The low dimensional dynamic and integrative core of cognition in the human brain

J.M. Shine[1*], M. Breakspear[2,3], P.T. Bell[4], K. Ehgoetz Martens[1], R. Shine[1], O. Koyejo[5], O. Sporns[6] and R.A. Poldrack[7]

Affiliations

[1] The University of Sydney, Sydney, NSW, Australia

[2] QIMR Berghofer, Brisbane, QLD, Australia

[3]Metro North Mental Health Service, QLD, Australia

[4]University of Queensland, QLD, Australia

[5]University of Illinois, Champagne, IL, USA

[6]Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN, USA

[7]Department of Psychology, Stanford University, Stanford, CA, USA

* Correspondence to: mac.shine@sydney.edu.au

Summary

The human brain seamlessly integrates innumerable cognitive functions into a coherent whole, shifting with fluidity between changing task demands. To test the hypothesis that the brain contains a core dynamic network that integrates specialized regions across a range of unique task demands, we investigated whether brain activity across multiple cognitive tasks could be embedded within a relatively low dimensional, dynamic manifold. Analysis of task-related fMRI data from the Human Connectome Project revealed a core brain system that fluctuates in accordance with cognitive demands, bringing new brain systems on-line in accordance with changing task demands, while maximizing temporal information processing complexity. Regional differences in noradrenergic neurotransmitter receptor density align with this integrative core, providing a biologically plausible mechanism for the control of global brain dynamics. Our results advance a unique window into functional brain organization that emphasizes the confluence between low dimensional neural activity, network topology, neuromodulator systems and cognitive function.

Main Text

The human brain is a complex biological system that is capable of remarkably diverse behavior, arising from the coordinated activity of billions of neurons across a range of spatial and temporal scales. From this symphony of distributed activity emerges a dynamic, large-scale spatiotemporal structure that facilitates a vast cognitive repertoire. Linking activity within this large-scale neural architecture to the emergence of diverse cognitive states remains an important goal for systems neuroscience. Nonetheless, we do not yet know whether cognition can be decomposed into largely independent, modular components or instead derives from a flexible, interconnected architecture that is constrained by a structural backbone.

Based on substantial redundancy in complex systems such as the brain [1], recent work has utilized dimensionality reduction techniques[1,2] to clarify functional brain circuitry underlying the behavioral repertoire of organisms such as the nematode (*Caenorhabditis elegans*), the fruit fly (*Drosophila melanogaster*) and the ferret (*Mustela putorius furo*)[3-6]. By consolidating components according to their spatiotemporal similarity, these approaches aim to unearth latent functional patterns in complex datasets[2]. Dimensionality reduction techniques have also been used to describe the functional organization of the human brain through space and time. For example, a substantial body of work shows that the brain's structural and functional architecture strikes a balance between the opposing forces of segregation and integration[7], a signature common to many complex systems[8]. In addition, computational modeling studies suggest that low dimensional dynamics constrain whole-brain activity[9]. Consistent with these dynamic attributes, network models of brain function hypothesize a 'dynamic core' of regions that flexibly guide the flow of activity within the brain to facilitate cognition[10,11]. These models predict that the functional signature of the brain should contain a distributed set of core regions that integrate more specialized systems[12], fluidly altering their interactions with the rest of the brain to process information for exogenous task demands[10,13]. While conceptually appealing, this prediction has not yet been definitively tested with empirical data.

Here, we use whole-brain functional neuroimaging data across a suite of cognitive tasks to test this prediction. We show that the dynamic functional organization of the brain describes a flow along a low-dimensional phase space that aligns with unique cognitive

functions. We then demonstrate that temporal progression along this low-dimensional axis reflects an 'integrative core' that arranges distributed, specialized regions into flexible coalitions responsible for distinct cognitive tasks. Finally, we establish that this low-dimensional axis is closely related to spatial patterns of α_{1A} and α_{2A} noradrenergic receptor expression, providing a plausible biological mechanism for the modulation of system-level brain dynamics. Together, our results present a novel view of brain function based on the coordinated interaction of functional brain networks over time.

Low dimensional global brain activity recurs across multiple cognitive tasks

We used high temporal resolution 3T fMRI data (TR = 0.72s) from the Human Connectome Project consortium to examine BOLD activity from a large cohort of unrelated individuals across seven unique cognitive tasks, each of which engages putatively distinct cognitive functions: emotional processing, gambling, mathematical calculation, motor execution, working memory performance, relational matching and social inference[14]. Preprocessed BOLD time series were extracted from cortical and subcortical regions[46] and concatenated across all seven tasks and across all subjects. To ensure reproducibility, analyses were initially developed within a 100-subject *Discovery* dataset and then replicated within a 100-subject *Replication* cohort.

We performed principal component analysis (PCA) on the BOLD time series from each region, concatenated across all tasks and subjects. Briefly, PCA estimates spatially orthogonal patterns in the data that maximally explain the variance across the entire dataset, and produces a set of spatial weights that represent the regions that are most relevant to the particular dimension (Figure 1A). Despite analyzing data from a collection of several highly divergent cognitive tasks, we found a dominant, low-dimensional neural signal[2]: the first five principal components (PCs) accounted for 67.9% of the variance and resolved greater than 90% of the embedding space[15] (Figure 1B). The *Discovery* and *Replication* cohorts showed strong correspondence across the first 5 PCs (mean r across $PC_{1\text{-}5}$: 0.85 ± 0.12).

