

1      High-level cognition during story listening is reflected in  
2      high-order dynamic correlations in neural activity patterns

3      Lucy L. W. Owen<sup>1</sup>, Thomas H. Chang<sup>1,2</sup>, and Jeremy R. Manning<sup>1,†</sup>

<sup>1</sup>Department of Psychological and Brain Sciences,  
Dartmouth College, Hanover, NH

<sup>2</sup>Amazon.com, Seattle, WA

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

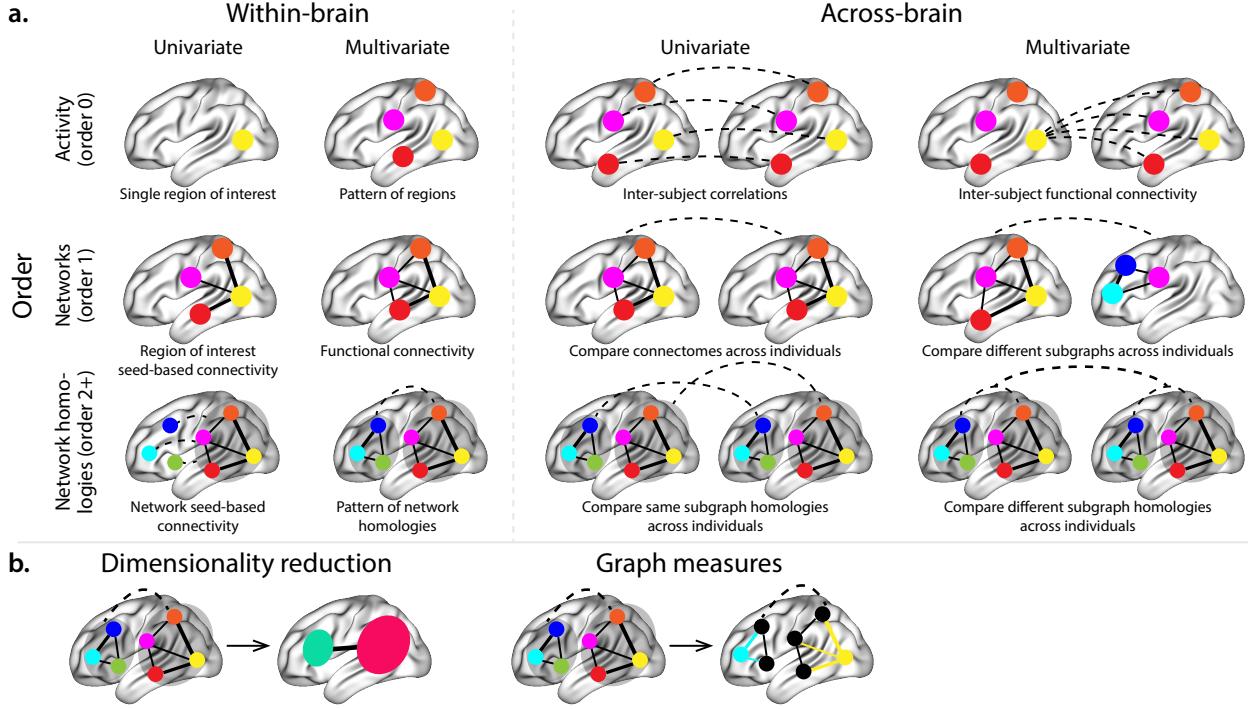
4      November 20, 2020

5      **Abstract**

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

21     **Introduction**

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



**Figure 1: Neural patterns.** **a. A space of neural features.** 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 activity 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 or non-linear transformations of the dimensions in the original space. Graph measures characterize each unit's participation in its associated network.

31 neural representations (e.g., Etzel et al., 2009), or to compare the relative contributions to the neural code of  
 32 multivariate activity patterns versus dynamic correlations between neural activity patterns (e.g., Fong et al.,  
 33 2019; Manning et al., 2018). An emerging theme in this literature is that cognition is mediated by dynamic  
 34 interactions between brain structures (Bassett et al., 2006; Demertzi et al., 2019; Friston, 2000; Grossberg,  
 35 1988; Lurie et al., 2018; Mack et al., 2017; Preti et al., 2017; Solomon et al., 2019; Sporns & Honey, 2006;  
 36 Turk-Browne, 2013; Zou et al., 2019).

37 Studies of the neural code to date have primarily focused on univariate or multivariate neural pat-  
 38 terns (for review see Norman et al., 2006), or (more recently) on patterns of dynamic first-order correla-  
 39 tions (i.e., interactions between pairs of brain structures; Demertzi et al., 2019; Fong et al., 2019; Lurie et al.,  
 40 2018; Manning et al., 2018; Preti et al., 2017; Zou et al., 2019). We wondered what What might the future

41 of this line of work ~~might hold~~. hold? For example, is the neural code ~~mediated by~~ implemented through  
42 higher-order interactions between brain structures (e.g., see Reimann et al., 2017)? Second-order correlations  
43 reflect *homologous* patterns of correlation. In other words, if the dynamic patterns of correlations between  
44 two regions, *A* and *B*, are similar to those between two other regions, *C* and *D*, this would be reflected  
45 in the second-order correlations between (*A*-*B*) and (*C*-*D*). In this way, second-order correlations identify  
46 similarities and differences between subgraphs of the brain’s connectome. Analogously, third-order cor-  
47 relations reflect homologies between second-order correlations—i.e., homologous patterns of homologous  
48 interactions between brain regions. More generally, higher-order correlations reflect homologies between  
49 patterns of lower-order correlations. We can then ask: which “orders” of interaction are most reflective of  
50 high-level cognitive processes?

51 One reason one might expect to see homologous networks in a dataset is related to the notion that  
52 network dynamics reflect ongoing neural computations or cognitive processing (e.g., Beatty et al., 2016). If  
53 the nodes in two brain networks are interacting (within each network) in similar ways then, according to  
54 our characterization of network dynamics, we refer to the similarities between those patterns of interaction  
55 as higher-order correlations. When higher-order correlations are themselves changing over time, we can  
56 also attempt to capture and characterize those high-order dynamics.

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

68 To gain insights into the above aspects of the neural code, we developed a computational framework  
69 for estimating dynamic high-order correlations in timeseries data. This framework provides an important  
70 advance, in that it enables us to examine patterns of higher-order correlations that are computationally  
71 intractable to estimate via conventional methods. Given a multivariate timeseries, our framework provides  
72 timepoint-by-timepoint estimates of the first-order correlations, second-order correlations, and so on. Our  
73 approach combines a kernel-based method for computing dynamic correlations in timeseries data with a di-

74 mensionality reduction step (Fig. 1b) that projects the resulting dynamic correlations into a low-dimensional  
75 space. We explored two dimensionality reduction approaches: principle components analysis (PCA; Pear-  
76 son, 1901), which preserves an approximately invertible transformation back to the original data (e.g., this  
77 follows related approaches taken by Gonzalez-Castillo et al., 2019; McIntosh & Jirsa, 2019; Toker & Sommer,  
78 2019); and a second non-invertible algorithm ~~that explored~~ for computing dynamic patterns in eigenvector  
79 centrality (Landau, 1895). This latter approach characterizes correlations between each feature dimension's  
80 relative *position* in the network in (at each moment in time) in favor of the specific activity histories of  
81 different features (also see Betzel et al., 2019; Reimann et al., 2017; Sizemore et al., 2018).

82 We validated our approach using synthetic data where the underlying correlations were known. We  
83 then applied our framework to a neuroimaging dataset collected as participants listened to either an audio  
84 recording of a ten-minute story, listened to a temporally scrambled version of the story, or underwent  
85 a resting state scan (Simony et al., 2016). Temporal scrambling has been used in a growing number of  
86 studies, largely by Uri Hasson's group, to identify brain regions that are sensitive to higher-order and  
87 longer-timescale information (e.g., cross-sensory integration, rich narrative meaning, complex situations,  
88 etc.) versus regions that are primarily sensitive to low-order (e.g., sensory) information. For example,  
89 Hasson et al. (2008) argues that when brain areas are sensitive to fine versus coarse temporal scrambling,  
90 this indicates that they are "higher order" in the sense that they process contextual information pertaining  
91 to further-away timepoints. By contrast, low-level regions, such as primary sensory cortices, do not  
92 meaningfully change their responses (after correcting for presentation order) even when the stimulus is  
93 scrambled at fine timescales.

94 We used a subset of the story listening and rest data to train across-participant classifiers to decode  
95 listening times (of groups of participants) using a blend of neural features (comprising neural activity  
96 patterns, as well as different orders of dynamic correlations between those patterns that were inferred  
97 using our computational framework). We found that both the PCA-based and eigenvector centrality-based  
98 approaches yielded neural patterns that could be used to decode accurately (i.e., well above chance). Both  
99 approaches also yielded the best decoding accuracy for data collected during (intact) story listening when  
100 high-order (PCA: second-order; eigenvector centrality: fourth-order) dynamic correlation patterns were  
101 included as features. When we trained classifiers on the scrambled stories or resting state data, only  
102 (relatively) lower-order dynamic patterns were informative to the decoders. Taken together, our results  
103 indicate that high-level cognition is supported by high-order dynamic patterns of communication between  
104 brain structures.

105 **Results**

106 We sought to understand whether high-level cognition is ~~supported by reflected in~~ dynamic patterns of high-  
107 order correlations. To that end, we developed a computational framework for estimating the dynamics of  
108 stimulus-driven high-order correlations in multivariate timeseries data (see *Dynamic inter-subject functional*  
109 *connectivity (DISFC)* and *Dynamic higher-order correlations*). We evaluated the efficacy of this framework at  
110 recovering known patterns in several synthetic datasets (see *Synthetic data: simulating dynamic first-order*  
111 *correlations and Synthetic data: simulating dynamic higher-order correlations*). We then applied the framework  
112 to a public fMRI dataset collected as participants listened to an ~~auditorily~~ auditorily presented story,  
113 listened to a temporally scrambled version of the story, or underwent a resting state scan (see *Functional*  
114 *neuroimaging data collected during story listening*). We used the relative decoding accuracies of classifiers  
115 trained on different sets of neural features to estimate which types of features reflected ongoing cognitive  
116 processing.

117 **Recovering known dynamic correlations from synthetic data**

118 **Recovering dynamic first-order correlations**

119 We generated synthetic datasets that differed in how the underlying first-order correlations changed over  
120 time. For each dataset, we applied Equation 4 with a variety of kernel shapes and widths. We assessed how  
121 well the true underlying correlations at each timepoint matched the recovered correlations (Fig. 2). For every  
122 kernel and dataset we tested, our approach recovered the correlation dynamics we embedded into the data.  
123 However, the quality of these recoveries varied across different synthetic datasets in a kernel-dependent  
124 way.

125 In general, wide monotonic kernel shapes (Laplace, Gaussian), and wider kernels (within a shape),  
126 performed best when the correlations varied gradually from moment-to-moment (Figs. 2a, c, and d). In the  
127 extreme, as the rate of change in correlations approaches 0 (Fig. 2a), an infinitely wide kernel would exactly  
128 recover the Pearson's correlation (e.g., compare Eqns. 1 and 4).

129 When the correlation dynamics were unstructured in time (Fig. 2b), a Dirac  $\delta$  kernel (infinitely narrow)  
130 performed best. This is because, when every timepoint's correlations are independent of the correlations at  
131 every other timepoint, averaging data over time dilutes the available signal. Following a similar pattern,  
132 holding kernel shape fixed, narrower kernel parameters better recovered randomly varying correlations.

