

Human cognition involves the dynamic integration of neural activity and neuromodulatory systems

James M. Shine^{1*}, Michael Breakspear^{2,3,4}, Peter T. Bell⁵, Kayla Ehgoetz Martens¹, Richard Shine^{1,6}, Oluwasanmi Koyejo⁷, Olaf Sporns⁸ and Russell A. Poldrack⁹

The human brain integrates diverse cognitive processes into a coherent whole, shifting fluidly as a function of changing environmental demands. Despite recent progress, the neurobiological mechanisms responsible for this dynamic system-level integration remain poorly understood. Here we investigated the spatial, dynamic, and molecular signatures of system-wide neural activity across a range of cognitive tasks. We found that neuronal activity converged onto a low-dimensional manifold that facilitates the execution of diverse task states. Flow within this attractor space was associated with dissociable cognitive functions, unique patterns of network-level topology, and individual differences in fluid intelligence. The axes of the low-dimensional neurocognitive architecture aligned with regional differences in the density of neuromodulatory receptors, which in turn relate to distinct signatures of network controllability estimated from the structural connectome. These results advance our understanding of functional brain organization by emphasizing the interface between neural activity, neuromodulatory systems, and cognitive function.

The human brain is a complex adaptive system in which a vast array of behaviors arises from coordinated neural activity across diverse spatial and temporal scales. Linking activity within this large-scale neural architecture to cognitive function remains an important goal for neuroscience.

The analysis of complex brain networks can be used to interrogate the organizational properties of the brain that are crucial for its functional dynamics¹. Network accounts of brain function have hypothesized that a ‘dynamic core’ of regions flexibly guides the flow of activity in the brain to facilitate cognition². These frameworks predict that a distributed set of core regions is active across multiple tasks³ and integrates more specialized regions⁴, altering baseline communication dynamics in service of task-specific computations. Although computational approaches have investigated these large-scale patterns⁵, little is currently known about the mechanisms that facilitate system-wide brain state dynamics as a function of cognition.

One tractable approach to this problem is to exploit the redundancy within complex systems by utilizing dimensionality reduction techniques^{6,7}. These approaches uncover latent functional patterns in complex data sets by distilling brain activity patterns into spatio-temporally similar components⁷. The dynamics of the system can then be interrogated within this low-dimensional space, offering insights into the mechanisms that underlie the system’s function. These approaches have successfully been used in the past to elucidate the functional brain circuitry that underlies the behavioral repertoires of a diverse range of organisms, including the nematode (*Caenorhabditis elegans*), the fruit fly (*Drosophila melanogaster*), and the ferret (*Mustela putorius furo*)^{8–11}. Computational modeling suggests that similar dynamic low-dimensional processes should exist within the human brain¹.

Here, we analyzed whole-brain functional neuroimaging data across a suite of cognitive tasks to identify the low-dimensional

dynamic core of cognition in the human brain. We found that the dynamic functional organization of the brain across a suite of cognitive tasks describes a flow along a low-dimensional state space. This dynamic flow aligns with unique cognitive brain states that recur across distinct cognitive tasks. We next found that the flow of activity reflects an ‘integrative core’ that maximizes information processing complexity over relatively long timescales. Finally, we showed that the axes of this low-dimensional space are closely related to spatial patterns of gene expression for specific families of neuromodulatory receptors, and to unique signatures of structural network controllability, which together provide a plausible biological medium for the modulation of system-level brain dynamics. These results present a novel view of brain function based on the coordinated dynamics of functional brain networks over time and highlight the association between a suite of distinct neuromodulatory systems and cognitive function.

Low-dimensional global brain activity recurs across multiple cognitive tasks

We used 3T functional magnetic resonance imaging (fMRI) data (repetition time (TR) = 0.72 s) from the Human Connectome Project (HCP) to examine blood-oxygen-level-dependent (BOLD) activity from 200 unrelated individuals across seven cognitive tasks, each of which engages distinct cognitive functions: emotional processing, gambling, mathematical calculation, language processing, motor execution, working memory performance, relational matching, and social inference¹². Preprocessed BOLD time series were extracted from 375 cortical and subcortical parcels¹³ and concatenated across all seven tasks and across all subjects. To ensure reproducibility, we initially developed these analyses in a 100-subject discovery data set and then replicated in a 100-subject replication cohort.

Principal component analysis (PCA) was applied to the multi-task BOLD time series to reorganize the regional BOLD data into

¹Brain and Mind Center, The University of Sydney, Sydney, New South Wales, Australia. ²QIMR Berghofer, Brisbane, Queensland, Australia. ³Metro North Mental Health Service, Brisbane, Queensland, Australia. ⁴Hunter Medical Research Institute, University of Newcastle, Callaghan, New South Wales, Australia. ⁵University of Queensland, Brisbane, Queensland, Australia. ⁶Macquarie University, Sydney, New South Wales, Australia. ⁷University of Illinois, Champagne, IL, USA. ⁸Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN, USA. ⁹Department of Psychology, Stanford University, Stanford, CA, USA. *e-mail: mac.shine@sydney.edu.au

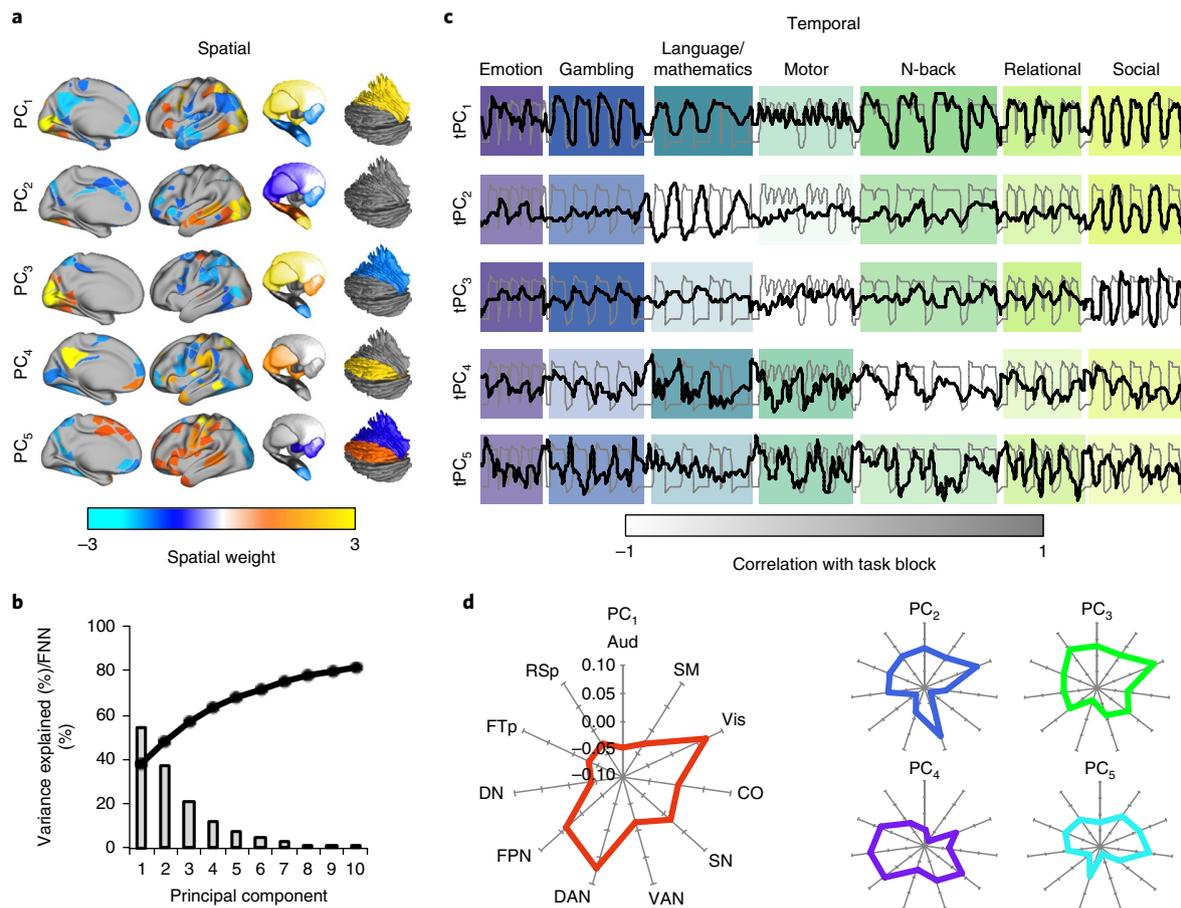


Fig. 1 | Spatiotemporal PCA across multiple cognitive tasks. **a**, Spatial maps for the first five principal components (colored according to spatial weight; thresholded for visualization). **b**, Line plot representing the percentage of variance explained by first ten principal components; bar plot depicting the percentage (single value per component) of false nearest neighbors for first ten principal components. FNN, false nearest neighbors. **c**, Correspondence between convolved, concatenated task block regressor (gray) and the time course of the first five tPCs (black); color intensities of the blocks reflect the Pearson's correlation between tPC₁₋₅ and each of the unique task blocks ($n=100$ subjects). **d**, Mean spatial loading of first five PCs, organized according to a set of predefined networks. DAN, dorsal attention; Vis, visual; FPN, frontoparietal; SN, salience; CO, cingulo-opercular; VAN, ventral attention; SM, somatomotor; RSp, retrosplenial; FTP, frontotemporal; DN, default mode; Aud, auditory.

a smaller set of spatially orthogonal principal components (PCs; Fig. 1a). The spatial maps identified by this approach were similar, but distinct to those identified using an intersubject correlation approach¹⁴ (Supplementary Fig. 1). Across divergent cognitive tasks, we found a dominant, low-dimensional neural signal: the first five PCs accounted for 67.9% of the variance. Furthermore, an analysis of the number of 'false nearest neighbors' demonstrated that the first five PCs resolved greater than 90% of the embedding space¹⁵ (Fig. 1b).