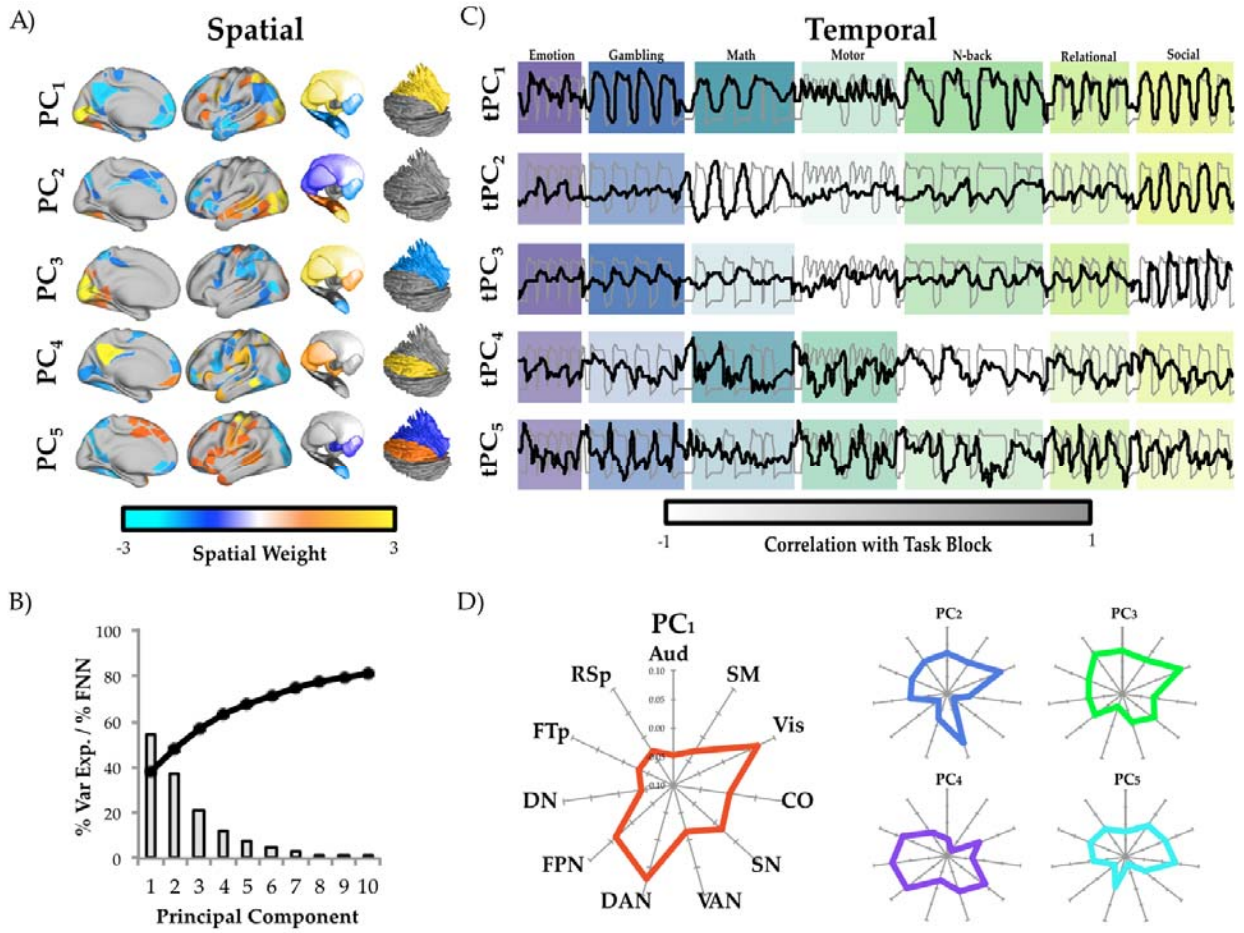

Figure 1 – Results of the spatiotemporal principal component analysis across multiple cognitive tasks. A) spatial maps for the first five principal components (colored according to spatial weight); B) line plot representing the percentage of variance explained by first 10 principal components; bar plot depicting the percentage of false nearest neighbors for first 10 principal components; C) correspondence between convolved, concatenated task block regressor (grey) and the time course of the first five tPCs (black) – color intensity of the blocks reflect the correlation between $tPC_{1\text{-}5}$ and each of the unique task blocks; D) mean spatial loading of first 5 PCs, organized according to pre-defined networks. Key: DAN – dorsal attention; VIS – visual; FPN – frontoparietal; SN – salience; CO – cingulo-opercular; CPar – cingulo-parietal; VAN – ventral attention; SM – somatomotor; RSp – Retrosplenial; FTP – frontotemporal; DN – default mode; AUD – auditory.

Next, we created a time series of each PC (tPC) by calculating the mean activity of the group-level BOLD time series weighted by the regional coefficients associated with each spatial component. The tPCs can each be thought of as linear subspaces along which the system traverses, and represent signals shared by regions with similar time courses. 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, $p < 0.01$; Figure 1C) and was strongly replicable across cohorts (r = 0.84). The magnitude of the tPC_1 time series also varied across tasks, suggesting that tPC_1 may represent the extent of brain engagement in each task as a function of task demands. Consistent with this notion, PC_1 showed strong spatial loading in frontal, parietal and occipital cortices, along with striatum, thalamus and lateral cerebellum (Figure 1D); many of these regions have been implicated in processing general task demands[16]. Importantly, fitting a PCA to each task separately did not recover the same underlying principal component, but instead recovered spatial maps that aligned with the idiosyncratic demands of each task: the mean correlation between task-specific BOLD effects and task-specific PC maps was 0.72 ± 0.2, but was 0.18 ± 0.2 when compared to PC_1 ($ZI^* > 2.0$; $p < 0.01$ for all task comparisons). Finally, we used a bootstrapping approach to demonstrate that at least four concatenated tasks were required to rediscover the same recurrent underlying component (Figure S1).

The observed relationship between tPC_1 and multiple tasks is inconsistent with a purely modular view of brain function. A modular view predicts that the most highly explanatory tPCs are associated with sets of (or individual) tasks, grouped according to conceptual similarity, rather than a single component that is expressed across all tasks (e.g., tPC_1). The next three components ($tPC_{2\text{-}4}$) more closely reflected this segmental organization (Figure 1C), suggesting a relationship with specific cognitive constructs. For example, tPC_2 was associated with the social task and language processing (during the 'rest' blocks of the mathematics task), tPC_3 with gambling and emotion, and tPC_4 with the math task (Figure 1C). Thus, the low-dimensional organization of brain activity contained signatures of both integration (tPC_1) and segregation ($tPC_{2\text{-}4}$).

Interestingly, in a similar fashion to tPC_1, the next component (tPC_5; 4.8% of variance explained) was associated with a consistent cross-task effect ($r = 0.20$, $p < 0.01$). However, the time course of tPC_5 showed sharp increases at the beginning and end of task blocks ($r = 0.11$, $p < 0.01$), a pattern that was not observed in tPC_1 ($r = -0.05$, $p > 0.05$). This suggests that the time course of tPC_5 was uniquely associated with the transition into and out of unique task states. Importantly, tPC_5 was associated with activity across a right-lateralized system of cingulate, parietal and opercular cortical regions that play a crucial role in cognitive task engagement and error monitoring[17].

Global brain states exist along a low dimensional manifold

To describe the temporal evolution of the global brain state, we next reconstructed the

state-space trajectory of the dominant low-dimensional signal (see Movie 1). Importantly, other popular dimensionality reduction techniques that do not require orthogonality (such as independent component analysis) naturally mix different amounts of each component, and as such are not suitable for the identification and analysis of state space attractors[3]. The tPC_1 time course was then partitioned into relative phase segments to demarcate the embedding space[3]. A trough in tPC_1 defined the LOW state (blue in Figure 2A), an increase in tPC_1 defined the RISE state (red), a plateau in the tPC_1 signal defined the HIGH state (orange), and a decrease in tPC_1 defined the FALL state (light blue). The resulting phase portrait describes the temporal evolution of the low-dimensional signal shared across all behavioral tasks (Figure 2B and Movie 1; see Figure S2 for projection of $tPC_{4\text{-}5}$). The different phases of tPC_1 were distinctly related to exogenous task demands, with a greater frequency of HIGH states during task blocks across all seven tasks, and LOW states during interleaved rest blocks ($p < 0.01$). By calculating the mean activity across the trajectories in 2B, and projecting these into the embedding space, we were able to recover a canonical low dimensional manifold that transcends multiple cognitive task states (Figure 2C)[18].