133 **Recovering dynamic higher-order correlations**

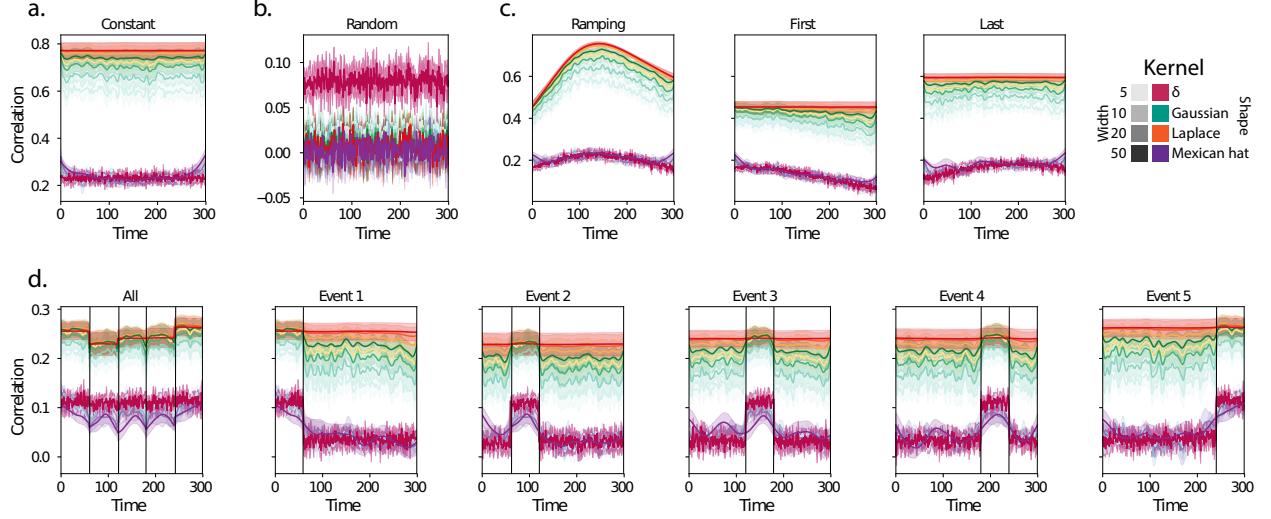


Figure 2: **Recovering known dynamic correlations from synthetic data.** **Recovering known dynamic first-order correlations from synthetic data.** Each panel displays the average correlations between the vectorized upper triangles of the recovered correlation matrix at each timepoint and either the 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, and the shading within each color family denotes the kernel width parameter. 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 the 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. [The vertical lines denote event boundaries.](#)

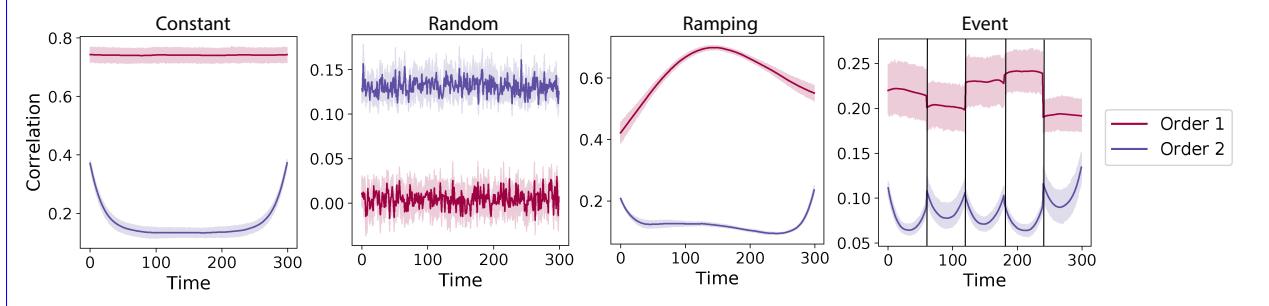


Figure 3: **Recovery of simulated first-order and second-order dynamic correlations.** Each panel displays the average correlations between the vectorized upper triangles of the recovered first-order and second-order correlation matrices and the true (simulated) first-order and second order correlation matrices at each timestep and for each synthetic dataset. (The averages are taken across 10 different randomly generated synthetic datasets of the given category.) Error ribbons denote 95% confidence intervals (taken across datasets). For a complete description of each synthetic dataset, see *Synthetic data: simulating dynamic higher-order correlations*. All estimates represented in this figure were computed using a Laplace kernel (width = 20). **a. Constant correlations.** These datasets have stable (unchanging) underlying second-order correlation matrices. **b. Random correlations.** These datasets are generated using a new independently drawn second-order correlation matrix at each timestep. **c. Ramping correlations.** These datasets are generated by smoothly varying the underlying second-order correlations between the randomly drawn correlation matrices at the first and last timepoints. **d. Event-based correlations.** These datasets are each generated using five randomly drawn second-order correlation matrices that each remain stable for a fifth of the total timecourse. The vertical lines denote event boundaries.

Following our approach to evaluating our ability to recover known dynamic first-order correlations from synthetic data, we generated an analogous second set of synthetic datasets that we designed to exhibit known dynamic first-order *and* second-order correlations (see *Synthetic data: simulating dynamic higher-order correlations*). We generated a total of 40 datasets that varied in how the first-order and second-order correlations changed over time. We then repeatedly applied Equation 4 using the overall best-performing kernel from our first-order tests (a Laplace kernel with a width of 20; Fig. 2) to assess how closely the recovered dynamic correlations matched the dynamic correlations we had embedded into the datasets.

Overall, we found that we could reliably recover both first-order and second-order correlations from the synthetic data (Fig. 3). When the correlations were stable for longer intervals, or changed gradually (constant, ramping, and event datasets), recovery performance was relatively high, and we were better able to recover dynamic first-order correlations than second-order correlations. We expected that this would happen, given that errors in our estimation procedure at lower orders necessarily propagate to higher orders (since lower-order correlations are used to estimate higher-order correlations). Interestingly, we also found that when the correlations were particularly *unstable* (random datasets), we better recovered second-order correlations.

Taken together, our explorations using synthetic data indicated that we are able to partially, but not perfectly, recover ground truth dynamic first-order and second-order correlations. This suggests that our

151 modeling approach provides a meaningful (if noisy) estimate of high-order correlations. We next turned  
152 to analyses of human fMRI data to examine whether the recovered dynamics might reflect the dynamics of  
153 human cognition during a naturalistic story-listening task.

## 154 Cognitively relevant dynamic high-order correlations in fMRI data

155 We used across-participant temporal decoders to identify cognitively relevant neural patterns in fMRI data  
156 (see *Forward inference and decoding accuracy*). The dataset we examined (collected by Simony et al., 2016)  
157 comprised four experimental conditions that exposed participants to stimuli that varied systematically in  
158 how cognitively engaging they were. The *intact* experimental condition had participants listen to an audio  
159 recording of a 10-minute story. The *paragraph*-scrambled experimental condition had participants listen to a  
160 temporally scrambled version of the story, where the paragraphs occurred out of order (but where the same  
161 total set of paragraphs were presented over the full listening interval). All participants in this condition  
162 experienced the scrambled paragraphs in the same order. The *word*-scrambled experimental condition had  
163 participants listen to a temporally scrambled version of the story where the words in the story occurred in a  
164 random order. All participants in the word condition experienced the scrambled words in the same order.  
165 Finally, in a *rest* experimental condition, participants lay in the scanner with no overt stimulus, with their  
166 eyes open (blinking as needed). This [public](#) dataset provided a convenient means of testing our hypothesis  
167 that different levels of cognitive processing and engagement are [supported by reflected in](#) different orders  
168 of brain activity dynamics.

169 In brief, we computed timeseries of dynamic high-order correlations that were similar across participants  
170 in each of two randomly assigned groups: a training group and a test group. We then trained classifiers  
171 on the training group's data to match each sample from the test group with a stimulus timepoint. Each  
172 classifier comprised a weighted blend of neural patterns that reflected up to  $n^{\text{th}}$ -order dynamic correlations  
173 (see *Feature weighting and testing*, Fig. 10). We repeated this process for  $n \in \{0, 1, 2, \dots, 10\}$ . Our examinations  
174 of synthetic data suggested that none of the kernels we examined were “universal” in the sense of optimally  
175 recovering underlying correlations regardless of the temporal structure of those correlations. We found a  
176 similar pattern in the (real) fMRI data, whereby different kernels yielded different decoding accuracies, but  
177 no single kernel emerged as the clear “best.” In our analyses of neural data, we therefore averaged our  
178 decoding results over a variety of kernel shapes and widths in order to identify results that were robust to  
179 specific kernel parameters (see *Identifying robust decoding results*).

180 Our approach to estimating dynamic high-order correlations entails mapping the high-dimensional  
181 feature space of correlations ([represented by](#) a  $T$  by  $O(K^2)$  matrix) onto a lower-dimensional [feature space](#)