The time series of each PC (tPC) was created by weighting the original BOLD time series from the replication data set with the parcel loading for each spatial component from the discovery data set at each time point of the experiment. This allowed us to track the engagement of each PC over time. The first tPC, which explained 38.1% of signal variance across all tasks, reflected a task-dominant signal that was strongly correlated with the overall task block structure across all seven tasks ($r=0.64$; $P < 0.01$; Fig. 1c) and was strongly replicable across cohorts ($r=0.84$). This result is consistent with previous work that has demonstrated a distributed network of task-invariant brain regions that persists across the performance of multiple unique cognitive tasks³. In addition, the analysis demonstrated prominent spatial overlap between PC₁ and dorsal attention, frontoparietal, and visual networks, along with striatum, thalamus, and lateral cerebellum (Fig. 1d). Notably, fit-

ting a PCA to each of the seven individual tasks separately did not recover the same underlying principal component, but instead identified spatial maps that aligned with the idiosyncratic demands of each task (Supplementary Fig. 2), thus providing evidence that the first PC represented a domain-general core dimension for cross-task cognitive function.

The next three components (tPC₂₋₄) reflected a closer relationship with specific cognitive tasks (Fig. 1c). For example, tPC₂ (10.3% variance explained) was associated with the social task and language processing; tPC₃ (8.4%) with gambling and emotion; and tPC₄ (6.3%) with arithmetic functions (Fig. 1c,d). In contrast, tPC₅ (4.8%) was associated with engagement across multiple tasks ($r=0.20$; $P < 0.01$). Specifically, the time course of tPC₅ was correlated with the absolute value of the first derivative of the task design ($r=0.11$; $P < 0.01$), suggesting that tPC₅ was uniquely associated with the transition into and out of unique task states. In accordance with this finding, tPC₅ was associated with activity across a right-lateralized system of cingulate, parietal, and opercular cortical regions (Fig. 1a) that have previously been shown to have a crucial role in cognitive task engagement and error monitoring¹⁶.

Global brain state dynamics

A substantial benefit of PCA over other analytic approaches is that the technique imposes orthogonality onto the components, which

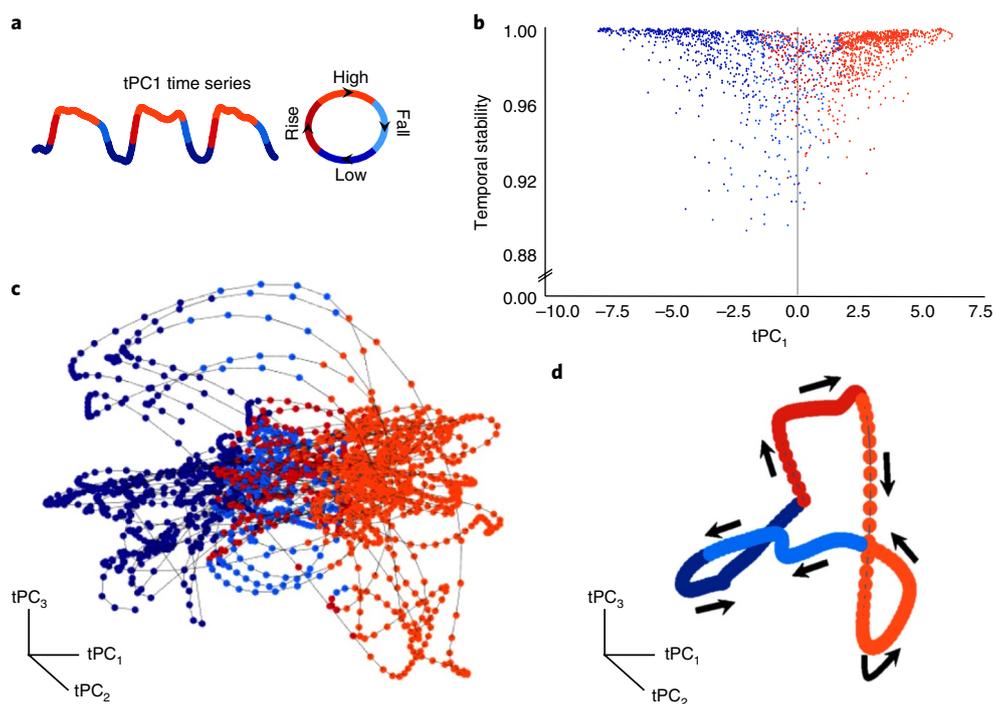


Fig. 2 | The low-dimensional signature across cognitive tasks. **a**, The procedure used to partition tPC_1 into unique phases: low (blue), rise (red), high (orange), and fall (light blue). **b**, Scatter plot comparing the loading of tPC_1 (colored according to the partition defined in **a**) with a temporal stability measure (defined by the similarity of the BOLD response at adjacent time points); we observed a significant positive Pearson's correlation ($r=0.58$) between $|tPC_1|$ and temporal stability ($n=1,939$ time points), providing heuristic evidence for attractor basins at the extremes of tPC_1 engagement. **c**, A three-dimensional scatter plot comparing the first three tPC s; each node represents one time point (colored according to the phase of tPC_1), with time implicitly unfolding across the embedding space (contiguous points connected by black line). **d**, The low-dimensional manifold traversed by the global brain state across the first three dimensions, with arrows depicting the direction of flow along the manifold.

is crucial for providing a low-dimensional subspace in which to embed the state space manifold (Fig. 2). Other popular methods (such as independent component analysis) find a different set of optimal solutions (such as maximal statistical independence), but these are not, in general, linearly independent¹⁷. Notably, state space attractors are invariant to linear transformations of their embedding phase space as long as the dimensions remain orthogonal¹⁸. Hence, PCA enables analysis of the state space trajectory (or flow) of the dominant low-dimensional signal, which in turn reflects the temporal evolution of the global brain state. Any residual variance (that is, from those components not included in the reconstruction) represents a stochastic influence on the ensuing flow.

To facilitate further analysis of the low-dimensional embedding space, the tPC_1 time course was partitioned into relative phase segments⁸: a trough in tPC_1 defined the low phase (blue in Fig. 2a); an increase in tPC_1 defined the rise phase (red); a plateau in the tPC_1 signal defined the high phase (orange); and a decrease in tPC_1 defined the fall phase (light blue) of the low-dimensional flow of tPC_1 (see Methods for details). The resulting phase portrait describes the temporal evolution of the low-dimensional signal shared across all behavioral tasks (Fig. 2c; see Supplementary Fig. 3 for projection of tPC_{4-5}). The different phases of tPC_1 were distinctly related to exogenous task demands, with a greater frequency of high phases during task blocks across all seven tasks, and low phases during interleaved rest blocks across tasks ($P < 0.01$).

To test the hypothesis that the extremes of tPC_1 engagement represented task-driven attractor states¹⁹, we correlated the absolute value of the tPC_1 signal with a temporal stability measure, which was calculated by comparing the BOLD response across the brain at adjacent time points. We found that tPC_1 signal was strongly engaged (both positively and negatively) during periods of relative

stability ($r=0.58$; $P < 0.001$; Fig. 2b), which is consistent with the presence of state space attractors at the extremes of the tPC_1 gradient. This yields a novel picture of brain functioning in which large-scale brain dynamics switch between trajectories that are driven by task-oriented and internally generated attractors according to pre-valuing external constraints.

By calculating the mean activity across the trajectories in Fig. 2c, and projecting these into the embedding space, we were able to recover a canonical low-dimensional manifold¹⁹ that transcends multiple cognitive task states (Fig. 2d)²⁰ and that was similar following removal of task-mediated effects before PCA²¹ (Supplementary Fig. 4). After accounting for task effects, the distribution of tPC_1 dwell times was best described by an exponential distribution, which is consistent with a noise-driven multistable process²², in which the global brain state transitions between relatively shallow attractor basins (Fig. 2b). That is, even though the average flow is smooth, the naive (that is, non-averaged) flow bears the imprint of noise-driven excursions. However, it is important to clarify that the specific nature of the flow through this embedding space invariably reflects a combination of the timing of the tasks, along with the particular cognitive processes driven by each task, and thus would not be invariant to changes in external context²³.

Finally, engagement of the low-dimensional architecture was significantly correlated with individual differences in fluid intelligence²⁴. Higher scores on the Raven's progressive matrices task (measured outside the scanner) were associated with stronger positive loadings on tPC_1 during high phases of tPC_1 ($\rho=0.25$; $P < 0.01$) and stronger negative loadings on tPC_1 during the low phases ($\rho=-0.24$; $P < 0.01$; Supplementary Fig. 5). That is, those individuals with greater fluid intelligence displayed more effective flow through the embedding space, both into and out of the state space attractor basins.