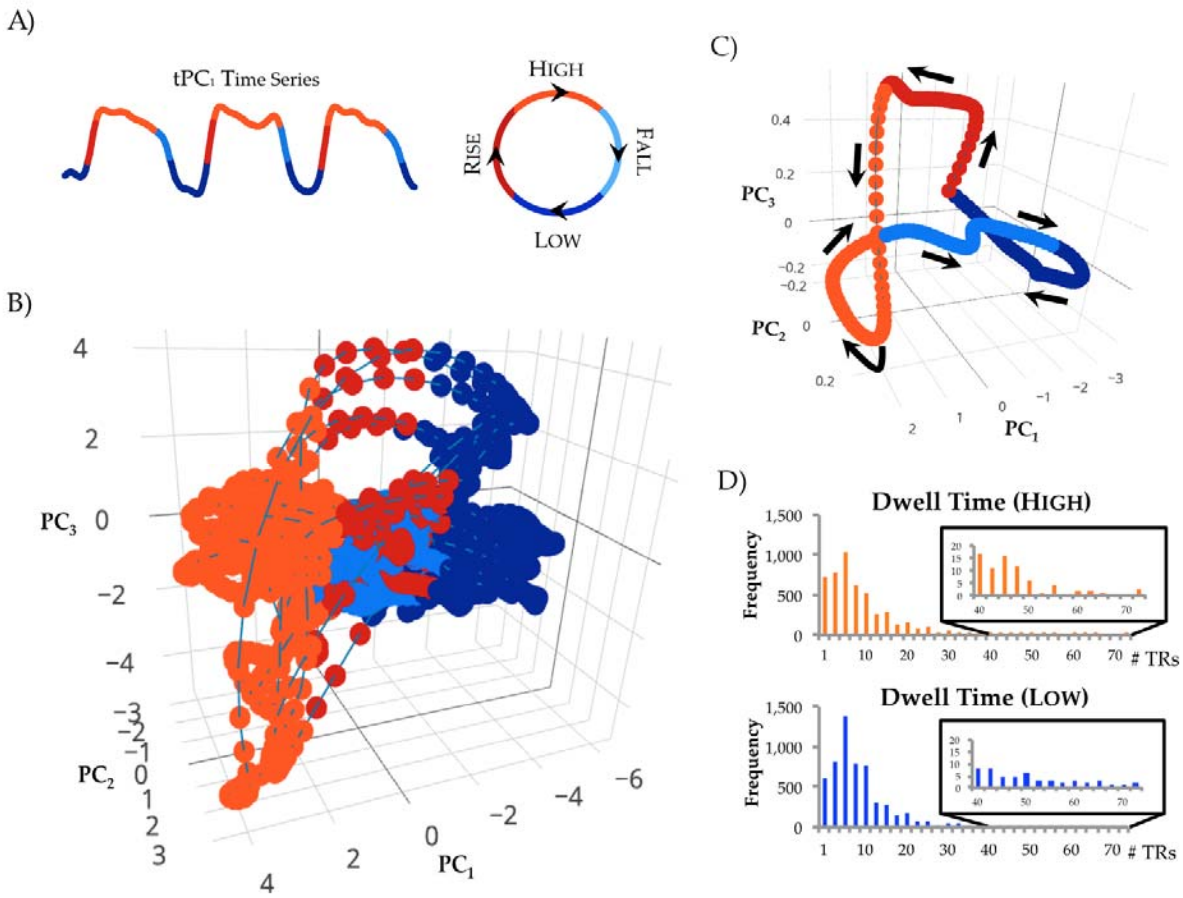

Figure 2 – A low-dimensional embedding space characterizes the temporal organization of the brain across cognitive tasks. A) The procedure used to partition tPC1 into unique phases – Low (blue), Rise (Red), High (Orange) and Fall (Light Blue); B) a 3-dimensional scatter plot comparing the first three tPCs, colored according to the character of tPC_1; C) the low dimensional manifold traversed by the global brain state across the first three dimensions – arrows depict the direction of 'flow' along the manifold; D) the dwell times for the 'high' and 'low' sections of tPC_1, which were best fit by exponential distributions – inserts demonstrate the tails of each distribution.

Next we investigated the characteristics of the dynamic process underlying progression along the manifold[19]. To disambiguate these processes, we investigated the dwell times of each state. The distribution of unique dwell time signatures can reveal the types of models that plausibly generated the system's dynamics[20]. To this end, we used the tPC_1 phase partition to determine the dwell time in each state (Figure 2D), and found that the distribution of low dimensional state dynamics was best described by an exponential distribution (lower Bayesian Information Criterion than for either gamma, Weibull or

power-law distributions). These results are consistent with a noise-driven multistable process[20], in which the global brain state transitions between relatively shallow attractors, presumably due to weak entrainment imposed by the tasks' temporal structure.

The cognitive relevance of the global brain state

To examine the cognitive relevance of the embedding space, we next utilized a topic-mapping approach to define spatial brain maps that represent the underlying co-activation structure of mental processing and associated brain functions[21]. Briefly, this method identified the latent structure present across 5,809 functional neuroimaging studies, linking spatial activity maps to the 'topics' investigated in the original fMRI experiments. Using a clustering approach on data from the NeuroSynth repository, we identified four cohesive 'topic families', representing 'Motor', 'Cognitive', 'Language' and 'Memory' capacities (Figure 3A). We used a weighted average between the topic maps and concatenated BOLD time series data to calculate a similarity index between topic family and the flow of principal components over time. A clear relationship between the tPC time series and latent cognitive processes was revealed – for instance, 'Motor' and 'Cognitive' functions were jointly separated from 'Memory' and 'Language' function by tPC_1, but were separated from each other by tPC_5 (Figure 3B).

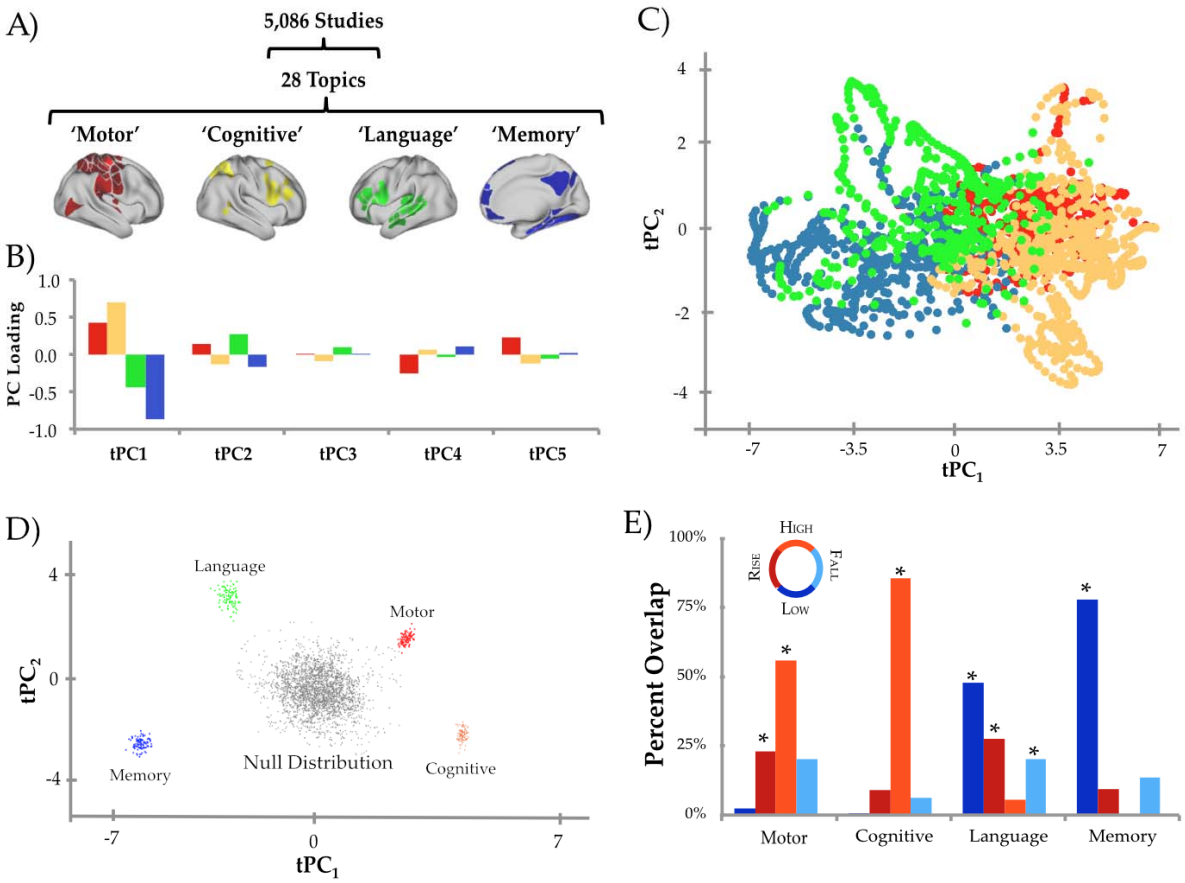

Figure 3 – The cognitive relevance of the low-dimensional embedding space. A) Four NeuroSynth "topic families": Motor (red), Cognition (yellow), Language (green) and Memory (blue); B) bar plot demonstrating loading of Topic Families onto top five principal components; C) scatter plot of time points of the first two tPCs, colored according to their loading onto each of the four NeuroSynth topic families; D) mean value (bootstrapped 100 times) of $tPC_{1\text{-}2}$ for each topic family, compared to a block bootstrapped null distribution (5,000 iterations); E) temporal conjunction between the topic families and the four phases of the tPC1 manifold – asterisks denote $p < 0.01$ (block bootstrapped null model).

To project the topic families into the low-dimensional embedding space, the regional BOLD pattern at each time point was classified according to the topic family to which it demonstrated greatest strongest spatial correspondence using a 'winner-take-all' approach (Figure 3C). To test for significance, we used a block-randomization bootstrapping procedure that arbitrarily shifts data over time, scrambling the alignment of the data to the task structure but preserving the autocorrelation (smoothness) to construct an ensemble of a null dataset. By comparison to this null distribution, the four topic families were partitioned into unique subspaces of the low dimensional manifold in a highly specific manner (Figure 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 ($p < 0.01$; Figure 3E). Thus, our results demonstrate a clear relationship between cognitive state and flow on a low-dimensional manifold.