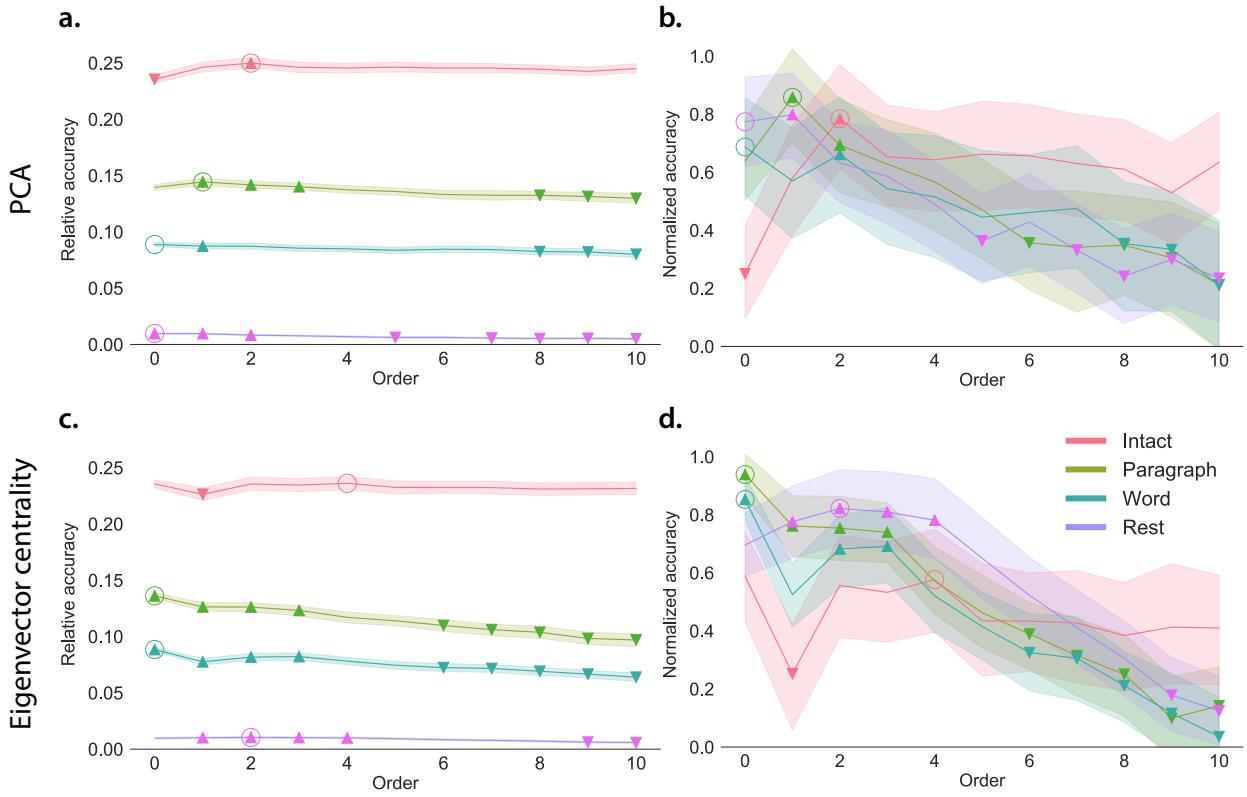
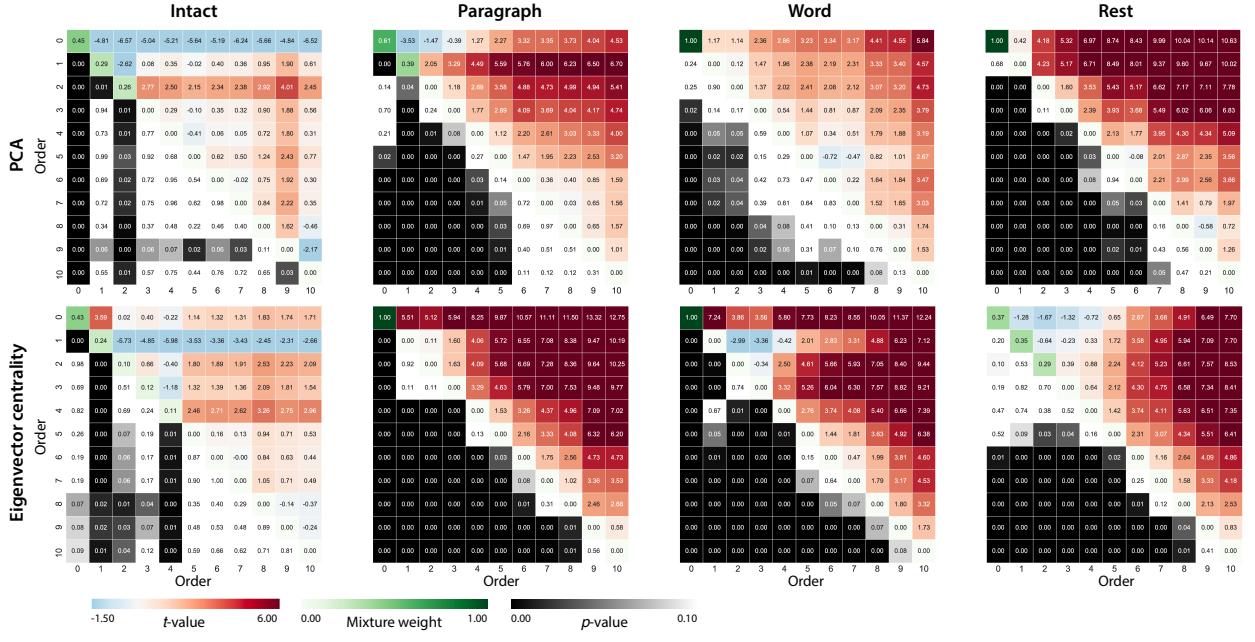


Figure 4: **Across-participant decoding accuracy varies with correlation order and cognitive engagement.** **Across-participant timepoint decoding accuracy varies with correlation order and cognitive engagement.**

**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:  $\frac{1}{400}$ ; **these chance accuracies were subtracted from the observed accuracies to obtain the relative accuracies reported on the y-axis**). The error ribbons 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 *t-test*, thresholded at  $p < 0.05$ ). Figure 5 displays additional comparisons between the decoding accuracies achieved using different sets of neural features. The circled values represent the maximum decoding accuracy within each experimental condition. **b. Normalized decoding accuracy as a function of order: PCA.** **b. Normalized timepoint 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.** **c. Timepoint 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.** **d. Normalized timepoint 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. **See Figures S1 and S2 for decoding results broken down by kernel shape and width, respectively.**



**Figure 5: Statistical summary of decoding accuracies for different neural features.** Each column of matrices displays decoding results for one experimental condition (intact, paragraph, word, and rest). We considered dynamic activity patterns (order 0) and dynamic correlations at different orders (order  $> 0$ ). We used two-tailed  $t$ -tests to compare the distributions of decoding accuracies obtained using each pair of features. The distributions for each feature reflect the set of average decoding accuracies (across all kernel parameters), obtained for each random assignment of training and test groups. In the upper triangles of each map, warmer colors (positive  $t$ -values) indicate that the neural feature indicated in the given row yielded higher accuracy than the feature indicated in the given column. Cooler colors (negative  $t$ -values) indicate that the feature in the given row yielded lower decoding accuracy than the feature in the given column. The lower triangles of each map denote the corresponding  $p$ -values for the  $t$ -tests. The diagonal entries display the relative average optimized weight given to each type of feature  $\tau$  in a decoder that included all feature types (see *Feature weighting and testing*).

(represented by a  $T$  by  $K$  matrix). We carried out two sets of analyses that differed in how this mapping was computed. The first set of analyses used PCA to find a low-dimensional embedding of the original dynamic correlation matrices (Fig. 4a,b). The second set of analyses characterized correlations in dynamics of each feature's eigenvector centrality, but did not preserve the underlying activity dynamics (Fig. 4c,d).

Both sets of temporal decoding analyses yielded qualitatively similar results for the auditory (non-rest) conditions of the experiment (Fig. 4: pink, green, and teal lines; Fig. 5: three leftmost columns). The highest decoding accuracy 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. 4: purple lines; Fig. 5: rightmost column). We note that while the resting state times could be decoded reliably, the accuracies were only very

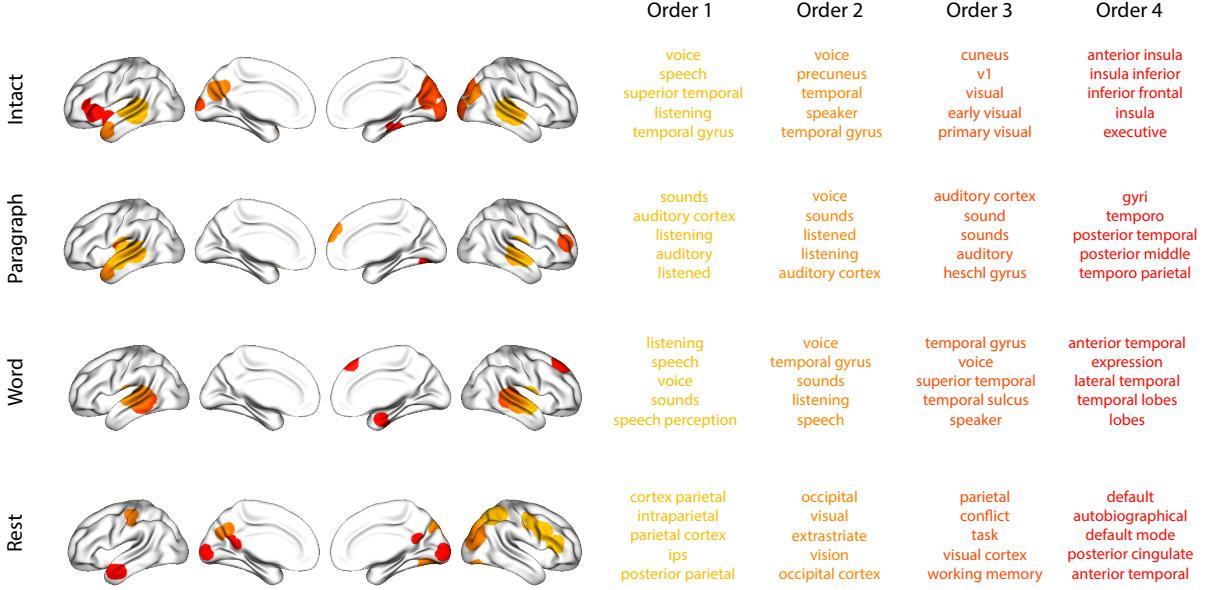


Figure 6: **Top terms associated with the endpoints of the strongest correlations.** **Top terms associated with the most strongly correlated nodes at each order.** Each color corresponds to one order of inter-subject functional correlations. To calculate the dynamic correlations, eigenvector centrality has been used to project the high-dimensional patterns of dynamic correlations onto a lower-dimensional space at each previous order, which allows us to map the brain regions at each order by retaining the features of the original space. The inflated brain plots display the locations of the endpoints of the 10 strongest (absolute value) correlations at each order, thresholded at 0.999, and 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: Fig. S3; paragraph: Fig. S4; word: Fig. S5; rest: Fig. S6).

194 slightly above chance. We speculate that the decoders might have picked up on attentional drift, boredom,  
 195 or tiredness; we hypothesize that these all increased throughout the resting state scan. The decoders might  
 196 be picking up on aspects of these loosely defined cognitive states that are common across individuals. The  
 197 PCA-based approach achieved the highest resting state decoding accuracy using order zero features (non-  
 198 correlational, activation-based), whereas the eigenvector centrality-based approach achieved the highest  
 199 resting state decoding accuracy using second-order correlations. Taken together, these analyses indicate  
 200 that high-level cognitive processing (while listening to the intact story) is reflected in the dynamics of high-  
 201 order correlations in brain activity, whereas lower-level cognitive processing (while listening to scrambled  
 202 versions of the story that lack rich meaning) is reflected in the dynamics of lower-order correlations and  
 203 non-correlational activity dynamics. Further, these patterns are associated both with the underlying activity  
 204 patterns (characterized using PCA) and also with the changing relative positions that different brain areas  
 205 occupy in their associated networks (characterized using eigenvector centrality).

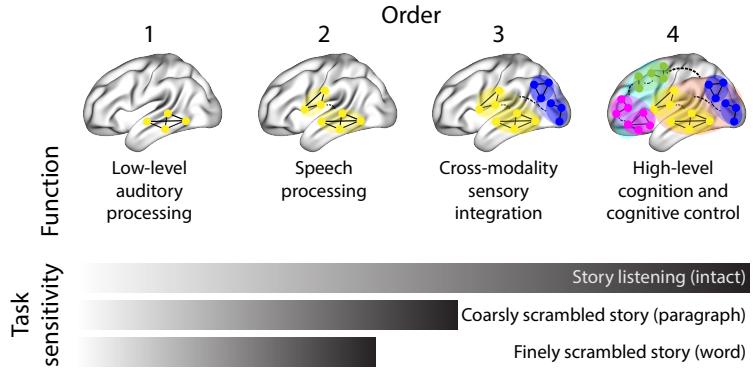
206 Having established that patterns of high-order correlations are informative to decoders, we next won-

207 dered which specific networks of brain regions contributed most to these patterns. As a representative  
208 example, we selected the kernel parameters that yielded decoding accuracies that ~~best matched the were~~  
209 the most strongly correlated (across conditions and orders) with the average accuracies across all of the  
210 kernel parameters we examined. Using Figure 4c as a template, the best-matching kernel was a Laplace  
211 kernel with a width of 50 (Fig. 9d; also see Fig. S7). We used this kernel to compute a single  $K$  by  $K$   $n^{\text{th}}$ -order  
212 DISFC matrix for each experimental condition. We then used Neurosynth (Rubin et al., 2017) to compute  
213 the terms most highly associated with the most strongly correlated pairs of regions in each of these matrices  
214 (Fig. 6; see *Reverse inference*).

