

1 **High-level cognition is supported by at least second order**
2 **dynamic correlations in neural activity patterns**

3 Lucy L. W. Owen¹, Thomas H. Chang^{1,2}, and Jeremy R. Manning^{1,†}

¹Department of Psychological and Brain Sciences,
Dartmouth College, Hanover, NH

³Amazon.com, Seattle, WA

[†]Address correspondence to jeremy.r.manning@dartmouth.edu

4 September 4, 2019

5 **Abstract**

6 Our thoughts arise from coordinated patterns of interactions between brain structures that change
7 with our ongoing experiences. High-order dynamic correlations in brain activity patterns reflect different
8 subgraphs of the brain’s connectome that display homologous lower-level dynamic correlations. We tested
9 the hypothesis that high-level cognition is supported by high-order dynamic correlations in brain activity
10 patterns. We developed an approach to estimating high-order dynamic correlations in timeseries data,
11 and we applied the approach to neuroimaging data collected as human participants either listened to a
12 ten-minute story or a temporally scrambled version of the story, or underwent a resting state scan. We
13 trained across-participants pattern classifiers to decode (in held-out data) when in the session each activity
14 snapshot was collected. We found that classifiers trained to decode from high-order dynamic correlations
15 yielded better performance on data collected as participants listened to the (unscrambled) story. By
16 contrast, classifiers trained to decode data from scrambled versions of the story or during the resting
17 state scan yielded the best performance when they were trained using first-order dynamic correlations
18 or raw activity patterns. We suggest that as our thoughts become more complex, they are supported by
19 higher-order patterns of dynamic network interactions throughout the brain.

20 **Introduction**

21 A central goal in cognitive neuroscience is to elucidate the *neural code*: the mapping between (a) mental
22 states or cognitive representations and (b) neural activity patterns. One means of testing models of the
23 neural code is to ask how accurately that model is able to “translate” neural activity patterns into known
24 (or hypothesized) mental states or cognitive representations (e.g., Haxby et al., 2001; Huth et al., 2016, 2012;
25 Kamitani & Tong, 2005; Mitchell et al., 2008; Nishimoto et al., 2011; Norman et al., 2006; Pereira et al.,
26 2018; Tong & Pratte, 2012). Training decoding models on different types of neural features (Fig. 1a) can
27 also help to elucidate which specific aspects of neural activity patterns are informative about cognition—
28 and, by extension, which types of neural activity patterns might comprise the neural code. For example,
29 prior work has used region of interest analyses to estimate the anatomical locations of specific neural
30 representations (e.g., Etzel et al., 2009), or to compare the relative contributions to the neural code of

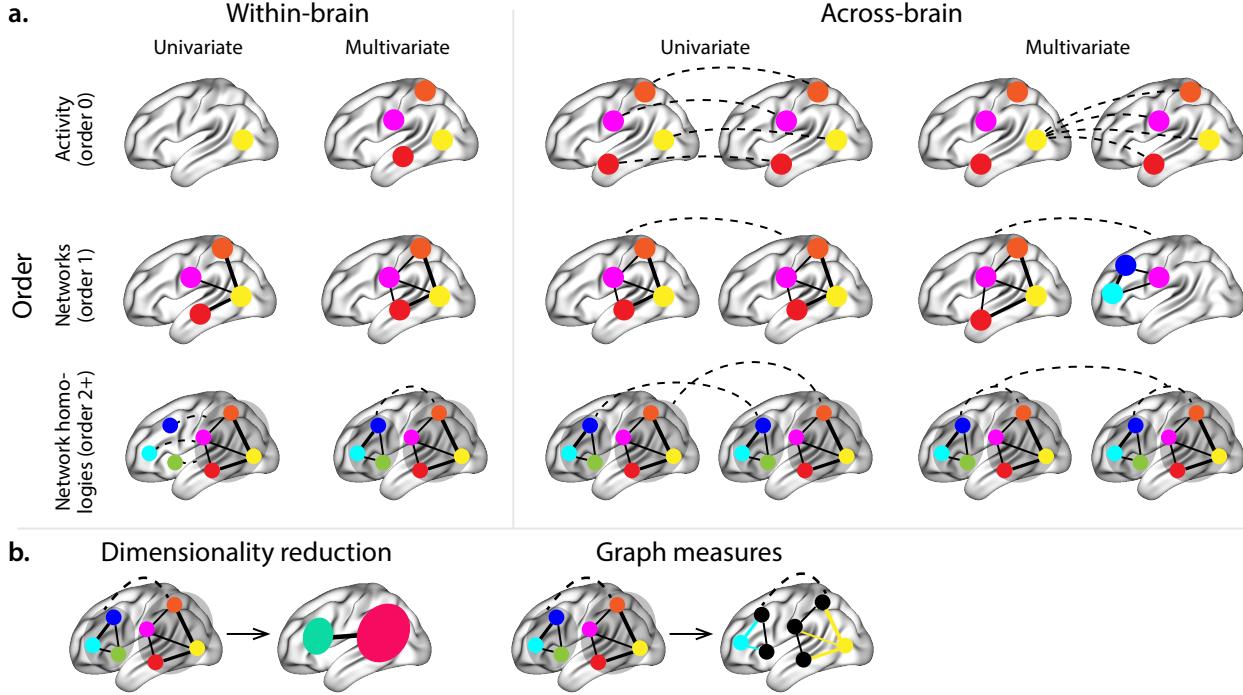


Figure 1: Neural patterns. a. A space of neural analyses Within-brain analyses are carried out within a single brain, whereas across-brain analyses compare neural patterns across two or more individuals' brains. Univariate analyses characterize the activities of individual units (e.g., nodes, small networks, hierarchies of networks, etc.), whereas multivariate analyses characterize the patterns of activities across units. Order 0 patterns involve individual nodes; order 1 patterns involve node-node interactions; order 2 (and higher) patterns relate to interactions between homologous networks. Each of these patterns may be static (e.g., averaging over time) or dynamic. **b. Summarizing neural patterns.** To efficiently compute with complex neural patterns, it can be useful to characterize the patterns using summary measures. Dimensionality reduction algorithms project the patterns onto lower-dimensional spaces whose dimensions reflect weighted combinations of the dimensions in the original space. Graph measures characterize each unit's participation in its associated network.

31 multivariate activity patterns versus patterns of dynamic correlations between neural activity patterns (e.g.,
 32 Fong et al., 2019; Manning et al., 2018). An emerging theme in this literature is that cognition is mediated
 33 by complex dynamic interactions between brain structures (Bassett et al., 2006; Demertzi et al., 2019; Lurie
 34 et al., 2018; Mack et al., 2017; Preti et al., 2017; Solomon et al., 2019; Sporns & Honey, 2006; Turk-Browne,
 35 2013; Zou et al., 2019).

36 Studies of the neural code to date have primarily focused on univariate or multivariate neural pat-
 37 terns (for review see Norman et al., 2006), or (more recently) on patterns of dynamic first-order corre-
 38 lations (i.e., interactions between pairs of brain structures; Demertzi et al., 2019; Fong et al., 2019; Lurie et al.,
 39 2018; Manning et al., 2018; Preti et al., 2017; Zou et al., 2019). We wondered what the future of this line
 40 of work might hold. For example, is the neural code mediated by higher-order interactions between brain
 41 structures (e.g., see Reimann et al., 2017)? Second-order correlations reflect *homologous* patterns of corre-

42 lation. In other words, if the changing patterns of correlations between two regions, A and B , are similar
43 to those between two other regions, C and D , this would be reflected in the second-order correlations be-
44 tween ($A-B$) and ($C-D$). In this way, second-order correlations identify similarities and differences between
45 subgraphs of the brain's connectome. Analogously, third-order correlations reflect homologies between
46 second-order correlations— i.e., homologous patterns of homologous interactions between brain regions.
47 More generally, higher-order correlations reflect homologies between patterns of lower-order correlations.
48 We can then ask: which “orders” of interaction are most reflective of high-level cognitive processes?

49 Another central question pertains to the extent to which the neural code is carried by activity patterns
50 that directly reflect ongoing cognition (e.g., following Haxby et al., 2001; Norman et al., 2006), versus the
51 dynamic properties of the network structure itself, independent of specific activity patterns in any given
52 set of regions (e.g., following Bassett et al., 2006). For example, graph measures such as centrality and
53 degree (Bullmore & Sporns, 2009) may be used to estimate how a given brain structure is “communicating”
54 with other structures, independently of the specific neural representations carried by those structures.
55 If one considers a brain region’s position in the network (e.g., its eigenvector centrality) as a dynamic
56 property, one can compare how the positions of different regions are correlated, and/or how those patterns
57 of correlations change over time. We can also compute higher-order patterns in these correlations to
58 characterize homologous subgraphs in the connectome that display similar changes in their constituent
59 brain structures’ interactions with the rest of the brain.

60 To gain insights into the above aspects of the neural code, we developed a computational framework
61 for estimating dynamic high-order correlations in timeseries data. This framework provides an important
62 advance, in that it enables us to examine patterns in higher-order correlations that are computationally
63 intractable to estimate via conventional methods. Given a multivariate timeseries, our framework pro-
64 vides timepoint-by-timepoint estimates of the first-order correlations, second-order correlations, and so
65 on (up to tenth-order correlations in this manuscript). Our approach combines a kernel-based method
66 for computing dynamic correlations in timeseries data with a dimensionality reduction step (Fig. 1b) that
67 projects the resulting dynamic correlations into a low-dimensional space. We explored two dimensionality
68 reduction approaches: principle components analysis (PCA; Pearson, 1901), which preserves an approxi-
69 mately invertible transformation back to the original data (e.g., this follows related approaches taken by
70 Gonzalez-Castillo et al., 2019; McIntosh & Jirsa, 2019; Toker & Sommer, 2019); and a second non-invertible
71 algorithm that explored patterns in eigenvector centrality (Landau, 1895). This latter approach characterizes
72 correlations between each feature dimension’s relative *position* in the network in favor of the specific activity
73 histories of different features (also see Betzel et al., 2019; Reimann et al., 2017; Sizemore et al., 2018).