The integrative core

On the basis of the above results, we hypothesized that flow along the tPC_1 dimension represented an integrative core that balances the competing requirements of integration (adaptively modifying the functional network signature of the brain in response to task demands) and differentiation (ensuring the distinctive nature of each cognitive state)[10,22]. To test this hypothesis, we calculated time-varying functional connectivity from the concatenated BOLD time series and applied graph theoretical analyses to the resultant temporal connectivity matrices. After first removing task effects in each time series, we used a general linear model to examine the relationship between the $tPC_{1\text{-}5}$ time series and time-resolved network architecture to identify the topological signature of each principal component. Expression of tPC_1 was associated with a distributed and integrated network topology with strong connections across specialist modules (Figure 4A). In contrast, expression of other components showed distinct topological signatures (Figure 4B). Specifically, $tPC_{2\text{-}4}$ demonstrated a trade-off between integration and segregation, whereas tPC_5 displayed a relatively segregated signature (Figure 4C), suggesting that different low-dimensional components may reflect constraints on the balance between integration and segregation in the brain.

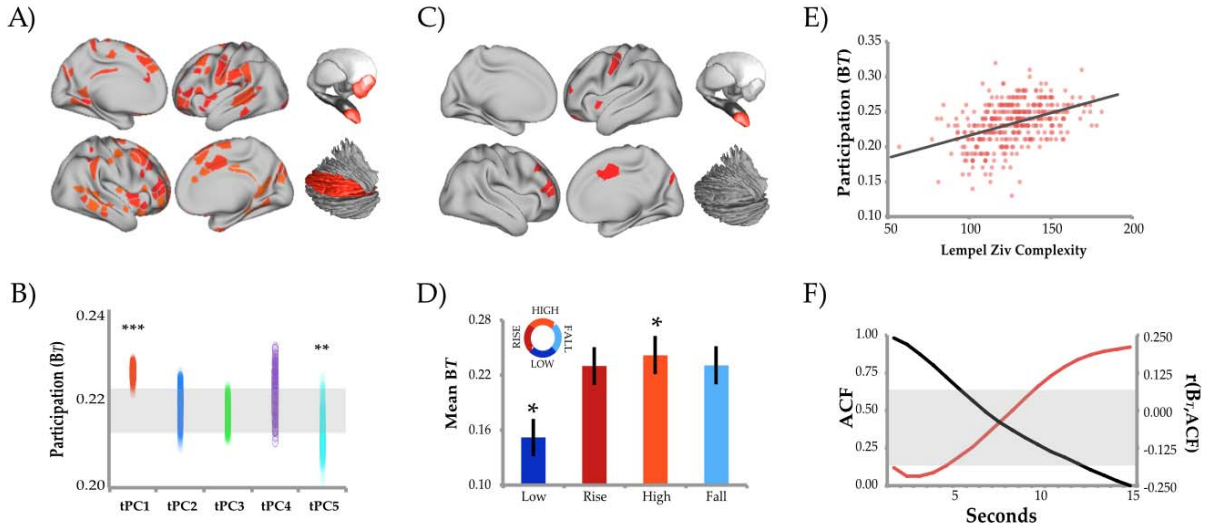

Figure 4 – The low-dimensional integrative core of the brain across cognitive tasks. A) surface plot of the most integrative regions in tPC_1; B) parcel-wise B_T across top 5 PCs, with block resampling null values

displayed as grey region; stars designate significance of mean B_T across the whole brain (*** - $p < 0.01$; ** - $p < 0.05$); C) a surface plot of the most integrative regions in tPC_5; D) mean B_T for each of the four phases of the tPC_1 manifold – asterisks denote $p < 0.01$ (block bootstrapped null model); E) correlation between regional B_T and Lempel Ziv Complexity ($r = 0.424$; $p < 0.001$); F) black line: mean autocorrelation function (ACF) for all regions; red line: correlation between regional ACF and BT; grey rectangle – values for which null hypothesis was not rejected.

The most integrative regions associated with tPC_1 were diffusely distributed across the majority of canonical 'resting state' networks, involving regions within the frontal, parietal and temporal cortex, along with bilateral amygdala and the lateral cerebellum (Figure 4A), similar to the pattern observed in the structural backbone of the brain[23]. The tPC_1 time series also exhibited a distinct topological relationship with the manifold (Figure 4D), as the connective architecture of the brain was more integrated during the High (median B_T across all parcels = 0.24 ± 0.1) phases ($p < 0.01$), but segregated during the Low phases (0.15 ± 0.1) of the manifold when compared to the block resampling null model. 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; $p < 0.01$), whereas the 'Memory' topic family was associated with relative segregation (0.15 ± 0.1; $p < 0.01$).

To further characterize the putative functions of the integrative core, we next estimated the mean time-resolved functional connectivity between the core regions (regions with top 10% B_T) and all other regions of the brain. This analysis provided evidence for 'functional clustering'[10], in which regions intermittently connected to the integrative core were more tightly interconnected with one another than were regions that were disconnected from the core (mean r-to-Z score core+: 2.04 ± 1.4; core-: 1.89 ± 1.2; $p < 0.01$). An inverse correlation between assortativity and tPC_1 ($r = -0.20$; $p < 0.01$) suggested that this result could not be explained by the transitive structure of the connectivity matrix over time[24]. Finally, we were able to replicate and extend previous work[16,25] by showing that the integrative core was connected to unique regions as a function of each task (connectivity signatures were robustly different across each task block; all $r < 0.05$). This result supports the hypothesis of a integrative core that binds specialist regions of the brain as a function of task demands, simultaneously augmenting information integration and differentiation[10].

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 differentiation[10], 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 series[13]. Briefly, this measure estimates the number of distinct binary sequences required to recapitulate a test sequence – information-rich time series contain higher complexity and thus require a larger dictionary of sequences in order to recreate them[13]. As predicted[10], the integrative regional signature of tPC_1 was positively correlated with elevated Lempel Ziv complexity ($r = 0.42$; $p < 0.01$; Figure 4E).

Finally, previous work has demonstrated a heterogeneity of time-scales across the brain[26,27], in which sensory regions process information relatively quickly (i.e., on the order of milliseconds-to-seconds), whereas more integrated hubs attune to information on much slower time-scales (i.e., on the order of seconds-to-minutes). The integrative signature of tPC_1 demonstrates a relationship with regional autocorrelation consistent with slow time-scales, with an inverse relationship for short (0.72 – 5.02 seconds) and a positive relationship for long time scales (9.36 – 13.68 seconds; $p < 0.05$; Figure 4F).

The neurochemical signature of integrative cognitive function

A candidate mechanism underlying alterations in higher-level network integration is activity within ascending neuromodulatory nuclei of the brainstem[28], which project widely throughout the brain to modulate the 'gain' of receptive neurons and hence, modify inter-regional communication[29]. In particular, the diversely projecting noradrenergic system[30] has been strongly linked to alterations in functional network topology in rodents[31], non-human primates[32] and humans[33]. These studies suggest that manipulating neural gain via ascending noradrenergic projections may shift a functional brain state towards higher levels of integration[34], perhaps through the recruitment of highly-connected structural hubs[23].

To investigate the link between neurochemistry and the network dynamics of the integrative core, the Allen Brain Micro-Array Atlas (http://human.brain-map.org/) was used to identify regional variation in the genetic expression of two classes of metabotropic adrenergic receptors that have a well-demonstrated relationship with cognition: activation of α_{2A} receptors promote cognitive behaviors such as working

memory[28,30,35], whereas the opposite effects are observed for stimulation of α_{1A} receptors[30,36]. Using a similar approach to that described above, we first identified the spatial topography of each receptor sub-type (Figure 5A), and then used these maps to estimate the similarity of regional BOLD activity to each spatial receptor map over time. This analysis revealed a positive relationship between the time series of tPC_1 and brain regions with a high density of α_{2A} adrenergic receptors (Figure 5B). In contrast, there was a negative relationship with tPC_1 for regions with a high density of α_{1A} receptors, suggesting that the recruitment of distinct neurotransmitter classes may influence the trajectory of the global brain state[37]. In addition, engagement of regions expressing α_{2A} was selectively elevated during the High phases of the manifold, with 'Cognitive' and 'Motor' topic families, and also with a more integrated network topology (all $p < 0.01$), whereas in contrast, regions expressing α_{1A} were more prominent in the Low phases of the manifold; with 'Memory' and 'Language' topic families; and with a segregated network architecture (all $p < 0.01$).

