Emergence of neural dynamics within a co-ordinate space of large-scale neural hierarchies

Theodoros Karapanagiotidis[1,*], Diego Vidaurre[2], Andrew J. Quinn[2], Deniz Vatansever[3], Giulia L. Poerio[4], Adam Turnbull[1], Robert Leech[5], Boris Bernhardt[6], Elizabeth Jefferies[1], Daniel S. Margulies[7] Thomas E. Nichols[8], Mark W. Woolrich[2], Jonathan Smallwood[1]

1 Department of Psychology, York Neuroimaging Centre, University of York, York YO10 5DD, United Kingdom
2 Oxford Centre for Human Brain Activity, Wellcome Centre for Integrative Neuroimaging, Department of Psychiatry, University of Oxford, Oxford OX3 7JX, United Kingdom
3 Institute of Science and Technology for Brain-inspired Intelligence, Fudan University, Shanghai 200433, PR China
4 Department of Psychology, University of Essex, Colchester, Essex CO4 3SQ, United Kingdom
5 Centre for Neuroimaging Science, Kings College, London SE5 8AF, United Kingdom
6 McConnell Brain Imaging Centre, Montréal Neurological Institute, McGill University, Montréal, QC H3A 2B4, Canada
7 Brain & Spine Institute (ICM), National Center for Scientific Research, Paris 75013, France
8 Oxford Centre for Functional MRI of the Brain, Wellcome Centre for Integrative Neuroimaging, Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford OX3 9DU, United Kingdom

* theodoros.karapanagiotidis@york.ac.uk

Abstract

Contemporary views of brain function emphasise that neural processing is organised along multiple macro-scale hierarchies that describe the relationship between large-scale networks. Examples include the organisation of networks on a continuum from unimodal to transmodal function, as well as the characteristic response of the brain to tasks. Our study examines the possibility that these hierarchies represent a functional landscape that describes how organised states of neural activity emerge. To address this possibility, we applied machine learning to recordings of task-free neural data to identify the distribution of naturally occurring network states at rest. We examined the distribution of these states in a lower dimensional space reflecting three well-described neurocognitive hierarchies (unimodal-transmodal, visual-sensorimotor and task positive). We found spontaneous neural states were generally absent from the centre of this lower dimensional space, and instead were located at the extreme ends of one or more hierarchies. Furthermore, states most discriminative of participants' experience at rest were not well-described by hierarchies reflecting the unimodal-transmodal architecture of the brain, and instead reflected both extremes of a hierarchy describing how the brain responds to task demands. Our analyses demonstrate that neural hierarchies provide a flexible co-ordinate space to understand the hidden dynamics of unconstrained thought and highlight that neurocognitive states characteristic of the brain's response to task demands capture important aspects of our self-generated mental lives during wakeful rest.

Introduction

Contemporary neuroscience recognises that brain organisation can be understood as a composite of large-scale hierarchies describing the complex relationships between different brain networks (Margulies et al., 2016; Mesulam, 1998). These macro-scale hierarchies reflect features of cortical organisation including the structure of unimodal cortex (Vázquez-Rodríguez et al., 2019; Paquola et al., 2019), a functional gradient between specialised areas of cortex and transmodal regions with more abstract functions (Margulies et al., 2016), and patterns of neural responses observed when tasks are performed (Duncan, 2010). We have a growing understanding of how these different cortical hierarchies are organised in time (Vidaurre et al., 2017b, 2018), and how they change during tasks (Shine et al., 2019a,b) and in disease (Hong et al., 2019); however, it remains unclear how these features of brain organisation impact upon naturally occurring dynamic cognitive states that emerge readily in daily life, such as during wakeful rest. In particular, there are two important questions that remain unanswered: (1) Do naturally occurring states at rest emerge within a framework that reflects constraints implied by different neural hierarchies? And (2) if so, which of these neural hierarchies are most important for how cognitive features of spontaneous states emerge?

To address these questions, we leveraged developments in machine learning that allow the identification of dynamic neural network states, and contemporary descriptions of neural organisation, to explore whether the functional architecture of the cortex constrains how naturally occurring states emerge at rest. We measured neural function during wakeful rest in a large cohort of healthy individuals using functional magnetic resonance imaging (fMRI). These participants subsequently completed a set of questions related to their experiences during the scan, and a subset completed questionnaire measures of physical and mental health at a later session. We applied hidden Markov modelling (HMM) to the neural data to reveal distinct brain states corresponding to specific periods of temporally reoccurring patterns of neural mean activity and functional connectivity (Vidaurre et al., 2017a). Our first analysis examined whether these states can be considered to be emergent properties of the hierarchical organisation of the cortex. To this end, we investigated how these states are distributed across a multi-dimensional space formed by three well-described neural hierarchies based on the differentiation of whole-brain connectivity patterns (Margulies et al., 2016). Second, we explored whether the organisation of states by these neural hierarchies helps reveal how self-generated experiences emerge at rest, by examining how different states related to the reports of our participants' experience at rest. The study workflow is presented in Figure 1.