215 For all of the story listening conditions (intact, paragraph, and word; top three rows of Fig. 6), we  
216 found that first- and second-order correlations were most strongly associated with auditory and speech  
217 processing areas. During intact story listening, third-order correlations reflected integration with visual  
218 areas, and fourth-order correlations reflected integration with areas associated with high-level cognition and  
219 cognitive control, such as the ventrolateral prefrontal cortex. However, ~~during listening when participants~~  
220 listened to temporally scrambled stories, these higher-order correlations instead involved interactions with  
221 additional regions associated with speech and semantic processing ~~—(second and third rows of Fig. 6)~~.  
222 By contrast, we found a much different set of patterns in the resting state data ~~—(Fig. 6, bottom row)~~.  
223 First-order resting state correlations were most strongly associated with regions involved in counting and  
224 numerical understanding. Second-order resting state correlations were strongest in visual areas; third-order  
225 correlations were strongest in task-positive areas; and fourth-order correlations were strongest in regions  
226 associated with autobiographical and episodic memory. We carried out analogous analyses to create maps  
227 (and decode the top associated Neurosynth terms) for up to fifteenth-order correlations (Figs. S3, S4, S5, and  
228 S6). Of note, examining fifteenth-order correlations between 700 nodes using conventional methods would  
229 have required storing roughly  $\frac{700^{2 \times 15}}{2} \approx 1.13 \times 10^{85}$  floating point numbers—assuming single-precision (32  
230 bits each), this would require roughly 32 times as many bits as there are molecules in the known universe!  
231 Although these fifteenth-order correlations do appear (visually) to have some well-formed structure, we  
232 provide this latter example primarily as a demonstration of the efficiency and scalability of our approach.

## 233 Discussion

234 We tested the hypothesis that high-level cognition is ~~supported by~~ reflected in high-order brain network  
235 dynamics (e.g., see Reimann et al., 2017; Solomon et al., 2019). We examined high-order network dynamics  
236 in functional neuroimaging data collected during a story listening experiment. When participants listened  
237 to an auditory recording of the story, participants exhibited similar high-order brain network dynamics. By



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

238 contrast, when participants instead listened to temporally scrambled recordings of the story, only lower-  
 239 order brain network dynamics were similar across participants. Our results indicate that higher orders of  
 240 network interactions support higher-level aspects of cognitive processing (Fig. 7).

241 The notion that cognition is reflected in (and possibly mediated by) patterns of first-order network  
 242 dynamics has been suggested by or proposed in myriad empirical studies and reviews (e.g., Chang &  
 243 Glover, 2010; Demertzi et al., 2019; Fong et al., 2019; Gonzalez-Castillo et al., 2019; Liégeois et al., 2019; Lurie  
 244 et al., 2018; Manning et al., 2018; Park et al., 2018; Preti et al., 2017; Roy et al., 2019; Turk-Browne, 2013;  
 245 Zou et al., 2019). Our study extends this line of work by finding cognitively relevant *higher-order* network  
 246 dynamics that reflect ongoing cognition. Our findings also complement other work that uses graph theory  
 247 and topology to characterize how brain networks reconfigure during cognition (e.g., Bassett et al., 2006;  
 248 Betzel et al., 2019; McIntosh & Jirsa, 2019; Reimann et al., 2017; Sizemore et al., 2018; Toker & Sommer, 2019;  
 249 Zheng et al., 2019).

250 An open question not addressed by our study pertains to how different structures integrate incom-  
 251 ing information with different time constants. For example, one line of work suggests that the cortical  
 252 surface comprises a structured map such that nearby brain structures process incoming information at  
 253 similar timescales. Low-level sensory areas integrate information relatively quickly, whereas higher-level  
 254 regions integrate information relatively slowly (Baldassano et al., 2017; Chien & Honey, 2019; Hasson et  
 255 al., 2015, 2008; Honey et al., 2012; Lerner et al., 2014, 2011). A similar hierarchy appears to play a role in  
 256 predicting future events (C. S. Lee et al., 2020). Other related work in human and mouse brains indicates  
 257 that the temporal response profile of a given brain structure may relate to how strongly connected that  
 258 structure is with other brain areas (Fallon et al., 2019). Further study is needed to understand the role of  
 259 temporal integration at different scales of network interaction, and across different anatomical structures.

260 Importantly, our analyses do not speak to the physiological basis of higher-order dynamics, and could  
261 reflect nonlinearities, chaotic patterns, non-stationarities, and/or multistability, etc. However, our decoding  
262 analyses do indicate that higher-order dynamics are consistent across individuals, and therefore unlikely to  
263 reflect non-stimulus-driven dynamics that are unlikely to be similar across individuals.

264 Another potential limitation of our approach relates to recent work suggesting that the brain undergoes  
265 rapid state changes, for example across event boundaries (e.g., Baldassano et al., 2017). Shappell et al.  
266 (2019) used hidden semi-Markov models to estimate state-specific network dynamics (also see Vidaurre et  
267 al., 2018). Our general approach might be extended by considering putative state transitions. For example,  
268 rather than weighting all timepoints using a similar kernel (Eqn. 4), the kernel function could adapt on a  
269 timepoint-by-timepoint basis such that only timepoints determined to be in the same “state” were given  
270 non-zero weight.

271 Identifying high-order network dynamics associated with high-level cognition required several impor-  
272 tant methods advances. First, we used kernel-based dynamic correlations to extended the notion of (static)  
273 inter-subject functional connectivity (Simony et al., 2016) to a dynamic measure of inter-subject functional  
274 connectivity (DISFC) that does not rely on sliding windows (e.g., as in Manning et al., 2018), and that may  
275 be computed at individual timepoints. This allowed us to precisely characterize stimulus-evoked network  
276 dynamics that were similar across individuals. Second, we developed a computational framework for  
277 efficiently and scalably estimating high-order dynamic correlations. Our approach uses dimensionality  
278 reduction algorithms and graph measures to obtain low-dimensional embeddings of patterns of network  
279 dynamics. Third, we developed an analysis framework for identifying robust decoding results by carrying  
280 out our analyses using a range of parameter values and then identifying which results were robust to specific  
281 parameter choices.

## 282 **Concluding remarks**

283 The complex hierarchy of dynamic interactions that underlie our thoughts is perhaps the greatest mystery in  
284 modern science. Methods for characterizing the dynamics of high-order correlations in neural data provide  
285 a window into the neural basis of cognition. By showing that high-level cognition is reflected in high-order  
286 network dynamics, we have elucidated the next step on the path towards understanding the neural basis  
287 of cognition.

288 **Methods**

289 Our general approach to efficiently estimating high-order dynamic correlations comprises four general  
290 steps (Fig. 8). First, we derive a kernel-based approach to computing dynamic pairwise correlations in  
291 a  $T$  (timepoints) by  $K$  (features) multivariate timeseries,  $\mathbf{X}_0$ . This yields a  $T$  by  $O(K^2)$  matrix of dynamic  
292 correlations,  $\mathbf{Y}_1$ , where each row comprises the upper triangle and diagonal of the correlation matrix at  
293 a single timepoint, reshaped into a row vector (this reshaped vector is  $(\frac{K^2-K}{2} + K)$ -dimensional). Second,  
294 we apply a dimensionality reduction step to project the matrix of dynamic correlations back onto a  $K$ -  
295 dimensional space. This yields a  $T$  by  $K$  matrix,  $\mathbf{X}_1$ , that reflects an approximation of the dynamic correlations  
296 reflected in the original data. Third, we use repeated applications of the kernel-based dynamic correlation  
297 step to  $\mathbf{X}_n$  and the dimensionality reduction step to the resulting  $\mathbf{Y}_{n+1}$  to estimate high-order dynamic  
298 correlations. Each application of these steps to a  $T$  by  $K$  time series  $\mathbf{X}_n$  yields a  $T$  by  $K$  matrix,  $\mathbf{X}_{n+1}$ , that  
299 reflects the dynamic correlations between the columns of  $\mathbf{X}_n$ . In this way, we refer to  $n$  as the *order* of the  
300 timeseries, where  $\mathbf{X}_0$  (order 0) denotes the original data and  $\mathbf{X}_n$  denotes (approximated)  $n^{\text{th}}$ -order dynamic  
301 correlations between the columns of  $\mathbf{X}_0$ . Finally, we use a cross-validation-based decoding approach to  
302 evaluate how well information contained in a given order (or weighted mixture of orders) may be used  
303 to decode relevant cognitive states. If including a given  $\mathbf{X}_n$  in the feature set yields higher classification  
304 accuracy on held-out data, we interpret this as evidence that the given cognitive states are reflected in  
305 patterns of  $n^{\text{th}}$ -order correlations.

306 All of the code used to produce the figures and results in this manuscript, along with links to the  
307 corresponding datasets, may be found at [github.com/ContextLab/timecorr-paper](https://github.com/ContextLab/timecorr-paper). In addition, we have  
308 released a Python toolbox for computing dynamic high-order correlations in timeseries data; our toolbox  
309 may be found at [timecorr.readthedocs.io](https://timecorr.readthedocs.io).

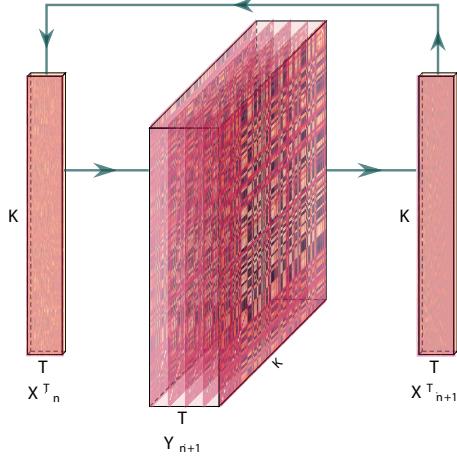


Figure 8: **Estimating dynamic high-order correlations.** Given a  $T$  by  $K$  matrix of multivariate timeseries data,  $\mathbf{X}_n$  (where  $n \in \mathbb{N}, n \geq 0$ ), we use Equation 4 to compute a timeseries of  $K$  by  $K$  correlation matrices,  $\mathbf{Y}_{n+1}$ . We then approximate  $\mathbf{Y}_{n+1}$  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  $\mathbf{X}_n$  and  $\mathbf{X}_{n+1}$  are displayed in the figure for compactness.

### 310 Kernel-based approach for computing dynamic correlations

Given a  $T$  by  $K$  matrix of observations,  $\mathbf{X}$ , 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):

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

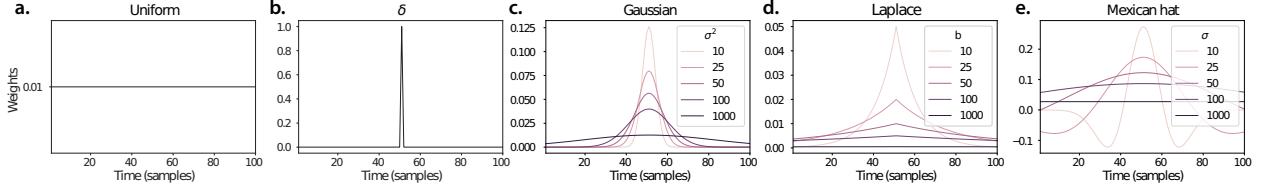
$$\bar{\mathbf{X}}(\cdot, k) = \frac{1}{T} \sum_{t=1}^T \mathbf{X}(t, k), \text{ and} \quad (2)$$

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

- 311 We can generalize this formula to compute time-varying correlations by incorporating a *kernel function* that
- 312 takes a time  $t$  as input, and returns how much the observed data at each timepoint  $\tau \in [-\infty, \infty]$  contributes
- 313 to the estimated instantaneous correlation at time  $t$  (Fig. 9; also see Allen et al., 2012, for a similar approach).