These results link noradrenergic system heterogeneity to the distinct neural states required for diverse cognitive tasks, however the tenets of biological degeneracy suggest that the spatial gradients of receptors from other neuromodulatory receptor systems should demonstrate similar effects. For example, neurotransmitter systems (such as dopamine, acetylcholine and serotonin) also modulate cognitive function[28]. As a final analysis, we provide evidence for this heterogeneity by examining a suite of cognitively relevant neurotransmitter receptors across neuromodulatory classes. The different neurotransmitter systems demonstrated consistent, yet distinct, associations with the spatiotemporal functional trajectory of the brain as a function of cognitive processing (Figure 5C).

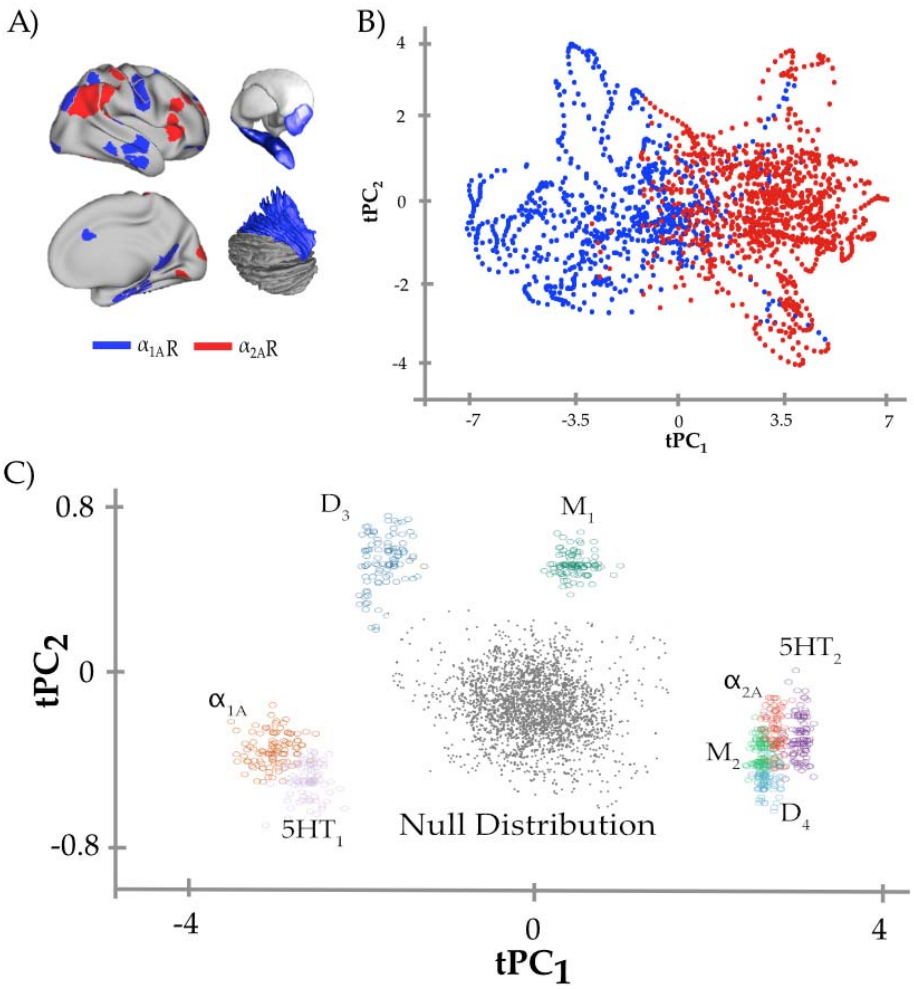

Figure 5 – The neurochemical signature of integrated cognitive function. a) spatial surface map relating relative density of α_{1A} (blue) and α_{2A} (red) parcels; b) low-dimensional embedding colored according to each adrenergic receptor family with maximal similarity to the data at each time point; c) bootstrapped mean density of $tPC_{1\text{-}2}$ for eight cognitively-relevant neurotransmitter receptor maps from four distinct neurotransmitter families: dopamine (D_3/D_4) – blue; serotonin ($5HT_1$/$5HT_2$) – purple; acetylcholine (M_1/M_2) – green; noradrenaline (α_{1A}/α_{2A}) – orange/red, with a null distribution in grey.

Discussion

We analyzed multi-task neuroimaging data to demonstrate a low-dimensional dynamic substrate for cognition in the human brain. An integrative core of brain regions appears to maximize information-processing complexity and dynamically mediate the connectional architecture of the brain in concert with task demands. By intermittently and transiently coordinating specialist regions of the brain, the integrative core

facilitates cognitive performance. Our data suggest that this core manipulates the low-dimensional architecture of the brain across an attractor landscape via highly conserved modulatory neurotransmitter systems.

Our primary discovery is the existence of a low-dimensional, dynamic, integrated component that recurs across multiple unique tasks (Figure 2) and demarcates a common cognitive architecture within the human brain (Figure 3). Fluctuations in this core network architecture were associated with maximal information processing complexity across relatively long time-scales (Figure 4), suggesting that the temporal signature of the integrative core is information-rich and accumulates information over such time-scales[26,27]. Although our results were robust to permutations and observed in an independent dataset, we do not claim that our results represent the global state of the brain *per se*, as our analyses are limited to the sub-space of cognitive processes that were available in the HCP task fMRI battery that predominantly utilize visual input and motor responses. Further, the organizational structure brain may not necessarily respect purely orthogonal dimensions. Future experiments could usefully examine the low-dimensional architecture of the brain across a broader range of psychological capacities, preferably in ecologically valid contexts[38], such as naturalistic stimulation paradigms.

Our work provides a conceptual bridge between neurochemistry, network topology and state dynamics (Figure 5). It has been known for some time that ascending neuromodulatory systems provide important constraints on cognitive function [28,39], but the systems-level mechanisms responsible for these capacities have remained obscure. The strong relationships between noradrenergic receptor density and the global brain state suggest a mechanism to modulate the flow of brain activity across the low-dimensional manifold. For instance, stimulation of adrenergic receptors can facilitate regionally specific alterations in synaptic potentials[40], modulating the effectiveness of synaptic inputs. This in turn could help to direct the system across a critical bifurcation[34], intermittently stabilizing then destabilizing a task-relevant low-energy attractor, thus linking the multistable dynamics[20] observed in our data to the facilitation of the neural interactions that underlie cognition[34]. Interestingly, α_{2A} and α_{1A} receptors have markedly different affinity for noradrenaline (high and low, respectively)[30], suggesting that both the spatial pattern of receptor density (Figure 5b) as well as the concentration of noradrenaline in the brain should affect cognitive performance[28]. In

addition, given the complex functional anatomy of the ascending arousal system, we predict that other neuromodulatory systems (such as the ascending cholinergic, serotonergic and dopaminergic projection nuclei; Figure 5c) will also influence dynamic network topology.

The brain systems that underlie cognition exist along multiple hierarchical timescales that span many orders of magnitude[41]. In contrast to electrophysiological modalities, the BOLD signal is only sensitive to neural activity on timescales that reflect slower components of cognition[27]. However, many of the cognitive processes that we investigate with typical cognitive neuroscience experiments rely on the recruitment of brain regions with slow information-processing timescales (e.g., on the order of multiple seconds)[42]. Indeed, the state-level dynamics in our experiment likely partly reflect the timescales imposed by the task block structure, a hypothesis that could be tested by using paradigms that manipulate task demands across other time scales or, as with film stimuli, embody more ecologically-valid time scale structures[42]. Future studies that use cross-modal data and/or methods to relate slow alterations in the integrative core to neurophysiological processes at higher temporal scales[43-45] will help to link descriptions of the complex architecture of the brain across temporal scales.