Our results show that many of the states identified by our analyses had associations with pre-established neural hierarchies, suggesting that these hierarchies do constrain the dynamic patterns seen at rest. Moreover, the time spent in those states which mimicked either the task positive or task negative features of how the brain responds to a task (e.g. Duncan, 2010; Power and Petersen, 2013) had reliable associations with retrospective self-reports of the pattern of cognition during the scan. This suggests that patterns seen in both complex tasks and during periods of low external demand can make important contributions to unconstrained patterns of neurocognitive activity.

Results

Identifying reoccurring neural states at rest

We applied hidden Markov modelling to the resting state fMRI data (see Methods). This method uses Bayesian inference to identify reoccurring patterns of brain network

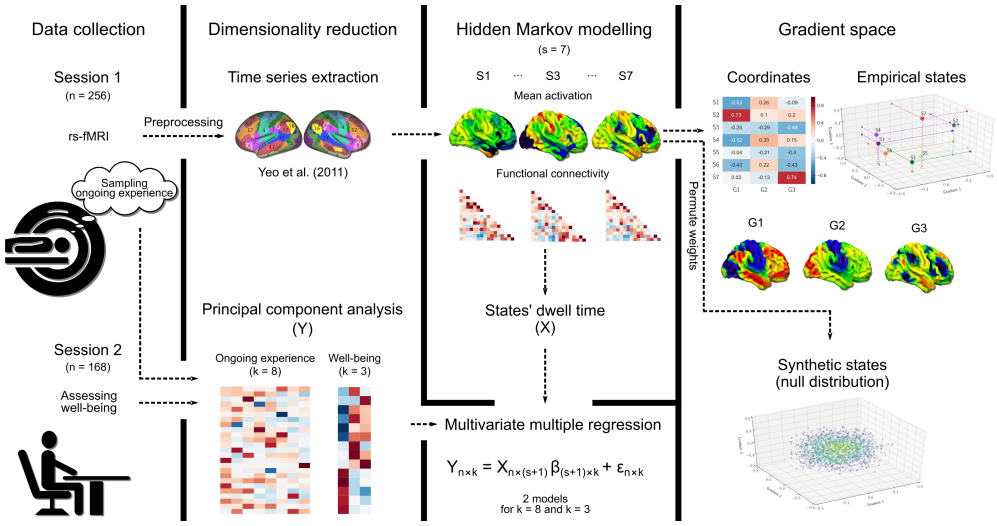

Figure 1. *Schematic of the workflow of the current study.* We measured neural activity at rest in a sample of n=256 individuals who then completed a retrospective measure of the experiences they had during this period. In a subset of the individuals we also acquired measures of well-being. We then reduced the dimensionality of both data sets and applied hidden Markov modelling to the neuroimaging time series. We calculated the association of the identified brain network states to the first three neurocognitive hierarchies identified by Margulies et al. (2016), as well as to a distribution of synthetic states generated through permutation. Finally, we performed a multiple regression analysis in which we calculated the association between the mean dwell-time of each state to the patterns of experiences calculated at rest, as well as to trait measures of well-being.

activity; see Vidaurre et al. (2017b,a) for an empirical validation of this approach. Here, we generated HMM solutions capturing mean activity and functional connectivity patterns within the neural data recorded from our sample. We opted for a 7-state solution, however, our results highlighting associations with behaviour remained broadly similar for a decomposition of 9 states as well. We ran the algorithm 10 times in each case and found that the 7-state solution produced the same decomposition across all iterations and was relatively more stable and had greater split-half reliability than the 9-state solution (see Fig. S3). Accordingly, we focus on the 7-state solution and report a parallel analysis of the 9-state solution in the supplementary Materials (Fig. S6 and S7). Figure S2 shows the spatial description of neural mean activity in each state and the associated cognitive terms produced by a meta-analysis using Neurosynth (Yarkoni et al., 2011), displayed in the form of word clouds.

The relationship between naturally occurring states at rest and neural hierarchies

Having identified a set of relatively stable recurring patterns within our data, we next explored their association with three well-described neural hierarchies that highlight gradients in connectivity patterns over space. We chose the first three hierarchies from Margulies et al. (2016) which represent the division of unimodal and transmodal systems (gradient 1), the distinction between vision and motor (gradient 2) and the patterns seen during the brain's response to tasks relative to rest (gradient 3). These were generated using diffusion embedding, a non-linear dimensionality reduction technique, on the Human Connectome Project Data, and identified components describing the maximum variance in functional connectivity patterns (see Methods and