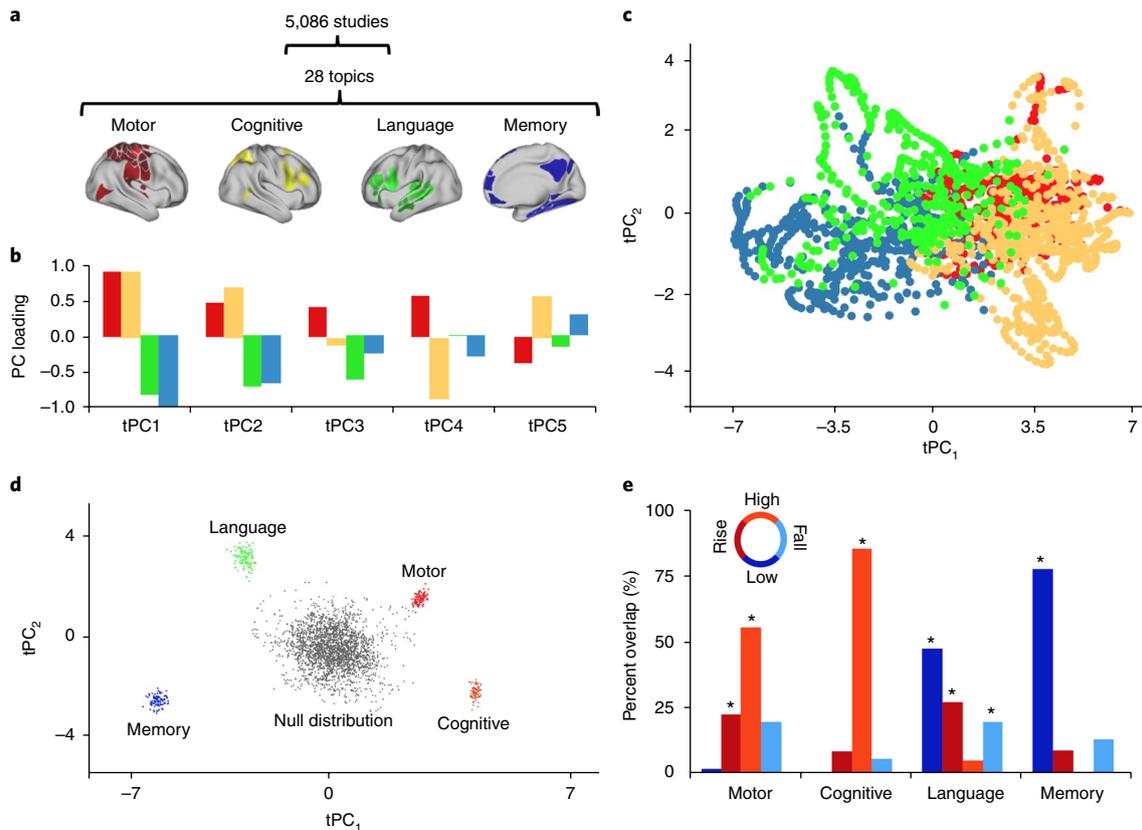


Fig. 3 | The cognitive relevance of the low-dimensional embedding space. **a**, Four NeuroSynth ‘topic families’: motor (red), cognition (yellow), language (green), and memory (blue). **b**, Bar plot demonstrating loading (single-value) of topic families onto top five principal components. **c**, Scatter plot of time points of the first two tPCs, colored according to each of the four NeuroSynth topic families. **d**, Mean value (resampled 100 times) of tPC_{1-2} for each topic family compared with a block resampled null distribution (5,000 iterations). **e**, Temporal conjunction between the topic families and the four phases of the tPC1 manifold; bar plots designate a single value (%) and asterisks denote $P < 0.01$ (block resampled null model; $n = 5,000$ iterations).

The cognitive relevance of the global brain state

Having demonstrated that brain state dynamics can be effectively described by the temporal evolution along a low-dimensional trajectory, we were next interested in understanding the potential cognitive relevance of the brain’s low-dimensional dynamics. Although the seven tasks were designed to probe specific cognitive functions, each also consists of coordinated activity along a series of different specialist dimensions. For example, the N-back task was designed to examine working memory maintenance and updating¹², but also involves visual processing, response inhibition, and motor coordination (among other processes). As such, we predicted that each of the tasks should recruit similar cognitive capacities, albeit to varying degrees, that were defined by idiosyncratic task challenges and complexity.

To test this hypothesis, we utilized meta-analytic data from an existing ‘topic-modeling’ analysis that identified the latent structure present across 5,809 functional neuroimaging studies. This approach links spatial BOLD activation patterns to the ‘topics’ investigated in the original fMRI experiments²⁵. Four ‘topic families’, representing ‘motor’, ‘cognitive’, ‘language’, and ‘memory’, were identified by clustering the activation patterns from a 50-topic solution using the NeuroSynth repository (Fig. 3a; each family represented a number of subtopics identified in the meta-analysis; results were confirmed using a ‘reverse inference’ approach; see Supplementary Table 1)²⁵. A time series was then created for each topic family by weighting the original BOLD data with the spatial activation pattern of each topic family over time. Comparison with the PC time series

revealed a clear relationship between the tPC time series and latent cognitive processes; for instance, ‘motor’ and ‘cognitive’ functions were jointly separated from ‘memory’ and ‘language’ function by tPC_1 , but were separated from one other by tPC_5 (Fig. 3b).

We then projected the topic families back into the low-dimensional embedding space by assigning the regional BOLD pattern at each time point to the topic family to which it demonstrated the strongest spatial correspondence (Fig. 3c). To test for statistical significance, we constructed a null data set (5,000 iterations) using a block-randomization resampling procedure that arbitrarily (and repeatedly) splits and reorganizes data over time, similar to the way a dealer would ‘cut’ a deck of cards. This approach scrambles the alignment of the data to the task structure, but largely preserves temporal autocorrelation, which can have important influences on the relative degrees of freedom in the data. By comparison to this null distribution, the four topic families occupied unique subspaces of the low-dimensional manifold (Fig. 3d): the ‘motor’ and ‘cognitive’ families were active during high phases, whereas the ‘memory’ and ‘language’ families were active during low phases of the manifold ($P < 0.01$; Fig. 3e). These results highlight the clear relationship between flow on the low-dimensional manifold and recruitment of specific cognitive processes.

Complex cognitive brain state dynamics

Based on earlier work^{2,26,27}, we hypothesized that flow along the tPC_1 dimension represented an integrative core that balances the competing requirements of global integration (adaptively modifying

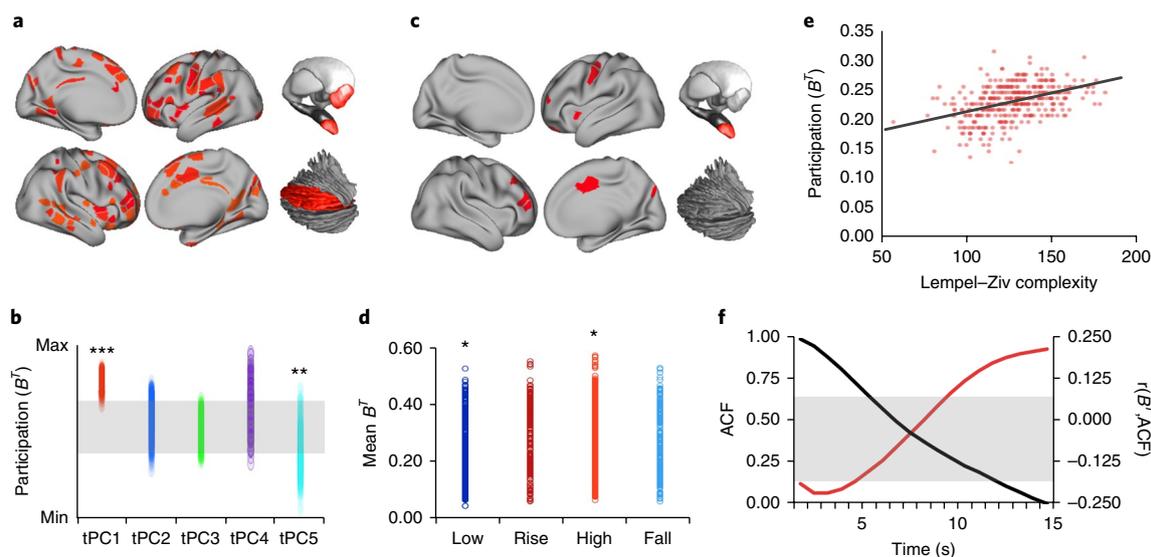


Fig. 4 | The low-dimensional integrative core of the brain across cognitive tasks. **a**, Surface plot of the most integrative regions in tPC_1 (regions significantly more integrative than null model). **b**, Parcel-wise B_T across the top five PCs, with block resampling null values displayed as gray rectangle; asterisks designate significance of mean B_T across the whole brain ($***P < 0.01$; $**P < 0.05$; two-sided t test; $n = 5,000$ iterations of permuted null); min: $B_T = 0.22$; max: $B_T = 0.26$). **c**, A surface plot of the most integrative regions in tPC_5 . **d**, B_T for each of the four phases of the tPC_5 manifold; asterisks denote $P < 0.01$ ($n = 5,000$ iterations of block resampled null model). **e**, Pearson's correlation between regional B_T and Lempel–Ziv complexity ($r = 0.424$; $P = 8.5 \times 10^{-18}$; $n = 375$ regions). **f**, Black line, mean autocorrelation function (ACF) for all regions; red line, correlation between regional ACF and B_T ; gray rectangle, values for which null hypothesis was not rejected.

the functional network signature of the brain in response to task demands) and differentiation (ensuring the distinctive configuration of neural systems required of each cognitive state)^{23,28}. To test this hypothesis, we calculated time-varying functional connectivity from the concatenated BOLD time series (after first regressing task effects from each time series) and applied graph theoretical analyses to the resultant temporal connectivity matrices (Supplementary Fig. 6). We used a general linear model to examine the relationship between the tPC_{1-5} time series and time-resolved network architecture, allowing the identification of the topological signature of each PC. In our analysis, a functional network was taken to be integrated when the connectivity matrix demonstrated low modularity (Q), which is a measure that quantifies how easily a network can be partitioned into modules, and a high median participation coefficient (B_T), which is elevated when regions connect across multiple distinct modules (segregation was associated with opposite topological signature; that is, low median B_T and high Q). Expression of tPC_1 was associated with a distributed and integrated network topology with strong connections across functionally specialized modules (Fig. 4a). In contrast, the topological signatures of lower components were more heterogeneous (Fig. 4b). Specifically, tPC_{2-4} demonstrated a trade-off between integration and segregation, whereas tPC_5 displayed a relatively segregated signature (Fig. 4c). These patterns suggest that different low-dimensional components may reflect unique constraints on the balance between integration and segregation in the brain.

The most integrative regions associated with PC_1 were diffusely distributed across the majority of canonical ‘resting state’ networks, involving regions in the frontal, parietal, and temporal cortex, along with the bilateral amygdala and the lateral cerebellum (Fig. 4a). There was a distinct relationship between time-resolved network topology and the low-dimensional manifold (Fig. 4d): the topological architecture of the brain was more integrated (that is, had higher median B_T and lower Q) during the high phases (median B_T across all parcels = 0.24 ± 0.1 ; $P < 0.01$; median $Q = 0.59 \pm 0.1$; $P < 0.01$), but segregated during the low phases (median $B_T = 0.15 \pm 0.1$; $P < 0.01$;

median $Q = 0.65 \pm 0.1$; $P < 0.01$), further supporting our hypothesis. Consistent with previous work²⁷, there was also a significant relationship between global network integration (mean B_T) and the ‘cognitive’ topic family (mean B_T across all parcels = 0.24 ± 0.1 ; $P < 0.01$). In contrast, the ‘memory’ topic family was associated with a more segregated network topology (0.15 ± 0.1 ; $P < 0.01$). The same relationship was reflected in network modularity ($P < 0.01$). Each of these topological patterns is consistent with previous studies—cognitive performance is commonly associated with an integrated network topology²⁷, whereas effective visual semantic memory has been shown to benefit from a more segregated architecture²⁹.

