Supplementary Data

Permutation Test for Statistical Significance of the Anticorrelation Between State A and State B

Methods

We sought to confirm the statistical significance of the observed anticorrelation by performing a permutation test with 10,000 iterations. For every iteration cycle, each subject's Infomap clustering solution was remapped into a two-state label solution based on the group-level clustering solution. The two-state labeling solution was then permuted for each time point before averaging all timepoints with the same label to create a ''permuted'' State A and State B. The permuted State A and State B activation patterns were then correlated. The 10,000 *r*-values across all iterations formed the null distribution for comparing the observed anticorrelation between State A and State B.

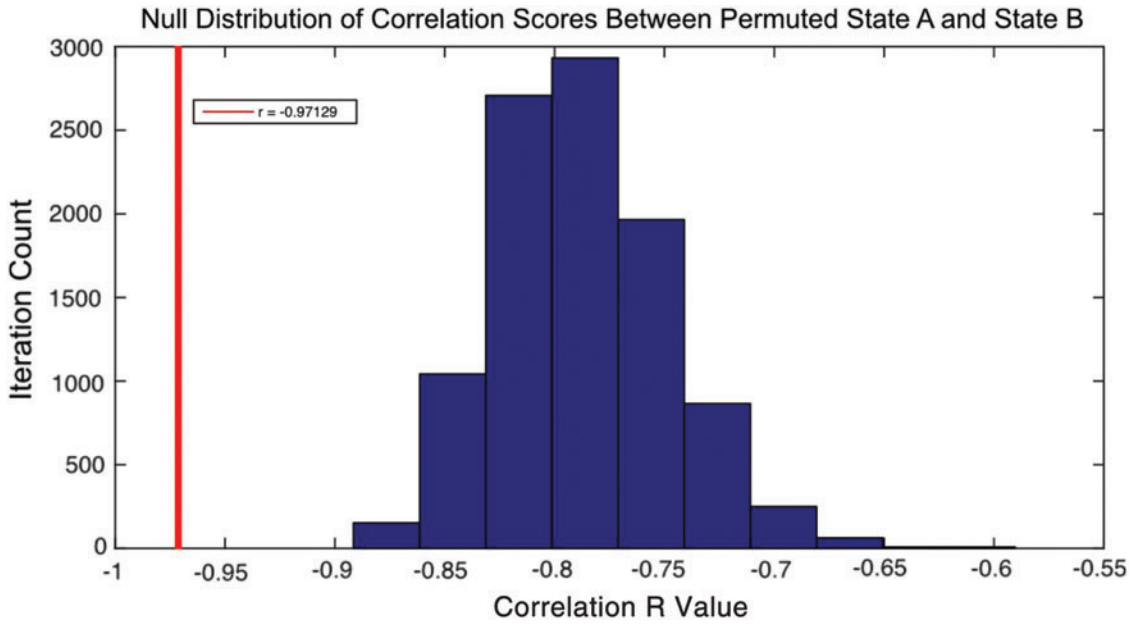
Results

The null distribution derived from the permutation test followed a relatively normal distribution. Pseudo-randomly dividing the resting-state data into two states tended to favor anticorrelated states, as seen in the negative mean of the null distribution. However, the observed anti-correlation State A and State B is statistically significant relative to this null distribution ($p < 0.0001$), supporting the distinctness of State A and State B.