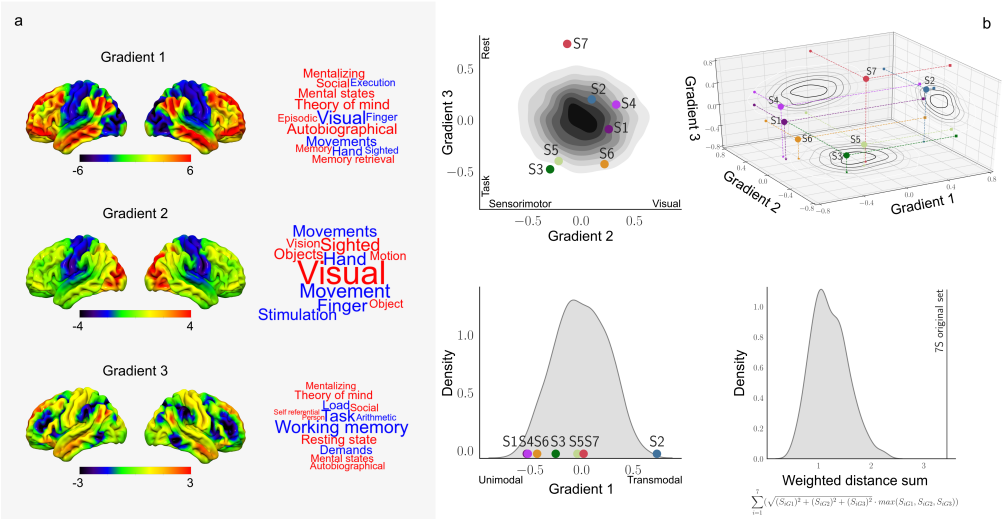

Figure 2. *Relationship between naturally occurring states at rest and three well-established cognitive hierarchies.* **a**) This panel shows the spatial distribution of the three neurocognitive hierarchies identified by Margulies et al. (2016) as well as a meta-analysis of these spatial maps using Neurosynth, which are displayed in the form of word clouds. Font size represents the strength of the association and font colour its sign (red for positive and blue for negative values). **b**) This panel shows the spatial similarity between the spatial maps describing these neural hierarchies and the set of spatial maps generated from the HMM decomposition of the real data. In these figures the contour plots projected onto each plane describe the distribution of values generated through permutation. The lower right-hand histogram shows the distribution of synthetic states compared to the real data, in terms of the sum of the weighted distance from the origin of each point in each set.

initial paper for further details). The spatial distribution of each of these gradients is presented in Figure 2a along with a meta-analysis of these maps using Neurosynth.

The aim of our next analysis was to understand whether there is a relationship between the states inferred by the HMM (see Figure S2) and the spatial maps which represent these pre-established hierarchies. To achieve this goal, we calculated the pair-wise similarity between each of the hierarchies and each of the states (see Methods). We compared the associations between the real states inferred in our analyses to a null distribution of synthetic states. The null distribution was generated by randomly permuting the weights describing the relative contribution of each brain region to the mean neural activity captured by each state. The results of these analyses are presented on Figure 2b. The contour plot and histogram on the left show where the real data fall on a subset of the dimensions, while the three-dimensional scatter plot show the same data in a combined manner. In these plots, the null distribution generated through permutation is represented as the shaded area in grey. We also generated a video that shows the data in a 3-d form (`https://github.com/tkarapan/hmm-large-scale-hierarchies/blob/master/7S_Null.mp4`).

The synthetic states clustered towards the middle of the three dimensions, indicating that states generated in a random manner tend not to be very similar in spatial terms to the three pre-established neurocognitive hierarchies (Margulies et al., 2016). In contrast, the states generated based on real data fell, on average, away from the centre and towards the outer edge of the distribution of synthetic states. To quantify whether the apparent difference between the real and permuted data was robust, we calculated the sum of the weighted distance from the origin for each of the sets of states generated through permutation of this space, and compared this to the same value from the real