Our results support a novel view of functional brain organization that emphasizes the dynamical and emergent nature of cognitive function. Model-based analyses inevitably rest upon system approximations – for instance, the assumptions of orthogonality may not be the most parsimonious means for estimating the dimensionality of complex biological systems – yet these experiments still yield novel predictions that can guide future work. For example, similar low-dimensional analyses of brain function on 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 psychological manifestations of neuropsychiatric and neurodegenerative disorders. If so, detailed mapping of individual differences in the integrative core may suggest novel therapeutic interventions. More generally, we hope

that our work will provide a platform for future insights into the modular and integrative processes at work in cognition.

Acknowledgments: 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 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 (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. The funding for the study was provided by an NHMRC CJ Martin Fellowship (GNT1072403). The study was designed by all authors and conducted primarily by JMS, who also wrote the first draft. All authors provided critical feedback on the manuscript, including editing of the final manuscript.

References

1. Tononi, G., Sporns, O. & Edelman, G. M. Measures of degeneracy and redundancy in biological networks. *Proceedings of the National Academy of Sciences* **96,** 3257–3262 (1999).
2. Cunningham, J. P. & Yu, B. M. Dimensionality reduction for large-scale neural recordings. *Nat Neurosci* **17,** 1500–1509 (2014).
3. Kato, S. *et al.* Global brain dynamics embed the motor command sequence of Caenorhabditis elegans. *Cell* **163,** 656–669 (2015).
4. Nichols, A. L. A., Eichler, T., Latham, R. & Zimmer, M. A global brain state underlies C. elegans sleep behavior. *Science* **356,** eaam6851 (2017).
5. Lemon, W. C. *et al.* Whole-central nervous system functional imaging in larval Drosophila. *Nat Commun* **6,** 7924 (2015).
6. Stitt, I. *et al.* Dynamic reconfiguration of cortical functional connectivity across brain states. *Sci Rep* **7,** 8797 (2017).
7. Sporns, O. Network attributes for segregation and integration in the human brain. *Current Opinion in Neurobiology* **23,** 162–171 (2013).
8. Tononi, G. & Sporns, O. A measure for brain complexity: relating functional segregation and integration in the nervous system. *Proc. Natl. Acad. Sci. U.S.A.* **91,** 5033–5037 (1994).
9. Breakspear, M. Dynamic models of large-scale brain activity. *Nat Neurosci* **20,** 340–352 (2017).
10. Tononi, G. & Edelman, G. M. Consciousness and complexity. *Science* **282,** 1846–1851 (1998).
11. Dehaene, S. & Changeux, J.-P. Experimental and theoretical approaches to conscious processing. *Neuron* **70,** 200–227 (2011).
12. 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).
13. Boly, M. *et al.* Stimulus set meaningfulness and neurophysiological differentiation: a functional magnetic resonance imaging study. *PLoS ONE* **10,** e0125337 (2015).
14. Barch, D. M. *et al.* Function in the human connectome: task-fMRI and individual differences in behavior. *NeuroImage* **80,** 169–189 (2013).
15. Kennel, M. B. & Abarbanel, H. D. I. False neighbors and false strands: a reliable minimum embedding dimension algorithm. *Phys Rev E Stat Nonlin Soft Matter Phys* **66,** 026209 (2002).
16. Duncan, J. The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. *Trends in Cognitive Sciences* **14,** 172–179 (2010).
17. Dosenbach, N. U. F. *et al.* A Core System for the Implementation of Task Sets. *Neuron* **50,** 799–812 (2006).
18. Ott, E. & Antonsen, T. M. Low dimensional behavior of large systems of globally coupled oscillators. *Chaos* **18,** 037113 (2008).
19. Rabinovich, M. I., Huerta, R., Varona, P. & Afraimovich, V. S. Transient cognitive dynamics, metastability, and decision making. *PLoS Comput. Biol.* **4,** e1000072 (2008).
20. Cocchi, L., Gollo, L. L., Zalesky, A. & Breakspear, M. Criticality in the brain: A synthesis of neurobiology, models and cognition. *Progress in Neurobiology* **158,** 132–152 (2017).
21. Poldrack, R. A. *et al.* Discovering relations between mind, brain, and mental disorders using topic mapping. *PLoS Comput. Biol.* **8,** e1002707 (2012).
22. Le Van Quyen, M. Disentangling the dynamic core: a research program for a neurodynamics at the large-scale. *Biol. Res.* **36,** 67–88 (2003).
23. van den Heuvel, M. P. & Sporns, O. An anatomical substrate for integration among functional networks in human cortex. *J. Neurosci.* **33,** 14489–14500 (2013).
24. Zalesky, A., Fornito, A. & Bullmore, E. On the use of correlation as a measure of network

connectivity. *NeuroImage* **60,** 2096–2106
25. Krienen, F. M., Yeo, B. T. T. & Buckner, R. L. Reconfigurable task-dependent functional coupling modes cluster around a core functional architecture. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* **369,** 20130526–20130526 (2014).
26. 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. *Philosophical Transactions of the Royal Society B: Biological Sciences* **370,** 20140165–20140165 (2015).
27. Honey, C. J. *et al.* Slow Cortical Dynamics and the Accumulation of Information over Long Timescales. *Neuron* **76,** 423–434 (2012).
28. Robbins, T. W. & Arnsten, A. F. T. The Neuropsychopharmacology of Fronto-Executive Function: Monoaminergic Modulation. *Annu. Rev. Neurosci.* **32,** 267–287 (2009).
29. 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).
30. Ramos, B. P. & Arnsten, A. F. T. Adrenergic pharmacology and cognition: focus on the prefrontal cortex. *Pharmacology & Therapeutics* **113,** 523–536 (2007).
31. McGinley, M. J., David, S. V. & McCormick, D. A. Cortical Membrane Potential Signature of Optimal States for Sensory Signal Detection. *Neuron* **87,** 179–192 (2015).
32. Engel, T. A., Steinmetz, N. A. & Gieselmann, M. A. Selective modulation of cortical state during spatial attention. (2016).
33. Shine, J. M. *et al.* The Dynamics of Functional Brain Networks: Integrated Network States during Cognitive Task Performance. *Neuron* **92,** 544–554 (2016).
34. 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* (2018).
35. Wang, M. *et al.* Alpha2A-adrenoceptors strengthen working memory networks by inhibiting cAMP-HCN channel signaling in prefrontal cortex. *Cell* **129,** 397–410 (2007).
36. Doze, V. A. *et al.* Long-term α1A-adrenergic receptor stimulation improves synaptic plasticity, cognitive function, mood, and longevity. *Mol. Pharmacol.* **80,** 747–758 (2011).
37. Shine, J. M. & Poldrack, R. A. Principles of dynamic network reconfiguration across diverse brain states. *NeuroImage* (2017). doi:10.1016/j.neuroimage.2017.08.010
38. 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).
39. Moran, R. J. *et al.* Free energy, precision and learning: the role of cholinergic neuromodulation. *J. Neurosci.* **33,** 8227–8236 (2013).
40. Lee, S.-H. & Dan, Y. Neuromodulation of brain states. *Neuron* **76,** 209–222 (2012).
41. Kiebel, S. J., Daunizeau, J. & Friston, K. J. A hierarchy of time-scales and the brain. *PLoS Comput. Biol.* **4,** e1000209 (2008).
42. Hasson, U., Yang, E., Vallines, I., Heeger, D. J. & Rubin, N. A hierarchy of temporal receptive windows in human cortex. *J. Neurosci.* **28,** 2539–2550 (2008).
43. Baker, A. P. *et al.* Fast transient networks in spontaneous human brain activity. *Elife* **3,** e01867 (2014).
44. Engel, A. K., Gerloff, C., Hilgetag, C. C. & Nolte, G. Intrinsic coupling modes: multiscale interactions in ongoing brain activity. *Neuron* **80,** 867–886 (2013).
45. de Pasquale, F., Penna, Della, S., Sporns, O., Romani, G. L. & Corbetta, M. A Dynamic Core Network and Global Efficiency in the Resting Human Brain. *Cerebral Cortex* **26,** 4015–4033 (2016).
46. Gordon, E. M., Lauman, T. O., Adeyemo, B., Huckins, J. F., Kelley, W. M. & Petersen, S. E. Generation and Evaluation of a Cortical Area Parcellation from Resting-State Correlations. *Cerebral Cortex* **26**, 288-303.