314

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



**Figure 9: Examples of kernel functions.** Each panel displays per-timepoint weights for a kernel centered at  $t = 50$ , evaluated at 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  $\tau$ . This is a special case kernel function that reduces dynamic correlations to static correlations. **b. Dirac  $\delta$  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 ( $\sigma^2$ ) 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 ( $\sigma$ ) are displayed.

static correlation formula in Equation 1 to estimate the *instantaneous correlation* at timepoint  $t$ :

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

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

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

315 Here  $\text{timecorr}_{\kappa_t}(\mathbf{X}(\cdot, i), \mathbf{X}(\cdot, j))$  reflects the correlation at time  $t$  between columns  $i$  and  $j$  of  $\mathbf{X}$ , estimated using  
 316 the kernel  $\kappa_t$ . We evaluate Equation 4 in turn for each pair of columns in  $\mathbf{X}$  and for kernels centered on each  
 317 timepoint in the timeseries, respectively, to obtain a  $T$  by  $K$  by  $K$  timeseries of dynamic correlations,  $\mathbf{Y}$ . For  
 318 convenience, we then reshape the upper triangles and diagonals of each timepoint's symmetric correlation  
 319 matrix into a row vector to obtain an equivalent  $T$  by  $\left(\frac{K^2-K}{2} + K\right)$  matrix.

### 320 Dynamic inter-subject functional connectivity (DISFC)

Equation 4 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 & Chang, 2020; Simony et al., 2016) to compute the stimulus-driven correlations reflected in the multi-participant dataset at a given timepoint  $t$  using:

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

where  $M$  extracts and vectorizes the upper triangle and diagonal 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  $\mathbf{Y}_{n+1}^p(t)$  denotes the correlation matrix at timepoint  $t$  (Eqn. 4) between each column of  $\mathbf{X}_n^p$  and each column of the average  $\mathbf{X}_n$  from all *other* participants,  $\bar{\mathbf{X}}_n^p$ :

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

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

### 323 Low-dimensional representations of dynamic correlations

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

#### 329 Dimensionality reduction-based approaches to computing $\mathbf{X}_n$

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

332 Spearman, 1904), Independent Components Analysis (ICA; Comon et al., 1991; Jutten & Herault, 1991),  
333 *t*-Stochastic Neighbor Embedding (*t*-SNE; van der Maaten & Hinton, 2008), Uniform Manifold Approximation and Projection (UMAP; McInnes et al., 2018), non-negative matrix factorization (NMF; D. D. Lee  
334 & Seung, 1999), Topographic Factor Analysis (TFA; Manning et al., 2014), Hierarchical Topographic Factor analysis (HTFA; Manning et al., 2018), Topographic Latent Source Analysis (TLSA; Gershman et al.,  
335 2011), dictionary learning (J. Mairal et al., 2009; J. B. Mairal et al., 2009), and deep auto-encoders (Hinton  
336 & Salakhutdinov, 2006), among others. While complete characterizations of each of these algorithms is  
337 beyond the scope of the present manuscript, the general intuition driving these approaches is to compute  
338 the  $T$  by  $K$  matrix,  $\mathbf{X}$ , that is closest to the original  $T$  by  $J$  matrix,  $\mathbf{Y}$ , where (typically)  $K \ll J$ . The different  
339 approaches place different constraints on what properties  $\mathbf{X}$  must satisfy and which aspects of the data are  
340 compared (and how) in order to optimize how well  $\mathbf{X}$  approximates  $\mathbf{Y}$ .

341 Applying dimensionality reduction algorithms to  $\mathbf{Y}$  yields an  $\mathbf{X}$  whose columns reflect weighted com-  
342 binations (or nonlinear transformations) of the original columns of  $\mathbf{Y}$ . This has two main consequences.  
343 First, with each repeated dimensionality reduction, the resulting  $\mathbf{X}_n$  has lower and lower fidelity (with  
344 respect to what the “true”  $\mathbf{Y}_n$  might have looked like without using dimensionality reduction to maintain  
345 **sealability****tractability**). In other words, computing  $\mathbf{X}_n$  is a lossy operation. Second, whereas each column  
346 of  $\mathbf{Y}_n$  may be mapped directly onto specific pairs of columns of  $\mathbf{X}_{n-1}$ , the columns of  $\mathbf{X}_n$  reflect weighted  
347 combinations and/or nonlinear transformations of the columns of  $\mathbf{Y}_n$ . Many dimensionality reduction algo-  
348 rithms are invertible (or approximately invertible). However, attempting to map a given  $\mathbf{X}_n$  back onto the  
349 original feature space of  $\mathbf{X}_0$  will usually require  $\mathcal{O}(TK^{2n})$  $\mathcal{O}(TK^2)$  space and therefore becomes intractable as  
350  $n$  or  $K$  grow large.

### 353 Graph measure approaches to computing $\mathbf{X}_n$

354 The above dimensionality reduction approaches to approximating a given  $\mathbf{Y}_n$  with a lower-dimensional  
355  $\mathbf{X}_n$  preserve a (potentially recombined and transformed) mapping back to the original data in  $\mathbf{X}_0$ . We also  
356 explore graph measures that instead characterize each feature’s relative *position* in the broader network of  
357 interactions and connections. To illustrate the distinction between the two general approaches we explore,  
358 suppose a network comprises nodes  $A$  and  $B$ , and  $C$  along with several other nodes. If  $A$  and  $B$  exhibit  
359 uncorrelated activity patterns, then by definition the functional connection (correlation) between them will  
360 be **(by definition)** close to 0. However, if  $A$  and  $B$  each interact with  $C$  other nodes in similar ways, we might  
361 attempt to capture those similarities using a measure that reflects how between  $A$ ’s and  $B$  interact with other  
362 ’s interactions with those other members of the network.

363 In general, graph measures take as input a matrix of interactions (e.g., using the above notation, a  $K$   
 364 by  $K$  correlation matrix or binarized correlation matrix reconstituted from a single timepoint's row of  $\mathbf{Y}$ ),  
 365 and return as output a set of  $K$  measures describing how each node (feature) sits within that correlation  
 366 matrix with respect to the rest of the population. Widely used measures include betweenness centrality (the  
 367 proportion of shortest paths between each pair of nodes in the population that involves the given node  
 368 in question; e.g., Barthélemy, 2004; Freeman, 1977; Geisberger et al., 2008; Newman, 2005; Opsahl et al.,  
 369 2010); diversity and dissimilarity (characterizations of how differently connected a given node is from others  
 370 in the population; e.g., Lin, 2009; Rao, 1982; Ricotta & Szeidl, 2006); eigenvector centrality and pagerank  
 371 centrality (measures of how influential a given node is within the broader network; e.g., Bonacich, 2007;  
 372 Halu et al., 2013; Lohmann et al., 2010; Newman, 2008); transfer entropy and flow coefficients (a measure of  
 373 how much information is flowing from a given node to other nodes in the network; e.g., Honey et al., 2007;  
 374 Schreiber, 2000);  $k$ -coreness centrality (a measure of the connectivity of a node within its local subgraph; e.g.,  
 375 Alvarez-Hamelin et al., 2005; Christakis & Fowler, 2010); within-module degree (a measure of how many  
 376 connections a node has to its close neighbors in the network; e.g., Rubinov & Sporns, 2010); participation  
 377 coefficient (a measure of the diversity of a node's connections to different subgraphs in the network; e.g.,  
 378 Rubinov & Sporns, 2010); and subgraph centrality (a measure of a node's participation in all of the network's  
 379 subgraphs; e.g., Estrada & Rodríguez-Velázquez, 2005); among others.

380 For a given graph measure,  $\eta : \mathbb{R}^{K \times K} \rightarrow \mathbb{R}^K$ , we can use  $\eta$  to transform each row of  $\mathbf{Y}_n$  in a way that  
 381 characterizes the corresponding graph properties of each column. This results in a new  $T$  by  $K$  matrix,  
 382  $\mathbf{X}_n$ , that reflects how the features reflected in the columns of  $\mathbf{X}_{n-1}$  participate in the network during each  
 383 timepoint (row).

### 384 **Dynamic higher-order correlations**

385 Because  $\mathbf{X}_n$  has the same shape as the original data  $\mathbf{X}_0$ , approximating  $\mathbf{Y}_n$  with a lower-dimensional  $\mathbf{X}_n$  enables  
 386 us to estimate high-order dynamic correlations in a scalable way. Given a  $T$  by  $K$  input matrix, the output  
 387 of Equation 4 requires  $O(TK^2)$  space to store. Repeated applications of Equation 4 (i.e., computing dynamic  
 388 correlations between the columns of the outputted dynamic correlation matrix) each require exponentially  
 389 more space; in general the  $n^{\text{th}}$ -order dynamic correlations of a  $T$  by  $K$  timeseries occupies  $\cancel{O(TK^{2n})} O(\underline{TK^{2^n}})$   
 390 space. However, when we approximate or summarize the output of Equation 4 with a  $T$  by  $K$  matrix (as  
 391 described above), it becomes feasible to compute even very high-order correlations in high-dimensional  
 392 data. Specifically, approximating the  $n^{\text{th}}$ -order dynamic correlations of a  $T$  by  $K$  timeseries requires only  
 393  $O(TK^2)$  additional space— the same as would be required to compute first-order dynamic correlations. In

<sup>394</sup> other words, the space required to store  $n + 1$  multivariate timeseries reflecting up to  $n^{\text{th}}$  order correlations  
<sup>395</sup> in the original data scales linearly with  $n$  using our approach (Fig. 8).

## <sup>396</sup> Data

<sup>397</sup> We examined two types of data: synthetic data and human functional neuroimaging data. We constructed  
<sup>398</sup> and leveraged the synthetic data to evaluate our general approach (for a related validation approach see  
<sup>399</sup> Thompson et al., 2018). Specifically, we tested how well Equation 4 could be used to recover known dynamic  
<sup>400</sup> correlations using different choices of kernel ( $\kappa$ ; Fig. 9), for each of several synthetic datasets that exhibited  
<sup>401</sup> different temporal properties. We also simulated higher-order correlations and tested how well Equation 4  
<sup>402</sup> could recover these correlations using the best kernel from the previous synthetic data analyses. We then  
<sup>403</sup> applied our approach to a functional neuroimaging dataset to test the hypothesis that ongoing cognitive  
<sup>404</sup> processing is reflected in high-order dynamic correlations. We used an across-participant classification test  
<sup>405</sup> to estimate whether dynamic correlations of different orders contain information about which timepoint in  
<sup>406</sup> a story participants were listening to.