Although system-wide integration is an important signature of complex networks, biological systems also need to retain sufficient flexibility to cope with an array of adaptive challenges. That is, the global brain state should also demonstrate differentiation², reflecting the need for each state of the brain to be distinct from every other possible state. To test this prediction, we estimated the Lempel–Ziv complexity of each regional time series³⁰. Lempel–Ziv complexity estimates the number of distinct binary sequences required to recapitulate a test sequence: information-rich time series have higher complexity and thus require a larger dictionary of sequences to recreate them³⁰. As predicted², the mean regional signature of integration was positively correlated with Lempel–Ziv complexity ($r = 0.42$; $P < 0.01$; Fig. 4e).

The brain can also control information flow by modulating inter-regional interactions at different temporal scales. Previous work has demonstrated a heterogeneity of time scales across the brain^{31,32}, in which sensory regions process information quickly (that is, on the order of milliseconds to seconds), whereas more integrated hubs attune to information on slower time scales (that is, seconds to minutes). To determine whether the low-dimensional topological signature of tPC_1 was associated with a unique temporal signature, we correlated the extent of autocorrelation within each region with the loading between the time-resolved B_T and tPC_1 . This analysis revealed a negative correlation at shorter time scales (0.72–5.02 s) and a positive correlation at longer time scales (9.36–13.68 s;

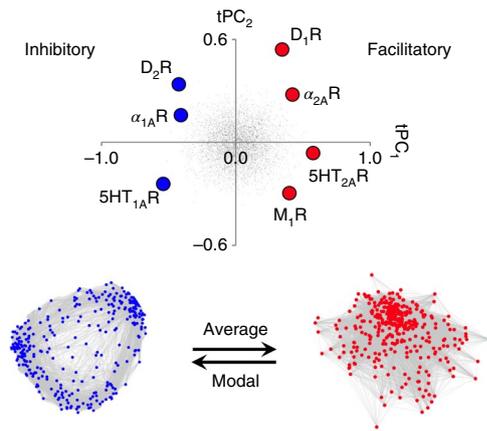


Fig. 5 | The neurochemical signature of integrated cognitive function.

The mean density of tPC_{-2} for two classes of neurotransmitter receptor maps: a group known to facilitate cognition (D_{1R} , α_{2AR} , $5HT_{2AR}$ and M_{1R} ; right) and a group known to inhibit cognition (D_{2R} , α_{1AR} , $5HT_{1AR}$; left). The spatiotemporal patterns in these two classes were associated with differential network topologies: the facilitatory group was associated with an integrated brain, whereas the inhibitory group was associated with a relatively segregated brain; force-directed plots reflect the mean time-resolved functional connectivity matrix when loading was positive for either the inhibitory (blue) or facilitatory (red) group (thresholded and binarized at 10% density for visualization purposes), and different classes of controllability: the facilitatory states were associated with high average controllability, whereas the inhibitory states were associated with high modal controllability.

$P < 0.05$; Fig. 4f), suggesting that tPC_1 processes information at relatively slow time scales.

Neurotransmitter receptor gradients gate cognitive brain dynamics

We next sought to understand the factors that control flow on the low-dimensional manifold outside of the particular sensory and cognitive constraints imposed by each task. Plausible candidates for orchestrating global control over brain state dynamics are the ascending neuromodulatory systems of the brainstem and fore-brain³³. These highly conserved nuclei project widely throughout the brain to modulate the ‘gain’ of receptive neuronal populations and, hence, alter inter-regional communication³⁴. That is, they are able to broadly modulate brain network connectivity in a context-sensitive fashion³⁵. Extensive literature links neuromodulatory systems to a broad range of cognitive functions³⁶, and receptors from several neurotransmitter families have been implicated in either facilitating or inhibiting cognitive processing³³. Interactions among these systems are also crucial, suggesting that the neuromodulatory system acts as a complex adaptive network that maintains nonlinear influence over brain network topology and dynamics³⁷.

To test this hypothesis, we used the Allen Brain Micro-Array Atlas (<http://human.brain-map.org/>) to identify the spatial coverage of a range of metabotropic neurotransmitter receptors using post-mortem data on variation in neurotransmitter receptor gene expression. We investigated two main classes of receptor with known opposing effects on cognitive function³⁷: a facilitatory group, including dopaminergic D_1 (which is encoded by the *DRD1* gene)³³, noradrenergic α_{2A} (*ADRA2A*), cholinergic M_1 (*CHRM1*), and serotonergic $5HT_{2A}$ (*HTR2A*) receptors; and an inhibitory group, including D_2 (*DRD2*), α_{1A} (*ADRA1A*), and $5HT_{1A}$ (*HTR1A*) receptors (as a result of inconsistent evidence in the literature, no muscarinic cholinergic receptors were included in the inhibitory group). Each of these receptors modulates the signal-to-noise ratio

in neurons by activating G-protein-coupled receptors³⁸, with effects typically most pronounced at the network level³⁹.

To compare brain state dynamics with neuromodulatory coverage, we related the spatial pattern of receptor density (as measured indirectly through gene transcription patterns) to each of the system-wide signatures identified in the preceding analyses (Supplementary Table 2). We observed opposing relationships between the neuromodulatory groups and the first two tPC s (see Supplementary Table 2): the facilitatory group loaded positively onto tPC_1 , whereas the inhibitory group loaded negatively (Fig. 5). In contrast, tPC_2 better delineated the receptor families, loading positively onto monoaminergic (that is, dopaminergic and noradrenergic), and negatively onto serotonergic and cholinergic receptors (Fig. 5). The neuromodulatory groups also demonstrated unique topology: the facilitatory group was associated with increased functional integration (Fig. 5; $Q=0.14$), whereas the inhibitory group was relatively segregated ($Q=0.55$). Our results thus provide a link between neuromodulatory system heterogeneity and the dynamic neural states required for diverse cognitive tasks³⁷.

The spatial overlap between low-dimensional system dynamics (estimated from task fMRI) and neurotransmitter receptor topography (estimated from post-mortem brain tissue) suggests that global brain state dynamics may be controlled by the recruitment of distinct neurotransmitter classes²⁶. To test this hypothesis, we compared the spatial maps for each receptor subtype with structural network signatures that mediate distinct control patterns in the human brain: some regions are able to drive the brain into many different states (known as ‘average’ controllability), whereas others facilitate the engagement of ‘hard to reach’ states (‘modal’ controllability)⁴⁰ (Fig. 5).

The regional signatures of these two control classes were estimated using structural diffusion data from the HCP⁴¹, and then related to the receptor maps from the Allen Brain Atlas. We observed strong positive correlations between the facilitatory group and ‘average’ controllability, and between the inhibitory group and ‘modal’ controllability (Fig. 5 and Supplementary Table 2)⁴⁰. Further analysis showed that the spatial loadings for PC_1 were selectively correlated to each region’s strength (weighted degree; $r=0.34$; $P < 0.001$) and showed predominant overlap with the structural rich-club³⁵ of the brain (41.7% of nodes from PC_1 were also within the rich-club; $P < 0.01$). In contrast, PC_3 was negatively correlated with regional strength ($r=-0.12$; $P < 0.05$), which is suggestive of a core-periphery architecture. Our results are therefore consistent with the notion that control over network dynamics is a relatively distributed capacity⁴² that may be mediated by highly conserved neuromodulatory systems that guide the flow of the brain within a low-dimensional manifold.

Discussion

The results of our multimodal analysis revealed that the neural activity required for the execution of cognitive tasks corresponds to flow within a low-dimensional state space⁴³. Across multiple, diverse cognitive tasks, the dynamics of large-scale brain activity engage an integrative core of brain regions that maximizes information-processing complexity and facilitates cognitive performance; only to then dissipate as the tasks conclude, flowing towards a more segregated architecture. Our findings further suggest that the brain’s dynamic trajectory may be shaped by a diverse set of highly conserved modulatory neurotransmitter systems that transition between distinct phases of the attractor landscape. Thus, any alteration in the activity of this diverse set of neurotransmitters (for example, as a result of prediction error signals, pharmacological manipulation, or neurodegenerative disease states) could plausibly enhance (or impair) neural signaling by raising the intrinsic potential of the regions that control global brain state dynamics (that is, the integrative core⁴⁰). In this way, the results of our analysis provide

a potential biological explanation for the control of brain state dynamics that relates to individual differences in fluid intelligence. Our results also provide a framework for studying cognitive neuroscience through the lens of dynamical systems, linking the flow of cognition to the dynamic reconfiguration of functional networks in the human brain.

Across multiple cognitive tasks with markedly different behavioral requirements, the dynamics of human brain activity were found to occupy a low-dimensional state space embedding that may form the functional backbone of cognition in the human brain. The dynamics of these brain states, and their close links with ongoing behavior, were reminiscent of the patterns observed in studies of neural dynamics in both *C. elegans*^{8,9} and *D. melanogaster*¹⁰. Given that the behaviors studied in these animal systems demonstrate far more constrained repertoires (for example, basic movements⁸ and sleep⁹ in *C. elegans*), it is perhaps surprising that comparable low-dimensional embeddings in humans are able to capture substantial amounts of variance (particularly given the breadth of behavioral tasks investigated in the HCP data set¹²). However, some considerations do argue for similarities across species. For instance, the interactions that comprise the low-dimensional functional organization of the brain are likely to be mediated via high-degree, white-matter structural connections, either through the cortex³⁵ or subcortical structures⁴⁴, that provide crucial constraints over ongoing energy dynamics⁴⁵. These structural properties are markedly consistent across species⁴⁶. Indeed, the complexity of human cognition may actually be constrained by much simpler underlying flows than we have described here⁴³. Together, these results suggest that low-dimensional analyses of brain organization may provide crucial clues into the inner workings of the brain, and that the subtle idiosyncratic functional dynamic signatures expressed by each species' connectome may hold the key to deciphering differences in behavioral repertoires across phylogeny.