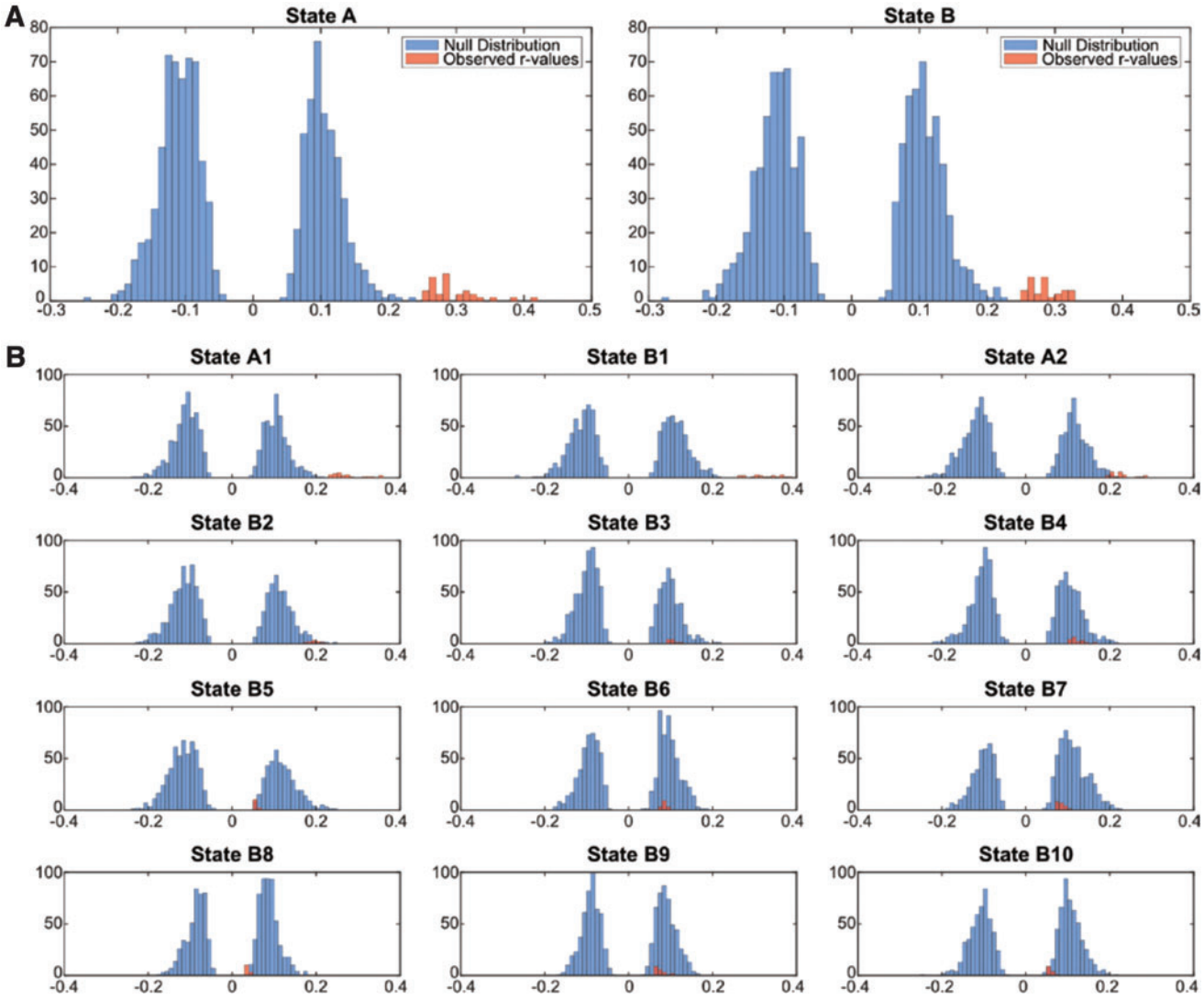
Permutation Test for Statistical Significance of Neurosynth Decoding for All States

Methods

We validated the Neurosynth results by running permutation tests (1000 iterations) for each brain state (State A, State B, and the 12 states at 20% density tier of the hierarchy). Specifically, within each subject, we permuted the order of the brain state temporal windows (consecutive time points of the same state) as returned by the Infomap community detection algorithm. This approach allowed us to preserve many of the temporal properties of brain states (e.g., duration) within each subject, while still permuting the data being averaged for each unique brain state prototype. The newly ''permuted'' brain state prototypes were averaged across subjects according to the group-level clustering solution to produce a ''permuted'' State A and State B. The ''permuted'' State A and State B were run through the Neurosynth decoder to obtain correlation scores for each of the 3406 terms. A total of 1000 permutations cycles were performed for each brain state, returning a total of 340,600 correlation scores per State A and State B. This allowed us to create a null distribution of maximum *r*-values across all 1000 permutations to test the statistical significance of the observed *r*-value for each cognitive term. Note that these permutation tests controlled for multiple comparisons, since all 3406 comparisons were computed during each permutation (this



SUPPLEMENTARY FIG. S1. Permutation test for State A versus State B anticorrelation. A total of 10,000 permutations were performed to create the plotted null distribution. The red vertical line represents the observed anticorrelation *r*-value between State A and State B ($r = -0.97$). The negative null distribution suggests that pseudo-randomly splitting the data into two ''states'' tends to create anticorrelated activation patterns. However, the permutation test indicates that State A and State B were more distinct from one another than they would be by chance ($p < 0.0001$).