74 We validated our approach using synthetic data where the underlying correlations were known. We

75 then applied our framework to a neuroimaging dataset collected as participants listened to either an audio
76 recording of a ten-minute story or a temporally scrambled version of the story, or underwent a resting state
77 scan (Simony et al., 2016). We used a subset of the data to train across-participant classifiers to decode
78 listening times using a blend of neural features (comprising neural activity patterns, as well as different
79 orders of correlations between those patterns that were inferred using our computational framework).
80 We found that both the PCA-based and eigenvector centrality-based approaches yielded neural patterns
81 that could be used to decode accurately. Both approaches also yielded the best decoding accuracy for
82 data collected during (intact) story listening when high-order (PCA: second-order; eigenvector centrality:
83 fourth-order) dynamic correlation patterns were included as features. When we trained classifiers on the
84 scrambled stories or resting state data, only lower-order dynamic patterns were informative to the decoders.
85 Taken together, our results indicate that high-level cognition is supported by high-order dynamic patterns
86 of communication between brain structures.

87 Methods

88 Our general approach to comprises four general steps (Fig. 2). First, we derive a kernel-based approach
89 to computing dynamic pairwise correlations in a T (timepoints) by K (features) multivariate timeseries,
90 \mathbf{X}_0 . This yields a T by $O(K^2)$ matrix of dynamic correlations, \mathbf{Y}_1 , where each row comprises the upper
91 triangle of the correlation matrix at a single timepoint, reshaped into a row vector (this reshaped vector is
92 $(\frac{K^2-K}{2})$ -dimensional). Second, we apply a dimensionality reduction step to project the matrix of dynamic
93 correlations back onto a K -dimensional space. This yields a T by K matrix, \mathbf{X}_1 , that reflects an approximation
94 of the dynamic correlations reflected in the original data. Third, we use repeated applications of the kernel-
95 based dynamic correlation step to \mathbf{X}_n and the dimensionality reduction step to the resulting \mathbf{Y}_{n+1} to estimate
96 high-order dynamic correlations. Each application of these steps to a T by K time series \mathbf{X}_n yields a T by K
97 matrix, \mathbf{X}_{n+1} , that reflects the dynamic correlations between the columns of \mathbf{X}_n . In this way, we refer to n as
98 the *order* of the timeseries, where \mathbf{X}_0 (order 0) denotes the original data and \mathbf{X}_n denotes n^{th} -order dynamic
99 correlations between the columns of \mathbf{X}_0 . Finally, we use a cross-validation-based decoding approach to
100 evaluate how well information contained in a given order (or weighted mixture of orders) may be used
101 to decode relevant cognitive states. If including a given \mathbf{X}_n in the feature set yields higher classification
102 accuracy on held-out data, we interpret this as evidence that the given cognitive states are reflected in
103 patterns of n^{th} -order correlations. All of the code used to produce the figures and results in this manuscript,
104 along with links to the corresponding datasets, may be found at github.com/ContextLab/timecorr-paper. In
105 addition, we have released a Python toolbox for computing dynamic high-order correlations in timeseries

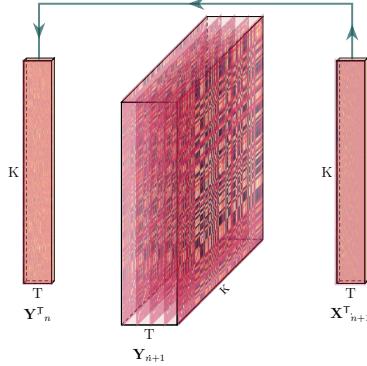


Figure 2: **Estimating dynamic high-order correlations.** Given a T by K matrix of multivariate timeseries data, \mathbf{Y}_n (where $n \in \mathbb{N}, n \geq 0$), we use Equation 5 to compute a timeseries of K by K correlation matrices, \mathbf{Y}_{n+1} . We then approximate \mathbf{Y}_{n+1} with the T by K matrix \mathbf{X}_{n+1} . This process may be repeated to scalably estimate iteratively higher-order correlations in the data. Note that the transposes of \mathbf{Y}_n and \mathbf{X}_{n+1} are displayed in the figure for compactness.

106 data; our toolbox may be found at timecorr.readthedocs.io. **JRM NOTE: CHECK LINK**

107 Kernel-based approach for computing dynamic correlations

Given a matrix of observations, we can compute the (static) Pearson's correlation between any pair of columns, $\mathbf{X}(\cdot, i)$ and $\mathbf{X}(\cdot, j)$ using (Pearson, 1901):

$$\text{corr}(\mathbf{X}(\cdot, i), \mathbf{X}(\cdot, j)) = \frac{\sum_{t=1}^T (\mathbf{X}(t, i) - \bar{\mathbf{X}}(\cdot, i))(\mathbf{X}(t, j) - \bar{\mathbf{X}}(\cdot, j))}{\sqrt{\sum_{t=1}^T \sigma_{\mathbf{X}(\cdot, i)}^2 \sigma_{\mathbf{X}(\cdot, j)}^2}}, \text{ where} \quad (1)$$

$$\bar{\mathbf{X}}(\cdot, k) = \sum_{t=1}^T \mathbf{X}(t, k), \text{ and} \quad (2)$$

$$\sigma_{\mathbf{X}(\cdot, k)}^2 = \sum_{t=1}^T (\mathbf{X}(t, k) - \bar{\mathbf{X}}(\cdot, k))^2 \quad (3)$$

108 We can generalize this formula to compute time-varying correlations by incorporating a *kernel function* that
109 takes a time t as input, and returns how much the observed data at each timepoint $\tau \in [-\infty, \infty]$ contributes
110 to the estimated instantaneous correlation at time t (Fig. 3; also see Allen et al., 2012, for a similar approach).

111

Given a kernel function $\kappa_t(\cdot)$ for timepoint t , evaluated at timepoints $\tau \in [1, \dots, T]$, we can update the

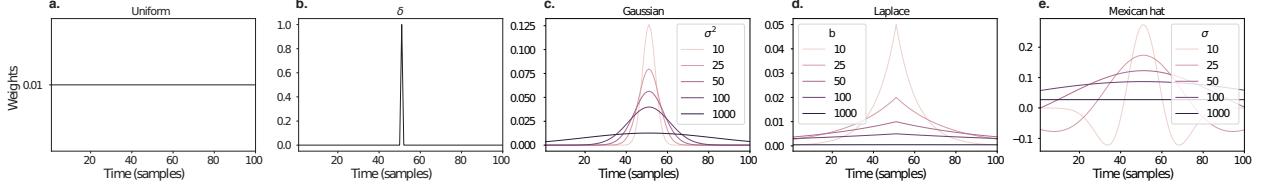


Figure 3: **Examples of kernel functions.** Each panel displays per-timepoint weights at $t = 50$, evaluated for 100 timepoints ($\tau \in [1, \dots, 100]$). **a. Uniform kernel.** The weights are timepoint-invariant; observations at all timepoints are weighted equally, and do not change as a function of τ . This is a special case of weight function that reduces dynamic correlations to static correlations. **b. Dirac δ kernel.** Only the observation at timepoint t is given a non-zero weight (of 1). **c. Gaussian kernels.** Each kernel's weights fall off in time according to a Gaussian probability density function centered on time t . Weights derived using several different example width parameters (σ^2) are displayed. **d. Laplace kernels.** Each kernel's weights fall off in time according to a Laplace probability density function centered on time t . Weights derived using several different example width parameters (b) are displayed. **e. Mexican hat (Ricker wavelet) kernels.** Each kernel's weights fall off in time according to a Ricker wavelet centered on time t . This function highlights the *contrasts* between local versus surrounding activity patterns in estimating dynamic correlations. Weights derived using several different example width parameters (σ) are displayed.

static correlation formula in Equation 2 to estimate the *instantaneous correlation* at timepoint t :

$$\text{timecorr}_{\kappa_t}(\mathbf{X}(\cdot, i), \mathbf{X}(\cdot, j)) = \frac{\sum_{\tau=1}^T (\mathbf{X}(\tau, i) - \tilde{\mathbf{X}}_{\kappa_t}(i))(\mathbf{X}(\tau, j) - \tilde{\mathbf{X}}_{\kappa_t}(j))}{\sqrt{\sum_{\tau=1}^T \tilde{\sigma}_{\kappa_t}^2(\mathbf{X}(\cdot, i))\tilde{\sigma}_{\kappa_t}^2(\mathbf{X}(\cdot, j))}}, \text{ where} \quad (4)$$

$$\tilde{\mathbf{X}}_{\kappa_t}(t, k) = \sum_{\tau=1}^T \kappa_t(\tau, k) \mathbf{X}(\tau, k), \quad (5)$$

$$\tilde{\sigma}_{\kappa_t}^2(\mathbf{X}(\cdot, k)) = \sum_{\tau=1}^T (\mathbf{X}(\tau, k) - \tilde{\mathbf{X}}_{\kappa_t}(t, k))^2. \quad (6)$$

112 Here $\text{timecorr}(\mathbf{X}(\cdot, i), \mathbf{X}(\cdot, j), \kappa_t)$ reflects the correlation at time t between columns i and j of \mathbf{X} , estimated using
 113 the kernel κ_t . We evaluate Equation 5 in turn each pair of columns in \mathbf{X} and for kernels centered on each
 114 timepoint in the timeseries, respectively, to obtain a T by K by K timeseries of dynamic correlations, \mathbf{Y} . For
 115 convenience, we then reshape the upper triangles of each timepoint's correlation matrix into a row vector
 116 to obtain an equivalent T by $\frac{K^2-K}{2}$ matrix.

117 **Dynamic inter-subject functional connectivity (DISFC)**

Equation 5 provides a means of taking a single observation matrix, \mathbf{X}_n and estimating the dynamic correlations from moment to moment, \mathbf{Y}_{n+1} . Suppose that one has access to a set of multiple observation matrices that reflect the same phenomenon. For example, one might collect neuroimaging data from several experimental participants, as each participant performs the same task (or sequence of tasks). Let $\mathbf{X}_n^1, \mathbf{X}_n^2, \dots, \mathbf{X}_n^P$ reflect the T by K observation matrices ($n = 0$) or reduced correlation matrices ($n > 0$) for each of

P participants in an experiment. We can use *inter-subject functional connectivity* (ISFC; Simony et al., 2016) to compute the degree of stimulus-driven correlations reflected in the multi-participant dataset at a given timepoint t using:

$$\bar{\mathbf{C}}(t) = M \left(R \left(\frac{1}{2P} \sum_{p=1}^P Z(Y_n^p(t))^T + Z(Y_n^p(t)) \right) \right), \quad (7)$$

where M extracts and vectorizes the diagonal and upper triangle of a symmetric matrix, Z is the Fisher z-transformation (Zar, 2010):

$$Z(r) = \frac{\log(1+r) - \log(1-r)}{2} \quad (8)$$

R is the inverse of Z :

$$R(z) = \frac{\exp(2z-1)}{\exp(2z+1)}, \quad (9)$$

and $\mathbf{Y}_n^p(t)$ denotes the correlation matrix (Eqn. 2) between each column of \mathbf{X}_n^p and each column of the average \mathbf{X}_n from all *other* participants, $\bar{\mathbf{X}}_n^{np}$:

$$\bar{\mathbf{X}}_n^{np} = R \left(\frac{1}{P-1} \sum_{q \in \setminus p} Z(\mathbf{X}_n^q) \right), \quad (10)$$

where $\setminus p$ denotes the set of all participants other than participant p . In this way, the T by $\frac{K^2-K}{2}$ DISFC matrix $\bar{\mathbf{C}}$ provides a time-varying extension of the ISFC approach developed by Simony et al. (2016).