### <sup>407</sup> Synthetic data: simulating dynamic first-order correlations

<sup>408</sup> We constructed a total of 40 different multivariate timeseries, collectively reflecting a total of 4 qualitatively  
<sup>409</sup> different patterns of dynamic first-order correlations (i.e., 10 datasets reflecting each type of dynamic pat-  
<sup>410</sup> tern). Each timeseries comprised 50 features (dimensions) that varied over 300 timepoints. The observations  
<sup>411</sup> at each timepoint were drawn from a zero-mean multivariate Gaussian distribution with a covariance matrix  
<sup>412</sup> defined for each timepoint as described below. We drew the observations at each timepoint independently  
<sup>413</sup> from the draws at all other timepoints; in other words, for each observation  $s_t \sim \mathcal{N}(\mathbf{0}, \Sigma_t)$  at timepoint  $t$ ,  
<sup>414</sup>  $p(s_t) = p(s_t | s_{\setminus t})$ .

**Constant.** We generated data with stable underlying correlations to evaluate how Equation 4 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 (indexed by  $m$ ), we constructed a random covariance matrix,  $\Sigma_m$ :

$$\underline{\Sigma_m} \approx \mathbf{C} \mathbf{C}^\top, \text{ where} \quad (11)$$

$$\underline{\mathbf{C}(i, j) \sim \mathcal{N}(0, 1)} \underline{\Sigma_m = \mathbf{C} \mathbf{C}^\top, \text{ where, and where}} \quad (12)$$

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

418 **Random.** We generated a second set of 10 synthetic datasets whose observations at each timepoint were  
 419 drawn from a Gaussian distribution with a new randomly constructed (using Eqn. 11) covariance matrix.  
 420 Because each timepoint's covariance matrix was drawn independently from the covariance matrices for all  
 421 other timepoints, these datasets provided a test of reconstruction accuracy in the absence of any meaningful  
 422 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* covariance matrices using Equation 11,  $\Sigma_{\text{start}}$  and  $\Sigma_{\text{end}}$ . 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

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

423 The gradually changing correlations underlying these datasets allow us to evaluate the recovery of dynamic  
 424 correlations when each timepoint's correlation matrix is unique (as in the random datasets), but where the  
 425 correlation dynamics are structured and exhibit first-order autocorrelations (as in the constant datasets).

426 **Event.** We generated a fourth set of 10 synthetic datasets whose underlying correlation matrices exhibited  
 427 prolonged intervals of stability, interspersed with abrupt changes. For each dataset, we used Equation 11  
 428 to generate 5 random covariance matrices. We constructed a timeseries where each set of 60 consecutive  
 429 samples was drawn from a Gaussian with the same covariance matrix. These datasets were intended to  
 430 simulate a system that undergoes exhibits periods of stability punctuated by occasional abrupt state changes.

### 431 Synthetic data: simulating dynamic high-order correlations

432 We developed an iterative procedure for constructing timeseries data that exhibits known dynamic high-order  
 433 correlations. The procedure builds on our approach to generating dynamic first-order correlations. Essentially,  
 434 once we generate a timeseries with known first-order correlations, we can use the known first-order  
 435 correlations as a template to generate a new timeseries of second-order correlations. In turn, we can  
 436 generate a timeseries of third-order correlations from the second-order correlations, and so on. In general,  
 437 we can generate order  $n$  correlations given a timeseries of order  $n-1$  correlations, for any  $n \geq 1$ . Finally,

438 given the order  $n$  timeseries, we can reverse the preceding process to generate an order  $n - 1$  timeseries, an  
439 order  $n - 2$  order timeseries, and so on, until we obtain an order 0 timeseries of simulated data that reflects  
440 the chosen high-order dynamics.

The central mathematical operations in our procedure are two bookkeeping functions,  $\text{vec}(\cdot)$  and  $\text{mat}(\cdot)$ . The  $\text{vec}(\cdot)$  function takes as input a  $K \times K$  symmetric matrix and returns as output a  $(\frac{K^2-K}{2} + K)$ -dimensional column vector containing the entries in the upper triangle and diagonal. The  $\text{mat}(\cdot)$  function inverts  $\text{vec}(\cdot)$  by taking as input a  $(\frac{K^2-K}{2} + K)$ -dimensional column vector and returning a  $K \times K$  symmetric matrix as output. We can then generate an order  $n$  correlation matrix (for one timepoint,  $t$ ) from an order  $n - 1$  template (from the same timepoint) as follows:

$$\Sigma_n(t) = \text{mat}(\text{vec}(\Sigma_{n-1}(t)) \otimes \text{vec}(\Sigma_{n-1}(t))^\top). \quad (14)$$

Given a timeseries of order  $n$  correlation matrices, we can draw an order  $n - 1$  correlation matrix for each timepoint  $t$  using

$$\sigma_{n-1}(t) \sim \mathcal{N}(\mathbf{0}, \Sigma_n(t)) \quad (15)$$

$$\Sigma_{n-1}(t) = \text{mat}(\sigma_{n-1}(t)). \quad (16)$$

441 We can then use repeated applications of Equations 15 and 16 in order to obtain a synthetic dataset.  
442 When the template first-order correlations are constructed to exhibit different temporal profiles (e.g.,  
443 using the constant, random, ramping, and event procedures described above), the resulting high-order  
444 correlations and synthetic data will exhibit the same category of temporal profile. Following our approach  
445 to generating synthetic data exhibiting known first-order correlations, we constructed a total of 40 additional  
446 multivariate timeseries, collectively reflecting a total of 4 qualitatively different patterns of dynamic  
447 correlations (i.e., 10 datasets reflecting each type of dynamic pattern: constant, random, ramping, and  
448 event).

#### 449 Functional neuroimaging data collected during story listening

450 We examined an fMRI dataset collected by Simony et al. (2016) that the authors have made publicly available  
451 at [arks.princeton.edu/ark:/88435/dsp015d86p269k](http://arks.princeton.edu/ark:/88435/dsp015d86p269k). The dataset comprises neuroimaging data collected as  
452 participants listened to an audio recording of a story (intact condition; 36 participants), listened to temporally  
453 scrambled recordings of the same story (17 participants in the paragraph-scrambled condition listened to  
454 the paragraphs in a randomized order and 36 in the word-scrambled condition listened to the words in a

455 randomized order), or lay resting with their eyes open in the scanner (rest condition; 36 participants). Full  
456 neuroimaging details may be found in the original paper for which the data were collected (Simony et al.,  
457 2016).

458 **Hierarchical topographic factor analysis (HTFA).** Following our prior related work, we used HTFA (Manning et al., 2018) to derive a compact representation of the neuroimaging data. In brief, this approach approximates the timeseries of voxel activations (44,415 voxels) using a much smaller number of radial basis function (RBF) nodes (in this case, 700 nodes, as determined by an optimization procedure described by Manning et al., 2018). This provides a convenient representation for examining full-brain network dynamics. All of the analyses we carried out on the neuroimaging dataset were performed in this lower-dimensional space. In other words, each participant's data matrix,  $X_0$ , was a number-of-timepoints by 700 matrix of HTFA-derived factor weights (where the row and column labels were matched across participants). Code for carrying out HTFA on fMRI data may be found as part of the BrainIAK toolbox (Capota et al., 2017), which may be downloaded at [brainiak.org](http://brainiak.org).

## 468 Temporal decoding

469 We sought to identify neural patterns that reflected participants' ongoing cognitive processing of incoming  
470 stimulus information. As reviewed by Simony et al. (2016), one way of homing in on these stimulus-driven  
471 neural patterns is to compare activity patterns across individuals (e.g., using ISFC analyses). In particular,  
472 neural patterns will be similar across individuals to the extent that the neural patterns under consideration  
473 are stimulus-driven, and to the extent that the corresponding cognitive representations are reflected in  
474 similar spatial patterns across people (also see Simony & Chang, 2020). Following this logic, we used an  
475 across-participant temporal decoding test developed by Manning et al. (2018) to assess the degree to which  
476 different neural patterns reflected ongoing stimulus-driven cognitive processing across people (Fig. 10).  
477 The approach entails using a subset of the data to train a classifier to decode stimulus timepoints (i.e.,  
478 moments in the story participants listened to) from neural patterns. We use decoding (forward inference)  
479 accuracy on held-out data, from held-out participants, as a proxy for the extent to which the inputted neural  
480 patterns reflected stimulus-driven cognitive processing in a similar way across individuals.

## 481 Forward inference and decoding accuracy

482 We used an across-participant correlation-based classifier to decode which stimulus timepoint matched each  
483 timepoint's neural pattern (Fig. 10). We first divided the participants into two groups: a template group,

484  $\mathcal{G}_{\text{template}}$  (i.e., training data), and a to-be-decoded group,  $\mathcal{G}_{\text{decode}}$  (i.e., test data). We used Equation 7 to  
485 compute a DISFC matrix for each group ( $\bar{\mathbf{C}}_{\text{template}}$  and  $\bar{\mathbf{C}}_{\text{decode}}$ , respectively). We then correlated the rows of  
486  $\bar{\mathbf{C}}_{\text{template}}$  and  $\bar{\mathbf{C}}_{\text{decode}}$  to form a number-of-timepoints by number-of-timepoints decoding matrix,  $\Lambda$ . In this  
487 way, the rows of  $\Lambda$  reflected timepoints from the template group, while the columns reflected timepoints  
488 from the to-be-decoded group. We used  $\Lambda$  to assign temporal labels to each row  $\bar{\mathbf{C}}_{\text{decode}}$  using the row of  
489  $\bar{\mathbf{C}}_{\text{template}}$  with which it was most highly correlated. We then repeated this decoding procedure, but using  
490  $\mathcal{G}_{\text{decode}}$  as the template group and  $\mathcal{G}_{\text{template}}$  as the to-be-decoded group. Given the true timepoint labels (for  
491 each group), we defined the *decoding accuracy* as the average proportion of correctly decoded timepoints,  
492 across both groups. We defined the *relative decoding accuracy* as the difference between the decoding accuracy  
493 and chance accuracy (i.e.,  $\frac{1}{T}$ ).

494 **Feature weighting and testing**

495 We sought to examine which types of neural features (i.e., activations, first-order dynamic correlations, and  
496 higher-order dynamic correlations) were informative to the temporal decoders. Using the notation above,  
497 these features correspond to  $\mathbf{X}_0, \mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3$ , and so on.

498 One challenge to fairly evaluating high-order correlations is that if the kernel used in Equation 4 is  
499 wider than a single timepoint, each repeated application of the equation will result in further temporal  
500 blur. Because our primary assessment metric is temporal decoding accuracy, this unfairly biases against  
501 detecting meaningful signal in higher-order correlations (relative to lower-order correlations). We attempted  
502 to mitigate temporal blur in estimating each  $\mathbf{X}_n$  by using a Dirac  $\delta$  function kernel (which places all of its  
503 mass over a single timepoint; Fig. 9b [10a](#)) to compute each lower-order correlation ( $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_{n-1}$ ). We  
504 then used a new (potentially wider, as described below) kernel to compute  $\mathbf{X}_n$  from  $\mathbf{X}_{n-1}$ . In this way,  
505 temporal blurring was applied only in the last step of computing  $\mathbf{X}_n$ . We note that, because each  $\mathbf{X}_n$  is a  
506 low-dimensional representation of the corresponding  $\mathbf{Y}_n$ , the higher-order correlations we estimated reflect  
507 true correlations in the data with lower-fidelity than estimates of lower-order correlations. Therefore, even  
508 after correcting for temporal blurring, our approach is still biased against finding meaningful signal in  
509 higher-order correlations.

