discovery rate (FDR)-corrected for multiple comparisons). In contrast, only 4 of 80 regions survive similar statistical testing for the escitalopram treatment arm (20 patients, $P < 0.05$, paired Wilcoxon test using 10,000 permutations and FDR-corrected for multiple comparisons). Please note that this difference in number of significant regions between the two treatment arms is also found in the classification.

Reconfiguration in responders and non-responders

Importantly, we also found significant differences between responders and non-responders in the hierarchical reconfigurations of brain dynamics following psilocybin and escitalopram treatment. Given the potential limitations of statistical testing in the small sample, we used machine learning (again with cross-validation; Methods) to quantify the significant hierarchical reorganization related to treatment response. We used the efficacy of treatment using the amended (revised) Beck Depression Inventory (BDI-1A), which captures a broad range of symptoms and places emphasis on the cognitive features of depression⁴⁴. For further justification for the use of this measure here, see ref. 45. Here we used the classification from Daws and colleagues²², where a patient is classified as a responder if they show a significant reduction in BDI-1A after treatment (details in ref. 22).

Figure 3a shows the results of using machine learning for testing significantly above chance level accuracies of pattern separation between responders versus non-responders for each treatment arm (0.67 ± 0.05 for psilocybin and 0.69 ± 0.04 for escitalopram, using the methods described above). The highest accuracy was obtained using five regions for psilocybin and six regions for escitalopram treatment.

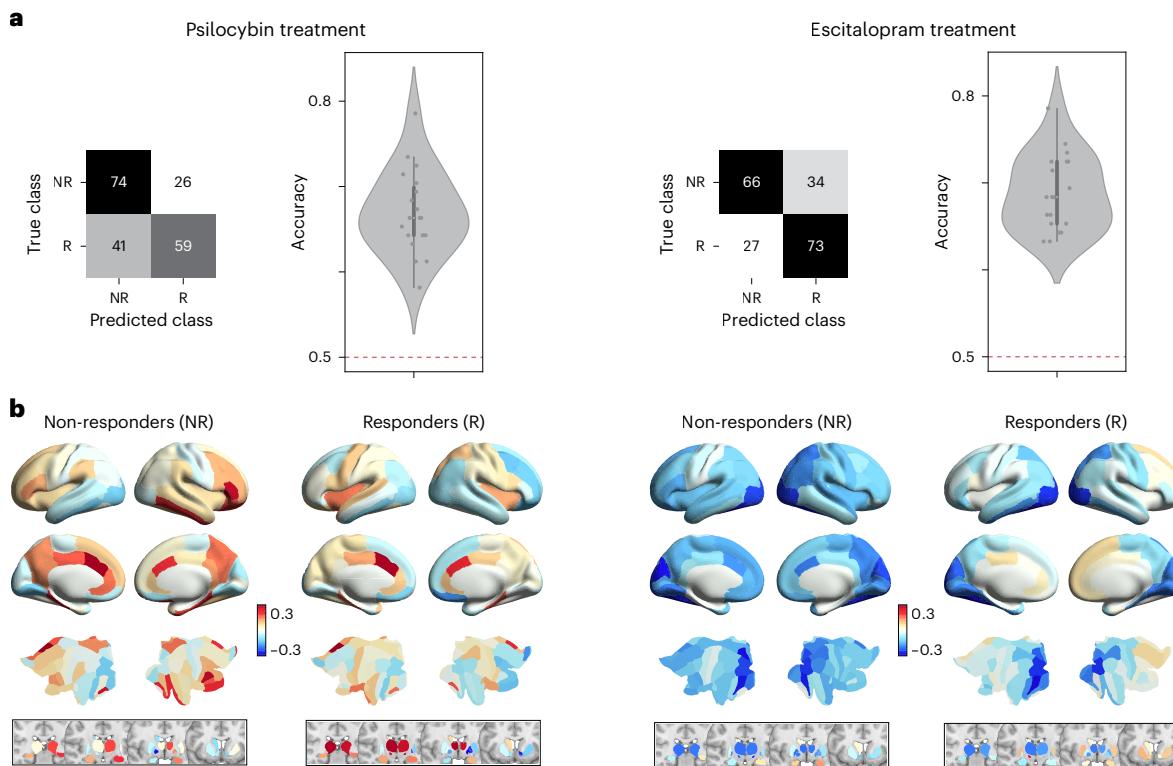


Fig. 3 | Significant differential hierarchical reconfiguration between responders and non-responders. Psilocybin results are shown in the left column and escitalopram results in the right column. **a**, The changes in hierarchical reorganization were used for machine-learning classification of responders versus non-responders. The confusion matrices show the percentage results of the 22 patients treated with psilocybin (of which there were 18 responders) and the 20 patients treated with escitalopram (of which there were eight responders). The results showed significantly above chance levels in hierarchical reconfigurations for responders and non-responders in both interventions (0.67 ± 0.05 (mean \pm s.d.) for psilocybin and 0.69 ± 0.04 for escitalopram). **b**, The renderings show the differences between before and after

for responders and non-responders. The subcortical regions are rendered on four coronal slices in MNI space ($y = [-26, -14, -2, -10]$ mm). As can be seen, there are significant differences in the hierarchical reconfiguration between these two groups, both between and within psilocybin and escitalopram. The renderings show lower precuneus and prefrontal hierarchical levels after treatment in the psilocybin responders compared to non-responders, and there is an intriguing increase in hierarchical levels of the inferior frontal gyrus for the non-responders. Similarly, the brains of the escitalopram responders compare to the full group in that regions of the cingulate and regions of the prefrontal cortex have stronger hierarchical levels in the responders than in non-responders.

We systematically investigated the regional changes underlying these different responses. Figure 3b shows renderings of the differences in trophic levels after and before treatment for psilocybin (left column) and escitalopram (right column) for non-responders (left) and responders (right). Confirming the significant machine-learning pattern separation results, within each treatment the responders and non-responders show differences in the regional trophic levels for responders versus non-responders. Supplementary Table 3 presents quantitative information reporting on the top 20% of regions involved in the overall hierarchical reorganization in each intervention arm. These results offer some added insights into the treatment mechanisms. In the case of the escitalopram treatment, the results show that patients improve if the treatment also affects the hierarchical reorganization of cortical regions that could be said to be part of the global workspace. Specifically, the following cortical regions of the global workspace moved up in hierarchy following treatment: left and right posterior cingulate, left rostral anterior cingulate and right rostral anterior cingulate cortices. In addition, the subcortical regions (left putamen, left nucleus accumbens and right amygdala) of the global workspace also moved back up the hierarchy. However, this will need to be confirmed in future larger studies.

In the case of the psilocybin, most of the patients responded and so show hierarchical changes similar to the general treatment response, which is both cortical and subcortical regions of the global workspace moving back up the hierarchy. As expected, non-responders do have a

slightly different response to responders, with prefrontal regions more likely to move up the hierarchy.

We note that, due to the small sample size, statistical testing at the regional level is difficult due to multiple comparisons. Still, for psilocybin responders (18/22), we found 41 regions that were significantly different before and after successful treatment ($P < 0.05$, paired Wilcoxon test using 10,000 permutations and FDR-corrected for multiple comparisons). In escitalopram responders (8/20), there was only one surviving region (left STN), which is at the top of the changing hierarchy in Supplementary Table 3.

Treatment response can be differentiated and predicted by the hierarchical reorganization

Further investigating the treatment response, we show that when comparing the directedness of the hierarchical organization in treatment, the after session gives significant differences across patients ($P < 0.001$, paired Wilcoxon test with 10,000 permutations; Fig. 4a). In other words, the directedness reveals which intervention a given patient has received, confirming the significantly different hierarchical reorganization linked to the two pharmacological interventions. Furthermore, as shown in Fig. 4b, we can also classify the treatment response using the regional trophic levels with a high accuracy of 0.83 ± 0.03 for the pattern-separation machine learning with cross-validation (Methods). Please note that the training set was carefully balanced in terms of number of examples for each class label, and