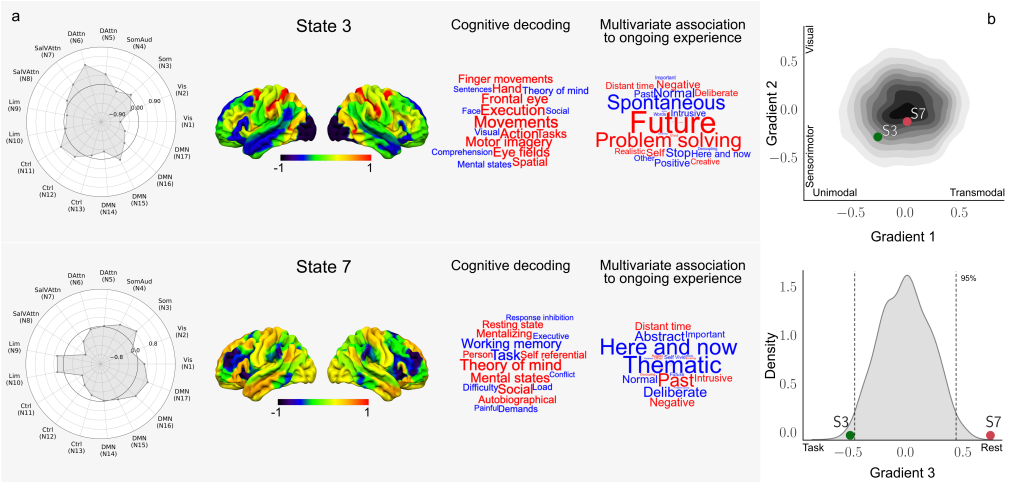

Figure 3. *Associations between the mean dwell-time of different neural states and experience, and their relationship to neurocognitive hierarchies.* **a**) This panel shows the spatial distribution of the two states as well as a meta-analysis of these spatial maps using Neurosynth, which are displayed in the form of word clouds. The word clouds on the right show the pattern of answers associated with each state. In addition, we show for each state, the relative contributions of each of the networks used in the HMM analyses, in the form of radar plots. **b**) This panel shows the spatial similarity between the spatial maps describing the large-scale neural hierarchies and the spatial maps generated from the real data, using the same representational scheme as in Figure 2.

data. The distribution of these values is presented in the right-hand histogram in Figure 2b, where it can be seen that the weighted distance sum of the real states from the centre of this space is higher than any of the states generated synthetically.

Together, these analyses suggest that states determined through the application of HMM to our real data have an association with the spatial maps describing neural hierarchies that is greater than chance. To identify the robustness of this effect we performed a parallel analysis based on the 9-state solution to the current data, as well as the published states inferred from applying HMM with the same model parameters on the Human Connectome Project data (Vidaurre et al., 2017b), finding comparable results (see Fig. S5). Together, these analyses show that the states that occur naturally at rest, as approximated through HMM, fall at the extremities of a co-ordinate space defined by well-established neural hierarchies. This is consistent with the hypothesis that macro-scale hierarchies can be thought of as constraining the state space from which the intrinsic network states emerge.

The relationship between neural states and experience

Having determined that network states found in real data distribute in such a manner that their emergence is associated to neural hierarchies constraints, we next explored whether this framework can also help to understand how these naturally occurring neural states are related to the experiences reported by our participants. To answer this question, we performed a multi-variate analysis of co-variance in which the mean dwell-time of each state was an explanatory variable and the self-reported data was the dependent variable (see Methods for details of the analyses; the specific questions can be found in Table S1). We included motion, age and gender as co-variates of no interest.

This analysis revealed that the mean dwell-time of states 3 and 7 had a significant multivariate association with differential patterns of reports made by our participants at

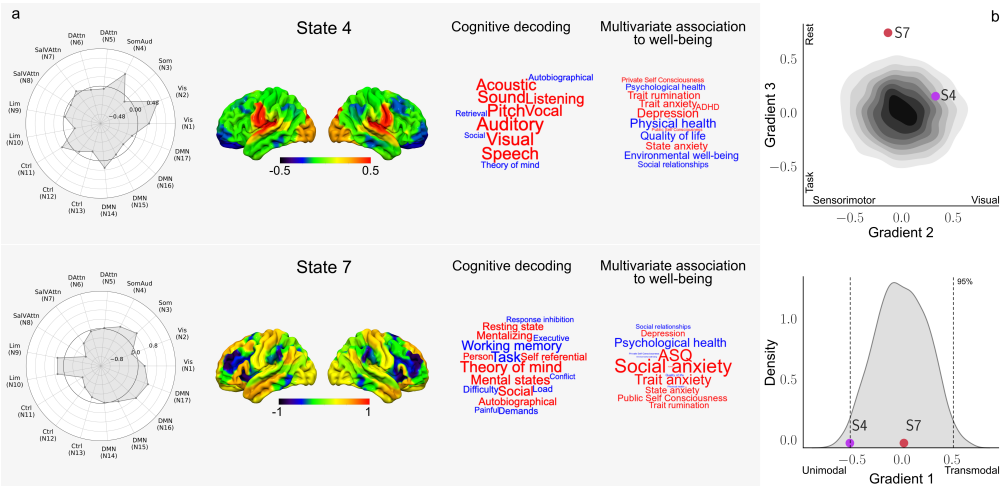

Figure 4. *Associations between the mean dwell-time of different neural states and well-being, and their relationship to neurocognitive hierarchies.* **a**) This panel shows the spatial distribution of the two states as well as a meta-analysis of these spatial maps using Neurosynth, which are displayed in the form of word clouds. The word clouds on the right show the pattern of answers associated with each state. In addition, we show for each state, the relative contributions of each of the networks used in the HMM analyses, in the form of radar plots. **b**) This panel shows the spatial similarity between the spatial maps describing the large-scale neural hierarchies and the spatial maps generated from the real data, using the same representational scheme as in Figure 2.

the end of the scan, while controlling for the dwell-times of all states (F(8, 238) = 2.22, p = 0.027, Wilks' $\Lambda = 0.931$, partial $\eta^2 = 0.069$ for state 3 and F(8, 238) = 2.18, p = 0.03, Wilks' $\Lambda = 0.932$, partial $\eta^2 = 0.068$ for state 7 accordingly). These results are presented in Figure 3. It can be seen that state 3 is dominated by task-positive systems (frontal parietal and dorsal attention) as well as the medial temporal subsystem of the default mode network (DMN). Experientially, it was associated with a state of autobiographical planning, reflecting high scores on "Future" and "Problem solving" thoughts. In contrast, state 7 was associated with somatomotor, the core and lateral temporal subsystems of the DMN and the limbic system. This state was associated with reports of negative, intrusive thoughts about the past. The contour plot shows that both of these states fall towards the mid-point of both gradient 1 and 2. In contrast, it can be seen that these states are maximally dissociated along the dimension describing how the brain responds to a task; state 3 is more similar to the task positive end of gradient 3, while state 7 shows the reverse.