510 After computing each  $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_{n-1}$  for each participant, we divided participants into two equally sized  
511 groups ( $\pm 1$  for odd numbers of participants):  $\mathcal{G}_{\text{train}}$  and  $\mathcal{G}_{\text{test}}$ . We then further subdivided  $\mathcal{G}_{\text{train}}$  into  $\mathcal{G}_{\text{train}_1}$   
512 and  $\mathcal{G}_{\text{train}_2}$ . We then computed  $\Lambda$  (temporal correlation) matrices for each type of neural feature, using  $\mathcal{G}_{\text{train}_1}$   
513 and  $\mathcal{G}_{\text{train}_2}$ . This resulted in  $n + 1$   $\Lambda$  matrices (one for the original timeseries of neural activations, and one  
514 for each of  $n$  orders of dynamic correlations). Our objective was to find a set of weights for each of these

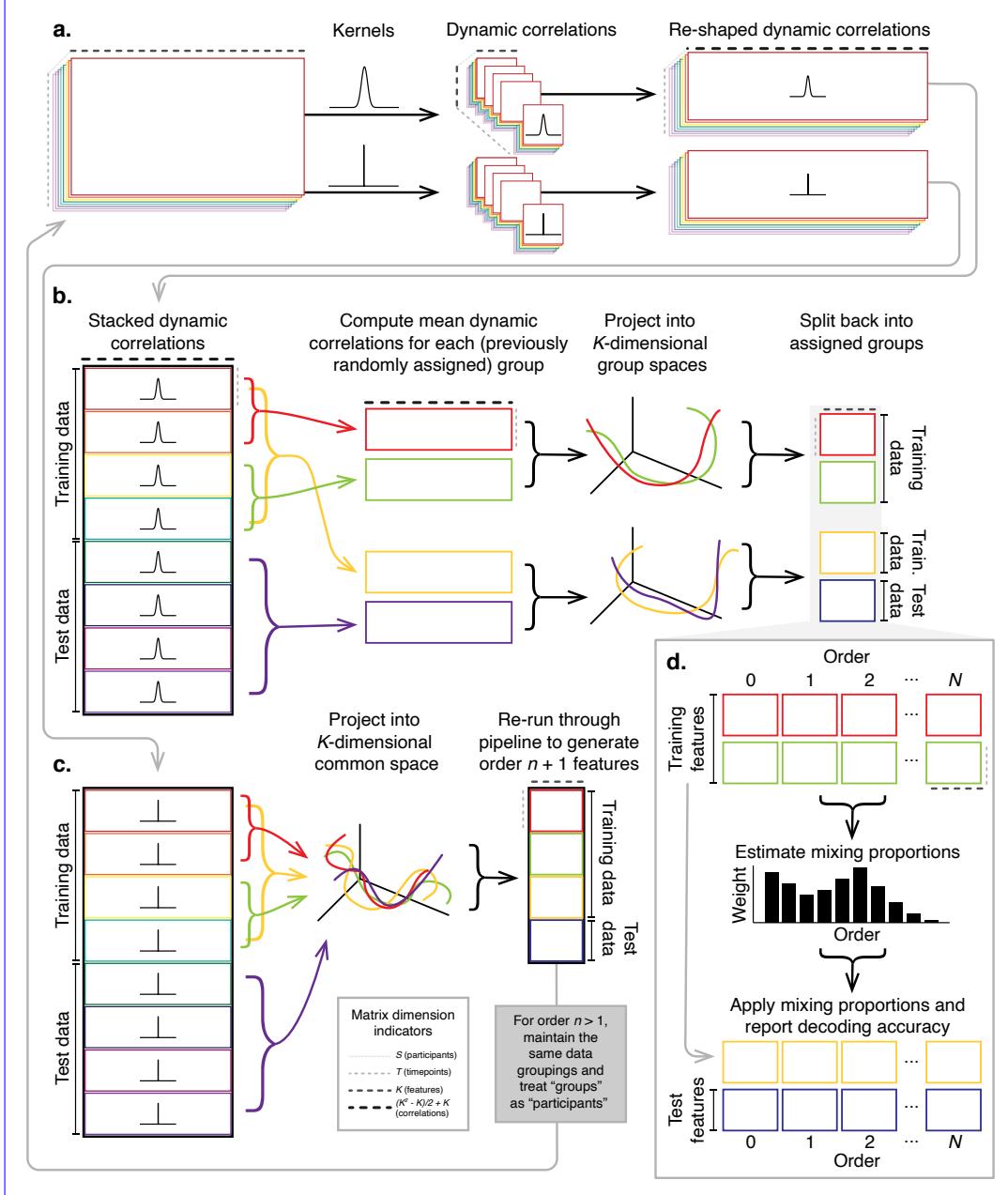


Figure 10: **Decoding analysis pipeline.** **a. Computing dynamic correlations from timeseries data.** Given a timeseries of observations as a  $T \times K$  matrix (or a set of  $S$  such matrices), we use Equation 4 to compute each participant's DISFC (relative to other participants in the training or test sub-group, as appropriate). We repeat this process twice—once using the analysis kernel (shown here as a Gaussian in the upper row of the panel), and once using a  $\delta$  function kernel (lower row of the panel). **b. Projecting dynamic correlations into a lower-dimensional space.** We project the training and test data into  $K$ -dimensional spaces to create compact representations of dynamic correlations at the given order (estimated using the analysis kernel). **c. Kernel trick.** We project the dynamic correlations computed using a  $\delta$  function kernel into a common  $K$ -dimensional space. These low-dimensional embeddings are fed back through the analysis pipeline in order to compute features at the next-highest order. **d. Decoding analysis.** We split the training data into two equal groups, and optimize the feature weights (i.e., dynamic correlations at each order) to maximize decoding accuracy. We then apply the trained classifier to the (held-out) test data.

515  $\Lambda$  matrices such that the weighted average of the  $n + 1$  matrices yielded the highest decoding accuracy.  
516 We used quasi-Newton gradient ascent (Nocedal & Wright, 2006), using decoding accuracy (for  $\mathcal{G}_{\text{train}_1}$  and  
517  $\mathcal{G}_{\text{train}_2}$ ) as the objective function to be maximized, to find an optimal set of training data-derived weights,  
518  $\phi_{0,1,\dots,n}$ , where  $\sum_{i=0}^n \phi_i = 1$  and where  $\phi_i \geq 0 \forall i \in [0, 1, \dots, n]$ .

519 After estimating an optimal set of weights, we computed a new set of  $n + 1$   $\Lambda$  matrices correlating the  
520 DISFC patterns from  $\mathcal{G}_{\text{train}}$  and  $\mathcal{G}_{\text{test}}$  at each timepoint. We use the resulting decoding accuracy of  $\mathcal{G}_{\text{test}}$   
521 timepoints (using the weights in  $\phi_{0,1,\dots,n}$  to average the  $\Lambda$  matrices) to estimate how informative the set of  
522 neural features containing up to  $n^{\text{th}}$  order correlations were.

523 We used a permutation-based procedure to form stable estimates of decoding accuracy for each set of  
524 neural features. In particular, we computed the decoding accuracy for each of 10 random group assignments  
525 of  $\mathcal{G}_{\text{train}}$  and  $\mathcal{G}_{\text{test}}$ . We report the mean accuracy (along with 95% confidence intervals) for each set of neural  
526 features.

## 527 Identifying robust decoding results

528 The temporal decoding procedure we use to estimate which neural features support ongoing cognitive  
529 processing is governed by several parameters. In particular, Equation 4 requires defining a kernel function,  
530 which can take on different shapes and widths. For a fixed set of neural features, each of these parameters  
531 can yield different decoding accuracies. Further, the best decoding accuracy for a given timepoint may be  
532 reliably achieved by one set of parameters, whereas the best decoding accuracy for another timepoint might  
533 be reliably achieved by a different set of parameters, and the best decoding accuracy across *all* timepoints  
534 might be reliably achieved by still another different set of parameters. Rather than attempting to maximize  
535 decoding accuracy, we sought to discover the trends in the data that were robust to classifier parameters  
536 choices. Specifically, we sought to characterize how decoding accuracy varied (under different experimental  
537 conditions) as a function of which neural features were considered.

538 To identify decoding results that were robust to specific classifier parameter choices, we repeated our  
539 decoding analyses after substituting into Equation 4 each of a variety of kernel shapes and widths. We  
540 examined Gaussian (Fig. 9c), Laplace (Fig. 9d), and Mexican Hat (Fig. 9e) kernels, each with widths of 5, 10,  
541 20, and 50 samples. We then report the average decoding accuracies across all of these parameter choices.  
542 This enabled us to (partially) factor out performance characteristics that were parameter-dependent, within  
543 the set of parameters we examined.

544 **Reverse inference**

545 The dynamic patterns we examined comprise high-dimensional correlation patterns at each timepoint. To  
546 help interpret the resulting patterns in the context of other studies, we created summary maps by computing  
547 the across-timepoint average pairwise correlations at each order of analysis (first order, second order, etc.).  
548 We selected the 10 strongest (absolute value) correlations at each order. Each correlation is between the  
549 dynamic activity patterns (or patterns of dynamic high-order correlations) measured at two RBF nodes  
550 (see *Hierarchical Topographic Factor Analysis*). Therefore, the 10 strongest correlations involved up to 20 RBF  
551 nodes. Each RBF defines a spatial function whose activations range from 0 to 1. We constructed a map  
552 of RBF components that denoted the endpoints of the 10 strongest correlations (we set each RBF to have a  
553 maximum value of 1). We then carried out a meta analysis using Neurosynth (Rubin et al., 2017) to identify  
554 the 10 terms most commonly associated with the given map. This resulted in a set of 10 terms associated  
555 with the average dynamic correlation patterns at each order.

556 **Acknowledgements**

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

563 **Author contributions**

564 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. Writing:  
565 L.L.W.O. and J.R.M.

566 **References**

567 Allen, E. A., Damaraju, E., Plis, S. M., Erhardt, E. B., Eichele, T., & Calhoun, V. D. (2012). Tracking  
568 whole-brain connectivity dynamics in the resting state. *Cerebral Cortex*, 24(3), 663–676.