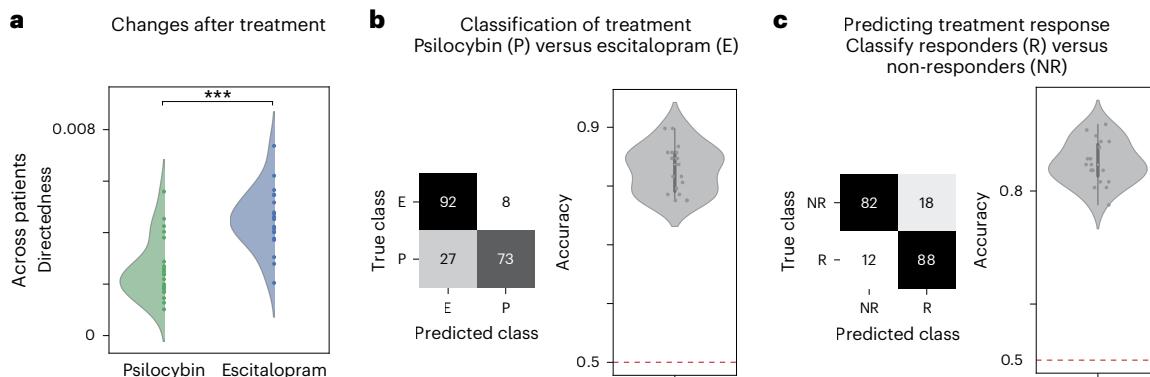


Fig. 4 | Treatment response can be differentiated and predicted by hierarchical organization. **a**, The directedness of the hierarchical organization in the session after treatment shows significant differences across patients ($***P < 0.001$, paired Wilcoxon test with 10,000 permutations). **b**, The machine-learning classification of the treatment response (using the trophic levels) also found that difference between psilocybin versus escitalopram after sessions can be classified with 0.83 ± 0.03 (mean \pm s.d.) accuracy. The confusion matrices show the percentage results of 22 patients treated with psilocybin and 20

patients treated with escitalopram. **c**, Equally, machine learning can be used to make a prediction of the treatment response for patients on escitalopram. We only do this for escitalopram, because the roughly equal numbers of responders (eight patients) and non-responders (12 patients) made this possible. The confusion matrices show the percentage results for the 20 patients treated with escitalopram. This resulted in a significant prediction above chance level and an accuracy of 0.85 ± 0.04 (mean \pm s.d.).

the patients in each class for each shuffling iteration were randomly selected. Finally, we were able to predict the treatment response using the baseline data (that is, the before escitalopram treatment data). This was only done for escitalopram, because this arm, unlike psilocybin, had the necessary power, with roughly equal numbers of responders and non-responders. Using machine learning with cross-validation, Fig. 4c shows a high accuracy of prediction (0.85 ± 0.04). This significant prediction is potentially interesting for future clinical applications.

Discussion

We have studied the effects of pharmacological interventions on the brain dynamics of MDD patients before and after treatment with either psilocybin or escitalopram. Directly addressing our main research question, we were able to show that—despite leading to equal improvements in depressive symptoms (as measured with BDI-1A)—the two drugs work in significantly different ways and show differential, opposite statistical effect responses, as indexed by how each reconfigures the global functional hierarchy of brain dynamics. In particular, we have demonstrated that hierarchical measures of both directedness and trophic levels associated with brain dynamics are significantly different before and after intervention, as well as different between responders and non-responders for both drug types. Our results complement previous research showing that there are shared and unique changes in brain connectivity following different types of intervention for depression⁴⁶. More generally, the framework presented here offers a principled route to evaluate the effects of any pharmacological intervention with before and after empirical brain dynamics data. To generalize the results, however, they should be replicated in larger populations and for other interventions, but there could be a path for fulfilling the great expectations of using neuroimaging for understanding the underlying mechanisms of neuropsychiatric disorders and successful treatments³².

Hierarchical reorganization for psilocybin and escitalopram

We found that psilocybin treatment leads to a decrease in the overall global measure of directedness of the hierarchy of brain dynamics when comparing brain dynamics after versus before treatment (Fig. 2b). In contrast, escitalopram leads to an increase in the overall global measure of directedness of the hierarchy of brain dynamics.

These differential effects are in line with the known differential effects of the pharmacology of the two drugs. Psilocybin is the pro-drug of psilocin (4-OH-dimethyltryptamine) and has been shown to act mainly through the serotonin 2A receptor (5-HT_{2A}R)^{18,19,47}, thereby

initiating a multi-level plasticity⁴⁸. Psychedelics works partly through 5-HT2A receptor agonism leading to an increase in the sensitivity of excitatory neurons expressing the receptor, which in turn causes dysregulation of spontaneous population-level activity and spike-wave decoupling⁴⁹. Attesting to this, dynamic sensitivity analysis of the systematic perturbation of whole-brain models has been used to identify brain networks that are part of the transition away from a depressive brain state following administration of psilocybin⁵⁰.

In contrast, escitalopram is a member of the most frequently prescribed antidepressant drug class, the so-called selective serotonin reuptake inhibitors (SSRIs). Unlike psilocybin, which has no appreciable affinity or action of the serotonin transporter (5-HTT), SSRIs are thought to rely on reuptake blockade at the 5-HTT.

This suggests an overall flattening of the hierarchy for psilocybin after administration in MDD patients, which is consistent with the evidence from healthy participants on psilocybin and other classic psychedelics^{51,52}. Specifically, psychedelics have been shown to broaden the repertoire of connectivity states^{53–57}, increase the entropy of resting-state activity^{58–61}, and enhance the connectivity between central high-level networks and the rest of the brain^{51,60,62,63}. More generally, the results are consistent with the REBUS theory²¹ and the anarchic brain hypothesis, integrating Friston's free-energy principle⁶⁴ with the entropic brain hypothesis of Carhart-Harris^{58,65}. Here, the main hypothesis is that psychedelics can bring about a relaxation of the precision of high-level priors or 'beliefs' (REBUS), allowing (anarchic) bottom-up information flow, consistent with 'anarchic' (that is, hierarchy-free) system dynamics. This theory is consistent with our results showing psilocybin treatment leading to a hierarchical reconfiguration and general flattening of the hierarchy. In contrast, the increase in directedness of the hierarchical reorganization after escitalopram treatment would imply that the responders show a more top-down hierarchical organization of their brain dynamics. It is tempting to speculate that this top-down organization may also be linked to dampened responsiveness in stress circuitry under these drugs⁶⁶, allowing easier top-down control. However, to test this speculation would require a new, carefully designed study, for which some progress has already been made⁶⁷. Even so, further work is required to better scrutinize this hypothesis.

The results show that psilocybin and escitalopram work in very different ways when used for rebalancing the hierarchical organization of the brain dynamics in depression. Both interventions may lead to a significant reduction in depressive symptoms, but do so in significantly different ways. This was first hypothesized by Carhart-Harris

and Nutt⁶⁶, and implied but not proven in the research of Daws and colleagues²², who used the same data analysed here but only found a change in integration–segregation for psilocybin but not for escitalopram, thus missing the differential, opposite statistical effect. Crucially, as we have demonstrated here, this differential, opposite statistical effect is not revealed using simple symmetric FC measures or even asymmetric time-shifted FC, which do not capture the relevant temporal dynamics. Fully revealing the hierarchical organization of the underlying spatiotemporal dynamics requires the sensitivity of our hierarchical method combined with a causal, generative whole-brain model.

The key argument for causality in connection with whole-brain modeling is that removing a ‘causal’ region from the model results in a significant inability to maintain the fit to empirical data. This is analogous to how lesions are used in animal models, but used now for ‘in silico’ models. Previously, Deco and colleagues lesioned the top and bottom of the hierarchy and showed that only the first lesioning significantly changed the fit of the model to the empirical data²⁸. Here, we first determine the GEC that allows the whole-brain model to generate the best fit to the empirical data. Hence, we know that the regions in this network are causally important for generating the functional dynamics.

Regional hierarchical reorganization for treatments

The regional results show that positive depression recovery is linked to when both cortical and subcortical regions of the global workspace are moving back up in the hierarchy following treatment. However, these changes differ in how psilocybin and escitalopram restore the orchestration of healthy brain function, and the main mechanisms underlying the treatment response for depression involve regional hierarchical reconfiguration.

In terms of hierarchical organization at the regional level, psilocybin treatment (Fig. 2c, left column brain rendering) leads to broad changes across the whole brain, with multiple regions in the cingulate cortex increasing their trophic level when comparing after with before the treatment. Similarly, subcortical regions such as the STN, hippocampus and amygdala are also moving up in the hierarchy and drive the orchestration to a higher degree than before. In contrast, large parts of the prefrontal and temporal cortices decrease their trophic level, suggesting that they are moving down the hierarchical organization.

These patterns of hierarchical reconfiguration are very different after escitalopram treatment for all patients (Fig. 2c, right column brain rendering); in all cortical regions the trophic levels decrease from after to before treatment for all patients, and many subcortical regions (such as amygdala, putamen, hippocampus, caudate and STN) show an increase in trophic levels. This drives the global increase in directedness from before to after treatment, irrespective of outcome. Importantly, comparing the regional trophic levels for responders with non-responders following escitalopram treatment (Fig. 3b, right column), the responders show a significantly different pattern, with increases of trophic level in parts of cortical as well as subcortical regions. These changes in the hierarchical position of brain regions are driving the treatment response. In fact, the patients that respond to either pharmacological intervention show an increase in the hierarchical position of both cortical and subcortical regions. Given that the trophic hierarchical levels are based on the GEC measure coming from a whole-brain model of the brain dynamics, this implicates these brain regions in the healthy transition away from treatment-resistant depression.