We next performed an analysis aimed at identifying how the mean dwell-time of the states varied with trait measures of well-being recorded in a subset of these participants. This analysis found that two of the states had a multivariate association with the questionnaire data (state 4, F(3,155) = 2.71, p = 0.047, Wilks' $\Lambda = 0.95$, partial $\eta^2 = 0.05$ and state 7, F(3, 155) = 3.74, p = 0.013, Wilks' $\Lambda = 0.933$, partial $\eta^2 = 0.067$). These relationships are presented in Figure 4 where it can be seen that state 4 falls towards the unimodal end of gradient 1 and is associated with higher levels of reports of depression and rumination. In contrast, state 7, identified in our prior analyses as falling at the task negative end of the task dimension, was most strongly linked to trait, state and social anxiety, as well as greater symptoms linked to autism.

Finally, we examined the robustness of our results linking the neural states to the psychological measures using the mean dwell-times of the 9-state solution. In the 9-state solution there were states that were similar to those in the 7-state solution (see Fig. S4).

A multivariate analysis showed that state 5, the homologue of state 3 in the 7-state solution, was associated with a similar multivariate pattern of experience, and state 6 (the homologue of state 7 in the 7-state solution) was related to the same pattern of trait measures (Fig. S6). Although no pattern of thoughts was significantly associated with state 6 in the 9-state solution, its experiential correlates were most similar across individuals to that seen in state 7 from the 7-state solution ($r = 0.6$, $p < .001$, see Fig. S7). These analyses suggest that associations of neural patterns with experience and well-being are partially preserved across small changes in the number of states generated through HMM.

Discussion

Our study set out to understand whether macro-scale neural hierarchies that describe the relationship between large-scale networks provide a set of constraints on how naturally occurring cognitive states emerge. Advanced machine learning techniques revealed a set of neurocognitive states, three of which (states 3, 4 and 7) related to either measures of experience, socio-affective traits or both. These states clustered at the extremities of a multidimensional space based on three well-described neurocognitive hierarchies, deviating substantially from a synthetic null distribution, which instead clustered at the origin of this space.

Prior studies have shown that the dimensional structure of the brain provides a compact description of how it reconfigures during tasks (Shine et al., 2019a; Telesford et al., 2016; Braun et al., 2015). Our study builds on these observations by showing that states emerging at rest, are themselves constrained by the specific shape of macro-scale functional hierarchies. A plausible account of this is that macro-scale hierarchies, identified through prior analyses exploring the functional organisation of the cortex, describe spatial constraints that determine the trajectories that naturally occurring whole-brain neural activity will take. One consequence of this discovery is that in the future it may be possible to validate hypotheses about similarity and difference in brain states through their similarity to motifs of large-scale organisation (Mars et al., 2018; Huntenburg et al., 2018).

We also shed new light into the relationships between specific macro-scale hierarchies and patterns of spontaneously occurring neurocognitive states. We found two states associated with descriptions of experience, providing the strongest evidence that they describe active cognitive states. One state was characterised by cognition focused on the future and problem solving, suggesting a mode of autobiographical planning that is common during self-generated thought (Baird et al., 2011) and is associated with activity in both the default mode and fronto-parietal systems (e.g. Spreng et al., 2010). The second state emphasised negative intrusive thoughts from the distant past and was associated with high levels of trait affective disturbance, suggesting it reflects a state of unpleasant negative rumination. Notably, these two states fell at opposite extremes of a neurocognitive hierarchy that reflects the neural response to tasks. Together, these data demonstrate that similar neural hierarchies that support general aspects of how individuals respond to the demands of an external task (Duncan, 2010) may continue to do so in the absence of any overt external behaviour. Importantly, this neural hierarchy discriminated between patterns of autobiographical planning, which are thought to be advantageous because they help consolidate personal goals (Medea et al., 2018), from those linked to patterns of rumination, which exacerbate unhappiness (Poerio et al., 2013). Our results thus suggest that the same neural hierarchy that determines effective external task performance may also discriminate beneficial and detrimental types of self-generated experiences.