Methods

Data acquisition

Data used in the preparation of this work were obtained from the Human Connectome Project (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% female[1]). For each participant, BOLD data from seven unique tasks were acquired using multiband gradient-echo EPI, amounting to 23 minutes 17 seconds of data (1940 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 degrees, field of view = 208x180 mm (matrix = 104 x 90), 2x2x2 isotropic voxels with 72 slices, alternated LR/RL phase encoding.

Data pre-processing

Bias field correction and motion correction (12 linear DOF using FSL's FLIRT) were applied to the HCP resting state data as part of the minimal preprocessing pipeline[1]. Temporal artifacts were identified in each dataset by calculating framewise displacement from the derivatives of the six rigid-body realignment parameters estimated during standard volume realignment[2], 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 12 linear head movement parameters (and their temporal derivatives), frame-wise displacement, DVARS, and anatomical masks from the CSF and deep cerebral WM were regressed from the data using the CompCor strategy[3]. 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.

Brain parcellation

Following pre-processing, the mean time series was extracted from 375 pre-defined regions-of-interest (ROI). To ensure whole-brain coverage, we extracted 333 cortical parcels (161 and 162 regions from the left and right hemispheres, respectively) using the

Gordon atlas[4], 14 subcortical regions from Harvard-Oxford subcortical atlas (bilateral thalamus, caudate, putamen, ventral striatum, globus pallidus, amygdala and hippocampus), and 28 cerebellar regions from the SUIT atlas[5].

Cognitive tasks

Seven unique tasks were utilized from the Human Connectome Project consortium: Emotion, Gambling, Language/Math, Motor, N-back, Relational and Social[6]. 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 Motor task involved an auditory, language-based task. The time points associated with each block were convolved with a canonical haemodynamic 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 grey lines in Figure 1B). All results were successfully replicated using a finite impulse response model.

Principal component analysis

Pre-processed data from each task were concatenated to form a multi-task time series per subject and a spatial PCA was performed on the resultant data[7]. Task block structure was not regressed from the data prior to PCA estimation. The time series of each PC (tPC) was then estimated by calculating the weighted mean of the group-level BOLD time series associated with each respective PC[8]. 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 data[9], we calculated the percentage of false nearest neighbors following the PCA decomposition (a measure of effective embedding across dimensions[9]), and found that there were < 10% present in the top five PCs, and <1% present in the top ten PCs (Figure 1B).

To ensure that the PCA results were robust to individual differences, we ran two subsequent analyses: i) the same analysis was conducted in a Replication cohort (N = 100); and ii) a bootstrap analysis was performed by estimating a PCA on 100 randomly chosen subjects (with replacement, from a larger pool of 200 unrelated 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: i) 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 ($r > 0.5$), 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*); ii) we used a bootstrapping approach in which we randomly selected between 2-6 tasks and re-ran the PCA, and then performed both a spatial and temporal correlation between the topography and time series for PC_1, respectively. In doing so, we found that at least 4 tasks were required to recreate the pattern found across the original 7 tasks (Figure S1); and iii) we ran a standard GLM using the concatenated task block structure, demonstrating a selective positive correlation between the spatial map of PC_1 and the resultant Beta map ($r = 0.937$; $p < 0.001$). A similar significant relationship was selectively observed between PC_5 and the Beta map estimated from a GLM comparing task-onsets to BOLD activity ($r = 0.335$; $p < 0.001$).

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 $tPC_{1\text{-}5}$: 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 $tPC_{1\text{-}5}$: 0.00 ± 0.1) or signals from the white matter and cerebrospinal fluid (mean r across $tPC_{1\text{-}5}$: 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' (i.e., censoring data with framewise displacement > 0.25 and DVARS > 2.5%)[2].

Low-dimensional manifold

To describe flow along the low-dimensional embedding space, we used an approach previously utilized in *C. elegans* calcium imaging[8], in which the global brain cycle is partitioned into four 'phases' using the time course of the tPC_1 as a reference signal: a trough in tPC_1 defined the LOW state ($tPC_1(t) < 33^{rd}$ percentile; blue in Figure 2A), an increase in tPC_1 defined the RISE state (33rd percentile < $tPC_1(t) < 67^{th}$ percentile & $dtPC_1' > 0$; red), a plateau in the tPC_1 signal defined the HIGH state ($tPC_1(t) > 67^{th}$ percentile; yellow), and a decrease in tPC_1 defined the FALL state (33rd percentile < $tPC_1(t) < 67^{th}$ percentile & $dtPC_1' < 0$; green). After classifying low-dimensional activity into these four states, we then performed a linear interpolation on each trajectory (i.e., to warp each segment into a set of identically-sized vectors). We were then able to estimate the trajectory of a low-dimensional 'manifold' by calculating the mean activity across the interpolated trajectories, which in turn could be projected into the embedding space to describe the manifold (Figure 2C).

Using the four tPC_1 categories (Low, Rise, High and Fall), we estimated the 'dwell time' for each of the states explored by the first PC by calculating the number of consecutive TRs in which each state 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 or metastable process respectively, see [10]). 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 (i.e. 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 literature[11], in which we excluded topics that were not explicitly related

to psychological constructs. These 28 maps were further collapsed into four tight-knit 'topic families' (see Table S1) by calculating the spatial similarity of each map and then clustering the matrix using a weighted version of the Louvain algorithm (i.e. the algorithm described in *Time-resolved network topology*). Topic Families were assigned labels according to the top 10 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 (Figure 3C). Finally, we used a non-linear, block bootstrapping permutation test, 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 ROIs, we used the Multiplication of Temporal Derivatives (*M*) technique[12]. *M* is computed by calculating the point-wise product of temporal derivative of pairwise time series (Equation 1). The resultant score is then averaged over a temporal window, *w*, in 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.

$$M_{ijt} = \frac{1}{w}\sum_{t}^{t+w}\frac{(t_{it}\prime \times t_{jt}\prime)}{(\sigma_{t_i}\prime \times \sigma_{t_j}\prime)} \qquad [1]$$

Where for each time point, *t*, the *M* for the pairwise interaction between region *i* and *j* is defined according to equation 1, where *t'* is the first temporal derivative of the ith 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.

Time-resolved network topology

The Louvain modularity algorithm from the Brain Connectivity Toolbox (BCT; [13]) 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.

$$Q_T = \frac{1}{v^+}\sum_{ij}\left(w_{ij}^+ - e_{ij}^+\right)\delta_{M_iM_j} - \frac{1}{v^+ + v^-}\sum_{ij}\left(w_{ij}^- - e_{ij}^-\right)\delta_{M_iM_j} \qquad [2]$$

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 δ_{MiMj} is set to 1 when regions are in the same community and 0 otherwise. '+' and '–' superscripts 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 (BCT, http://www.brain-connectivity-toolbox.net/). This afforded an estimate of both the time-resolved modularity (Q_T) and cluster assignment (Ci_T) within each temporal window for each participant in the study. We calculated all graph theoretical measures on un-thresholded, weighted and signed connectivity matrices[13] and the γ parameter was set to 1. Consistent with previous studies[14], 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)[15], 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 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 (i.e. between-module strength) and has previously been used to successfully characterize hubs within brain networks (e.g. see [16]). 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. The participation coefficient 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. Although synchronous sensory inputs do not necessarily force regions to couple together over time[17], 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 connectivity[12]. 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 bootstrapping 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 tPC_1, albeit with a peak in similarity at a window length of 20 TRs).

The integrative core

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 (i.e. mean regional-to-core connectivity ≤ 1.0). Prior to 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.

Next, in keeping with previous work[18], we used a generalized psychophysiological interaction approach to estimate task-based functional connectivity, separately for each task. Briefly, this approach controls for both block-related effects, as well as context-independent connectivity before modeling the effect of the task on each regions connectivity[18]. We then calculated the mean task-based connectivity from the integrative core (regions with significantly elevated B_T) to all other parcels. The resultant signatures were distinctly different across tasks (mean r = 0.03 ± 0.1).