These findings fit well with previous literature, which has implicated major depression with disturbances in multiple resting-state networks, including the SN and DMN, which are known to regulate cognitive control and attention^{68–71}. In particular, the findings also fit well with the changes in large-scale FC between networks^{72,73}. Specifically in remitted depressed patients, increased FC has been observed between the DMN and the dorsal attention network, and also within and between

networks such as the SN and executive control networks⁷⁴. There is also an emerging literature showing how an increase in activity in the DMN is associated with rumination and recurrence of depression^{75,76}.

Interestingly, many of the regions changing with treatment have been shown to be part of the global workspace²⁸, which orchestrates healthy brain function. This opens up the interesting hypothesis that depression is caused by a malfunctioning orchestration of brain dynamics, where there is a partial breakdown in one or more of the brain regions at the top of the hierarchy²⁸. This hypothesis may even hold for other neuropsychiatric disorders⁷⁰ and, as such, it would be of considerable interest to further test this hypothesis in other datasets.

Predicting outcome using machine learning

Even more relevant for treatment, we were able to demonstrate that the trophic hierarchical levels for an individual depressed patient at baseline before treatment can be used to predict the outcome of escitalopram treatment (Fig. 4c). However, there were too few non-responders for the psilocybin treatment arm to be able to use this for prediction. This finding for escitalopram is potentially exciting but will need replication in a much larger dataset to test the true predictive value of this finding. The importance of more data can be appreciated when comparing this predictive result with the lower accuracy results in Fig. 3a, where machine learning was used to measure the accuracy in separating the difference between responders and non-responders before and after treatment. This lower accuracy runs counter to the intuition that more information should lead to higher accuracy than the prediction accuracy. Still, in both cases, machine learning performs significantly better than chance, although the noise level is clearly having a major effect with the relatively low sample size. Hence, in future this problem should be further investigated in more appropriate large-scale studies with more non-responders.

Potential limitations

Importantly, it should be noted that there are several potential limitations to this study. First, the results here are based on neuroimaging using BOLD signals from fMRI and thus carry a number of potential limitations. These include the fact that BOLD signals are indirect measures of neural activity⁷⁷ at a coarse spatial scale of 1–2 mm³, corresponding to ~5 million neurons with diverse properties and functions^{78,79}. This heterogeneity of neuronal populations is especially true in the higher-order brain regions often associated with neuropsychiatric disorders⁸⁰. Equally, another limitation of resting-state fMRI is the significant intra-individual rest–retest variability, which has been linked to variations in a number of different factors including diet, diurnal changes, blood pressure and even cognitive load⁸¹. Nevertheless, BOLD signals are still useful given that they are highly correlated with local field potentials and multi-unit activity and thus accurately reflect the activity in local cortical circuits^{78,79}. Second, the study has a relatively low sample size, which limits the statistical power used for regional analyses. As such, the regional results mentioned above are not conclusive but will require further study and replication in larger studies.

Conclusion

Overall, the hierarchy of brain dynamics is clearly a very sensitive measure of change. This is made possible by recent advances in whole-brain modeling using a novel thermodynamics-inspired framework²⁶, providing the level of non-reversibility (or arrow of time) in brain signals^{35,37,43}. This provides the observable for a model-based, direct quantification of the asymmetric interactions, which is then quantified using the trophic framework. The thermodynamic framework provides measures of the non-reversibility of brain states and the identification of brain regions involved in breaking the balance as well as the net fluxes between underlying brain networks before and after interventions. In other words, the arrow of time of brain signals can provide a powerful new way to identify the cause and effect when the

brain is reorganized. Here we have used this framework to identify how different pharmacological interventions reorganize brain dynamics differently. This has provided insights into the underlying mechanisms of depression and may in time lead to even better interventions. The results also confirm the hypothesis that problems with the main regions of the global workspace orchestrating brain dynamics could be the main cause of neuropsychiatric disorders—consistent with previous findings and hypotheses^{70,82}. Future larger studies should further investigate this hypothesis. We also note that the present whole-brain modeling framework could be used for treatment studies using any kind of effective intervention, whether pharmacological, electrical or behavioral.

Methods

Empirical data

The trial's design (Fig. 1a) and primary clinical outcomes (clinicaltrials.gov: NCT03429075) have been documented previously^{22,23}. The clinical trial took place at the National Institute for Health Research Imperial Clinical Research Facility and received sponsorship from Imperial College London. It obtained ethical approval (ID 17/LO/0389) from the NHS Research and Imperial College Joint Research and Compliance Office, as well as approval from the Health Research Authority and Medicines and Healthcare Products Regulatory Agency. This study was carried out under a Schedule 1 Drug Licence granted by the UK Home Office. Written informed consent was provided by all participants who did not receive any financial compensation.

Participants

To be eligible for participation, individuals needed a confirmed diagnosis of unipolar MDD from a general practitioner, scoring 16 or higher on the 21-item Hamilton Depression Rating scale. Patients were also queried about any prior use of psychedelics. Within this trial, 31% of patients in the psilocybin group and 24% in the escitalopram group reported previous experience with psychedelics. Individuals were excluded from the trial if they had an immediate family or personal history of psychosis, a physician-assessed risky physical health condition, a history of serious suicide attempts, a positive pregnancy test or contraindications for undergoing an MRI. In addition, individuals with contraindications for selective serotonin reuptake inhibitors (SSRIs) or previous use of escitalopram were also excluded. It is important to note that treatment resistance was not considered as an inclusion or exclusion criterion. All eligible patients underwent telephone screening interviews, provided written informed consent, and underwent comprehensive evaluations of their mental and physical medical histories. Supplementary Table 8 provides the self-reported demographics for all patients.

Interventions

Of the 59 recruited patients with MDD, a random number generator was used to assign 30 patients to the psilocybin arm and 29 patients to the escitalopram arm. Similar to Daws and colleagues²², 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. The remaining 22 patients (mean age, 41.9 years, s.d. = 11.0, 14 men and 8 women) were included in the psilocybin imaging sample. 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. The remaining 20 patients (mean age, 38.7 years, s.d. = 11.0, 14 men and 6 women) were included in the escitalopram imaging sample. Before treatment, all patients underwent a baseline resting-state fMRI session with their eyes closed. On the

first dosing day (DD1), patients received either 25 mg of psilocybin (psilocybin arm) or a presumed negligible dose of 1 mg of psilocybin (escitalopram arm). Although all patients were informed that they would receive psilocybin, they were unaware of the specific dosage to ensure blinding. A second dosing day (DD2) took place three weeks after DD1, where patients received the same dosage as in the first session. There was no crossover in dosages between the two arms. Starting from the day after DD1, patients took daily capsules for a total of six weeks and one day. In both conditions, patients ingested one capsule per day during the initial three 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 three weeks and a total of 2×10 mg (20 mg) thereafter. Blinding was not done in the clinical trial upon which the data were based. However, there are recent data showing that psychedelic therapy may be less vulnerable to expectancy biases than previously suspected⁸³. In this study, the authors analyzed the clinical data with mixed linear models to investigate the association between both escitalopram and psilocybin and pre-treatment efficacy-related expectations, baseline trait suggestibility and absorption and therapeutic response. The results showed that patients overall had significantly higher expectancy for psilocybin relative to escitalopram. Yet, expectancy for escitalopram was associated with improved therapeutic outcomes to escitalopram, and expectancy for psilocybin was not predictive of the response to psilocybin. The BDI remission rates were 64% for the psilocybin arm and 30% for the escitalopram arm.

Measuring depression severity

The severity of depression was evaluated using BDI (BDI-1A) scores in this study. BDI-1A is a patient-rated assessment tool that encompasses a wider range of symptoms and places particular emphasis on the cognitive aspects of depression⁴⁴. Baseline BDI assessments were conducted before the first dosing day (DD1), and subsequent evaluations took place at two, four and six weeks after DD1. It is important to note that BDI was considered a secondary outcome measure for this study, as indicated by its registration on ClinicalTrials.gov (NCT03429075). The primary outcome measure (QIDS-SR-16) was found not to be different between the two intervention arms²³.

MRI acquisition

Imaging of the brain was carried out with a 3T Siemens Tim Trio set-up at Invicro. For the acquisition of brain anatomy, we used the recommended MPRAGE parameters from Alzheimer's Disease Neuroimaging Initiative, Grand Opportunity (ADNI-GO56): 1-mm isotropic voxels; 160 sagittal slices; 256×256 in-plane field of view; echo time (TE), 2.98 ms; TR, 2,300 ms; generalized autocalibrating partially parallel acquisitions (GRAPPA) acceleration, 2; flip angle, 9°; bandwidth, 240 Hz per pixel.