120 Low-dimensional representations of dynamic correlations

Given a T by $\frac{K^2-K}{2}$ matrix of dynamic correlations, \mathbf{Y}_n , we propose two general approaches to computing a T by K low-dimensional representation of these correlations, \mathbf{X}_n . The first approach uses dimensionality reduction algorithms to project \mathbf{Y}_n onto a K -dimensional space. The second approach uses graph measures to characterize the relative positions of each feature ($k \in [1, \dots, K]$) in the network defined by the correlation matrix at each timepoint.

126 Dimensionality reduction-based approaches to computing \mathbf{X}_n

The modern library of dimensionality reduction algorithms include Principal Components Analysis (PCA; Pearson, 1901), Probabilistic PCA (PPCA; Tipping & Bishop, 1999), Exploratory Factor Analysis (EFA;

129 Spearman, 1904), Independent Components Analysis (ICA; Comon et al., 1991; Jutten & Herault, 1991),
130 *t*-Stochastic Neighbor Embedding (*t*-SNE; van der Maaten & Hinton, 2008), Uniform Manifold Approximation and Projection (UMAP; McInnes & Healy, 2018), non-negative matrix factorization (NMF; Lee &
131 Seung, 1999), Topographic Factor Analysis (TFA) Manning et al. (2014), Hierarchical Topographic Factor
132 analysis (HTFA) Manning et al. (2018), Topographic Latent Source Analysis (TLSA) Gershman et al. (2011),
133 Dictionary learning (J. Mairal et al., 2009; J. B. Mairal et al., 2009), deep auto-encoders (Hinton & Salakhutdinov, 2006), among others. While complete characterizations of each of these algorithms is beyond the
134 scope of the present manuscript, the general intuition driving these approaches is to compute the T by I
135 matrix, \mathbf{X} , that is closest to the original T by J matrix, \mathbf{Y} , where (typically) $I \ll J$. The different approaches
136 place different constraints on what properties \mathbf{X} must satisfy and which aspects of the data are compared
137 (and how) to characterize the match between \mathbf{X} and \mathbf{Y} .

140 Applying dimensionality reduction algorithms to \mathbf{Y} yields a \mathbf{X} whose columns reflect weighted combinations
141 (or nonlinear transformations) of the original columns of \mathbf{Y} . This has two main consequences. First,
142 with each repeated dimensionality reduction, the resulting \mathbf{X}_n has lower and lower fidelity (with respect to
143 what the “true” \mathbf{Y}_n might have looked like without using dimensionality reduction to maintain scalability).
144 In other words, computing \mathbf{X}_n is a lossy operation. Second, whereas each columns of \mathbf{Y}_n may always be
145 mapped directly onto specific pairs of columns of \mathbf{Y}_{n-1} , the columns of \mathbf{X}_n reflect weighted combinations
146 and/or nonlinear transformations of the columns of \mathbf{Y}_n . Many dimensionality reduction algorithms are
147 invertible (or approximately invertible). However, attempting to map a given \mathbf{X}_n back onto the original
148 feature space of \mathbf{Y}_0 will usually require $O(TK^{2n})$ space and therefore quickly becomes intractable as n or K
149 grow large.

150 **Graph measure approaches to computing \mathbf{X}_n**

151 The above dimensionality reduction approaches to approximating a given \mathbf{Y}_n with a lower-dimensional
152 \mathbf{X}_n preserve a (potentially recombined and transformed) mapping back to the original data in \mathbf{Y}_0 . We
153 also explore graph measure approaches that forgo a preserved mapping back to the original data in favor
154 of preserving each feature’s relative *position* in the broader network of interactions and connections. To
155 illustrate the distinction between the two general approaches we explore, suppose a network comprises
156 nodes A , B , and C . If A and B exhibit uncorrelated activity patterns, the functional connection between
157 them will be (by definition) close to 0. However, if A and B each interact with C in similar ways, we might
158 attempt to capture those similarities using a measure that reflects the how A and B interact in the network.
159 In general, graph measures take as input a matrix of interactions (e.g., using the above notation, an K

160 by K correlation matrix or binarized correlation matrix reconstituted from a single timepoint's row of \mathbf{Y})
161 and return as output a set of K measures describing how each node (feature) sits within that correlation
162 matrix with respect to the rest of the population. Widely used measures include betweenness centrality (the
163 proportion of shortest paths between each pair of nodes in the population that involves the given node in
164 question; e.g., Barthélemy, 2004; Freeman, 1977; Geisberger et al., 2008; Newman, 2005; Opsahl et al., 2010);
165 diversity and dissimilarity (characterizations of how differently connected a given node is from others in
166 the population; e.g., Lin, 2009; Rao, 1982; Ricotta & Szeidl, 2006); Eigenvector centrality and pagerank
167 centrality (measures of how influential a given node is within the broader network; e.g., Bonacich, 2007;
168 Halu et al., 2013; Lohmann et al., 2010; Newman, 2008); transfer entropy and flow coefficients (a measure
169 of how much information is flowing from a given node to other nodes in the network; e.g., Honey et
170 al., 2007; Schreiber, 2000); k -coreness centrality (a measure of the connectivity of a node within its local
171 sub-graph; e.g., Alvarez-Hamelin et al., 2005; Christakis & Fowler, 2010); within-module degree (a measure
172 of how many connections a node has to its close neighbors in the network; e.g., Rubinov & Sporns, 2010);
173 participation coefficient (a measure of the diversity of a node's connections to different sub-graphs in the
174 network; e.g., Rubinov & Sporns, 2010); and sub-graph centrality (a measure of a node's participation in
175 all of the network's sub-graphs; e.g., Estrada & Rodríguez-Velázquez, 2005); among others.

176 For a given graph measure, $\eta : \mathbb{R}^{K \times K} \rightarrow \mathbb{R}^K$, we can use η to transform each row of \mathbf{Y}_n in a way that
177 characterizes the corresponding graph properties of each column. This results in a new T by K matrix, \mathbf{X}_n ,
178 that reflects how the features reflected in the columns of \mathbf{Y}_n participate in the network during each timepoint
179 (row).

180 Dynamic higher-order correlations

181 Because \mathbf{X}_n has the same shape as the original data \mathbf{X}_0 , approximating \mathbf{Y}_n with a lower-dimensional \mathbf{X}_n
182 enables us to estimate high-order dynamic correlations in a scalable way. Given a T by K input matrix, the
183 output of Equation 5 requires $O(TK^2)$ space to store. Repeated applications of Equation 5 (i.e., computing
184 dynamic correlations between the columns of the outputted dynamic correlation matrix) each require
185 exponentially more space; in general the n^{th} -order dynamic correlations of a T by K timeseries occupies
186 $O(TK^{2n})$ space. However, when we approximate or summarize the output of Equation 5 with a T by K matrix
187 (as described above), it becomes feasible to compute even very high-order correlations in high-dimensional
188 data. Specifically, approximating the n^{th} -order dynamic correlations of a T by K timeseries requires only
189 $O(TK^2)$ additional space— the same as would be required to compute first-order dynamic correlations. In
190 other words, the space required to store $n + 1$ multivariate timeseries reflecting up to n^{th} order correlations

191 in the original data scales linearly with n using our approach (Fig. 2).

192 Data

193 We examined two types of data: synthetic data and human functional neuroimaging data. We constructed
194 and leveraged the synthetic data to evaluate our general approach (for a related approach, see Thompson et
195 al., 2018). Specifically, we tested how well Equation 5 could be used to recover known dynamic correlations
196 using different choices of kernel (κ ; Fig. 3), for each of several synthetic datasets that exhibited different tem-
197 poral properties. We applied our approach to a functional neuroimaging dataset to test the hypothesis that
198 ongoing cognitive processing is reflected in high-order dynamic correlations. We used an across-participant
199 classification test to estimate whether dynamic correlations of different orders contain information about
200 which timepoint in a story participants were listening to.

201 Synthetic data

202 We constructed a total of 40 multivariate timeseries, collectively reflecting a total of 4 different patterns of
203 dynamic correlations (i.e., 10 datasets reflecting each type of dynamic pattern). Each timeseries comprised
204 50 features (dimensions) that varied over 300 timepoints. The observations at each timepoint were drawn
205 from a zero-mean multivariate Gaussian distribution with a covariance matrix defined for each timepoint
206 as described below. We drew the observations at each timepoint independently from the draws at all other
207 timepoints; in other words, for each observation $s_t \sim \mathcal{N}(\mathbf{0}, \Sigma_t)$ at timepoint t , $p(s_t) = p(s_t|p_{\setminus t})$.

Constant. We generated data with stable underlying correlations to evaluate how Equation 5 characterized correlation “dynamics” when the ground truth correlations were static. We constructed 10 multivariate timeseries, whose observations were each drawn from a single (stable) Gaussian distribution. For each dataset, we constructed a random covariance matrix, Σ_m :

$$\mathbf{C}(i, j) \sim \mathcal{N}(0, 1) \quad (11)$$

$$\Sigma_m = \mathbf{C}\mathbf{C}^\top, \text{ where} \quad (12)$$

208 $i, j \in [1, 2, \dots, 50]$. In other words, all of the observations (for each of the 300 timepoints) within each dataset
209 were drawn from a multivariate Gaussian distribution with the same covariance matrix, and the 10 datasets
210 each used a different covariance matrix.

211 **Random.** We generated a second set of 10 synthetic datasets whose observations at each timepoint were
212 drawn from a Gaussian distribution with a new randomly constructed (using Eqn. 12) covariance matrix.
213 Because each timepoint’s covariance matrix was drawn independently of the covariance matrices for all
214 other timepoints, these datasets provided a test of reconstruction accuracy in the absence of any meaningful
215 underlying temporal structure in the dynamic correlations underlying the data.