Finally, our study provides an important advance into how to conceptualise and

measure naturally occurring patterns of ongoing experience. Our study shows that the application of hidden Markov modelling to resting state data yields periods of quasi-stationary large-scale network activity that have correlates with self-reported descriptions of experience. Our study, therefore, confirms the value of examining periods of ongoing thought from a dynamical perspective that acknowledges a sequence of episodes with unique neural and experiential correlates (Kucyi, 2018). Furthermore, our analysis shows that time-varying neural states can be used as explanatory variables to determine qualitative patterns of ongoing thought. In the future, machine learning may be used to identify taxonomies of the most relevant features of experience based on objective features of brain activity (Lorenz et al., 2017).

Methods and Materials

Participants

277 healthy participants were recruited from the University of York. Informed consent was obtained for all participants and the study was approved by the York Neuroimaging Centre Ethics Committee. 21 participants were excluded from analyses, 1 due to technical issues during the neuroimaging data acquisition and 20 for excessive movement during the fMRI scan (mean framewise displacement (Power et al., 2014) $>$ 0.3 mm and/or more than 15% of their data affected by motion), resulting in a final cohort of n = 256 (169 females, μ_{age} = 20.7 years, σ_{age} = 2.4).

Behavioural methods

We sampled participants' experience during the resting state fMRI scan by asking them to retrospectively report their thoughts at the end of the scan. Experience was measured using a 4-scale Likert scale with the question order randomised (all 25 questions are shown in Table S1). In a subset (n=168) of the final cohort, we also assessed their physical and mental health by administering well-established questionnaire measures at a later separate session outside of the scanner. Details about each questionnaire are presented in SI.

MRI data acquisition

MRI data were acquired on a GE 3 Tesla Signa Excite HDxMRI scanner, equipped with an eight-channel phased array head coil at York Neuroimaging Centre, University of York. For each participant, we acquired a sagittal isotropic 3D fast spoiled gradient-recalled echo T1-weighted structural scan (TR = 7.8 ms, TE = minimum full, flip angle = 20°, matrix = 256x256, voxel size = 1.13x1.13x1 mm^3, FOV = 289x289 mm^2). Resting-state functional MRI data based on blood oxygen level-dependent contrast images with fat saturation were acquired using a gradient single-shot echo-planar imaging sequence with the following parameters; TE = minimum full (≈19 ms), flip angle = 90°, matrix = 64x64, FOV = 192x192 mm^2, voxel size = 3x3x3 mm^3, TR = 3000 ms, 60 axial slices with no gap and slice thickness of 3 mm. Scan duration was 9 minutes which allowed us to collect 180 whole-brain volumes.

fMRI data pre-processing

Functional MRI data pre-processing was performed using SPM12 (`http://www.fil.ion.ucl.ac.uk/spm`) and the CONN toolbox (v.18b) (`https://www.nitrc.org/projects/conn`) (Whitfield-Gabrieli and Nieto-Castanon, 2012) implemented in Matlab (R2018a)

(https://uk.mathworks.com/products/matlab). Pre-processing steps followed CONN's default pipeline and included motion estimation and correction by volume realignment using a six-parameter rigid body transformation, slice-time correction, and simultaneous grey matter (GM), white matter (WM) and cerebrospinal fluid (CSF) segmentation and normalisation to MNI152 stereotactic space (2 mm isotropic) of both functional and structural data. Following pre-processing, the following potential confounding effects were removed from the BOLD signal using linear regression: 6 motion parameters calculated at the previous step and their 1st and 2nd order derivatives, volumes with excessive movement (motion greater than 0.5 mm and global signal changes larger than z = 3), signal linear trend, and five principal components of the signal from WM and CSF (CompCor approach, Behzadi et al., 2007). Finally, data were band-pass filtered between 0.01 and 0.1 Hz. No global signal regression was performed.

Dimensionality reduction

In order to reduce the dimensionality structure of the neuroimaging data, we used the 17 functional network parcellation from Yeo et al. (2011), slightly eroded the parcels to avoid signal leakage from neighbouring parcels, masked them with subject specific grey matter masks, and calculated the average signal from the unsmoothed data within each parcel for each volume per participant. The acquired time series were then standardised and concatenated to form a "(256 participants x 180 volumes) x 17 parcels" matrix that was used as input to the HMM algorithm.

For the behavioural data, we applied a principal component analysis (PCA) with varimax rotation to the scores describing the participants' experience at the end of the resting state scan. This revealed eight components with eigenvalues greater than 1. We followed the same procedure for the trait measures of well-being that we acquired for a sub-sample of our cohort, which identified three principal components in their structure (see Figure S1 for the loadings of both decompositions in the form of word clouds). The component scores were used as dependent variables in two separate multivariate multiple regressions, run at a later stage in order to investigate the relationship between ongoing experience and well-being respectively, with the mean dwell-times of the dynamic neural states as described by our HMM decomposition.