For the acquisition of functional data, we collected eyes-closed resting-state fMRI data using T2*-weighted echo-planar images. We used a 32-channel head coil to acquire 480 volumes in ~10 min: 3-mm isotropic voxels; 44 axial slices; TE, 30 ms; TR, 1,250 ms; GRAPPA acceleration, 2; flip angle, 70°; bandwidth, 2,232 Hz per pixel.

Parcellations

All neuroimaging data were processed using the DK80 cortical parcellation²⁸, which is a combination of the Mindboggle-modified Desikan–Killiany parcellation⁸⁴ with a total of 62 cortical regions (31 regions per hemisphere)⁸⁵ and 18 subcortical regions, that is, nine regions per hemisphere: hippocampus, amygdala, subthalamic nucleus (STN), globus pallidus internal segment (GPi), globus pallidus external segment (GPe), putamen, caudate, nucleus accumbens and thalamus. We chose this parcellation given our previous work showing that this

is the best compromise between spatial accuracy and computational load for whole-brain modeling when measuring brain hierarchy in neuropsychiatric disorders²⁸. However, we note that there is no current consensus about what is the best spatial parcellation scheme, as shown by the paper by Eickhoff and colleagues reviewing the literature on the topographic organization of the brain⁸⁶.

fMRI data preprocessing

The imaging data were preprocessed using an in-house pipeline using the FMRIB Software Library (FSL)⁸⁷, Analysis of Functional NeuroImages (AFNI)⁸⁸, Freesurfer⁸⁹ and Advanced Normalization Tools⁹⁰ packages, as described in detail in a previous publication²². Briefly, this pipeline consists of de-spiking, slice time correction, motion correction, brain extraction, rigid body registration to anatomical scans, nonlinear template registration, scrubbing, bandpass filtering, regression with six realignment motion regressors, three tissue signal regressors, draining veins and local white matter. Daws and colleagues were careful to rule out any systematic bias from akathisia or similar movement artifacts, as also shown in their supplementary materials and in particular their supplementary table 1 showing the head motion descriptive statistics²². The DK80 parcellated timeseries used here for hierarchical analysis was extracted from this preprocessed data.

Hopf whole-brain model

In this study we utilized a Stuart–Landau oscillator model to represent the local dynamics of each brain region. This model, which corresponds to the normal form of a supercritical Hopf bifurcation, is widely used for investigating the transition from noisy to oscillatory dynamics⁹¹. Previous research has successfully employed whole-brain Hopf models to replicate important characteristics of brain dynamics observed in electrophysiology^{92,93}, magnetoencephalography⁹⁴ and fMRI⁹⁵. The full mathematical description of the Hopf model and the linearization of this model are available in Supplementary Information.

Model optimization—GEC

To fit the model to the empirical data (BOLD fMRI of each participant in each brain state), we used a pseudo-gradient procedure to optimize the coupling connectivity matrix C , where the starting point was the standard structural connectivity matrix computed from the diffusion MRI data from a specially optimized state-of-the-art Human Connectome Project (HCP) protocol²⁸. The final optimized matrix comprises the effective conductivity values for each anatomical existing pair connections instead of just the diffusion MRI (dMRI)-based density of fibers. Specifically, we iteratively compared the output of the model with the empirical measures of the functional correlation matrix ($FC^{\text{empirical}}$), that is, the normalized covariance matrix of the functional neuroimaging data.

Additionally, to fit the whole-brain level of non-reversibility, we define the forward and reversal matrices of time-shifted correlations for the forward version and respective reversed backward version of a multidimensional timeseries for the different brain regions. This allows us to compute the forward and reversal matrices, expressing the functional causal dependencies between the different variables for the forward and artificially generated reversed backward version of a multidimensional system. We compared the output of the model with the forward normalized τ time-shifted covariances ($FS_{\text{forward}}^{\text{empirical}}(\tau)$). These normalized time-shifted covariances were derived by shifting the empirical covariance matrix $KS_{\text{forward}}^{\text{empirical}}(\tau)$ and dividing each pair (i,j) by $\sqrt{KS_{\text{forward},ii}^{\text{empirical}}(0)KS_{\text{forward},jj}^{\text{empirical}}(0)}$. It is worth noting that these normalized time-shifted covariances break the symmetry of the couplings, resulting in an enhanced fitting quality⁹⁶. To fully capture the asymmetry, we fit the non-reversibility by performing the same procedure on the reversed normalized τ time-shifted covariances ($FS_{\text{reversal}}^{\text{empirical}}(\tau)$).

Using a heuristic pseudo-gradient algorithm, we set $\alpha = \varsigma = 0.00001$ and updated C until achieving a fully optimized fit. More specifically, the updating uses the following form:

$$\begin{aligned} C_{ij} &= C_{ij} + \alpha (FC_{ij}^{\text{empirical}} - FC_{ij}^{\text{model}}) \\ &\quad + \varsigma ((FS_{\text{forward},ij}^{\text{empirical}}(\tau) - FS_{\text{reversal},ij}^{\text{empirical}}(\tau)) \\ &\quad - (FS_{\text{forward},ij}^{\text{model}}(\tau) - FS_{\text{reversal},ij}^{\text{model}}(\tau))) \end{aligned} \quad (1)$$

where $FS_{\text{forward},ij}^{\text{model}}(\tau)$ is defined similarly for forward $FS_{\text{forward},ij}^{\text{empirical}}(\tau)$ and $FS_{\text{reversal},ij}^{\text{model}}(\tau)$ is defined for reversal $FS_{\text{reversal},ij}^{\text{empirical}}(\tau)$. In other words, for the forward version, it is given by the first N rows and columns of the simulated τ time-shifted covariances $KS_{\text{forward}}^{\text{model}}(\tau)$ normalized by dividing each pair (i,j) by $\sqrt{KS_{\text{forward},ii}^{\text{model}}(0)KS_{\text{forward},jj}^{\text{model}}(0)}$, where $KS_{\text{forward}}^{\text{model}}(\tau)$ is the shifted simulated covariance matrix computed as

$$KS_{\text{forward}}^{\text{model}}(\tau) = \exp(\tau J) K \quad (2)$$

where the J matrix is the Jacobian of the linearized Hopf model evaluated at the fixed point (Supplementary Information). It is important to note that $KS_{\text{forward}}^{\text{model}}(0) = K$. The same procedure was applied to the reversal version of $FS_{\text{reversal},ij}^{\text{model}}(\tau)$. The model is executed iteratively with the updated C until a stable and convergent fit is achieved. We use this method in two stages: (1) at the group level (for the before and after for each of the two treatment types); (2) this is then used as the starting point for individual optimization. At the group level, we initialize C as the anatomical connectivity data obtained through probabilistic tractography from dMRI (Methods). The update process only modifies the known existing connections from this matrix within each hemisphere, following the anatomical connections. However, an exception is made for homolog connections between corresponding regions in both hemispheres, as tractography tends to be less accurate in capturing this type of connectivity. For the Stuart–Landau model, as mentioned above, we set $\alpha = \varsigma = 0.00001$ and continue the algorithm until convergence is attained. In each iteration, we compute the model results by averaging over multiple simulations corresponding to the number of participants. In summary, we refer to the optimized matrix C as the GEC²⁶.

Briefly, the main difference between the GEC and other examples of effective connectivity such as dynamic causal modeling (DCM)²⁷, Granger causality⁹⁷ and transfer entropy^{28,98–102} is the anatomically constrained nature of GEC. In most implementations of DCM, Granger and transfer entropy, the underlying brain anatomy is not constraining the results. Furthermore, both methods measure some form of causality from the timeseries but do not use a generative form of causality. In our GEC, this generative aspect is explicitly given by brain anatomical constraints and whole-brain modeling. Importantly, this modeling typically allows for an estimation across the whole brain rather than only for the limited amount of regions given by DCM (up to ten regions in recent implementations), although recent work has applied the framework to the whole brain^{103–105}.

Defining GBC

For each session in each patient, we computed the GBC, given by

$$GBC_i = \frac{1}{N} \sum_{j=1}^N FC_{ij} \quad (3)$$

Functional hierarchy metrics

To investigate functional hierarchical organization following pharmacological treatments in MDD patients, we adapted the hierarchy measures of directedness and trophic levels in directed networks²⁴.

This provides both the hierarchical node-level information (trophic level) and the global information (directedness, or trophic coherence), and here we apply this to the directed graph obtained from the GEC matrix, the individualized optimized matrix C given by equation (1).