Ramping. We generated a third set of 10 synthetic datasets whose underlying correlations changed gradually over time. For each dataset, we constructed two *anchor* correlation matrices using Equation 12, Σ_{start} and Σ_{end} . For each of the 300 timepoints in each dataset, we drew the observations from a multivariate Gaussian distribution whose covariance matrix at each timepoint $t \in [0, \dots, 299]$ was given by

$$\Sigma_t = \left(1 - \frac{1-t}{299}\right)\Sigma_{\text{start}} + \frac{t}{299}\Sigma_{\text{end}}. \quad (13)$$

216 The gradually changing correlations underlying these datasets allow us to evaluate the recovery of dynamic
217 correlations when each timepoint’s correlation matrix is unique (as in the random datasets), but where the
218 correlation dynamics are structured.

219 **Event.** We generated a fourth set of 10 synthetic datasets whose underlying correlation matrices exhibited
220 prolonged intervals of stability, interspersed with abrupt changes. For each dataset, we used Equation 12
221 to generate 5 random covariance matrices. We constructed a timeseries where each set of 60 consecutive
222 samples was drawn from a Gaussian with the same covariance matrix. These datasets were intended to
223 simulate a system that undergoes occasional abrupt state changes.

224 Functional neuroimaging data collected during story listening

225 We examined an fMRI dataset collected by Simony et al. (2016) that the authors have made publicly available
226 at arks.princeton.edu/ark:/88435/dsp015d86p269k. The dataset comprises neuroimaging data collected as
227 participants listened to an audio recording of a story (intact condition; 36 participants), listened to time
228 scrambled recordings of the same story (17 participants in the paragraph-scrambled condition listened to
229 the paragraphs in a randomized order and 36 in the word-scrambled condition listened to the words in a
230 randomized order), or lay resting with their eyes open in the scanner (rest condition; 36 participants). Full
231 neuroimaging details may be found in the original paper for which the data were collected (Simony et al.,
232 2016).

233 **Hierarchical topographic factor analysis (HTFA).** Following our prior related work, we used HTFA (Manning et al., 2018) to derive a compact representation of the data. In brief, this approach approximates the
234 timeseries of voxel activations (44,415 voxels) using a much smaller number of radial basis function (RBF)
235 nodes (in this case 700 nodes). This provides a convenient representation for examining full-brain network
236 dynamics. All of the analyses we carried out on the neuroimaging dataset were performed in this lower-
237 dimensional space. In other words, each participant's data matrix, Y_0 , was a number-of-timepoints by 700
238 matrix of HTFA-derived factor weights (where the row and column labels were matched across partici-
239 pants). Code for carrying out HTFA on fMRI data may be found as part of the BrainIAK toolbox (Capota et
240 al., 2017), which may be downloaded at brainiak.org.

242 Temporal decoding

243 We sought to identify neural patterns that reflected participants' ongoing cognitive processing of incoming
244 stimulus information. As reviewed by Simony et al. (2016), one way of homing in on these stimulus-driven
245 neural patterns is to compare activity patterns across individuals (e.g., using ISFC analyses). In particular,
246 neural patterns will be similar across individuals, to the extent that the neural patterns under consideration
247 are stimulus driven, and to the extent that the corresponding cognitive representations are reflected in similar
248 spatial patterns across people. Following this logic, we used an across-participants temporal decoding test
249 developed by Manning et al. (2018) to assess the degree to which different neural patterns reflected ongoing
250 stimulus-driven cognitive processing across people. The approach entails using a subset of the data to
251 train a classifier to decode which stimulus timepoint (i.e., moment in the story participants listened to). We
252 use decoding (forward inference) accuracy on held-out data, from held-out participants, as a proxy for the
253 extent to which the inputted neural patterns reflected stimulus-driven cognitive processing in a similar way
254 across individuals.

255 Forward inference and decoding accuracy

256 We used an across-participants correlation-based classifier to decode which stimulus timepoint matched a
257 given neural pattern. We first divided the participants into two groups: a template group, $\mathcal{G}_{\text{template}}$, and a
258 to-be-decoded group, $\mathcal{G}_{\text{decode}}$. We used Equation 7 to compute a DISFC matrix for each group ($\bar{\mathbf{C}}_{\text{template}}$ and
259 $\bar{\mathbf{C}}_{\text{decode}}$, respectively). We then correlated the rows of $\bar{\mathbf{C}}_{\text{template}}$ and $\bar{\mathbf{C}}_{\text{decode}}$ to form a number-of-timepoints by
260 number-of-timepoints decoding matrix, Λ . In this way, the rows of Λ reflected timepoints from the template
261 group, while the columns reflected timepoints from the to-be-decoded group. We assigned temporal labels
262 to each row $\bar{\mathbf{C}}_{\text{decode}}$ using the row of $\bar{\mathbf{C}}_{\text{template}}$ to which it was most highly correlated. We then repeated

263 this decoding procedure, but using $\mathcal{G}_{\text{decode}}$ as the template group and $\mathcal{G}_{\text{template}}$ as the to-be-decoded group.
264 Given the true timepoint labels (for each group), we defined the *decoding accuracy* as the proportion of
265 correctly decoded timepoints, across both groups.

266 **Feature weighting and testing**

267 We sought to examine which types of neural features (i.e., activations, first-order dynamic correlations, and
268 higher-order dynamic correlations) were informative to the temporal decoders. Using the notation above,
269 these features correspond to $\mathbf{Y}_0, \mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3$, and so on (we examined up to tenth order correlations, or \mathbf{X}_{10}).

270 One challenge to fairly evaluating high-order correlations is that if the kernel used in Equation 5 is
271 wider than a single timepoint, each repeated application of the equation will result in further temporal
272 blur. Because our primary assessment metric is temporal decoding accuracy, this unfairly biases against
273 detecting meaningful signal in higher-order correlations (relative to lower-order correlations). We attempted
274 to mitigate temporal blur in estimating each \mathbf{X}_n by using a Dirac δ function kernel (which places all of its
275 mass over a single timepoint; Fig. 3b) to compute each lower-order correlation ($\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_{n-1}$). We then used
276 a (potentially wider, as described below) kernel to compute \mathbf{X}_n from \mathbf{X}_{n-1} . In this way, temporal blurring
277 was applied only in the last step of computing \mathbf{X}_n . We note that, because each \mathbf{X}_n is a low-dimensional
278 representation of the corresponding \mathbf{Y}_n , the higher-order correlations we estimated reflect true correlations
279 in the data with lower-fidelity than estimates of lower-order correlations.

280 After computing each $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_{n-1}$ for each participant, we divided participants into two equally sized
281 groups (± 1 for odd numbers of participants): $\mathcal{G}_{\text{train}}$ and $\mathcal{G}_{\text{test}}$. We then further subdivided $\mathcal{G}_{\text{train}}$ into $\mathcal{G}_{\text{train}_1}$
282 and $\mathcal{G}_{\text{train}_2}$. We then computed Λ temporal correlation matrices for each type of neural feature, using $\mathcal{G}_{\text{train}_1}$
283 and $\mathcal{G}_{\text{train}_2}$. This resulted in $n + 1$ Λ matrices (one for the original timeseries of neural activations, and one
284 for each of n orders of dynamic correlations). Our objective was to find a set of weights of each of these Λ
285 matrices such that the weighted average of the $n + 1$ matrices yielded the highest decoding accuracy. We
286 used quasi-Newton gradient ascent (Nocedal & Wright, 2006), using decoding accuracy as the objective
287 function to be maximized, to find an optimal set of training data-derived weights, $\phi_{0,1,\dots,n}$, where $\sum_{i=0}^n \phi_i = 1$
288 and where $\phi_i \geq 0 \forall i \in [0, 1, \dots, n]$.

289 After estimating an optimal set of weights, we computed a new set of $n + 1$ Λ matrices correlating the
290 DISFC patterns from $\mathcal{G}_{\text{train}}$ and $\mathcal{G}_{\text{test}}$ at each timepoint. We use the resulting decoding accuracy of $\mathcal{G}_{\text{test}}$
291 timepoints to estimate how informative the set up neural features containing up to n^{th} order correlations
292 were.

293 We used a permutation-based procedure to form a stable estimate of decoding accuracy for each set of

294 neural features. In particular, we computed the decoding accuracy for each of 10 random group assignments
295 of $\mathcal{G}_{\text{train}}$ and $\mathcal{G}_{\text{test}}$. We report the mean accuracy (and 95% confidence intervals) for each set of neural features.

296 **Identifying robust decoding results**

297 The temporal decoding procedure we use to estimate which neural features support ongoing cognitive
298 processing is governed by many parameters. For example, Equation 5 requires defining a kernel function,
299 which can take on different shapes and widths. For a fixed set of neural features, each of these parameters
300 can yield different decoding accuracies. Further, the best decoding accuracy for a given timepoint may be
301 reliably achieved by one set of parameters, whereas the best decoding accuracy for another timepoint might
302 be reliably achieved by a different set of parameters, and the best decoding accuracy across *all* timepoints
303 might be reliably achieved by still another different set of parameters. Rather than attempting to maximize
304 decoding accuracy, we sought to discover the trends in the data that were robust to specific classifier
305 parameters choices. Specifically, we sought to characterize how decoding accuracy varied (under different
306 experimental conditions) as a function of which neural features were considered.

307 To identify decoding results that were robust to specific classifier parameter choices, we repeated our
308 decoding analyses that substituted in a variety of kernel shapes and widths for Equation 5. We examined
309 Gaussian (Fig. 3c), Laplace (Fig. 3d), and Mexican Hat (Fig. 3e) kernels, each with widths of 5, 10, 20, and
310 50 samples. We then report the average decoding accuracies across all of these parameter choices. This
311 enabled us to (roughly) factor out performance characteristics that were parameter dependent (within the
312 space of parameters we examined).

313 **Reverse inference**

314 The dynamic patterns we examine comprise high-dimensional correlation patterns at each timepoint. To
315 help interpret the resulting patterns in the context of other studies, we created summary maps by computing
316 the across-timepoint average pairwise correlations at each order of analysis (first order, second order, etc.,
317 up to fifteenth order correlations). We selected the 10 strongest (absolute value) correlations at each order.
318 Each correlation is between the dynamic activity patterns (or patterns of dynamic high-order correlations)
319 measured at two RBF nodes (see *Hierarchical Topographic Factor Analysis*). Therefore, the 10 strongest
320 correlations involved up to 20 RBF nodes. Each RBF defines a spatial function whose activations range
321 from 0 to 1. We thresholded each RBF at 0.999 to construct a map of spherical components that denoted the
322 endpoints of the 10 strongest correlations. We then carried out a meta analysis using Neurosynth (Rubin et
323 al., 2017) to identify the 10 terms most commonly associated with the given map. This resulted in a set of

324 10 terms associated with the average dynamic correlation patterns at each order.