SUPPLEMENTARY FIG. S2. Permutation test for Neurosynth decoding of State A and State B. The order of the state time windows was permuted per subject before averaging to obtain ''permuted'' prototypical states. A total of 1000 permutations per state were performed before Neurosynth decoding. The maximum (absolute value) *r*-value across the 3400 terms was selected for each permutation. The observed *r*-value for each cognitive term of interest was then compared to the null distribution formed from the 1000 permutations. **(A)** Results for State A and State B are shown. Note that the blue histogram represents the null distribution and the red histogram represents correlation values for all terms included in the word cloud. All cognitive terms were significant ($p < 0.001$) for both states. **(B)** Results for the 12 states are displayed in the same manner as **(A)**. All terms for A1, A2, B1, and B2 were significant ($p < 0.05$), while no terms from B3 to B10 passed the permutation test. Note that multiple comparisons were corrected as part of the permutation test procedure (see Methods section).

is similar in concept to maxT statistic, see Nichols and Holmes, 2002). The same analysis was also performed for each of the 12 states at the 20% density tier of the hierarchy.

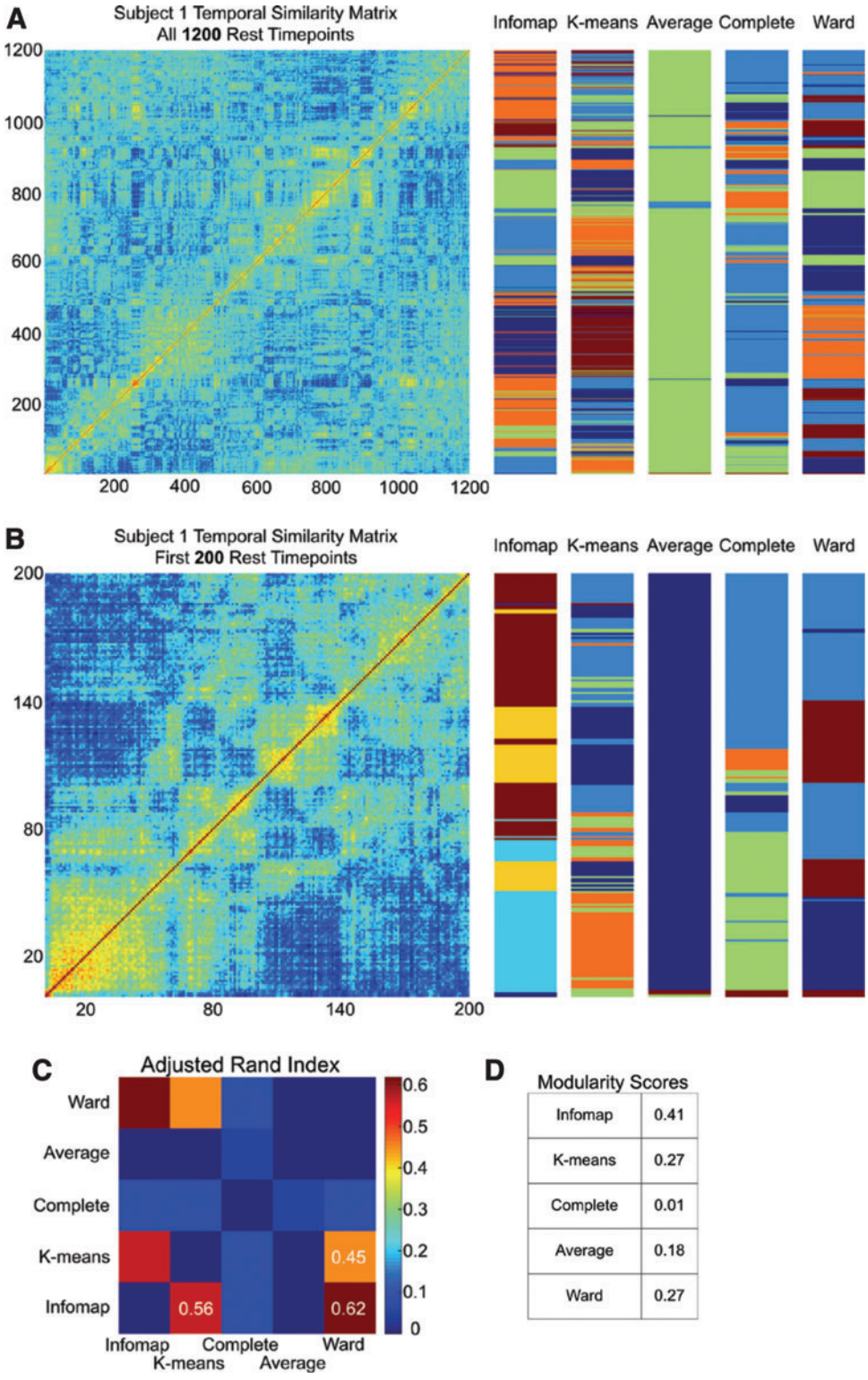
Results

The null distribution for each state was formed by taking the highest correlation score across 3400 comparisons for every permutation cycle ($n = 1000$). Each cognitive term of interest included in the respective word cloud (Fig. 6) was tested against this null distribution for significance. Across all panels in Supplementary Figure S2, the blue histograms represent the null distribution ($n = 1000$) derived from the 1000 permutations and the red histograms illustrates all observed correlation scores for each cognitive term. All terms for State A and State B were significant ($p < 0.001$). In addition, all terms for State A1, A2, B1, and B2 were significant ($p < 0.05$), but no terms from State B3 to B10 passed the permutation test.