Trophic levels

The GEC matrix defines a graph of N nodes connected by weighted edges determined by the elements of C . For each node n in the graph, we introduce the concepts of in-weight (d_n^{in}), by summing the m columns of C , and out-weight (d_n^{out}), by summing the m rows of C , defined as follows:

$$d_n^{\text{in}} = \sum_m C_{nm} \quad (4)$$

$$d_n^{\text{out}} = \sum_m C_{mn} \quad (5)$$

We define the total weight of node n as u_n by

$$u_n = d_n^{\text{in}} + d_n^{\text{out}} \quad (6)$$

Furthermore, we define the imbalance for node n as v_n , representing the difference between the flow into and out of the node by

$$v_n = d_n^{\text{in}} - d_n^{\text{out}} \quad (7)$$

The (weighted) graph-Laplacian operator Λ on vector \mathbf{h} is given by

$$\Lambda = \text{diag}(u) - C - C^T \quad (8)$$

Consequently, the enhanced concept of trophic level corresponds to the solution \mathbf{h} of the linear system of equations

$$\Lambda \mathbf{h} = \mathbf{v} \quad (9)$$

Here each component of the vector \mathbf{h} corresponds to the trophic level in a given brain region. Importantly, while the operator Λ is symmetric, the asymmetry of the network is evident in the imbalance vector \mathbf{v} .

Directedness (trophic coherence) of a network

Once the hierarchy level h has been established, we can assess the network's global directionality by computing its directedness (or trophic coherence) using the equation

$$F_0 = 1 - \frac{\sum_{mn} C_{mn} (h_n - h_m - 1)^2}{\sum_{mn} C_{mn}} \quad (10)$$

A network is considered maximally coherent when $F_0 = 1$, whereas it is regarded as incoherent when $F_0 = 0$. The trophic coherence is a graph theoretical measure of hierarchical organization. Elevated values of trophic coherence indicate a greater degree of hierarchical organization. Using the GEC obtained through the whole-brain model-based framework, we computed hierarchical levels for each region in our parcellation and trophic coherence characterizing the global level of hierarchical organization.

SVM for pattern separation and classification

Both pattern separation and classification were conducted using a support vector machine (SVM) with Gaussian kernels as implemented by the function fitcecoc in MATLAB (2022b), returning a full, trained, two-class, error-correcting output codes model with the predictors in the input with class labels. We used the one-versus-one coding design with $K = 2$ as the number of unique class labels in $K(K - 1)/2$ binary SVM models.

The output of the SVM was two classes, which correspond to (1) after versus before, (2) responder versus non-responder or (3) psilocybin versus escitalopram treatment. The input features used for classification were hierarchical trophic levels (for each patient and condition). We used a selection of inputs by statistically comparing hierarchical trophic levels across patients, region by region, using a Wilcoxon 10,000 permutation tests and sorting these. We selected the minimum number of inputs that yielded the largest accuracy. The SVM was trained using the leave-one-out cross-validation procedure; that is, we randomly chose one patient for generalization and used the rest for training. This was repeated and shuffled 1,000 times. Furthermore, the training set was balanced in terms of number of examples for each class label, while randomly selecting the patients in each class for each shuffling iteration.

To estimate the highest possible accuracy, we systematically computed the accuracy by sequentially selecting different numbers of regions from the sorted list of statistically significant trophic levels mentioned above. In terms of the minimum number of regions used for the pattern separation in Fig. 2a, the highest accuracy was obtained for 69 regions for psilocybin and six regions for escitalopram treatment before versus after. In Fig. 3a, for the pattern separation between responders versus non-responders, the highest accuracy was obtained using five regions for psilocybin and six regions for escitalopram treatment. In Fig. 4b, the highest accuracy comparing treatment was obtained using 33 regions. In Fig. 4c, the highest accuracy was obtained using five regions for the prediction of treatment response.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All requests for raw and analysed data and materials are promptly reviewed by R.L.C.-H., the chief investigator on the original work. Patient-related data not included in the paper were generated as part of clinical trials and may be subject to patient confidentiality. All data needed to evaluate the conclusions in the paper are present in the paper and/or Supplementary Information.

Code availability

Code used to analyze the data is available from <https://github.com/decolab/Psilodep2>.

References

1. Vigo, D., Thornicroft, G. & Atun, R. Estimating the true global burden of mental illness. *Lancet Psychiatry* **3**, 171–178 (2016).
2. Vigo, D., Jones, L., Atun, R. & Thornicroft, G. The true global disease burden of mental illness: still elusive. *Lancet Psychiatry* **9**, 98–100 (2022).
3. World Health Organization. *Depression and Other Common Mental Disorders: Global Health Estimates* (World Health Organization, 2017).
4. Cipriani, A. et al. Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. *Lancet* **391**, 1357–1366 (2018).
5. Hofmann, S. G., Curtiss, J., Carpenter, J. K. & Kind, S. Effect of treatments for depression on quality of life: a meta-analysis. *Cogn. Behav. Ther.* **46**, 265–286 (2017).
6. Locher, C. et al. Efficacy and safety of selective serotonin reuptake inhibitors, serotonin-norepinephrine reuptake inhibitors, and placebo for common psychiatric disorders among children and adolescents: a systematic review and meta-analysis. *JAMA Psychiatry* **74**, 1011–1020 (2017).