325 Results

326 We sought to understand whether high-level cognition is supported by dynamic patterns of high-order
327 correlations. To that end, we developed a computational framework for estimating the dynamics of high-
328 order correlations in multivariate timeseries data (see *Dynamic inter-subject functional connectivity (DISFC)*
329 and *Dynamic higher-order correlations*). We evaluated the efficacy of this framework at recovering known
330 patterns in several synthetic datasets (see *Synthetic data*). We then applied the framework to a public fMRI
331 dataset collected as participants listened to an auditorily presented story, a temporally scrambled version
332 of the story, or underwent a resting state scan (see *Functional neuroimaging data collected during story listening*).
333 We used the relative decoding accuracies of classifiers trained on different sets of neural features to estimate
334 which types of features reflected ongoing cognitive processing.

335 Recovering known dynamic correlations from synthetic data

336 We generated synthetic datasets that differed in how the underlying correlations changed over time. For
337 each dataset, we applied Equation 5 with a variety of kernel shapes and widths. We assessed how well
338 the true underlying correlations at each timepoint matched the recovered correlations (Fig. 4). For every
339 kernel and dataset we tested, our approach recovered the correlation dynamics we embedded into the data.
340 However, the quality of these recoveries varied across different synthetic datasets in a kernel-dependent
341 way.

342 In general, wide monotonic kernels (Laplace, Gaussian) performed best when the correlations varied
343 gradually from moment-to-moment (Figs. 4a, c, and d). **TODO: Say something about kernel widths within**
344 **a shape.** In the extreme, as the rate of change in correlations approaches 0 (Fig. 4a), an infinitely wide kernel
345 would exactly recover the Pearson's correlation (e.g., compare Eqns. 2 and 5).

346 When the correlation dynamics were unstructured in time (Fig. 4b), a Dirac δ kernel (infinitely narrow)
347 performed best. This is because, when every timepoint's correlations are independent of the correlations in
348 every other timepoint, averaging data over time dilutes the available signal. **TODO: Say something about**
349 **kernel widths within a shape.**

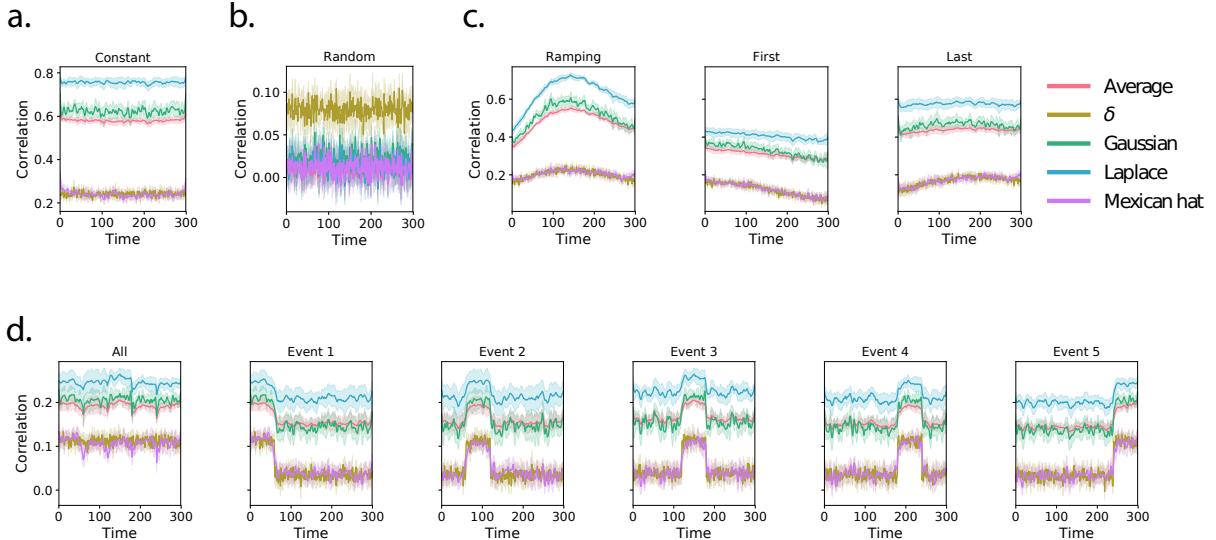


Figure 4: Recovering known dynamic correlations from synthetic data. Each panel displays the average correlations between the vectorized upper triangles of the recovered correlation matrix at each timepoint and either the true underlying correlation at each timepoint or a reference correlation matrix. (The averages are taken across 10 different randomly generated synthetic datasets of the given category.) Error ribbons denote 95% confidence intervals (taken across datasets). Different colors denote different kernel shapes, whereas the shading within each color type denotes kernel widths. For a complete description of each synthetic dataset, see *Synthetic data*. **a. Constant correlations.** These datasets have a stable (unchanging) underlying correlation matrix. **b. Random correlations.** These datasets are generated using a new independently drawn correlation matrix at each new timepoint. **c. Ramping correlations.** These datasets are generated by smoothly varying the underlying correlations between the randomly drawn correlation matrices at first and last timepoints. The left panel displays the correlations between the recovered dynamic correlations and the underlying ground truth correlations. The middle panel compares the recovered correlations with the *first* timepoint’s correlation matrix. The right panel compares the recovered correlations with the *last* timepoint’s correlation matrix. **d. Event-based correlations.** These datasets are each generated using five randomly drawn correlation matrices that each remain stable for a fifth of the total timecourse. The left panel displays the correlations between the recovered dynamic correlations and the underlying ground truth correlations. The right panels compare the recovered correlations with the correlation matrices unique to each event.

350 **Cognitively relevant dynamic high-order correlations in fMRI data**

351 We used across-participant temporal decoders to identify cognitively relevant neural patterns in fMRI data
352 (see *Forward inference and decoding accuracy*). The dataset we examined (collected by Simony et al., 2016)
353 comprised four experimental conditions that exposed participants to stimuli that varied systematically in
354 how cognitively engaging they were. The *intact* experimental condition had participants listen to an audio
355 recording of a 10-minute story. The *paragraph*-scrambled experimental condition had participants listen to a
356 temporally scrambled version of the story, where the paragraphs occurred out of order (but where the same
357 total set of paragraphs were presented over the full listening interval). All participants in this condition
358 experienced the scrambled paragraphs in the same order. The *word*-scrambled experimental condition had
359 participants listen to a temporally scrambled version of the story where the words in the story occurred in a
360 random order. All participants in the word conditions experienced the scrambled words in the same order.
361 Finally, in a *rest* experimental condition participants lay in the scanner with no overt stimulus, with their
362 eyes open (blinking as needed). This dataset provided a convenient means of testing our hypothesis that
363 different levels of cognitive engagement might be supported by different orders of complex brain activity
364 dynamics.

365 In brief, we computed timeseries of dynamic high-order correlations that were similar across participants
366 in each of two randomly assigned groups: a training group and a test group. We then trained classifiers
367 on the training group's data to match each sample from the test group with a stimulus timepoint. Each
368 classifier comprised a weighted blend of neural patterns that reflected up to n^{th} -order dynamic correlations
369 (see *Feature weighting and testing*). We repeated this process for $n \in \{0, 1, 2, \dots, 10\}$. Our examinations of
370 synthetic data suggested that none of the kernels we examined were "universal" in the sense of optimally
371 recovering underlying correlations regardless of the structure of those correlations. In our analyses of neural
372 data, we therefore averaged our decoding results over a variety of kernel shapes and widths in order to
373 identify results that were robust to specific kernel parameters (also see *Identifying robust decoding results*).

374 Our approach to estimating dynamic high-order correlations requires mapping the high-dimensional
375 feature space of correlations (a T by $O(K^2)$ matrix) onto a lower-dimensional T by K matrix. We carried out
376 two sets of analyses that differed in how this mapping was computed. The first set of analyses used PCA
377 to find a low-dimensional embedding of the original dynamic correlation matrices (Fig. 5a,b). The second
378 set of analyses characterized correlations in dynamics of each feature's eigenvector centrality, but did not
379 preserve the underlying activity dynamics (Fig. 5c,d).

380 Both sets of temporal decoding analyses yielded qualitatively similar results for the auditory (non-
381 rest) conditions of the experiment (Fig. 5: pink, yellow, and teal lines). The highest decoding accuracy

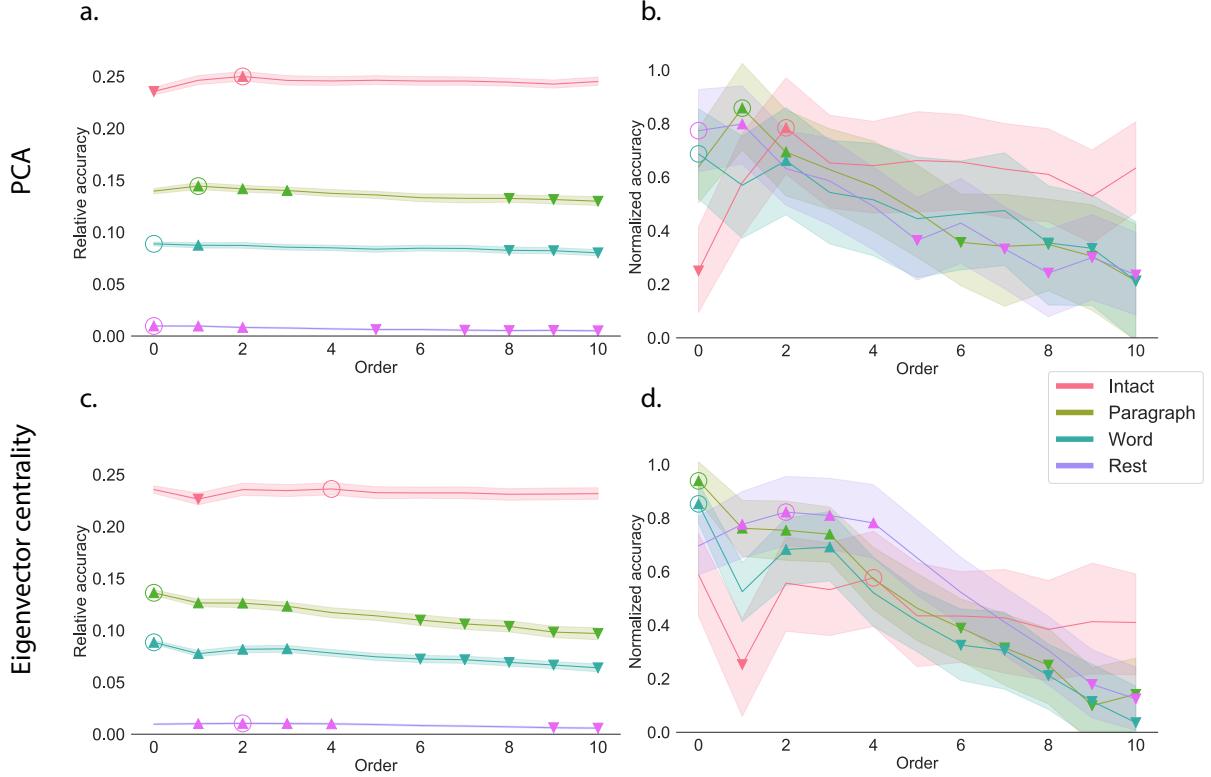


Figure 5: Across-participant decoding accuracy varies with correlation order and cognitive engagement.