Hidden Markov model

To characterise the dynamics of neural activity, we applied hidden Markov modelling to the concatenated time series of the 17 parcels (Yeo et al., 2011). The inference of the model parameters was based on variational Bayes and the minimisation of free energy, as implemented in the HMM-MAR toolbox (Vidaurre et al., 2016). The HMM's inference assigns state probabilities to each time point of the time series (i.e. reflecting how likely each time point is to be explained by each state) and estimates the parameters of the states, where each state has its own model of the observed data. Each state was represented as a multivariate Gaussian distribution (Vidaurre et al., 2017a), described by its mean and covariance. Inference was run at the group level, such that the state spatial descriptions are defined across subjects, whereas the temporal activation of the states is defined at the subject level. This allowed us to discover dynamic temporal patterns of whole-brain activity and functional interactions along with their occurrence (state time series) and transition probabilities for the duration of the whole resting state fMRI scan. Detailed information about the HMM implementation and the variational Bayes inference can be found in (Vidaurre et al., 2016, 2017b,a). As with other probabilistic unsupervised learning methods (e.g. independent component analysis), HMM is sensitive to initial values. In order to

account for HMM run-to-run variability, we ran the algorithm 10 times and selected the iteration with the lowest free-energy at the end of the inference. The summary statistic describing the states' dwell-times was computed after hard classifying the states as being active or not by using Viterbi decoding. The mean dwell-time of each state was then used as an explanatory variable in subsequent multivariate generalised linear models. Dwell-times were standardised and any values $> 2.5\sigma$ were substituted with the mean ($\mu = 0$) to minimise the effect of potential outliers. All analyses controlled for age, gender and motion during the resting state fMRI scan.

Projection of state maps to a 3-dimensional space of neural hierarchies

Following the HMM inference, each state S can be characterised by parameters (μ_s, Σ_s), where μ_s is a vector containing the activation/weights (with respect to the average) of each of the 17 parcels used in our analyses and Σ_s is the covariance matrix describing their functional interactions. Using the weights of each parcel, we obtained each state's "mean activity" spatial map (Fig. S2) and calculated its spatial similarity to the first three gradient maps from Margulies et al. (2016). These maps were produced by a diffusion embedding algorithm applied on data from the Human Connectome Project, they are orthogonal, and highlight topographies of information exchange based on the differentiation of connectivity patterns. Similarity was calculated as the pairwise correlation between each state spatial map and gradient and was used as the state's co-ordinate for the corresponding dimension in gradient space. Aiming to construct a summary metric describing the topography of the states in this space, we Fisher-z transformed the correlation values (used as gradient space co-ordinates), calculated the distance of each state from the origin weighted by its maximum co-ordinate (as a way to differentiate between points lying on the surface of the same sphere), and added the weighted distances to produce a sum for the set of all states

$$\sum_{i=1}^{7}\left(\sqrt{(S_{iG1})^2+(S_{iG2})^2+(S_{iG3})^2}\cdot max\left(S_{iG1},S_{iG2},S_{iG3}\right)\right)$$

where S_{iG1}, S_{iG2}, S_{iG3} are the Fisher's z correlation values between the spatial maps of state i and gradient 1, 2, and 3 respectively.

Null distribution

In order to compute a null distribution, we randomly permuted the parcel weights, constructed the synthetic states' spatial maps and projected them into gradient space by calculating their similarity to the gradients in the same way as before. We ran 300 permutations, producing 2100 synthetic states and calculated the weighted distance sum of each set in a similar manner as with the set of empirical states.

Acknowledgments

JS was supported by European Research Council (WANDERINGMINDS - 646927). The authors would like to thank Mladen Sormaz, Charlotte Murphy and Hao-Ting Wang for their contribution to data acquisition.

References

Baird, B., Smallwood, J., and Schooler, J. W. (2011). Back to the future: autobiographical planning and the functionality of mind-wandering. *Consciousness and cognition*, 20(4):1604–1611.

Behzadi, Y., Restom, K., Liau, J., and Liu, T. T. (2007). A component based noise correction method (compcor) for bold and perfusion based fmri. *Neuroimage*, 37(1):90–101.

Braun, U., Schäfer, A., Walter, H., Erk, S., Romanczuk-Seiferth, N., Haddad, L., Schweiger, J. I., Grimm, O., Heinz, A., Tost, H., et al. (2015). Dynamic reconfiguration of frontal brain networks during executive cognition in humans. *Proceedings of the National Academy of Sciences*, 112(37):11678–11683.

Duncan, J. (2010). The multiple-demand (md) system of the primate brain: mental programs for intelligent behaviour. *Trends in cognitive sciences*, 14(4):172–179.

Hong, S.-J., de Wael, R. V., Bethlehem, R. A., Lariviere, S., Paquola, C., Valk, S. L., Milham, M. P., Di Martino, A., Margulies, D. S., Smallwood, J., et al. (2019). Atypical functional connectome hierarchy in autism. *Nature communications*, 10(1):1022.

Huntenburg, J. M., Bazin, P.-L., and Margulies, D. S. (2018). Large-scale gradients in human cortical organization. *Trends in cognitive sciences*, 22(1):21–31.

Kucyi, A. (2018). Just a thought: how mind-wandering is represented in dynamic brain connectivity. *Neuroimage*, 180:505–514.