7. Steinert, C., Hofmann, M., Kruse, J. & Leichsenring, F. Relapse rates after psychotherapy for depression—stable long-term effects? A meta-analysis. *J. Affect. Disord.* **168**, 107–118 (2014).
8. Holtzheimer, P. E. & Mayberg, H. S. Stuck in a rut: rethinking depression and its treatment. *Trends Neurosci.* **34**, 1–9 (2011).
9. Nutt, D. & Carhart-Harris, R. The current status of psychedelics in psychiatry. *JAMA Psychiatry* **78**, 121–122 (2021).
10. Hamilton, J. P. et al. Default-mode and task-positive network activity in major depressive disorder: implications for adaptive and maladaptive rumination. *Biol. Psychiatry* **70**, 327–333 (2011).
11. Goodman, Z. T. et al. Whole-brain functional dynamics track depressive symptom severity. *Cereb. Cortex* **31**, 4867–4876 (2021).
12. Lydon-Staley, D. M. et al. Repetitive negative thinking in daily life and functional connectivity among default mode, fronto-parietal and salience networks. *Transl. Psychiatry* **9**, 234 (2019).
13. Turnbull, A. et al. Reductions in task positive neural systems occur with the passage of time and are associated with changes in ongoing thought. *Sci. Rep.* **10**, 9912 (2020).
14. Wilkinson, P. O. & Goodyer, I. M. Attention difficulties and mood-related ruminative response style in adolescents with unipolar depression. *J. Child Psychol. Psychiatry* **47**, 1284–1291 (2006).
15. Jamieson, A. J., Harrison, B. J., Razi, A. & Davey, C. G. Rostral anterior cingulate network effective connectivity in depressed adolescents and associations with treatment response in a randomized controlled trial. *Neuropsychopharmacology* **47**, 1240–1248 (2022).
16. Sacu, S. et al. Effective connectivity during face processing in major depression—distinguishing markers of pathology, risk and resilience. *Psychol. Med.* **53**, 4139–4151 (2023).
17. Rolle, C. E. et al. Cortical connectivity moderators of antidepressant vs placebo treatment response in major depressive disorder: secondary analysis of a randomized clinical trial. *JAMA Psychiatry* **77**, 397–408 (2020).
18. Kringelbach, M. L. et al. Dynamic coupling of whole-brain neuronal and neurotransmitter systems. *Proc. Natl Acad. Sci. USA* **117**, 9566–9576 (2020).
19. Nichols, D. E. Psychedelics. *Pharmacol. Rev.* **68**, 264–355 (2016).
20. Beliveau, V. et al. A high-resolution in vivo atlas of the human brain's serotonin system. *J. Neurosci.* **37**, 120–128 (2017).
21. Carhart-Harris, R. L. & Friston, K. J. REBUS and the anarchic brain: toward a unified model of the brain action of psychedelics. *Pharmacol. Rev.* **71**, 316–344 (2019).
22. Daws, R. E. et al. Increased global integration in the brain after psilocybin therapy for depression. *Nat. Med.* **28**, 844–851 (2022).
23. Carhart-Harris, R. et al. Trial of psilocybin versus escitalopram for depression. *N. Engl. J. Med.* **384**, 1402–1411 (2021).
24. MacKay, R. S., Johnson, S. & Sansom, B. How directed is a directed network? *R. Soc. Open Sci.* **7**, 201138 (2020).
25. Rodgers, N., Tino, P. & Johnson, S. Strong connectivity in real directed networks. *Proc. Natl Acad. Sci. USA* **120**, e2215752120 (2023).
26. Kringelbach, M. L., Sanz Perl, Y., Tagliazucchi, E. & Deco, G. Toward naturalistic neuroscience: mechanisms underlying the flattening of brain hierarchy in movie-watching compared to rest and task. *Sci. Adv.* **9**, eade6049 (2023).
27. Friston, K. J., Harrison, L. & Penny, W. Dynamic causal modelling. *NeuroImage* **19**, 1273–1302 (2003).
28. Deco, G., Vidaurre, D. & Kringelbach, M. L. Revisiting the Global Workspace orchestrating the hierarchical organisation of the human brain. *Nat. Human Behav.* **5**, 497–511 (2021).
29. Levine, S. Several measures of trophic structure applicable to complex food webs. *J. Theor. Biol.* **83**, 195–207 (1980).
30. Antràs, P., Chor, D., Fally, T. & Hillberry, R. Measuring the upstreamness of production and trade flows. *Am. Econ. Rev.* **102**, 412–416 (2012).
31. Soramäki, K. & Cook, S. SinkRank: an algorithm for identifying systemically important banks in payment systems. *Economics* **7**, 20130028 (2013).
32. Deco, G. & Kringelbach, M. L. Great expectations: using whole-brain computational connectomics for understanding neuropsychiatric disorders. *Neuron* **84**, 892–905 (2014).
33. Kringelbach, M. L. & Deco, G. Brain states and transitions: insights from computational neuroscience. *Cell Rep.* **32**, 108128 (2020).
34. Breakspear, M. Dynamic models of large-scale brain activity. *Nat. Neurosci.* **20**, 340–352 (2017).
35. Deco, G., Sanz Perl, Y., Tagliazucchi, E. & Kringelbach, M. L. The INSIDEOUT framework provides precise signatures of the balance of intrinsic and extrinsic dynamics in brain states. *Commun. Biol.* **5**, 572 (2022).
36. Guzman, E. et al. The lack of temporal brain dynamics asymmetry as a signature of impaired consciousness states. *Interface Focus* **13**, 20220086 (2023).
37. Deco, G. et al. The arrow of time of brain signals in cognition: potential intriguing role of parts of the default mode network. *Netw. Neurosci.* **7**, 966–998 (2023).
38. Deco, G. et al. One ring to rule them all: the unifying role of prefrontal cortex in steering cognitive brain dynamics in review. *Prog. Neurobiol.* **227**, 102468 (2023).
39. de la Fuente, L. A. et al. Temporal irreversibility of neural dynamics as a signature of consciousness. *Cereb. Cortex* **33**, 1856–1865 (2023).
40. Sanz Perl, Y. et al. Non-equilibrium brain dynamics as a signature of consciousness. *Phys. Rev. E* **104**, 014411 (2021).
41. Lynn, C. W., Holmes, C. M., Bialek, W. & Schwab, D. J. Decomposing the local arrow of time in interacting systems. *Phys. Rev. Lett.* **129**, 118101 (2022).
42. Lynn, C. W., Holmes, C. M., Bialek, W. & Schwab, D. J. Emergence of local irreversibility in complex interacting systems. *Phys. Rev. E* **106**, 034102 (2022).
43. Lynn, C. W., Cornblath, E. J., Papadopoulos, L., Bertolero, M. A. & Bassett, D. S. Broken detailed balance and entropy production in the human brain. *Proc. Natl Acad. Sci. USA* **118**, e2109889118 (2021).
44. Fried, E. I. The 52 symptoms of major depression: lack of content overlap among seven common depression scales. *J. Affect. Disord.* **208**, 191–197 (2017).
45. Weiss, B., Erritzoe, D., Giribaldi, B., Nutt, D. J. & Carhart-Harris, R. L. A critical evaluation of QIDS-SR-16 using data from a trial of psilocybin therapy versus escitalopram treatment for depression. *J. Psychopharmacol.* **37**, 717–732 (2023).
46. Dunlop, B. W. et al. Shared and unique changes in brain connectivity among depressed patients after remission with pharmacotherapy versus psychotherapy. *Am. J. Psychiatry* **180**, 218–229 (2023).
47. Vargas, M. V. et al. Psychedelics promote neuroplasticity through the activation of intracellular 5-HT2A receptors. *Science* **379**, 700–706 (2023).
48. Carhart-Harris, R. L. How do psychedelics work? *Curr. Opin. Psychiatry* **32**, 16–21 (2019).
49. Nutt, D., Erritzoe, D. & Carhart-Harris, R. Psychedelic psychiatry's brave new world. *Cell* **181**, 24–28 (2020).
50. Vohryzek, J. et al. Brain dynamics predictive of response to psilocybin for treatment-resistant depression. *Brain Commun.* **6**, fcae049 (2024).
51. Girn, M. et al. Serotonergic psychedelic drugs LSD and psilocybin reduce the hierarchical differentiation of unimodal and transmodal cortex. *NeuroImage* **256**, 119220 (2022).