Fluctuations in the low-dimensional core network architecture are associated with maximal information processing complexity across relatively long time scales, suggesting that the temporal signature of the integrative core is information rich and accumulates information over long time scales^{31,32}. This slow, integrative core—consistent with previous multi-task analyses³—contrasts with the architecture during epochs in which there was no task, in which the brain occupied a segregated topology and was associated with a shorter time scale of information processing. We also found a low-dimensional component corresponding to task onset and offset (tPC₅). Notably, these results were robust to permutation testing and resampling, suggesting that results were relatively impervious to individual differences (Supplementary Fig. 7). However, the specific form of the flow that we describe probably reflects idiosyncrasies of imaging modality—namely, the low temporal resolution of the BOLD response and specifics of the task fMRI battery, such as the predominant use of visual stimuli and reliance on motor responses. In addition, the assumption of linearity inherent in the PCA approach could also potentially be improved through the use of nonlinear dimensionality approaches, so long as they also allow for the orthogonalized coordinate system that enables tracking of the state space over time. Nonetheless, we propose that a flow between the integrated and segregated phases will persist beyond the specific tasks used here, and probably constrains cognitive capacities across a variety of psychological contexts. Future experiments could usefully examine the low-dimensional architecture of the brain across a broader range of psychological capacities, including ecologically valid contexts⁴⁷ such as naturalistic stimulation paradigms.

It has been known for some time that ascending neuromodulatory systems provide important constraints on cognitive function^{33,48}, but the systems-level mechanisms responsible for these capacities have remained relatively obscure. Here, we provide a

description of the association between these distinct neuromodulatory systems and cognitive function. Specifically, a diverse set of modulatory neurotransmitter receptors occupy privileged spatial locations in the cortex that provide opportunities for mediating the flow of cognitive brain states over time²³. Neuromodulatory receptors stimulate G-protein-coupled receptors, which alter trans-membrane ion gradients, and thus make neurons more (or less) likely to fire in response to glutamatergic inputs³⁸. This process has been interpreted as altering the signal-to-noise ratio within neural circuits³⁷—that is, neuromodulatory receptors have an information-gating role in the brain. Notably, computational work has also linked neuromodulatory activity to the alteration of the current attractor state¹⁹, which by influencing the low-dimensional flow could facilitate cognitive function. Our results provide empirical evidence for these concepts, and support the notion that neuromodulation may affect network-level effects in the brain^{39,49}. That is, neuromodulatory systems can alter the amount of integration and segregation in the brain by mediating communication among neural regions⁴⁷. These nonlinear, competitive, and cooperative dynamic interactions between neuromodulatory systems may at least partially imbue the nervous system with its notable flexibility³⁷, enabling the hard-wired 'backbone' of the brain to dynamically facilitate the neural coalitions required to navigate an evolving cognitive landscape.

Our observations yield novel predictions that can guide future work. For example, similar low-dimensional analyses of brain function in other species, notably other primates, might clarify whether or not shifts in this organizational framework underpin some of the distinctive cognitive abilities of humans. Strong phylogenetic conservatism in neuromodulatory systems suggests that evolution has modified pre-existing structures to shape human cognitive function, and the neural architecture that we have described offers a plausible substrate for such changes. Within our own species, variation in the low-dimensional core of the brain may also underlie some of the psychological manifestations of neuropsychiatric and neurodegenerative disorders. If so, detailed mapping of individual differences in the factors that shape the integrative core may suggest novel therapeutic interventions. Notably, given the multi-scale temporal organization inherent to the brain, the results of our study should be compared to similar low-dimensional derivations of functional brain architecture obtained by neuroimaging techniques with distinct spatiotemporal 'apertures', such as those based on electrophysiological signals⁵⁰. Future work could use the form of the attractor to enable model-based hemodynamic deconvolution, hence uncovering the form of the generative processes. More generally, we hope that this work will provide a platform for future insights into the modular and integrative processes that form the infrastructure for cognition in the human brain.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at <https://doi.org/10.1038/s41593-018-0312-0>.

Received: 29 August 2018; Accepted: 26 November 2018;

Published online: 21 January 2019

References

1. Breakspear, M. Dynamic models of large-scale brain activity. *Nat. Neurosci.* **20**, 340–352 (2017).
2. Tononi, G. & Edelman, G. M. Consciousness and complexity. *Science* **282**, 1846–1851 (1998).
3. Krienen, F. M., Yeo, B. T. T. & Buckner, R. L. Reconfigurable task-dependent functional coupling modes cluster around a core functional architecture. *Phil. Trans. R. Soc. Lond. B* **369**, 20130526 (2014).

4. Cole, M. W., Bassett, D. S., Power, J. D., Braver, T. S. & Petersen, S. E. Intrinsic and task-evoked network architectures of the human brain. *Neuron* **83**, 238–251 (2014).
5. Ghosh, A., Rho, Y., McIntosh, A. R., Kötter, R. & Jirsa, V. K. Noise during rest enables the exploration of the brain's dynamic repertoire. *PLoS Comput. Biol.* **4**, e1000196 (2008).
6. Tononi, G., Sporns, O. & Edelman, G. M. Measures of degeneracy and redundancy in biological networks. *Proc. Natl Acad. Sci. USA* **96**, 3257–3262 (1999).
7. Cunningham, J. P. & Yu, B. M. Dimensionality reduction for large-scale neural recordings. *Nat. Neurosci.* **17**, 1500–1509 (2014).
8. Kato, S. et al. Global brain dynamics embed the motor command sequence of *Caenorhabditis elegans*. *Cell* **163**, 656–669 (2015).
9. Nichols, A. L. A., Eichler, T., Latham, R. & Zimmer, M. A global brain state underlies *C. elegans* sleep behavior. *Science* **356**, eaam6851 (2017).
10. Lemon, W. C. et al. Whole-central nervous system functional imaging in larval *Drosophila*. *Nat. Commun.* **6**, 7924 (2015).
11. Stitt, I. et al. Dynamic reconfiguration of cortical functional connectivity across brain states. *Sci. Rep.* **7**, 8797 (2017).
12. Barch, D. M. et al. Function in the human connectome: task-fMRI and individual differences in behavior. *Neuroimage* **80**, 169–189 (2013).
13. Gordon, E. M. et al. Generation and evaluation of a cortical area parcellation from resting-state correlations. *Cereb. Cortex* **26**, 288–303 (2016).
14. Hasson, U., Nir, Y., Levy, I., Fuhrmann, G. & Malach, R. Intersubject synchronization of cortical activity during natural vision. *Science* **303**, 1634–1640 (2004).
15. Kennel, M. B. & Abarbanel, H. D. I. False neighbors and false strands: a reliable minimum embedding dimension algorithm. *Phys. Rev. E* **66**, 026209 (2002).
16. Dosenbach, N. U. F. et al. A core system for the implementation of task sets. *Neuron* **50**, 799–812 (2006).
17. Broomhead, D. S. & King, G. P. Extracting qualitative dynamics from experimental data. *Physica D* **20**, 217–236 (1986).
18. Friston, K. J., Frith, C. D., Liddle, P. F. & Frackowiak, R. S. Functional connectivity: the principal-component analysis of large (PET) data sets. *J. Cereb. Blood Flow Metab.* **13**, 5–14 (1993).
19. Woodman, M. M. & Jirsa, V. K. Emergent dynamics from spiking neuron networks through symmetry breaking of connectivity. *PLoS ONE* **8**, e64339 (2013).
20. Ott, E. & Antonsen, T. M. Low dimensional behavior of large systems of globally coupled oscillators. *Chaos* **18**, 037113 (2008).
21. Al-Aidroos, N., Said, C. P. & Turk-Browne, N. B. Top-down attention switches coupling between low-level and high-level areas of human visual cortex. *Proc. Natl Acad. Sci. USA* **109**, 14675–14680 (2012).
22. Cocchi, L., Gollo, L. L., Zalesky, A. & Breakspear, M. Criticality in the brain: a synthesis of neurobiology, models and cognition. *Prog. Neurobiol.* **158**, 132–152 (2017).
23. Pillai, A. S. & Jirsa, V. K. Symmetry breaking in space-time hierarchies shapes brain dynamics and behavior. *Neuron* **94**, 1010–1026 (2017).
24. Gray, J. R., Chabris, C. F. & Braver, T. S. Neural mechanisms of general fluid intelligence. *Nat. Neurosci.* **6**, 316–322 (2003).
25. Poldrack, R. A. et al. Discovering relations between mind, brain, and mental disorders using topic mapping. *PLoS Comput. Biol.* **8**, e1002707 (2012).
26. Shine, J. M. & Poldrack, R. A. Principles of dynamic network reconfiguration across diverse brain states. *Neuroimage* **180**, 396–405 (2018).
27. Shine, J. M. et al. The dynamics of functional brain networks: integrated network states during cognitive task performance. *Neuron* **92**, 544–554 (2016).
28. Le Van Quyen, M. Disentangling the dynamic core: a research program for a neurodynamics at the large-scale. *Biol. Res.* **36**, 67–88 (2003).
29. DeSalvo, M. N., Douw, L., Takaya, S., Liu, H. & Stufflebeam, S. M. Task-dependent reorganization of functional connectivity networks during visual semantic decision making. *Brain Behav.* **4**, 877–885 (2014).
30. Boly, M. et al. Stimulus set meaningfulness and neurophysiological differentiation: a functional magnetic resonance imaging study. *PLoS ONE* **10**, e0125337 (2015).
31. Gollo, L. L., Zalesky, A., Hutchison, R. M., van den Heuvel, M. & Breakspear, M. Dwelling quietly in the rich club: brain network determinants of slow cortical fluctuations. *Phil. Trans. R. Soc. Lond. B* **370**, 20140165 (2015).
32. Honey, C. J. et al. Slow cortical dynamics and the accumulation of information over long timescales. *Neuron* **76**, 423–434 (2012).
33. Robbins, T. W. & Arnsten, A. F. T. The neuropsychopharmacology of fronto-executive function: monoaminergic modulation. *Annu. Rev. Neurosci.* **32**, 267–287 (2009).
34. Aston-Jones, G. & Cohen, J. D. An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annu. Rev. Neurosci.* **28**, 403–450 (2005).
35. van den Heuvel, M. P. & Sporns, O. An anatomical substrate for integration among functional networks in human cortex. *J. Neurosci.* **33**, 14489–14500 (2013).
36. Puig, M. V., Gullede, A. T., Lambe, E. K. & Gonzalez-Burgos, G. Editorial: neuromodulation of executive circuits. *Front. Neural Circuits* **9**, 58 (2015).
37. Brezina, V. Beyond the wiring diagram: signalling through complex neuromodulator networks. *Phil. Trans. R. Soc. Lond. B* **365**, 2363–2374 (2010).
38. Avery, M. C. & Krichmar, J. L. Neuromodulatory systems and their interactions: a review of models, theories, and experiments. *Front. Neural Circuits* **11**, 108 (2017).
39. Shine, J. M., Aburn, M. J., Breakspear, M. & Poldrack, R. A. The modulation of neural gain facilitates a transition between functional segregation and integration in the brain. *eLife* **7**, e31130 (2018).
40. Gu, S. et al. Controllability of structural brain networks. *Nat. Commun.* **6**, 8414 (2015).
41. Yeh, F.-C. et al. Population-averaged atlas of the macroscale human structural connectome and its network topology. *Neuroimage* **178**, 57–68 (2018).
42. Tu, C. et al. Warnings and caveats in brain controllability. *Neuroimage* **176**, 83–91 (2018).
43. Huys, R., Perdikis, D. & Jirsa, V. K. Functional architectures and structured flows on manifolds: a dynamical framework for motor behavior. *Psychol. Rev.* **121**, 302–336 (2014).
44. Bell, P. T. & Shine, J. M. Subcortical contributions to large-scale network communication. *Neurosci. Biobehav. Rev.* **71**, 313–322 (2016).
45. Fulcher, B. D. & Fornito, A. A transcriptional signature of hub connectivity in the mouse connectome. *Proc. Natl Acad. Sci. USA* **113**, 1435–1440 (2016).
46. van den Heuvel, M. P., Bullmore, E. T. & Sporns, O. Comparative connectomics. *Trends Cogn. Sci.* **20**, 345–361 (2016).
47. Poldrack, R. A. & Yarkoni, T. From brain maps to cognitive ontologies: informatics and the search for mental structure. *Annu. Rev. Psychol.* **67**, 587–612 (2016).
48. Moran, R. J. et al. Free energy, precision and learning: the role of cholinergic neuromodulation. *J. Neurosci.* **33**, 8227–8236 (2013).
49. Cohen, J. D., Braver, T. S. & Brown, J. W. Computational perspectives on dopamine function in prefrontal cortex. *Curr. Opin. Neurobiol.* **12**, 223–229 (2002).
50. Vidaurre, D., Smith, S. M. & Woolrich, M. W. Brain network dynamics are hierarchically organized in time. *Proc. Natl Acad. Sci. USA* **114**, 12827–12832 (2017).