Lorenz, R., Hampshire, A., and Leech, R. (2017). Neuroadaptive bayesian optimization and hypothesis testing. *Trends in cognitive sciences*, 21(3):155–167.

Margulies, D. S., Ghosh, S. S., Goulas, A., Falkiewicz, M., Huntenburg, J. M., Langs, G., Bezgin, G., Eickhoff, S. B., Castellanos, F. X., Petrides, M., et al. (2016). Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proceedings of the National Academy of Sciences*, page 201608282.

Mars, R. B., Passingham, R. E., and Jbabdi, S. (2018). Connectivity fingerprints: from areal descriptions to abstract spaces. *Trends in cognitive sciences*, 22(11):1026–1037.

Medea, B., Karapanagiotidis, T., Konishi, M., Ottaviani, C., Margulies, D., Bernasconi, A., Bernasconi, N., Bernhardt, B. C., Jefferies, E., and Smallwood, J. (2018). How do we decide what to do? resting-state connectivity patterns and components of self-generated thought linked to the development of more concrete personal goals. *Experimental brain research*, 236(9):2469–2481.

Mesulam, M. M. (1998). From sensation to cognition. *Brain: a journal of neurology*, 121(6):1013–1052.

Paquola, C., De Wael, R. V., Wagstyl, K., Bethlehem, R. A., Hong, S.-J., Seidlitz, J., Bullmore, E. T., Evans, A. C., Misic, B., Margulies, D. S., et al. (2019). Microstructural and functional gradients are increasingly dissociated in transmodal cortices. *PLoS biology*, 17(5):e3000284.

Poerio, G. L., Totterdell, P., and Miles, E. (2013). Mind-wandering and negative mood: Does one thing really lead to another? *Consciousness and cognition*, 22(4):1412–1421.

Power, J. D., Mitra, A., Laumann, T. O., Snyder, A. Z., Schlaggar, B. L., and Petersen, S. E. (2014). Methods to detect, characterize, and remove motion artifact in resting state fmri. *Neuroimage*, 84:320–341.

Power, J. D. and Petersen, S. E. (2013). Control-related systems in the human brain. *Current opinion in neurobiology*, 23(2):223–228.

Shine, J. M., Breakspear, M., Bell, P. T., Martens, K. A. E., Shine, R., Koyejo, O., Sporns, O., and Poldrack, R. A. (2019a). Human cognition involves the dynamic integration of neural activity and neuromodulatory systems. *Nature neuroscience*, 22(2):289.

Shine, J. M., Hearne, L. J., Breakspear, M., Hwang, K., Müller, E. J., Sporns, O., Poldrack, R. A., Mattingley, J. B., and Cocchi, L. (2019b). The low-dimensional neural architecture of cognitive complexity is related to activity in medial thalamic nuclei. *Neuron*.

Spreng, R. N., Stevens, W. D., Chamberlain, J. P., Gilmore, A. W., and Schacter, D. L. (2010). Default network activity, coupled with the frontoparietal control network, supports goal-directed cognition. *Neuroimage*, 53(1):303–317.

Telesford, Q. K., Lynall, M.-E., Vettel, J., Miller, M. B., Grafton, S. T., and Bassett, D. S. (2016). Detection of functional brain network reconfiguration during task-driven cognitive states. *NeuroImage*, 142:198–210.

Vázquez-Rodríguez, B., Suárez, L. E., Markello, R. D., Shafiei, G., Paquola, C., Hagmann, P., van den Heuvel, M. P., Bernhardt, B. C., Spreng, R. N., and Misic, B. (2019). Gradients of structure–function tethering across neocortex. *Proceedings of the National Academy of Sciences*, 116(42):21219–21227.

Vidaurre, D., Abeysuriya, R., Becker, R., Quinn, A. J., Alfaro-Almagro, F., Smith, S. M., and Woolrich, M. W. (2017a). Discovering dynamic brain networks from big data in rest and task. *Neuroimage*.

Vidaurre, D., Hunt, L. T., Quinn, A. J., Hunt, B. A., Brookes, M. J., Nobre, A. C., and Woolrich, M. W. (2018). Spontaneous cortical activity transiently organises into frequency specific phase-coupling networks. *Nature communications*, 9(1):2987.

Vidaurre, D., Quinn, A. J., Baker, A. P., Dupret, D., Tejero-Cantero, A., and Woolrich, M. W. (2016). Spectrally resolved fast transient brain states in electrophysiological data. *Neuroimage*, 126:81–95.

Vidaurre, D., Smith, S. M., and Woolrich, M. W. (2017b). Brain network dynamics are hierarchically organized in time. *Proceedings of the National Academy of Sciences*, 114(48):12827–12832.

Whitfield-Gabrieli, S. and Nieto-Castanon, A. (2012). Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks. *Brain connectivity*, 2(3):125–141.

Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., and Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nature methods*, 8(8):665.

Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., Roffman, J. L., Smoller, J. W., Zöllei, L., Polimeni, J. R., et al. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, 106(3):1125.