a. Decoding accuracy as a function of order: PCA. Order (x-axis) refers to the maximum order of dynamic correlations that were available to the classifiers (see *Feature weighting and testing*). The reported across-participant decoding accuracies are averaged over all kernel shapes and widths (see *Identifying robust decoding results*). The y-values are displayed relative to chance accuracy (intact: $\frac{1}{300}$; paragraph: $\frac{1}{272}$; word: $\frac{1}{300}$; rest: $\frac{1}{400}$). The error bars denote 95% confidence intervals across cross-validation folds (i.e., random assignments of participants to the training and test sets). The colors denote the experimental condition. Arrows denote sets of features that yielded reliably higher (upwards facing) or lower (downward facing) decoding accuracy than the mean of all other features (via a two-tailed test, thresholded at $p < 0.05$). The circled values represent the maximum decoding accuracy within each experimental condition.

b. Normalized decoding accuracy as a function of order: PCA. This panel displays the same results as Panel a, but here each curve has been normalized to have a maximum value of 1 and a minimum value of 0 (including the upper and lower bounds of the respective 95% confidence intervals). Panels a and b used PCA to project each high-dimensional pattern of dynamic correlations onto a lower-dimensional space.

c. Decoding accuracy as a function of order: eigenvector centrality. This panel is in the same format as Panel a, but here eigenvector centrality has been used to project the high-dimensional patterns of dynamic correlations onto a lower-dimensional space.

d. Normalized decoding accuracy as a function of order: eigenvector centrality. This panel is in the same format as Panel b, but here eigenvector centrality has been used to project the high-dimensional patterns of dynamic correlations onto a lower-dimensional space.

for participants who listened to the intact (unscrambled) story was achieved using high-order dynamic correlations (PCA: second-order; eigenvector-centrality: fourth-order). Scrambled versions of the story were best decoded by lower-order correlations (PCA/paragraph: first-order; PCA/word: order zero; eigenvector centrality/paragraph: order zero; eigenvector centrality/word: order zero). The two sets of analyses yielded different decoding results on resting state data (Fig. 5: purple lines). We note could be decoded only very slightly above chance; we speculate that the decoders might have picked up on attentional drift, boredom, or tiredness; we hypothesize these all increase throughout the resting state scan. Our decoders might be picking up on aspects of these loosely defined cognitive states that are common across individuals. The PCA-based approach achieved the highest resting state decoding accuracy using order zero features (non-correlational activation-based), whereas the eigenvector centrality-based approach achieved the highest resting state decoding accuracy using second-order correlations. Taken together, these analyses indicate that high-level cognitive processing (while listening to the intact story) is reflected in the dynamics of high-order correlations in brain activity, whereas lower-level cognitive processing (while listening to scrambled versions of the story that lack rich meaning) is reflected in the dynamics of lower-order correlations and non-correlational activity dynamics. Further, these patterns are associated both with the underlying activity patterns (characterized using PCA) and also with the changing relative positions that different brain areas occupy in their associated networks (characterized using eigenvector centrality).

Having established that patterns of high-order correlations are informative to decoders, we next wondered which specific networks of brain regions contributed most to these patterns. As a representative example, we selected the kernel parameters that yielded decoding accuracies that best matched the average accuracies across all of the kernel parameters we examined. Using Figure 5c as a template, the best-matching kernel was a Laplace kernel with a width of 50 (Fig. 3d). We used this kernel to compute a single K by K n^{th} -order correlation matrix for each experimental condition, along with the Neurosynth (Rubin et al., 2017) terms most highly associated with each of these matrices (Fig. 6; see *Reverse inference*).

For all of the story listening conditions (intact, paragraph, and word), we found that first- and second-order correlations were most strongly associated with auditory and speech processing. During intact story listening, third-order correlations reflected integration with visual areas, and fourth-order correlations reflected integration with areas associated with high-level cognition and cognitive control, such as the ventrolateral prefrontal cortex. However, during listening to temporally scrambled stories, these higher-order correlations instead involved interactions with additional regions associated with speech and semantic processing. By contrast, we found a much different set of patterns in the resting state data. First-order resting state correlations were most strongly associated with regions involved in counting and numerical understanding. Second-order resting state correlations were strongest in visual areas; third-order correlations

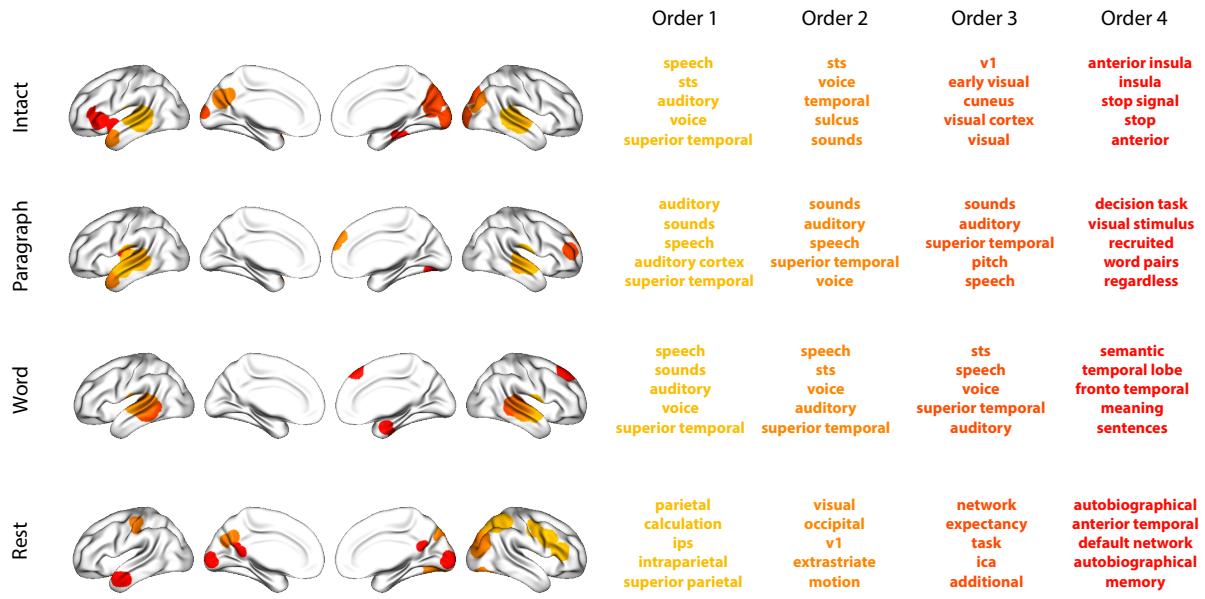


Figure 6: Top terms associated with the endpoints of the strongest correlations. Each color corresponds to one order of correlations, averaged across participants and time. The inflated brain plots display the locations of the endpoints of the 10 strongest (absolute value) correlations at each order, projected onto the cortical surface (Combrisson et al., 2019). The lists of terms on the right display the top five Neurosynth terms (Rubin et al., 2017) decoded from the corresponding brain maps for each order. Each row displays data from a different experimental condition. Additional maps and their corresponding Neurosynth terms may be found in the *Supplementary materials* (intact: Figure S1; paragraph: Figure S2; word: Figure S3; rest: Figure S4).

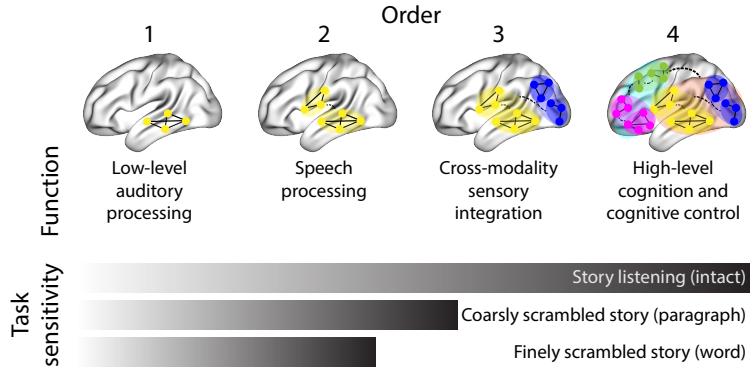


Figure 7: Proposed high-order network dynamics underlying high-level cognition. Higher orders of network interactions support higher-level aspects of cognitive processing. When tasks evoke richer, deeper, and/or higher-level processing, this is reflected in higher-order network interactions.

were strongest in task-positive areas; and fourth-order correlations were strongest in regions associated with autobiographical and episodic memory. We carried out analogous analyses to create maps (and decode the top associated Neurosynth terms) for up to fifteenth-order correlations (Figs. S1, S2, S3, and S4). Of note, examining fifteenth-order correlations between 700 nodes using conventional methods would have required storing roughly $\frac{700^{2 \times 15}}{2} \approx 1.13 \times 10^{85}$ floating point numbers—assuming single-precision (32 bits each), this would require roughly 32 times as many bits as there are molecules in the known universe! Although these fifteenth-order correlations do appear (visually) to have some well-formed structure, we provide this latter example primarily as a demonstration of the efficiency and scalability of our approach.

Discussion

We tested the hypothesis that high-level cognition is supported by high-order brain network dynamics (e.g., see Reimann et al., 2017; Solomon et al., 2019). We examined high-order network dynamics in functional neuroimaging data collected during a story listening experiment. When participants listened to an auditory recording of the story, participants exhibited similar high-order brain network dynamics. By contrast, when participants instead listened to temporally scrambled recordings of the story, only lower-order brain network dynamics were similar across participants. Our results indicate that higher orders of network interactions support higher-level aspects of cognitive processing (Fig. 7).