Acknowledgements

We would like to thank T. Verstynen for the diffusion data, and D. Bassett for the controllability code. The funding for the study was provided by an NHMRC CJ Martin Fellowship (GNT1072403) and a University of Sydney SOAR Fellowship (J.M.S.).

Author contributions

J.M.S. and O.S. conceived of the idea. J.M.S., M.B., O.S., and R.A.P. designed the analysis plan. J.M.S. ran the analyses and wrote the first draft of the manuscript. M.B., P.T.B., K.A.E.M., O.S., R.S., and R.A.P. provided critical methodological and conceptual input. All authors provided critical feedback on the manuscript, including editing of the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41593-018-0312-0>.

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

Correspondence and requests for materials should be addressed to J.M.S.

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

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2019

Methods

Data acquisition. Data used in the preparation of this work were obtained from the HCP database. For both the discovery and replication analyses, minimally preprocessed fMRI data were acquired from 100 unrelated participants (mean age 29.5 years, 55% female⁵¹). For each participant, BOLD data from the left-right encoding session from seven unique tasks were acquired using multiband gradient-echo echo planar imaging, amounting to 23 min 17 s of data (1,940 individual time points) per subject. The following parameters were used for data acquisition: TR = 720 ms, echo time = 33.1 ms, multiband factor = 8, flip angle = 52°, field of view = 208 × 180 mm² (matrix = 104 × 90), 2 × 2 × 2 isotropic voxels with 72 slices, alternated left-right/right-left phase encoding. No statistical methods were used to predetermine sample sizes, but our sample sizes are similar to those reported in previous publications²⁷. Ethics approval and randomization of task presentation were performed by the authors of the original studies (see ref. ⁵¹ for further details). All data analyses were conducted in accordance with the ethical guidelines mandated by the University of Sydney. The authors have complied with all local institutional and national ethical regulations. See the Life Sciences Reporting Summary for further details.

Data preprocessing. Bias field and motion correction (12 linear degrees of freedom using FSL's FLIRT) were applied to the HCP resting state data as part of the minimal preprocessing pipeline⁵¹. Temporal artifacts were identified in each data set by calculating framewise displacement from the derivatives of the six rigid-body realignment parameters estimated during standard volume realignment⁵², as well as the root mean square change in BOLD signal from volume to volume (DVARS). Abnormal frames were not excluded from the data. However, we observed no significant relationship between any of the tPC time series and framewise displacement (estimated from the temporal head motion parameters) at the individual subject level ($P > 0.5$). Following artifact detection, nuisance covariates associated with the six linear head movement parameters (and their temporal derivatives), frame-wise displacement, DVARS, and anatomical masks from the cerebrospinal fluid and deep cerebral WM were regressed from the data using the CompCor strategy⁵³. To ensure equivalence across tasks, the data were also normalized within each temporal window, which effectively controlled for the global signal, while also equilibrating the data across independent subjects. Finally, a temporal low-pass filter ($f < 0.125$ Hz) was applied to the data. Subjects were chosen according to the quality of the data (that is, low head motion and high temporal signal-to-noise ratio), and, as such, no subjects were removed after the initial inclusion. Data collection and analysis were not performed blind to the conditions of the experiments.

Brain parcellation. Following preprocessing, the mean time series was extracted from 375 predefined regions of interest. To ensure whole-brain coverage, we extracted 333 cortical parcels (161 and 162 regions from the left and right hemispheres, respectively) using the Gordon atlas¹³, 14 subcortical regions from the Harvard–Oxford subcortical atlas (bilateral thalamus, caudate, putamen, ventral striatum, globus pallidus, amygdala, and hippocampus), and 28 cerebellar regions from the SUIT atlas⁵⁴.

Cognitive tasks. Seven unique tasks were utilized from the HCP consortium: emotion, gambling, language/mathematics, motor, N-back, relational, and social¹². For each task, a block regressor was created by partitioning the time series into time points in which subjects were actively performing the task, and those in which they were 'resting' (note: not all tasks contained designated 'rest' blocks). Of note, the 'rest' blocks in the mathematical task involved an auditory, language-based task. The time points associated with each block were convolved with a canonical hemodynamic response function (using the *spm_hrf.m* from SPM12) and then concatenated over time to create a single task block regressor. These served as reference time series for comparison to the tPC time series (see gray lines in Fig. 1b). All results were successfully replicated using a finite impulse response model.

PCA. Preprocessed data from each task were concatenated to form a multi-task time series per subject and a spatial PCA was performed on the resultant data⁵⁵. Task block structure was not regressed from the data before PCA estimation. The time series of each PC was then estimated by calculating the weighted mean of the group-level BOLD time series associated with each respective PC⁹. To aid inference, group-level tPC time series were calculated by taking the mean for each PC time series across all subjects. To estimate the appropriate dimensionality of the data¹⁵, we calculated the percentage of false nearest neighbors following the PCA decomposition (a measure of effective embedding across dimensions¹⁵), and found that there were <10% present in the top 5 PCs, and <1% present in the top 10 PCs (Fig. 1b). This and all subsequent code are freely available at http://github.com/macshine/state_space/.

To ensure that the low-dimensional embedding space was not adversely affected by the task block structure, we replicated our analysis using the residuals of an ordinary least squares regression in which we modeled regional BOLD data according to the task structure present across all seven tasks (with independent blocks modeled as unique regressors). Each of the major results in our study was replicated following this procedure, suggesting that, although the low-dimensional

signature of the brain was related to the temporal structure imposed by the tasks, this factor was not solely responsible for the psychophysiological relationships observed in our study.

To ensure that the PCA results were robust to individual differences, we ran 2 subsequent analyses: (1) the same analysis was conducted in a replication cohort ($N = 100$); and (2) a resampling analysis was performed by estimating a PCA on 100 randomly chosen subjects. Each of these analyses was associated with robustly similar results to the group mean analysis ($r > 0.85$).

To determine the importance of running the PCA across all concatenated tasks, we performed three subsequent analyses: (1) We re-ran separate PCAs for each task individually and found that, although one of the first five components for each task was strongly related to task block structure ($r > 0.50$), the spatial weightings were dissimilar to the pattern observed when all tasks were concatenated (mean $r = 0.18 \pm 0.2$) and more similar to the main effect of each task (mean $r = 0.72 \pm 0.2$). A difference score was then calculated between the two groups, which allowed us to estimate statistical significance using the Dunn and Clark statistic (ZI^*). (2) We used a resampling approach in which we randomly selected between two and six tasks and re-ran the PCA, and then performed both a spatial and a temporal correlation between the topography and time series for PC₁, respectively. In doing so, we found that at least four tasks were required to recreate the pattern found across the original seven tasks (Supplementary Fig. 2). (3) We ran a standard general linear model (GLM) using the concatenated task block structure, demonstrating a selective positive correlation between the spatial map of PC₁ and the resultant beta map ($r = 0.94$; $P < 0.001$). A similar significant relationship was selectively observed between PC₂ and the beta map estimated from a GLM comparing task onsets to BOLD activity ($r = 0.34$; $P < 0.001$).