52. Erritzoe, D. et al. Exploring mechanisms of psychedelic action using neuroimaging. *Nat. Mental Health* **2**, 141–153 (2024).
53. Tagliazucchi, E., Carhart-Harris, R., Leech, R., Nutt, D. & Chialvo, D. R. Enhanced repertoire of brain dynamical states during the psychedelic experience. *Hum. Brain Mapping* **35**, 5442–5456 (2014).
54. Atasoy, S., Vohryzek, J., Deco, G., Carhart-Harris, R. L. & Kringelbach, M. L. Common neural signatures of psychedelics: frequency-specific energy changes and repertoire expansion revealed using connectome-harmonic decomposition. *Prog. Brain Res.* **242**, 97–120 (2018).
55. Varley, T. F., Carhart-Harris, R., Roseman, L., Menon, D. K. & Stamatakis, E. A. Serotonergic psychedelics LSD & psilocybin increase the fractal dimension of cortical brain activity in spatial and temporal domains. *NeuroImage* **220**, 117049 (2020).
56. Cruzat, J. et al. Effects of classic psychedelic drugs on turbulent signatures in brain dynamics. *Netw. Neurosci.* **6**, 1104–1124 (2022).
57. Luppi, A. I. et al. Distributed harmonic patterns of structure-function dependence orchestrate human consciousness. *Commun. Biol.* **6**, 117 (2023).
58. Carhart-Harris, R. et al. The entropic brain: a theory of conscious states informed by neuroimaging research with psychedelic drugs. *Front. Hum. Neurosci.* **8**, 20 (2014).
59. Herzog, R. et al. A mechanistic model of the neural entropy increase elicited by psychedelic drugs. *Sci. Rep.* **10**, 17725 (2020).
60. Timmermann, C. et al. Human brain effects of DMT assessed via EEG-fMRI. *Proc. Natl Acad. Sci. USA* **120**, e2218949120 (2023).
61. Schartner, M. M., Carhart-Harris, R. L., Barrett, A. B., Seth, A. K. & Muthukumaraswamy, S. D. Increased spontaneous MEG signal diversity for psychoactive doses of ketamine, LSD and psilocybin. *Sci. Rep.* **7**, 46421 (2017).
62. Carhart-Harris, R. et al. Neural correlates of the LSD experience revealed by multimodal neuroimaging. *Proc. Natl Acad. Sci. USA* **113**, 4853–4858 (2016).
63. Tagliazucchi, E. et al. Increased global functional connectivity correlates with LSD-induced ego dissolution. *Curr. Biol.* **26**, 1043–1050 (2016).
64. Friston, K. The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* **11**, 127–138 (2010).
65. Carhart-Harris, R. The entropic brain-revisited. *Neuropharmacology* **142**, 167–178 (2018).
66. Carhart-Harris, R. L. & Nutt, D. J. Serotonin and brain function: a tale of two receptors. *J. Psychopharmacol.* **31**, 1091–1120 (2017).
67. Matthew, B. W. et al. Reduced brain responsiveness to emotional stimuli with escitalopram but not psilocybin therapy for depression. Preprint at medRxiv <https://doi.org/10.1101/2023.05.29.23290667> (2023).
68. Warlow, S. M., Naffziger, E. E. & Berridge, K. C. The central amygdala recruits mesocorticolimbic circuitry for pursuit of reward or pain. *Nat. Commun.* **11**, 2716 (2020).
69. Kaiser, R. H., Andrews-Hanna, J. R., Wager, T. D. & Pizzagalli, D. A. Large-scale network dysfunction in major depressive disorder: a meta-analysis of resting-state functional connectivity. *JAMA Psychiatry* **72**, 603–611 (2015).
70. Menon, V. Large-scale brain networks and psychopathology: a unifying triple network model. *Trends Cogn. Sci.* **15**, 483–506 (2011).
71. Mulders, P. C., van Eijndhoven, P. F., Schene, A. H., Beckmann, C. F. & Tendolkar, I. Resting-state functional connectivity in major depressive disorder: a review. *Neurosci. Biobehav. Rev.* **56**, 330–344 (2015).
72. Li, G. et al. Large-scale dynamic causal modeling of major depressive disorder based on resting-state functional magnetic resonance imaging. *Hum. Brain Mapp.* **41**, 865–881 (2020).
73. Liu, X. et al. Disrupted rich-club network organization and individualized identification of patients with major depressive disorder. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **108**, 110074 (2021).
74. Liu, J. et al. The neuroprogressive nature of major depressive disorder: evidence from an intrinsic connectome analysis. *Transl. Psychiatry* **11**, 102 (2021).
75. Lythe, K. E. et al. Self-blame-selective hyperconnectivity between anterior temporal and subgenual cortices and prediction of recurrent depressive episodes. *JAMA Psychiatry* **72**, 1119–1126 (2015).
76. Marchetti, I., Koster, E. H., Sonuga-Barke, E. J. & De Raedt, R. The default mode network and recurrent depression: a neurobiological model of cognitive risk factors. *Neuropsychol. Rev.* **22**, 229–251 (2012).
77. Logothetis, N. K., Pauls, J., Augath, M., Trinath, T. & Oeltermann, A. Neurophysiological investigation of the basis of the fMRI signal. *Nature* **412**, 150–157 (2001).
78. Bandettini, P. A. & Ungerleider, L. G. From neuron to BOLD: new connections. *Nat. Neurosci.* **4**, 864–866 (2001).
79. Logothetis, N. K. What we can do and what we cannot do with fMRI. *Nature* **453**, 869–878 (2008).
80. Price, J. L. & Drevets, W. C. Neural circuits underlying the pathophysiology of mood disorders. *Trends Cogn. Sci.* **16**, 61–71 (2012).
81. Specht, K. Current challenges in translational and clinical fMRI and future directions. *Front. Psychiatry* **10**, 924 (2019).
82. Carhart-Harris, R. L. et al. Canalization and plasticity in psychopathology. *Neuropharmacology* **226**, 109398 (2023).
83. Szegedi, B. et al. Assessing expectancy and suggestibility in a trial of escitalopram v. psilocybin for depression. *Psychol. Med.* **54**, 1717–1724 (2024).
84. Desikan, R. S. et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage* **31**, 968–980 (2006).
85. Klein, A. & Tourville, J. 101 labeled brain images and a consistent human cortical labeling protocol. *Front. Neurosci.* **6**, 171 (2012).
86. Eickhoff, S. B., Yeo, B. T. T. & Genon, S. Imaging-based parcellations of the human brain. *Nat. Rev. Neurosci.* **19**, 672–686 (2018).
87. Smith, S. M. et al. Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage* **23**, S208–S219 (2004).
88. Cox, R. W. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res.* **29**, 162–173 (1996).
89. Dale, A. M., Fischl, B. & Sereno, M. I. Cortical surface-based analysis. I. Segmentation and surface reconstruction. *NeuroImage* **9**, 179–194 (1999).
90. Avants, B. B., Tustison, N. & Song, G. Advanced normalization tools (ANTS). *Insight J.* **2**, 1–35 (2009).
91. Kuznetsov, Y. A. *Elements of Applied Bifurcation Theory* (Springer, 1998).
92. Freyer, F. et al. Biophysical mechanisms of multistability in resting-state cortical rhythms. *J. Neurosci.* **31**, 6353–6361 (2011).
93. Freyer, F., Roberts, J. A., Ritter, P. & Breakspear, M. A canonical model of multistability and scale-invariance in biological systems. *PLoS Comput. Biol.* **8**, e1002634 (2012).
94. Deco, G. et al. Single or multi-frequency generators in on-going MEG data: a mechanistic whole-brain model of empirical MEG data. *NeuroImage* **152**, 538–550 (2017).
95. Deco, G. et al. Awakening: predicting external stimulation forcing transitions between different brain states. *Proc. Natl Acad. Sci. USA* **116**, 18088–18097 (2019).

96. Gilson, M. et al. Effective connectivity inferred from fMRI transition dynamics during movie viewing points to a balanced reconfiguration of cortical interactions. *NeuroImage* **180**, 534–546 (2017).
97. Granger, C. Testing for causality. *J. Econ. Dyn. Control* **2**, 329–352 (1980).
98. Schreiber, T. Measuring information transfer. *Phys. Rev. Lett.* **85**, 461–464 (2000).
99. Vicente, R., Wibral, M., Lindner, M. & Pipa, G. Transfer entropy—a model-free measure of effective connectivity for the neurosciences. *J. Comput. Neurosci.* **30**, 45–67 (2011).
100. Brovelli, A., Chicharro, D., Badier, J. M., Wang, H. & Jirsa, V. Characterization of cortical networks and corticocortical functional connectivity mediating arbitrary visuomotor mapping. *J. Neurosci.* **35**, 12643–12658 (2015).
101. Chicharro, D. & Ledberg, A. Framework to study dynamic dependencies in networks of interacting processes. *Phys. Rev. E* **86**, 041901 (2012).
102. Wibral, M., Vicente, R. & Lindner, M. in *Directed Information Measures in Neuroscience* (eds. Wibral, M. et al.) 3–36 (Springer, 2014).
103. Frassle, S. et al. Regression DCM for fMRI. *NeuroImage* **155**, 406–421 (2017).
104. Prando, G. et al. Sparse DCM for whole-brain effective connectivity from resting-state fMRI data. *NeuroImage* **208**, 116367 (2020).
105. Razi, A. et al. Large-scale DCMs for resting-state fMRI. *Netw. Neurosci.* **1**, 222–241 (2017).

Acknowledgements

G.D. is supported by grant no. PID2022-136216NB-I00 funded by MICIU/AEI/10.13039/501100011033 and 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, and AGAUR research support grant (ref. 2021 SGR 00917) funded by the Department of Research and Universities of the Generalitat of Catalunya. Y.S.P. is supported by the project NEurological MEchanismS of Injury, and Sleep-like cellular dynamics (NEMESIS; ref. 101071900) funded by the EU ERC Synergy Horizon Europe. M.L.K. is supported by the Centre for Eudaimonia and Human Flourishing (funded by the Pettit and Carlsberg Foundations) and the Center for Music in the Brain (funded by the Danish National Research Foundation, DNRF117). 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.

Author contributions

All authors designed the study, developed the methods, performed the analyses, and wrote and edited the paper.

Competing interests

R.L.C.-H. reports receiving consulting fees from COMPASS Pathways, Entheon Biomedical, Mydecine, Synthesis Institute, Tryp Therapeutics and Usona Institute. The other authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s44220-024-00298-y>.

Correspondence and requests for materials should be addressed to Gustavo Deco or Morten L. Kringelbach.

Peer review information *Nature Mental Health* thanks Richard Dear, Alec Jamieson, Jakob Kaminski and Lucina Uddin for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024

¹Center for Brain and Cognition, Computational Neuroscience Group, Department of Information and Communication Technologies, Universitat Pompeu Fabra, Barcelona, Spain. ²Institució Catalana de la Recerca i Estudis Avançats (ICREA), Barcelona, Spain. ³Department of Physics, University of Buenos Aires, Buenos Aires, Argentina. ⁴School of Mathematics, University of Birmingham, Birmingham, UK. ⁵The Alan Turing Institute, London, UK. ⁶Centre for Eudaimonia and Human Flourishing, Linacre College, University of Oxford, Oxford, UK. ⁷Department of Psychiatry, University of Oxford, Oxford, UK. ⁸Centre for Psychedelic Research, Department of Brain Sciences, Imperial College London, London, UK. ⁹Psychedelics Division, Neuroscape, Department of Neurology, University of California, San Francisco, CA, USA. ¹⁰Center for Music in the Brain, Department of Clinical Medicine, Aarhus University, Aarhus, Denmark.  e-mail: gustavo.deco@upf.edu; morten.kringelbach@psych.ox.ac.uk

Corresponding author(s): Krriegelbach, Deco

Last updated by author(s): Jul 8, 2024

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection The MRI sequences used to collect the fMRI data 3T Siemens Tim Trio are part of the proprietary software provided by Siemens

Data analysis Custom-made MATLAB scripts are freely available at <https://github.com/decolab/Psilodep2>