The notion that cognition is reflected in (and possibly mediated by) patterns of network dynamics has been suggested by or proposed in myriad empirical studies and reviews (e.g., Chang & Glover, 2010; Demertzi et al., 2019; Fong et al., 2019; Gonzalez-Castillo et al., 2019; Liégeois et al., 2019; Lurie et al., 2018; Park et al., 2018; Preti et al., 2017; Roy et al., 2019; Turk-Browne, 2013; Zou et al., 2019). Our study

435 extends this line of work by finding cognitively relevant higher-order network dynamics that reflect ongoing
436 cognition. Our findings complement other work that uses graph theory and topology to characterize how
437 brain networks reconfigure during cognition (e.g., Bassett et al., 2006; Betzel et al., 2019; McIntosh & Jirsa,
438 2019; Reimann et al., 2017; Sizemore et al., 2018; Toker & Sommer, 2019; Zheng et al., 2019).

439 An open question not addressed by our study pertains to how different structures integrate incoming
440 information with different time constants. For example, one line of work suggests that the cortical surface
441 comprises a structured map such that nearby brain structures process incoming information at similar
442 timescales. Low-level sensory areas integrate information relatively quickly, whereas higher-level regions
443 integrate information relatively slowly (Baldassano et al., 2017; Hasson et al., 2015, 2008; Honey et al., 2012;
444 Lerner et al., 2014, 2011). Other related work in human and mouse brains indicates that the temporal
445 response profile of a given brain structure may relate to how strongly connected that structure is with other
446 brain areas (Fallon et al., 2019). Further study is needed to understand the role of temporal integration at
447 different scales of network interaction, and across different anatomical structures.

448 Another potential limitation of our approach relates to recent work suggesting that the brain undergoes
449 rapid state changes, for example across event boundaries (e.g., Baldassano et al., 2017). Shappell et al. (2019)
450 used hidden semi-Markov models to estimate state-specific network dynamics. Our general approach might
451 be extended by considering putative state transitions. For example, rather than weighting all timepoints
452 using a similar kernel (Eqn. 5), the kernel function could adapt on a timepoint-by-timepoint basis such that
453 only timepoints determined to be in the same “state” were given non-zero weight.

454 Identifying higher-order network dynamics associated with high-level cognition required several im-
455 portant methods advances. First, we used kernel-based dynamic correlations to extended the notion of
456 (static) inter-subject functional connectivity (Simony et al., 2016) to a dynamic measure of inter-subject func-
457 tional connectivity (DISFC) that does not rely on sliding windows, and that may be computed at individual
458 timepoints. This allowed us to precisely characterize stimulus-evoked network dynamics that were similar
459 across individuals. Second, we developed a computational framework for efficiently and scalably estimat-
460 ing high-order dynamic correlations. Our approach uses dimensionality reduction algorithms and graph
461 measures to obtain low-dimensional embeddings of patterns of network dynamics. Third, we developed
462 an analysis framework for identifying robust decoding results by carrying out our analyses using a range
463 of parameter values and then identifying which results were robust to specific parameter choices.

464 **Concluding remarks**

465 The complex hierarchy of dynamic interactions that underlie our thoughts is perhaps the greatest mystery in
466 modern science. Methods for characterizing the dynamics of high-order correlations in neural data provides
467 a window into the neural basis of cognition. **Need a good closing sentence.**

468 **Acknowledgements**

469 We acknowledge discussions with Luke Chang, Hany Farid, Paxton Fitzpatrick, Andrew Heusser, Eshin
470 Jolly, Qiang Liu, Matthijs van der Meer, Judith Mildner, Gina Notaro, Stephen Satterthwaite, Emily Whitaker,
471 Weizhen Xie, and Kirsten Ziman. Our work was supported in part by NSF EPSCoR Award Number 1632738
472 to J.R.M. and by a sub-award of DARPA RAM Cooperative Agreement N66001-14-2-4-032 to J.R.M. The
473 content is solely the responsibility of the authors and does not necessarily represent the official views of our
474 supporting organizations.

475 **Author contributions**

476 Concept: J.R.M. Implementation: T.H.C., L.L.W.O., and J.R.M. Analyses: L.L.W.O and J.R.M.

477 **References**

- 478 Allen, E. A., Damaraju, E., Plis, S. M., Erhardt, E. B., Eichele, T., & Calhoun, V. D. (2012). Tracking
479 whole-brain connectivity dynamics in the resting state. *Cerebral Cortex*, 24(3), 663–676.
- 480 Alvarez-Hamelin, I., Dall'Asta, L., Barrat, A., & Vespignani, A. (2005). *k*-corr decomposition: a tool for the
481 visualiztion of large scale networks. *arXiv*, cs/0504107v2.
- 482 Baldassano, C., Chen, J., Zadbood, A., Pillow, J. W., Hasson, U., & Norman, K. A. (2017). Discovering event
483 structure in continuous narrative perception and memory. *Neuron*, 95(3), 709–721.
- 484 Barthélémy, M. (2004). Betweenness centrality in large complex networks. *European Physical Journal B*, 38,
485 163–168.
- 486 Bassett, D., Meyer-Lindenberg, A., Achard, S., Duke, T., & Bullmore, E. (2006, December). Adaptive
487 reconfiguration of fractal small-world human brain functional networks. *Proceedings of the National
488 Academy of Sciences, USA*, 103(51), 19518-23.

- 489 Betzel, R. F., Byrge, L., Esfahlani, F. Z., & Kennedy, D. P. (2019). Temporal fluctuations in the brain's modular
490 architecture during movie-watching. *bioRxiv*, doi.org/10.1101/750919.
- 491 Bonacich, P. (2007). Some unique properties of eigenvector centrality. *Social Networks*, 29(4), 555–564.
- 492 Bullmore, E., & Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and
493 functional systems. *Nature Reviews Neuroscience*, 10(3), 186–198.
- 494 Capota, M., Turek, J., Chen, P.-H., Zhu, X., Manning, J. R., Sundaram, N., ... Shin, Y. S. (2017). *Brain imaging
495 analysis kit*. Retrieved from <https://doi.org/10.5281/zenodo.59780>
- 496 Chang, C., & Glover, G. H. (2010). Time-frequency dynamics of resting-state brain connectivity measured
497 with fMRI. *NeuroImage*, 50, 81–98.
- 498 Christakis, N. A., & Fowler, J. H. (2010). Social network sensors for early detection of contagious outbreaks.
499 *PLoS One*, 5(9), e12948.
- 500 Combrisson, E., Vallat, R., O'Reilly, C., Jas, M., Pascarella, A., l Saive, A., ... Jerbi, K. (2019). Visbrain: a
501 multi-purpose GPU-accelerated open-source suite for multimodal brain data visualization. *Frontiers in
502 Neuroinformatics*, 13(14), 1–14.
- 503 Comon, P., Jutten, C., & Herault, J. (1991). Blind separation of sources, part II: Problems statement. *Signal
504 Processing*, 24(1), 11 - 20.
- 505 Demertzi, A., Tagliazucchi, E., Dehaene, S., Deco, G., Barttfeld, P., Raimondo, F., ... Sitt, J. D. (2019). Human
506 consciousness is supported by dynamic complex patterns of brain signal coordination. *Science Advances*,
507 5(2), eaat7603.
- 508 Estrada, E., & Rodríguez-Velázquez, J. A. (2005). Subgraph centrality in complex networks. *Physical Review
509 E*, 71(5), 056103.
- 510 Etzel, J. A., Gazzola, V., & Keysers, C. (2009). An introduction to anatomical ROI-based fMRI classification.
511 *Brain Research*, 1282, 114–125.
- 512 Fallon, J., Ward, P., Parkes, L., Oldham, S., Arnatkevičiūtė, A., Fornito, A., & Fulcher, B. D. (2019).
513 Timescales of spontaneous activity fluctuations relate to structural connectivity in the brain. *bioRxiv*,
514 doi.org/10.1101/655050.
- 515 Fong, A. H. C., Yoo, K., Rosenberg, M. D., Zhang, S., Li, C.-S. R., Scheinost, D., ... Chun, M. M. (2019).
516 Dynamic functional connectivity during task performance and rest predicts individual differences in
517 attention across studies. *NeuroImage*, 188, 14–25.

- 518 Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry*, 40(1), 35–41.
- 519 Geisberger, R., Sanders, P., & Schultes, D. (2008). Better approximation of betweenness centrality. *Proceedings*
520 *of the meeting on Algorithm Engineering and Experiments*, 90–100.
- 521 Gershman, S., Blei, D., Pereira, F., & Norman, K. (2011). A topographic latent source model for fMRI data.
522 *NeuroImage*, 57, 89–100.
- 523 Gonzalez-Castillo, J., Caballero-Gaudes, C., Topolski, N., Handwerker, D. A., Pereira, F., & Bandettini, P. A.
524 (2019). Imaging the spontaneous flow of thought: distinct periods of cognition contribute to dynamic
525 functional connectivity during rest. *NeuroImage*, 202(116129).
- 526 Halu, A., Mondragón, R. J., Panzarasa, P., & Bianconi, G. (2013). Multiplex PageRank. *PLoS One*, 8(10),
527 e78293.
- 528 Hasson, U., Chen, J., & Honey, C. J. (2015). Hierarchical process memory: memory as an integral component
529 of information processing. *Trends in Cognitive Science*, 19(6), 304-315.
- 530 Hasson, U., Yang, E., Vallines, I., Heeger, D. J., & Rubin, N. (2008). A hierarchy of temporal receptive windows
531 in human cortex. *Journal of Neuroscience*, 28(10), 2539-2550. doi: 10.1523/JNEUROSCI.5487-07.2008
- 532 Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and
533 overlapping representations of faces and objects in ventral temporal cortex. *Science*, 293, 2425–2430.
- 534 Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks.
535 *Science*, 313(5786), 504–507.
- 536 Honey, C. J., Kotter, R., Breakspear, M., & Sporns, O. (2007). Network structure of cerebral cortex shapes
537 functional connectivity on multiple time scales. *Proceedings of the National Academy of Science USA*, 104(24),
538 10240–10245.
- 539 Honey, C. J., Thesen, T., Donner, T. H., Silbert, L. J., Carlson, C. E., Devinsky, O., ... Hasson, U. (2012). Slow
540 cortical dynamics and the accumulation of information over long timescales. *Neuron*, 76, 423-434. doi:
541 10.1016/j.neuron.2012.08.011
- 542 Huth, A. G., de Heer, W. A., Griffiths, T. L., Theunissen, F. E., & Gallant, J. L. (2016). Natural speech reveals
543 the semantic maps that tile human cerebral cortex. *Nature*, 532, 453–458.
- 544 Huth, A. G., Nisimoto, S., Vu, A. T., & Gallant, J. L. (2012). A continuous semantic space describes
545 the representation of thousands of object and action categories across the human brain. *Neuron*, 76(6),
546 1210-1224.