Relationship between principal components and task block structure. To estimate the relationship between the tPC time series and the task block structure, we ran a series of linear regression analyses comparing the temporal fluctuations in tPC time courses and the concatenated, convolved task block time series, both for the entire set of seven tasks and also for each task independently. In a similar fashion to the spatial PCs, tPCs were strongly replicable across the discovery and replication cohorts (mean r across tPC₁₋₅: 0.84 ± 0.1), confirming the specificity of low-dimensional temporal brain activity during cognitive task performance. Finally, none of the tPCs were significantly correlated with typical noise confounds, such as head motion (framewise displacement; mean r across tPC₁₋₅: 0.00 ± 0.01) or signals from the white matter and cerebrospinal fluid (mean r across tPC₁₋₅: 0.00 ± 0.1) at the individual subject level, nor the global signal over time ($r = -0.04$). In addition, there were no interactions between task blocks and head motion, and results were found to be replicable when performing moderate levels of 'scrubbing' (that is, censoring data with framewise displacement > 0.25 and DVARS $> 2.5\%$).

Low-dimensional manifold. To describe flow along the low-dimensional embedding space, we used an approach previously utilized in *C. elegans* calcium imaging⁸, in which the global brain cycle is partitioned into four 'phases' using the time course of the tPC₁ as a reference signal: a trough in tPC₁ defined the low phase (tPC₁(t) < 33rd percentile; blue in Fig. 2a), an increase in tPC₁ defined the rise phase (33rd percentile < tPC₁(t) < 67th percentile and dtPC₁' > 0; red), a plateau in the tPC₁ signal defined the high phase (tPC₁(t) > 67th percentile; yellow), and a decrease in tPC₁ defined the fall phase (33rd percentile < tPC₁(t) < 67th percentile and dtPC₁' < 0; green). After classifying low-dimensional activity into these four phases, we then performed a linear interpolation on each trajectory (that is, to warp each segment into a set of identically sized vectors). We were then able to estimate the trajectory of a low-dimensional 'manifold'¹⁹ by calculating the mean activity across the interpolated trajectories, which in turn could be projected into the embedding space to describe the manifold (Fig. 2c).

Using the four tPC₁ phases (low, rise, high, and fall), we estimated the 'dwell time' for each of the phases explored by the first PC by calculating the number of consecutive TRs in which each phase was present in the data. We then separately fit exponential, Weibull (stretched exponential), power law, and gamma distributions to these data (these reflect the likelihood of a multistable, critical, power law, or metastable process, respectively; see ref. ²²). The log likelihood of each fit was then used to compute the Bayesian information criterion (BIC) for each distribution—low values here represent stronger evidence for a particular fit. We found that exponential fits were better able to explain the data (that is, lower BIC = -34,637) than were the Weibull (BIC = -33,814), gamma (BIC = -33,620), or power-law (BIC = -7,012) distributions.

Topic mapping. To determine the potential cognitive relevance of the low-dimensional embedding space, we created regional estimates of 28 spatial maps that represented a curated selection of the 50 'topics' identified during a large-scale analysis of existing neuroimaging literature²⁵ (topics that were not explicitly related to psychological constructs were excluded from the 50-topic solution before the analysis). These 28 maps were further collapsed into four tight-knit 'topic families' (see Supplementary Table 1) by calculating the spatial similarity of each map and then clustering the matrix using a weighted version of the Louvain algorithm (that is, the algorithm described in the section headed Time-resolved network topology). Topic families were assigned labels according to the top ten terms

associated with the topic-word loading matrix that related study terms to brain mappings. We then created a weighted mean between each of these topic family spatial maps and the concatenated BOLD time series data. Using a ‘winner take all’ approach, we categorized each time point according to the topic family with the strongest spatial correspondence to the regional BOLD pattern present at that time, which allowed us to then project the topic families into the low-dimensional embedding space (Fig. 3c). Finally, we used a nonlinear, block resampling permutation test (see Statistical analysis below for more details), which preserves some of the autocorrelation structure, to demonstrate that the four topic families were selectively associated with unique aspects of the low-dimensional manifold (5,000 iterations; $P < 0.01$).

Time-resolved functional connectivity. To estimate functional connectivity between the 375 regions of interest, we used the multiplication of temporal derivatives (M) technique⁵⁶. M is computed by calculating the point-wise product of temporal derivative of pair-wise time series (Equation 1). The resultant score is then averaged over a temporal window, w , to reduce the contamination of high-frequency noise in the time-resolved connectivity data. A window length of 20 TRs was used in this study, though results were consistent across a range of w values (10–50 TRs). To ensure relatively smooth transitions between each task, connectivity analyses were performed on each individual task separately, and were subsequently concatenated. In addition, all analyses involving connectivity (or the resultant topological estimates) incorporated the junction between each task as a nuisance regressor.

$$M_{ijt} = \frac{1}{w} \sum_t^{t+w} \frac{(t'_i \times t'_j)}{(\sigma'_i \times \sigma'_j)} \quad (1)$$

where for each time point, t , the M for the pair-wise interaction between region i and j is defined according to equation 1, where t' is the first temporal derivative ($t+1-t$) of the i th or j th time series at time t , σ is the standard deviation of the temporal derivative time series for region i or j , and w is the window length of the simple moving average. This equation can then be calculated over the course of a time series to obtain an estimate of time-resolved connectivity between pairs of regions. Results were replicated using a jack-knife connectivity approach⁵⁷ that does not require the fitting of a window.

Time-resolved network topology. The Louvain modularity algorithm from the Brain Connectivity Toolbox⁵⁸ was used in combination with the M to estimate time-resolved community structure. The Louvain algorithm iteratively maximizes the modularity statistic, Q , for different community assignments until the maximum possible score of Q has been obtained (see Equation 2). The modularity of a given network is therefore a quantification of the extent to which the network may be subdivided into communities with stronger within-module than between-module connections.

$$Q_T = \frac{1}{v^+} \sum_{ij} (w_{ij}^+ - e_{ij}^+) \delta_{M_i M_j} - \frac{1}{v^+ + v^-} \sum_{ij} (w_{ij}^- - e_{ij}^-) \delta_{M_i M_j} \quad (2)$$

where v is the total weight of the network (sum of all negative and positive connections), w_{ij} is the weighted and signed connection between regions i and j , e_{ij} is the strength of a connection divided by the total weight of the network, and $\delta_{M_i M_j}$ is set to 1 when regions are in the same community and 0 otherwise. ‘+’ and ‘-’ super-scripts denote all positive and negative connections, respectively.

For each temporal window, we assessed the community assignment for each region 500 times and a consensus partition was identified using a fine-tuning algorithm from the Brain Connectivity Toolbox (<http://www.brain-connectivity-toolbox.net/>). This afforded an estimate of both the time-resolved modularity (Q_T) and cluster assignment (C_{iT}) within each temporal window for each participant in the study. We calculated all graph theoretical measures on un-thresholded, weighted, and signed connectivity matrices⁵⁸. The stability of the γ parameter was estimated by iteratively calculating the modularity across a range of γ values (0.5–2.0) on the time-averaged connectivity matrix for each subject—across iterations and subjects, a γ value of 1.0 was found to be the least variable, and hence was used for the resultant topological analyses. Consistent with previous studies²⁷, the average number of communities identified in each window was 2.74 ± 0.5 .

Based on time-resolved community assignments, we estimated within-module connectivity by calculating the time-resolved module-degree Z -score (W_T ; within module strength) for each region in our analysis (Equation 3)⁵⁹, where κ_{iT} is the strength of the connections of region i to other regions in its module, s_i , at time T , $\bar{\kappa}_{s_iT}$ is the average of κ over all of the regions in s_i at time T , and $\sigma_{\kappa_{s_iT}}$ is the standard deviation of κ in s_i at time T .

$$W_{iT} = \frac{\kappa_{iT} - \bar{\kappa}_{s_iT}}{\sigma_{\kappa_{s_iT}}} \quad (3)$$

The participation coefficient, B_T , quantifies the extent to which a region connects across all modules (that is, between-module strength) and has previously been used to successfully characterize hubs within brain networks (for example, see ref. 60).

The B_T for each region was calculated within each temporal window using Equation 4, where κ_{isT} is the strength of the positive connections of region i to regions in module s at time T , and κ_{iT} is the sum of strengths of all positive connections of region i at time T . Negative connections were discarded before calculation. The B_T of a region is therefore close to 1 if its connections are uniformly distributed among all the modules and 0 if all of its links are within its own module.

$$B_{iT} = 1 - \sum_{s=1}^{n_M} \left(\frac{\kappa_{isT}}{\kappa_{iT}} \right)^2 \quad (4)$$

To determine the topological signature of each tPC, we used a general linear model to fit the top five tPC time series to time-varying network topology (both mean B_T and Q). Although synchronous sensory inputs do not necessarily force regions to couple together over time³, we first regressed out all unique patterns associated with the task blocks from each of the seven tasks (with each unique block modeled as a separate regressor). In addition, the calculation of the temporal derivative also down-weights the task effects on overall activity that are presumed to drive spurious functional connectivity⁵⁶. We then fit a general linear model using the tPC time series as predictors, along with separate regressors that modeled the junction between corresponding tasks. Separate models were fit for unique regional B_T values. We subsequently ran a block resampling permutation test to determine whether there were significantly elevated values of B_T within predefined networks of the brain (5,000 iterations; $P < 0.01$). In addition, we also analyzed multiple separate window lengths (10–100 in steps of 10), and observed a similar relationship between mean B_T and tPC₁, albeit with a peak in similarity at a window length of 20 TRs.

Complex cognitive brain state dynamics. To explore the functional signature of the integrative core, we calculated the mean intra-regional connectivity for all regions connected to the integrative core (mean regional-to-core connectivity (transformed using Fisher’s r -to- Z) > 1.0) and compared this value to the mean connectivity for all regions outside of the core (that is, mean regional-to-core connectivity ≤ 1.0). Before calculating the difference score, we first applied a Fisher’s r -to- Z transform to each data point to increase Gaussianity. These values were compared using an independent samples t -test.

To estimate time series differentiation, we calculated the Lempel–Ziv complexity³⁰ of each region’s concatenated time series, binarized to values greater than or less than 0. We then ran a Pearson’s correlation comparing the Lempel–Ziv complexity scores with the B_T associated with tPC₁ (that is, the regional beta weights from a general linear model in which the tPC time series were regressed against time-resolved B_T values). The autocorrelation function was estimated for each region by calculating the time-delayed Pearson’s correlation between each region’s pre-processed BOLD time series, using a lag of 1–30 TRs (0.7–21.6 s). For each lag, a Pearson’s correlation was conducted between the integrative core and the autocorrelation function of each region. For each analysis, a block resampling permutation test was conducted to test statistical significance.