To estimate time series differentiation, we calculated the Lempel Ziv complexity [19] 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 LZ complexity scores with the participation coefficient (B_T) associated with tPC_1. The autocorrelation function was estimated for each region by calculating the time-delayed Pearson's correlation between each region's pre-preprocessed BOLD time series, using a lag of 1-30 TRs (0.7 – 21.6s). 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 noradrenergic neurotransmitter receptors. Similar to the approach outlined above, we identified neurotransmitter receptor maps for two major noradrenergic receptor subtypes (α_{1A} and α_{2A}) that have opposing effects on cognitive

performance (impair and improve, respectively). We then created a weighted mean between each of these neurotransmitter receptor maps and the concatenated BOLD time series data. Using a 'winner-take-all' approach, we categorized each time point according to the cluster with the strongest spatial correspondence to the regional BOLD pattern present at that time, which allowed us to then project the clusters into the low-dimensional embedding space (Figure 5B). We then applied a series of block-resampling permutation tests to test for temporal alignments between noradrenergic receptor maps and the manifold and topic family maps, separately (5,000 iterations; $p < 0.01$). Finally, to determine the specificity of the noradrenergic system as a whole, we used a non-linear permutation test[20] to demonstrate that the family of noradrenergic neurotransmitter receptors ($\alpha_{2A/B/D}$ and $\alpha_{2A/B/C}$) were selectively associated with unique aspects of the low-dimensional manifold (mean Euclidean distance across top 5 PCs greater than 95th percentile of iterated null distribution). Although other neurotransmitter receptor groups (such as cholinergic and dopaminergic receptors) also related to the manifold, the noradrenergic receptors were associated with a larger mean Euclidean distance separating the top 5 PCs (5.12 ± 0.5) than either cholinergic (4.89 ± 0.5) or dopaminergic receptor groups (4.41 ± 0.4). Despite this result, we ran an additional analysis in which the two most cognitively-relevant transmitter receptors from each of three major families (dopaminergic, cholinergic and serotonergic) were fit to the embedding space (Figure 5C).

Supplementary Figures S1-S2

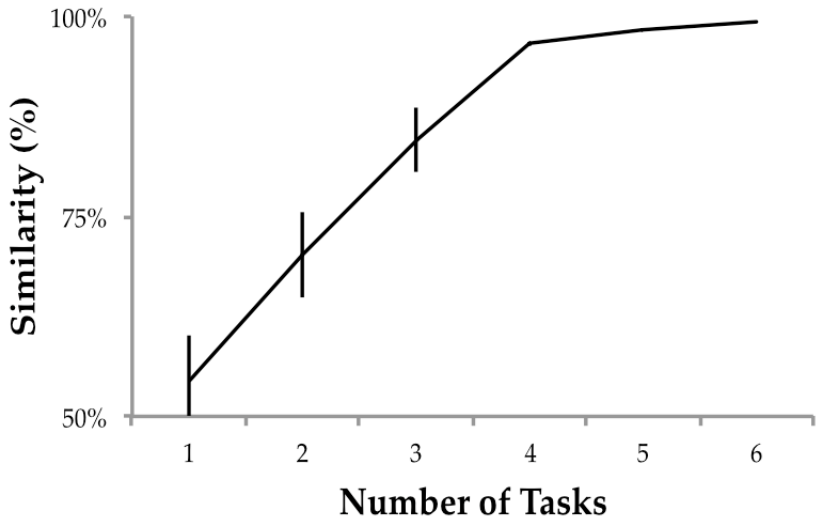

Figure S1 – results of the bootstrapping analysis that shows that 4-5 tasks are required to discover the same underlying principal component (tPC_1) that recurs across task blocks; error bars denote standard error across 100 iterations.

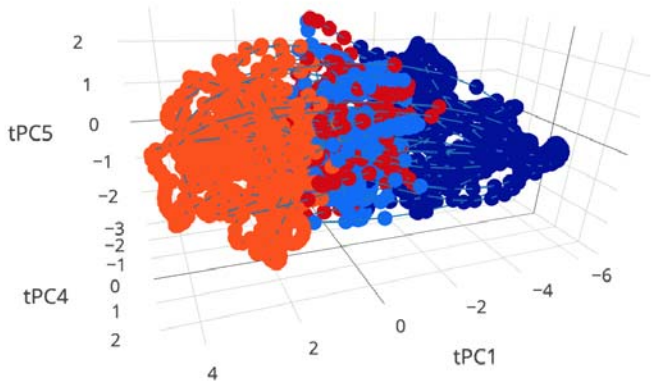

Figure S2 – embedding space, comparing tPC_1 with tPC_4 and tPC_5.

Tables

Family	Topic number	Main Term
Motor/Attention	Topic 000	Motor Observation
	Topic 011	Motion
	Topic 015	Spatial
	Topic 035	Sensory
	Topic 041	Attention
	Topic 049	Motor
Memory	Topic 001	Vision/Object
	Topic 005	Faces
	Topic 007	Language
	Topic 008	Inhibition
	Topic 010	Mind Wander
	Topic 012	Memory
	Topic 017	Social
	Topic 019	Fear
	Topic 023	Emotion +
	Topic 025	Learning
	Topic 029	Reward
	Topic 040	Emotion -
Cognition	Topic 002	Conflict/Control
	Topic 022	Working Memory
	Topic 033	Processing
	Topic 042	Number
	Topic 047	Conflict/Control
Language	Topic 003	Visual
	Topic 020	Language
	Topic 032	Speech/Aud
	Topic 036	Language
	Topic 044	Semantic

Table S1 – NeuroSynth Topic Families

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., Liau, J. & Liu, T. T. A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage* **37,** 90–101 (2007).
4. Gordon, E. M. *et al.* Generation and Evaluation of a Cortical Area Parcellation from Resting-State Correlations. *Cereb. Cortex* bhu239 (2014). doi:10.1093/cercor/bhu239
5. Diedrichsen, J., Balsters, J. H., Flavell, J., Cussans, E. & Ramnani, N. A probabilistic MR atlas of the human cerebellum. *NeuroImage* **46,** 39–46 (2009).
6. Barch, D. M. *et al*. Function in the human connectome: task-fMRI and individual differences in behavior. *NeuroImage* **80,** 169–189 (2013).
7. McIntosh, A. R. & Lobaugh, N. J. Partial least squares analysis of neuroimaging data: applications and advances. *NeuroImage* **23 Suppl 1,** S250–63 (2004).
8. Kato, S. *et al.* Global brain dynamics embed the motor command sequence of Caenorhabditis elegans. *Cell* **163,** 656–669 (2015).
9. Kennel, M. B. & Abarbanel, H. D. I. False neighbors and false strands: a reliable minimum embedding dimension algorithm. *Phys Rev E Stat Nonlin Soft Matter Phys* **66,** 026209 (2002).
10. Cocchi, L., Gollo, L. L., Zalesky, A. & Breakspear, M. Criticality in the brain: A synthesis of neurobiology, models and cognition. *Progress in Neurobiology* **158,** 132–152 (2017).
11. Poldrack, R. A. *et al.* Discovering relations between mind, brain, and mental disorders using topic mapping. *PLoS Comput. Biol.* **8,** e1002707 (2012).
12. Shine, J. M. *et al*. Estimation of dynamic functional connectivity using Multiplication of Temporal Derivatives. *NeuroImage* **122,** 399–407 (2015).
13. Rubinov, M. & Sporns, O. Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage* **52,** 1059–1069 (2010).
14. Shine, J. M. *et al*. The Dynamics of Functional Brain Networks: Integrated Network States during Cognitive Task Performance. *Neuron* **92,** 544–554 (2016).
15. Guimerà, R. & Nunes Amaral, L. A. Functional cartography of complex metabolic networks. *Nature* **433,** 895–900 (2005).
16. 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).
17. Krienen, F. M., Yeo, B. T. T. & Buckner, R. L. Reconfigurable task-dependent functional coupling modes cluster around a core functional architecture. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* **369,** 20130526–20130526 (2014).
18. Cole, M. W. *et al*. Multi-task connectivity reveals flexible hubs for adaptive task control. *Nat Neurosci* **16,** 1348–1355 (2013).
19. Boly, M. *et al.* Stimulus set meaningfulness and neurophysiological differentiation: a functional magnetic resonance imaging study. *PLoS ONE* **10,** e0125337 (2015).
20. Nichols, T. E. & Holmes, A. P. Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Hum Brain Mapp* **15,** 1–25 (2002).