Comparisons of Clustering Methods for Brain State Identification

Methods

To find the most suitable class of approach for clustering the temporal similarity matrix, we tested several possible approaches, with the goal of assessing several highly distinct



SUPPLEMENTARY FIG. S3. Comparison of clustering approaches. We tested three main clustering approaches: Infomap community detection, *k*-means clustering, and agglomerative clustering (with three different linkage criterions: complete, ward, average). The clustering solutions for Infomap, *k*-means, and ward clustering were highly similar in terms of visual inspection **(A, B)** and adjusted Rand index scores **(C)**. However, Infomap performed better based on modularity score **(D)**. The first rest run data of HCP subject 100307 (time points 1:200 for 5B) were used for this analysis.

clustering methods for clustering brain states with multiband fMRI. The clustering approaches included Infomap community detection, *k*-means clustering, and agglomerative/hierarchical clustering. Note that this analysis is not a comprehensive test across clustering choices (although similarity across several clustering approaches increased our confidence in the generalizability of approaches). We hypothesized community detection algorithms to provide the best clustering solutions given theoretical evidence, suggesting that community detection algorithms are better at handling boundary cases than other commonly used clustering approaches (Newman, 2006). In this analysis, we applied the clustering approaches on a single subject's resting state data. Infomap was applied as described in the main method. *k*-Means and agglomerative clustering were performed using Python's scikit-learn module. For *k*-means clustering, we chose $k = 5$ to parallel the results from Infomap. Similarly, for agglomerative clustering, we chose the same number of clusters to detect using three different linkage criterions (ward, complete, average) computed using Euclidean distance. To compare between the five clustering solutions, we calculated adjusted Rand index (scikit-learn Python module) and modularity score (Blondel et al., 2008; Rubinov and Sporns, 2011) for each clustering solution (Supplementary Fig. S3C, D). Adjusted Rand index determines how similar a particular clustering solution aligns with the ''ground truth.'' In this analysis, our ''ground truths'' were alternative clustering solutions (e.g., ground truth—Infomap, prediction labels—*k*-means). Modularity measures how well separated the clusters are, which is consistent with our hypothesis for predicting community detection algorithms providing the best clustering solution.

Results

The results of the clustering approaches are juxtaposed next to the first subject's temporal similarity matrix. For visual clarity, we included both the full rest run time series (Supplementary Fig. S3A) and a 200 time point snapshot (Supplementary Fig. S3B). Supplementary Figure S3A provides a clearer perspective of brain state repetition over time (represented by the off-diagonal correlation blocks). Supplementary Figure S3B highlights the temporal boundaries, or transitions, between each unique state cluster. By visual inspection, it seems that the clustering solutions for Infomap, *k*-means, and ward clustering are quite similar. Furthermore, the adjusted Rand index scores (range -1 to 1, $1 =$ perfect match) for the pairwise combination of comparisons of these three specific clustering approaches are also quite similar (Supplementary Fig. S3C), suggesting that all different clustering approaches provide fairly similar clustering solutions. However, consistent with our hypothesis, the modularity score for the Infomap clustering solution was much higher than alternative clustering solutions (Supplementary Fig. S3D). Higher modularity score indicates denser connections within clusters and sparser connections across clusters. Note that this result may be slightly biased in favor of Infomap's solution, as the modularity metric is optimized for testing community detection algorithms. While the similarity across clustering approaches are quite high (as indicated by similar adjusted Rand index scores), Infomap performed better on the modularity score. This led us to our choice of applying Infomap over the alternative choices for our main analyses.

Supplementary References

Blondel VD, Guillaume JL, Lambiotte R, Lefebvre E. 2008. Fast unfolding of communities in large networks. J Stat Mech Theory Exp 2008:10008.

Newman MEJ. 2006. Modularity and community structure in networks. Proc Natl Acad Sci U S A 103:8577–8582.

Nichols TE, Holmes AP. 2001. Nonparametric permutation tests for functional neuroimaging: a primer with examples. Hum Brain Mapp 15:1–25.

Rubinov M, Sporns O. 2011. Weight-conserving characterization of complex functional brain networks. Neuroimage 56:2068–2079.