- 547 Jutten, C., & Herault, J. (1991). Blind separation of sources, part I: An adaptive algorithm based on
548 neuromimetic architecture. *Signal Processing*, 24(1), 1–10.
- 549 Kamitani, Y., & Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nature*
550 *Neuroscience*, 8, 679–685.
- 551 Landau, E. (1895). Zur relativen Wertbemessung der Turnierresultate. *Deutsches Wochenschach*, 11, 366–369.
- 552 Lee, D. D., & Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature*,
553 401, 788–791.
- 554 Lerner, Y., Honey, C. J., Katkov, M., & Hasson, U. (2014). Temporal scaling of neural responses to compressed
555 and dilated natural speech. *Journal of Neurophysiology*, 111, 2433–2444.
- 556 Lerner, Y., Honey, C. J., Silbert, L. J., & Hasson, U. (2011). Topographic mapping of a hierarchy of
557 temporal receptive windows using a narrated story. *Journal of Neuroscience*, 31(8), 2906–2915. doi: 10.1523/
558 JNEUROSCI.3684-10.2011
- 559 Liégeois, R., Li, J., Kong, R., Orban, C., De Ville, D. V., Ge, T., ... Yeo, B. T. T. (2019). Resting brain
560 dynamics at different timescales capture distinct aspects of human behavior. *Nature Communications*,
561 10(2317), doi.org/10.1038/s41467-019-10317-7.
- 562 Lin, J. (2009). Divergence measures based on the Shannon entropy. *IEEE Transactions on Information Theory*,
563 37(1), 145–151.
- 564 Lohmann, G., Margulies, D. S., Horstmann, A., Pleger, B., Lepsien, J., Goldhahn, D., ... Turner, R. (2010).
565 Eigenvector centrality mapping for analyzing connectivity patterns in fMRI data of the human brain.
566 *PLoS One*, 5(4), e10232.
- 567 Lurie, D., Kessler, D., Bassett, D., Betzel, R., Breakspear, M., Keilholz, S., ... Calhoun, V. (2018). On the
568 nature of time-varying functional connectivity in resting fMRI. *PsyArXiv*, doi.org/10.31234/osf.io/xtzre.
- 569 Mack, M. L., Preston, A. R., & Love, B. C. (2017). Medial prefrontal cortex compresses concept representations
570 through learning. *bioRxiv*, doi.org/10.1101/178145.
- 571 Mairal, J., Ponce, J., Sapiro, G., Zisserman, A., & Bach, F. R. (2009). Supervised dictionary learning. *Advances*
572 *in Neural Information Processing Systems*, 1033–1040.
- 573 Mairal, J. B., Bach, F., Ponce, J., & Sapiro, G. (2009). Online dictionary learning for sparse coding. *Proceedings*
574 *of the 26th annual international conference on machine learning*, 689–696.

- 575 Manning, J. R., Ranganath, R., Norman, K. A., & Blei, D. M. (2014). Topographic factor analysis: a Bayesian
576 model for inferring brain networks from neural data. *PLoS One*, 9(5), e94914.
- 577 Manning, J. R., Zhu, X., Willke, T. L., Ranganath, R., Stachenfeld, K., Hasson, U., ... Norman, K. A. (2018).
578 A probabilistic approach to discovering dynamic full-brain functional connectivity patterns. *NeuroImage*,
579 180, 243–252.
- 580 McInnes, L., & Healy, J. (2018). UMAP: Uniform manifold approximation and projection for dimension
581 reduction. *arXiv*, 1802(03426).
- 582 McIntosh, A. R., & Jirsa, V. K. (2019). The hidden repertoire of brain dynamics and dysfunction. *Network
583 Neuroscience*, doi.org/10.1162/netn_a_00107.
- 584 Mitchell, T., Shinkareva, S., Carlson, A., Chang, K., Malave, V., Mason, R., & Just, M. (2008). Predicting
585 human brain activity associated with the meanings of nouns. *Science*, 320(5880), 1191.
- 586 Newman, M. E. J. (2005). A measure of betweenness centrality based on random walks. *Social Networks*, 27,
587 39–54.
- 588 Newman, M. E. J. (2008). The mathematics of networks. *The New Palgrave Encyclopedia of Economics*, 2, 1–12.
- 589 Nishimoto, S., Vu, A., Naselaris, T., Benjamini, Y., Yu, B., & Gallant, J. (2011). Reconstructing visual
590 experience from brain activity evoked by natural movies. *Current Biology*, 21, 1 - 6.
- 591 Nocedal, J., & Wright, S. J. (2006). *Numerical optimization*. New York: Springer.
- 592 Norman, K. A., Newman, E., Detre, G., & Polyn, S. M. (2006). How inhibitory oscillations can train neural
593 networks and punish competitors. *Neural Computation*, 18, 1577–1610.
- 594 Opsahl, T., Agneessens, F., & Skvoretz, J. (2010). Node centrality in weighted networks: generalizing degree
595 and shortest paths. *Social Networks*, 32, 245–251.
- 596 Park, H.-J., Friston, K. J., Pae, C., Park, B., & Razi, A. (2018). Dynamic effective connectivity in resting state
597 fMRI. *NeuroImage*, 180, 594–608.
- 598 Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. *The London, Edinburgh,
599 and Dublin Philosophical Magazine and Journal of Science*, 2, 559-572.
- 600 Pereira, F., Lou, B., Pritchett, B., Ritter, S., Gershman, S. J., Kanwisher, N., ... Fedorenko, E. (2018). Toward
601 a universal decoder of linguistic meaning from brain activation. *Nature Communications*, 9(963), 1–13.

- 602 Preti, M. G., Bolton, T. A. W., & Van De Ville, D. (2017). The dynamic functional connectome: state-of-the-art
603 and perspectives. *NeuroImage*, 160, 41–54.
- 604 Rao, C. R. (1982). Diversity and dissimilarity coefficients: a unified approach. *Theoretical Population Biology*,
605 21(1), 24–43.
- 606 Reimann, M. W., Nolte, M., Scolamiero, M., Turner, K., Perin, R., Chindemi, G., ... Markram, H. (2017).
607 Cliques of neurons bound into cavities provide a missing link between structure and function. *Frontiers*
608 in *Computational Neuroscience*, 11(48), 1–16.
- 609 Ricotta, C., & Szeidl, L. (2006). Towards a unifying approach to diversity measures: Bridging the gap
610 between the Shannon entropy and Rao's quadratic index. *Theoretical Population Biology*, 70(3), 237–243.
- 611 Roy, D. S., Park, Y.-G., Ogawa, S. K., Cho, J. H., Choi, H., Kamensky, L., ... Tonegawa, S. (2019). Brain-
612 wide mapping of contextual fear memory engram ensembles supports the dispersed engram complex
613 hypothesis. *bioRxiv*, doi.org/10.1101/668483.
- 614 Rubin, T. N., Kyoejo, O., Gorgolewski, K. J., Jones, M. N., Poldrack, R. A., & Yarkoni, T. (2017). Decoding
615 brain activity using a large-scale probabilistic functional-anatomical atlas of human cognition. *PLoS*
616 *Computational Biology*, 13(10), e1005649.
- 617 Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpreta-
618 tions. *NeuroImage*, 52, 1059–1069.
- 619 Schreiber, T. (2000). Measuring information transfer. *Physical Review Letters*, 85(2), 461–464.
- 620 Shappell, H., Caffo, B. S., Pekar, J. J., & Lindquist, M. A. (2019). Improved state change estimation in
621 dynamic functional connectivity using hidden semi-Markov models. *NeuroImage*, 191, 243–257.
- 622 Simony, E., Honey, C. J., Chen, J., & Hasson, U. (2016). Uncovering stimulus-locked network dynamics
623 during narrative comprehension. *Nature Communications*, 7(12141), 1–13.
- 624 Sizemore, A. E., Giusti, C., Kahn, A., Vettel, J. M., Betzel, R. F., & Bassett, D. S. (2018). Cliques and cavities
625 in the human connectome. *Journal of Computational Neuroscience*, 44(1), 115–145.
- 626 Solomon, S. H., Medaglia, J. D., & Thompson-Schill, S. L. (2019). Implementing a concept network model.
627 *Behavior Research Methods*, 51, 1717–1736.
- 628 Spearman, C. (1904). General intelligence, objectively determined and measured. *American Journal of*
629 *Psychology*, 15, 201–292.

- 630 Sporns, O., & Honey, C. J. (2006). Small worlds inside big brains. *Proceedings of the National Academy of*
631 *Science USA*, 103(51), 19219–19220.
- 632 Thompson, W. H., Richter, C. G., Plavén-Sigray, P., & Fransson, P. (2018). Simulations to benchmark
633 time-varying connectivity methods for fMRI. *PLoS Computational Biology*, 14(5), e1006196.
- 634 Tipping, M. E., & Bishop, C. M. (1999). Probabilistic principal component analysis. *Journal of Royal Statistical*
635 *Society, Series B*, 61(3), 611–622.
- 636 Toker, D., & Sommer, F. T. (2019). Information integration in large brain networks. *PLoS Computational*
637 *Biology*, 15(2), e1006807.
- 638 Tong, F., & Pratte, M. S. (2012). Decoding patterns of human brain activity. *Annual Review of Psychology*, 63,
639 483–509.
- 640 Turk-Browne, N. B. (2013). Functional interactions as big data in the human brain. *Science*, 342, 580–584.
- 641 van der Maaten, L. J. P., & Hinton, G. E. (2008). Visualizing high-dimensional data using t-SNE. *Journal of*
642 *Machine Learning Research*, 9, 2579-2605.
- 643 Zar, J. H. (2010). *Biostatistical analysis*. Prentice-Hall/Pearson.
- 644 Zheng, M., Allard, A., Hagmann, P., & Serrano, M. A. (2019). Geometric renormalization unravels self-
645 similarity of the multiscale human connectome. *arXiv*, 1904.11793.
- 646 Zou, Y., Donner, R. V., Marwan, N., Donges, J. F., & Kurths, J. (2019). Complex network approaches to
647 nonlinear time series analysis. *Physics Reports*, 787, 1–97.