Neurotransmitter receptor mapping. To investigate the potential pharmacological correlates of progressive evolution along the manifold, we interrogated the neurotransmitter receptor signature of each region of the brain. To do so, we used the Allen Brain Atlas micro-array atlas (<http://human.brain-map.org/>) to identify the regional signature of genetic expression of metabotropic neurotransmitter receptors that were a priori related to cognitive function. We identified neurotransmitter receptor maps for receptors from four major neurotransmitter families, which were grouped into two families: a facilitatory group, comprising dopaminergic (D_1), noradrenergic (α_{2A}), cholinergic (M_1), and serotonergic (5HT_{2A}) receptors; and an inhibitory group, comprising D_2 , α_{1A} , and 5HT_{1A} receptors. We first identified the spatial topography of each receptor subtype. We then created a weighted mean between each of these neurotransmitter receptor maps and the concatenated BOLD time series data. These time series were then related to: (1) the tPC time series; (2) the topic map time series; and (3) the time-resolved topological time series. We applied a series of block resampling permutation tests to test for temporal alignments between neurotransmitter receptor maps and the manifold and topic family maps, separately (5,000 iterations; $P < 0.01$).

Structural controllability. A structural connectome was created from diffusion magnetic resonance imaging data from 842 subjects (372 males and 470 females, age 22–36) from the HCP cohort using a deterministic fiber tracking algorithm that leverages information in spin distribution functions (for details, see ref. 41). The spatial resolution was 1.25 mm isotropic; TR was 5,500 ms; TE was 89.50 ms; the b-values were 1,000, 2,000, and 3,000 mm^{-2} ; and the total number of diffusion sampling directions was 90, 90, and 90 for each of the shells, in addition to 6 b0 images. A weighted connectivity matrix was quantified using the same cortical and subcortical parcellation used in the functional analysis. The strength (that is, weighted degree) of each region was collected for further analysis, and a simple randomized null model (5,000 permutations) was run to determine whether the core regions demonstrated greater structural interconnectivity than the rest of the brain.

To estimate regional controllability, we calculated the average and modal controllability of the weighted structural connectome (see ref.⁴⁰ for details). Briefly, average controllability is defined as the trace(W_k^{-1}), where $W_k = \sum_{\tau=0}^{\infty} A^{\tau} B_k B_k^T A^{\tau}$ is the controllability Gramian, A is the weighted connectivity matrix, and B is the input matrix that defines the control points in the network; and modal controllability is computed as the eigenvector matrix $V = [v_{ij}]$ of the network adjacency matrix A (if the entry v_{ij} is small, then the j th mode is poorly controllable from node i). While it is known that these measures relate to degree/strength, there is also evidence that they remain after controlling for degree, and hence may relate to other topological features of the structural connectome⁴². The regional patterns created from these analyses were then used to create a weighted mean between each of these control maps and the concatenated BOLD time series data. These time series were then related to the other outcome measures in our study, and we used a block resampling null model to determine statistical significance. We also correlated the spatial loading of the first five PCs with the strength (that is, weighted degree) of each region within the structural connectome.

Statistical analysis. For a number of the main analyses, we ran separate block resampling permutation tests to determine statistical significance. This approach scrambles the alignment of the data to the task structure but (relatively) preserves autocorrelation by randomly ‘cutting’ the data in two sections, and then re-analyzing the data. For each comparison, we constructed unique null data sets (5,000 iterations) and then determined statistical significance by calculating the proportion of the 5,000 iterations that our original test statistic was greater than (or lesser than) the 97.5th percentile (or 2.5th percentile) of the null distribution. Importantly, this approach is non-parametric, and thus the distribution of the data was not of concern.

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

Code availability

All code used to analyze the data in this study is available from http://github.com/macshane/state_space/.

Data availability

Data were provided by the Human Connectome Project (HCP); the Washington University, University of Minnesota, and Oxford University Consortium (Principal Investigators David Van Essen and Kamil Ugurbil; grant no. 1U54MH091657) funded by 16 NIH institutes and centers that support the NIH Blueprint for Neuroscience Research; and the McDonnell Center for Systems Neuroscience at Washington University. This project also made use of Connectome DB and Connectome Workbench, developed under the auspices of the HCP (HCP 1200 Subject Release, <http://www.humanconnectome.org/>). Neurotransmitter receptor data from the Allen Human Brain Atlas (2010 Allen Institute for Brain Science; available from: human.brain-map.org) were obtained from neurosynth.org.

References

1. Glasser, M. F. et al. The minimal preprocessing pipelines for the human connectome project. *Neuroimage* **80**, 105–124 (2013).
2. Power, J. D. et al. Methods to detect, characterize, and remove motion artifact in resting state fMRI. *Neuroimage* **84**, 320–341 (2014).
3. Behzadi, Y., Restom, K., Liao, J. & Liu, T. T. A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *Neuroimage* **37**, 90–101 (2007).
4. Diedrichsen, J., Balsters, J. H., Flavell, J., Cussans, E. & Ramnani, N. A probabilistic MR atlas of the human cerebellum. *Neuroimage* **46**, 39–46 (2009).
5. McIntosh, A. R. & Lobaugh, N. J. Partial least squares analysis of neuroimaging data: applications and advances. *Neuroimage* **23**, S250–S263 (2004).
6. Shine, J. M. et al. Estimation of dynamic functional connectivity using multiplication of temporal derivatives. *Neuroimage* **122**, 399–407 (2015).
7. Thompson, W. H., Richter, C. G., Plavén-Sigra, P. & Fransson, P. Simulations to benchmark time-varying connectivity methods for fMRI. *PLoS Comput. Biol.* **14**, e1006196 (2018).
8. Rubinov, M. & Sporns, O. Complex network measures of brain connectivity: uses and interpretations. *Neuroimage* **52**, 1059–1069 (2010).
9. Guimerà, R. & Nunes Amaral, L. A. Functional cartography of complex metabolic networks. *Nature* **433**, 895–900 (2005).
10. Power, J. D., Schlaggar, B. L., Lessov-Schlaggar, C. N. & Petersen, S. E. Evidence for hubs in human functional brain networks. *Neuron* **79**, 798–813 (2013).

Reporting Summary

Nature Research 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 Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. 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
- An indication of 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 statistics 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
- Clearly defined error bars
State explicitly what error bars represent (e.g. SD, SE, CI)

Our web collection on [statistics for biologists](#) may be useful.

Software and code

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

Data collection The data were collected by the Human Connectome Project and the Allen Brain Project, both of which are openly available datasets.

Data analysis The code used to create the analysis are freely available at http://github.com/macshine/state_space/. In addition, code used for the graph theoretical analysis is freely available from the Brain Connectivity Toolbox (<https://sites.google.com/site/bctnet/>).

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/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [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 list of figures that have associated raw data
- A description of any restrictions on data availability

Data were provided by the Human Connectome Project (HCP), and the Washington University, University of Minnesota, and Oxford University Consortium (Principal Investigators David Van Essen and Kamil Ugurbil; Grant 1U54MH091657) funded by 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience

Field-specific reporting

Please select 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/authors/policies/ReportingSummary-flat.pdf

Behavioural & social sciences study design

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

Study description	The study uses quantitative methods.
Research sample	200 unrelated individuals from the Human Connectome project were used for the main study sample. A further 642 subjects were used to create the structural connectome. Data from the Allen Brain Atlas was collected from 6 subjects post-mortem.
Sampling strategy	Two separate cohorts were identified by members of the HCP consortium. Effect sizes were not known prior to our analysis, however it was determined that two samples of 100 subjects would retain adequate power to detect small-to-moderate effects (neuropowertools.org).
Data collection	Data were collected by members of the Human Connectome Project and Allen Brain Atlas.
Timing	Please see Barch et al., 2014 (NeuroImage) for details.
Data exclusions	No data were excluded from the analysis
Non-participation	Please see Barch et al., 2014 (NeuroImage) for details.
Randomization	Please see Barch et al., 2014 (NeuroImage) for details.

Reporting for specific materials, systems and methods

Materials & experimental systems

- n/a Involved in the study
- Unique biological materials
 - Antibodies
 - Eukaryotic cell lines
 - Palaeontology
 - Animals and other organisms
 - Human research participants

Methods

- n/a Involved in the study
- ChIP-seq
 - Flow cytometry
 - MRI-based neuroimaging

Human research participants

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

Population characteristics	Please see Barch et al., 2014 (NeuroImage) for details.
Recruitment	Please see Barch et al., 2014 (NeuroImage) for details.

Magnetic resonance imaging

Experimental design

Design type	task block/event design
-------------	-------------------------

Design specifications

Behavioral performance measures

Acquisition

Imaging type(s)

Field strength

Sequence & imaging parameters

Area of acquisition

Diffusion MRI Used Not used

Parameters

Preprocessing

Preprocessing software

Normalization

Normalization template

Noise and artifact removal

Volume censoring

Statistical modeling & inference

Model type and settings

Effect(s) tested

Specify type of analysis: Whole brain ROI-based Both

Anatomical location(s)

Statistic type for inference (See [Eklund et al. 2016](#))

Correction

Models & analysis

n/a | Involved in the study

Functional and/or effective connectivity

Graph analysis

Multivariate modeling or predictive analysis

Functional and/or effective connectivity

order to reduce the contamination of high-frequency noise in the time-resolved connectivity data. A window length of 20 TRs was used in this study, though results were consistent across a range of w values (10-50 TRs). To ensure relatively smooth transitions between each task, connectivity analyses were performed on each individual task separately, and were subsequently concatenated. In addition, all analyses involving connectivity (or the resultant topological estimates) incorporated the junction between each task as a nuisance regressor. Results were replicated using a jack-knife connectivity approach that does not require the fitting of a window.

The Louvain modularity algorithm from the Brain Connectivity Toolbox (BCT60) was used in combination with the MTD to estimate time-resolved community structure. The Louvain algorithm iteratively maximizes the modularity statistic, Q , for different community assignments until the maximum possible score of Q has been obtained (see Equation 2). The modularity estimate for a given network is therefore a quantification of the extent to which the network may be subdivided into communities with stronger within-module than between-module connections.

Graph analysis