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

All requests for raw and analysed data and materials are promptly reviewed by R.C.H., the chief investigator on the original work. Patient-related data not included in the paper were generated as part of clinical trials and may be subject to patient confidentiality.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	The data is from a phase 2, double-blind RCT where 59 patients were recruited with major unipolar depression. The sample size was chosen to allow between trial-arm comparisons and was sufficient for detecting medium to large effect sizes.
Data exclusions	Of the 29 randomly assigned patients 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 Of the 30 randomly assigned patients in 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 and finally five patients were excluded due to excessive fMRI head motion. Together, this resulted in a final sample of n=20 in the escitalopram-arm and n=22 in the psilocybin-arm.
Replication	The machine learning trained the support vector machine (SVM) with the leave-one-out cross-validation procedure, that is we randomly chose one patient for generalisation and the whole rest for training, repeated and shuffled 1,000 times. Furthermore, we made sure that the training set was balanced in terms of number of examples for each class label, and randomly selecting the patients in each class for each shuffling iteration
Randomization	A random number generator was used to randomly assign patients to the psilocybin-arm or the escitalopram-arm.
Blinding	In this double-blind RCT, the patients and investigators were blind to the trial-arm, psilocybin therapy dosage (a negligible 1mg vs 25mg) and daily capsule contents (escitalopram vs placebo). More specifically, patients in the psilocybin-arm received 2 x 25mg psilocybin therapy sessions and daily placebo tablets. Patients in the escitalopram-arm received 2 x 1mg psilocybin therapy sessions and daily escitalopram tablets.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input type="checkbox"/>	<input checked="" type="checkbox"/> Human research participants
<input type="checkbox"/>	<input checked="" type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input type="checkbox"/>	<input checked="" type="checkbox"/> MRI-based neuroimaging

Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics

Psilocybin arm: 22 patients mean age, 41.9 years, s.d.=11.0, 8 women. Escitalopram-arm: 20 patients mean age, 38.7 years, s.d. = 11.0, 6 women.

To be eligible for participation, individuals needed a confirmed diagnosis of unipolar Major Depressive Disorder (MDD) from a general practitioner, scoring 16 or higher on the 21-item Hamilton Depression Rating scale.

Recruitment

Individuals were excluded from the trial if they had an immediate family or personal history of psychosis, a physician-assessed risky physical health condition, a history of serious suicide attempts, a positive pregnancy test, or contraindications

for undergoing an MRI. In addition, individuals with contraindications for selective serotonin reuptake inhibitors (SSRIs) or previous use of escitalopram were also excluded. It is important to note that treatment resistance was not considered as an inclusion or exclusion criterion. All eligible patients underwent telephone screening interviews, provided written informed consent, and underwent comprehensive evaluations of their mental and physical medical histories.

Ethics oversight

The RCT obtained ethical approval from the NHS research and Imperial College Joint Research and Compliance Office, as well as approval from the Health Research Authority and Medicines and Healthcare products Regulatory Agency. This study was carried out under a Schedule 1 Drug License granted by the UK Home Office.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about [clinical studies](#)

All manuscripts should comply with the ICMJE [guidelines for publication of clinical research](#) and a completed [CONSORT checklist](#) must be included with all submissions.

Clinical trial registration

[clinicaltrials.gov](#): NCT03429075

Study protocol

<https://clinicaltrials.gov/study/NCT03429075>, we have also uploaded the full set of protocols from the trial as supplied with the original publication

Data collection

Study recruitment and data collection: January 2019 to March 2000.

The data was collected at three sites:

Neuropsychopharmacology Unit
Centre for Psychiatry, Division of Brain Sciences, Imperial College London
Burlington Danes Building, Du Cane Rd,
London, W12 0NN

NIHR/Wellcome Trust Imperial Clinical Research Facility (CRF)

Hammersmith Hospital, London, W12 0HS

IMANOVA

Centre for Neuropsychopharmacology
Division of Brain Sciences
Burlington Danes Building, Du Cane Rd,
London
W12 0NN

Outcomes

In the original clinical trial, the primary outcome measure was the Quick Inventory of Depressive Symptoms (QIDS) and changes from baseline to 4-weeks after the (psilocybin) dosing day. The QIDS measures internal states including most of the diagnostic criteria for depression and is sensitive to treatment effects. The 14-item daily QIDS will be collected from study entry to the primary endpoint at 4 weeks post dosing day, using a version of this scale tailored to daily use. Follow-up occurred monthly at 2-6 months. Secondary outcomes include (but were not limited to): additional patient (BDI) and clinician (HAM-D) depression rating scales, well-being (WEMWBS), anxiety (STAI), optimism (LOT-R), personality (BIG-5) and others, plus all imaging outcomes (e.g. anatomical measures, including: morphometry, cortical thickness and tractography; functional measures, including: CBF, BOLD RSFC, signal variance and entropy/complexity, and activations to emotional faces). After the 4-week post dosing day primary end-point, subsequent follow-up was done remotely with the exception of a structured interview at 6 months.

For this study, the causal brain mechanisms were captured by generative effective connectivity, estimated from whole-brain modelling of resting state for each session and patient. Hierarchy was determined for each of these sessions using measures of directedness and trophic levels on the effective connectivity, which captures cycle structure, stability and percolation. Results showed that the two pharmacological interventions caused significantly different hierarchical reconfigurations of whole-brain dynamics with clear interaction effects. Treatment responses were found to depend on re-establishing cortical and subcortical regions in the so-called 'global workspace' closer to the top of the hierarchy. Further, using machine learning revealed significant differential reorganisation of brain hierarchy before and after the two treatments. Machine learning was also able to predict treatment response. Overall, the results demonstrate that psilocybin and escitalopram work in different ways for rebalancing brain dynamics in depression. This confirm the hypothesis that neuropsychiatric disorders could be caused by the breakdown in regions orchestrating brain dynamics from the top of the hierarchy.

Magnetic resonance imaging

Experimental design

Design type

Resting state blood oxygen level dependent fMRI (eyes closed)

Design specifications

Two rs-rMRI scans per patient, each with 480 measurements (TR=1250ms)

Behavioral performance measures

State number and/or type of variables recorded (e.g. correct button press, response time) and what statistics were used to establish that the subjects were performing the task as expected (e.g. mean, range, and/or standard deviation across subjects).

Acquisition

Imaging type(s)	Functional MRI	
Field strength	3T Siemens Tim Trio	
Sequence & imaging parameters	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 ~10min: TR, 1,250ms; TE, 30ms; 44 axial slices; flip angle, 70 degrees; bandwidth, 2,232Hz per pixel; and GRAPPA acceleration, 2).	
Area of acquisition	Whole brain	
Diffusion MRI	<input type="checkbox"/> Used	<input checked="" type="checkbox"/> Not used

Preprocessing

Preprocessing software	Standardized preprocessing methods used an in-house pipeline using FMRIB Software Library (FSL), Analysis of Functional Neuroimages (AFNI), Freesurfer and Advanced Normalization Tools packages, as described in details in Daws et al 2023. Briefly, this pipeline consists of de-spiking, slice time correction, motion correction, brain extraction, rigid body registration to anatomical scans, nonlinear template registration, scrubbing, bandpass filtering, regression with six realignment motion regressors, three tissue signal regressors, draining veins and local white matter.
Normalization	fMRI volumes were co-registered to the anatomical scans using a rigid body registration (BBR, FSL). Non-linear registration to the 2mm MNI brain was applied to the co-registered volumes (Symmetric Normalization (SyN), ANTS).
Normalization template	MNI305 2mm brain
Noise and artifact removal	Voxelwise nuisance regression was used with the six realignment motion regressors and three tissue signal regressors (Ventricles, Freesurfer, eroded in 2 mm space), draining veins (FSL's CSF minus Freesurfer's Ventricles, eroded in 1 mm space) and local white matter (WM) (FSL's WM minus Freesurfer's subcortical gray matter structures, eroded in 2 mm space). Regarding local WM regression, AFNI's 3dLocalstat was used to calculate the mean local WM time series for each voxel, using a 25-mm radius sphere centered on each voxel.
Volume censoring	Scrubbing using a framewise displacement threshold of 0.5 mm, scrubbed volumes were replaced with the mean of the neighboring volumes.

Statistical modeling & inference

Model type and settings	We used the whole-brain modelling framework to analyse the empirical data
Effect(s) tested	Spatiotemporal structure, hierarchy
Specify type of analysis:	<input checked="" type="checkbox"/> Whole brain <input type="checkbox"/> ROI-based <input type="checkbox"/> Both
Statistic type for inference (See Eklund et al. 2016)	Non-parametric permutation testing and Wilcoxon
Correction	False discovery rate (FDR) was applied

Models & analysis

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Functional and/or effective connectivity
<input type="checkbox"/>	<input checked="" type="checkbox"/> Graph analysis
<input checked="" type="checkbox"/>	<input type="checkbox"/> Multivariate modeling or predictive analysis
Functional and/or effective connectivity	Pearson correlation
Graph analysis	Trophic levels defined by the asymmetric generative effective connectivity