- 569 Alvarez-Hamelin, I., Dall'Asta, L., Barrat, A., & Vespignani, A. (2005). *k*-corr decomposition: a tool for the  
570 visualiztion of large scale networks. *arXiv*, cs/0504107v2.
- 571 Baldassano, C., Chen, J., Zadbood, A., Pillow, J. W., Hasson, U., & Norman, K. A. (2017). Discovering event  
572 structure in continuous narrative perception and memory. *Neuron*, 95(3), 709–721.
- 573 Barthélémy, M. (2004). Betweenness centrality in large complex networks. *European Physical Journal B*, 38,  
574 163–168.
- 575 Bassett, D., Meyer-Lindenberg, A., Achard, S., Duke, T., & Bullmore, E. (2006, December). Adaptive  
576 reconfiguration of fractal small-world human brain functional networks. *Proceedings of the National  
577 Academy of Sciences, USA*, 103(51), 19518–23.
- 578 Beaty, R. E., Benedek, M., Silvia, P. J., & Schacter, D. L. (2016). Creative cognition and brain network  
579 dynamics. *Trends in Cognitive Sciences*, 20(2), 87–95.
- 580 Betzel, R. F., Byrge, L., Esfahlani, F. Z., & Kennedy, D. P. (2019). Temporal fluctuations in the brain's modular  
581 architecture during movie-watching. *bioRxiv*, doi.org/10.1101/750919.
- 582 Bonacich, P. (2007). Some unique properties of eigenvector centrality. *Social Networks*, 29(4), 555–564.
- 583 Bullmore, E., & Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and  
584 functional systems. *Nature Reviews Neuroscience*, 10(3), 186–198.
- 585 Capota, M., Turek, J., Chen, P.-H., Zhu, X., Manning, J. R., Sundaram, N., . . . Shin, Y. S. (2017). *Brain imaging  
586 analysis kit*. Retrieved from <https://doi.org/10.5281/zenodo.59780>
- 587 Chang, C., & Glover, G. H. (2010). Time-frequency dynamics of resting-state brain connectivity measured  
588 with fMRI. *NeuroImage*, 50, 81–98.
- 589 Chien, H.-Y. S., & Honey, C. J. (2019). Constructing and forgetting temporal context in the human cerebral  
590 cortex. *bioRxiv*, doi.org/10.1101/761593.
- 591 Christakis, N. A., & Fowler, J. H. (2010). Social network sensors for early detection of contagious outbreaks.  
592 *PLoS One*, 5(9), e12948.
- 593 Combrisson, E., Vallat, R., O'Reilly, C., Jas, M., Pascarella, A., l Saive, A., . . . Jerbi, K. (2019). Visbrain: a  
594 multi-purpose GPU-accelerated open-source suite for multimodal brain data visualization. *Frontiers in  
595 Neuroinformatics*, 13(14), 1–14.

- 596 Comon, P., Jutten, C., & Herault, J. (1991). Blind separation of sources, part II: Problems statement. *Signal*  
597 *Processing*, 24(1), 11 - 20.
- 598 Demertzi, A., Tagliazucchi, E., Dehaene, S., Deco, G., Barttfeld, P., Raimondo, F., ... Sitt, J. D. (2019). Human  
599 consciousness is supported by dynamic complex patterns of brain signal coordination. *Science Advances*,  
600 5(2), eaat7603.
- 601 Estrada, E., & Rodríguez-Velázquez, J. A. (2005). Subgraph centrality in complex networks. *Physical Review*  
602 *E*, 71(5), 056103.
- 603 Etzel, J. A., Gazzola, V., & Keysers, C. (2009). An introduction to anatomical ROI-based fMRI classification.  
604 *Brain Research*, 1281, 114–125.
- 605 Fallon, J., Ward, P., Parkes, L., Oldham, S., Arnatkevičiūtė, A., Fornito, A., & Fulcher, B. D. (2019).  
606 Timescales of spontaneous activity fluctuations relate to structural connectivity in the brain. *bioRxiv*,  
607 doi.org/10.1101/655050.
- 608 Fong, A. H. C., Yoo, K., Rosenberg, M. D., Zhang, S., Li, C.-S. R., Scheinost, D., ... Chun, M. M. (2019).  
609 Dynamic functional connectivity during task performance and rest predicts individual differences in  
610 attention across studies. *NeuroImage*, 188, 14–25.
- 611 Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry*, 40(1), 35–41.
- 612 Friston, K. (2000). The labile brain. i. neuronal transients and nonlinear coupling. *Phil. Trans. Roy. Soc. Lon.*,  
613 355B, 215–236.
- 614 Geisberger, R., Sanders, P., & Schultes, D. (2008). Better approximation of betweenness centrality. *Proceedings*  
615 *of the meeting on Algorithm Engineering and Experiments*, 90–100.
- 616 Gershman, S., Blei, D., Pereira, F., & Norman, K. (2011). A topographic latent source model for fMRI data.  
617 *NeuroImage*, 57, 89–100.
- 618 Gonzalez-Castillo, J., Caballero-Gaudes, C., Topolski, N., Handwerker, D. A., Pereira, F., & Bandettini, P. A.  
619 (2019). Imaging the spontaneous flow of thought: distinct periods of cognition contribute to dynamic  
620 functional connectivity during rest. *NeuroImage*, 202(116129).
- 621 Grossberg, S. (1988). Nonlinear neural networks: Principles, mechanisms, and architectures. *Neural*  
622 *networks*, 1(1), 17–61.
- 623 Halu, A., Mondragón, R. J., Panzarasa, P., & Bianconi, G. (2013). Multiplex PageRank. *PLoS One*, 8(10),  
624 e78293.

- 625 Hasson, U., Chen, J., & Honey, C. J. (2015). Hierarchical process memory: memory as an integral component  
626 of information processing. *Trends in Cognitive Science*, 19(6), 304-315.
- 627 Hasson, U., Yang, E., Vallines, I., Heeger, D. J., & Rubin, N. (2008). A hierarchy of temporal receptive windows  
628 in human cortex. *Journal of Neuroscience*, 28(10), 2539-2550. doi: 10.1523/JNEUROSCI.5487-07.2008
- 629 Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and  
630 overlapping representations of faces and objects in ventral temporal cortex. *Science*, 293, 2425–2430.
- 631 Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks.  
632 *Science*, 313(5786), 504–507.
- 633 Honey, C. J., Kötter, R., Breakspear, M., & Sporns, O. (2007). Network structure of cerebral cortex shapes  
634 functional connectivity on multiple time scales. *Proceedings of the National Academy of Science USA*, 104(24),  
635 10240–10245.
- 636 Honey, C. J., Thesen, T., Donner, T. H., Silbert, L. J., Carlson, C. E., Devinsky, O., . . . Hasson, U. (2012). Slow  
637 cortical dynamics and the accumulation of information over long timescales. *Neuron*, 76, 423-434. doi:  
638 10.1016/j.neuron.2012.08.011
- 639 Huth, A. G., de Heer, W. A., Griffiths, T. L., Theunissen, F. E., & Gallant, J. L. (2016). Natural speech reveals  
640 the semantic maps that tile human cerebral cortex. *Nature*, 532, 453–458.
- 641 Huth, A. G., Nisimoto, S., Vu, A. T., & Gallant, J. L. (2012). A continuous semantic space describes  
642 the representation of thousands of object and action categories across the human brain. *Neuron*, 76(6),  
643 1210-1224.
- 644 Jutten, C., & Herault, J. (1991). Blind separation of sources, part I: An adaptive algorithm based on  
645 neuromimetic architecture. *Signal Processing*, 24(1), 1–10.
- 646 Kamitani, Y., & Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nature*  
647 *Neuroscience*, 8, 679–685.
- 648 Landau, E. (1895). Zur relativen Wertbemessung der Turnierresultate. *Deutsches Wochenschach*, 11, 366–369.
- 649 Lee, C. S., Aly, M., & Baldassano, C. (2020). Anticipation of temporally structured events in the brain.  
650 *bioRxiv*, 10.1101/2020.10.14.338145.
- 651 Lee, D. D., & Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature*,  
652 401, 788–791.

- 653 Lerner, Y., Honey, C. J., Katkov, M., & Hasson, U. (2014). Temporal scaling of neural responses to compressed  
654 and dilated natural speech. *Journal of Neurophysiology*, 111, 2433–2444.
- 655 Lerner, Y., Honey, C. J., Silbert, L. J., & Hasson, U. (2011). Topographic mapping of a hierarchy of  
656 temporal receptive windows using a narrated story. *Journal of Neuroscience*, 31(8), 2906-2915. doi: 10.1523/  
657 JNEUROSCI.3684-10.2011
- 658 Liégeois, R., Li, J., Kong, R., Orban, C., De Ville, D. V., Ge, T., ... Yeo, B. T. T. (2019). Resting brain  
659 dynamics at different timescales capture distinct aspects of human behavior. *Nature Communications*,  
660 10(2317), doi.org/10.1038/s41467-019-10317-7.
- 661 Lin, J. (2009). Divergence measures based on the Shannon entropy. *IEEE Transactions on Information Theory*,  
662 37(1), 145–151.
- 663 Lohmann, G., Margulies, D. S., Horstmann, A., Pleger, B., Lepsien, J., Goldhahn, D., ... Turner, R. (2010).  
664 Eigenvector centrality mapping for analyzing connectivity patterns in fMRI data of the human brain.  
665 *PLoS One*, 5(4), e10232.
- 666 Lurie, D., Kessler, D., Bassett, D., Betzel, R., Breakspear, M., Keilholz, S., ... Calhoun, V. (2018). On the  
667 nature of time-varying functional connectivity in resting fMRI. *PsyArXiv*, doi.org/10.31234/osf.io/xtzre.
- 668 Mack, M. L., Preston, A. R., & Love, B. C. (2017). Medial prefrontal cortex compresses concept representations  
669 through learning. *bioRxiv*, doi.org/10.1101/178145.
- 670 Mairal, J., Ponce, J., Sapiro, G., Zisserman, A., & Bach, F. R. (2009). Supervised dictionary learning. *Advances*  
671 *in Neural Information Processing Systems*, 1033–1040.
- 672 Mairal, J. B., Bach, F., Ponce, J., & Sapiro, G. (2009). Online dictionary learning for sparse coding. *Proceedings*  
673 *of the 26th annual international conference on machine learning*, 689–696.
- 674 Manning, J. R., Ranganath, R., Norman, K. A., & Blei, D. M. (2014). Topographic factor analysis: a Bayesian  
675 model for inferring brain networks from neural data. *PLoS One*, 9(5), e94914.
- 676 Manning, J. R., Zhu, X., Willke, T. L., Ranganath, R., Stachenfeld, K., Hasson, U., ... Norman, K. A. (2018).  
677 A probabilistic approach to discovering dynamic full-brain functional connectivity patterns. *NeuroImage*,  
678 180, 243–252.
- 679 McInnes, L., Healy, J., & Melville, J. (2018). UMAP: Uniform manifold approximation and projection for  
680 dimension reduction. *arXiv*, 1802(03426).

- 681 McIntosh, A. R., & Jirsa, V. K. (2019). The hidden repertoire of brain dynamics and dysfunction. *Network*  
682 *Neuroscience*, doi.org/10.1162/netn\_a\_00107.
- 683 Mitchell, T., Shinkareva, S., Carlson, A., Chang, K., Malave, V., Mason, R., & Just, M. (2008). Predicting  
684 human brain activity associated with the meanings of nouns. *Science*, 320(5880), 1191.
- 685 Newman, M. E. J. (2005). A measure of betweenness centrality based on random walks. *Social Networks*, 27,  
686 39–54.
- 687 Newman, M. E. J. (2008). The mathematics of networks. *The New Palgrave Encyclopedia of Economics*, 2, 1–12.
- 688 Nishimoto, S., Vu, A., Naselaris, T., Benjamini, Y., Yu, B., & Gallant, J. (2011). Reconstructing visual  
689 experience from brain activity evoked by natural movies. *Current Biology*, 21, 1 - 6.
- 690 Nocedal, J., & Wright, S. J. (2006). *Numerical optimization*. New York: Springer.
- 691 Norman, K. A., Newman, E., Detre, G., & Polyn, S. M. (2006). How inhibitory oscillations can train neural  
692 networks and punish competitors. *Neural Computation*, 18, 1577–1610.
- 693 Opsahl, T., Agneessens, F., & Skvoretz, J. (2010). Node centrality in weighted networks: generalizing degree  
694 and shortest paths. *Social Networks*, 32, 245–251.
- 695 Park, H.-J., Friston, K. J., Pae, C., Park, B., & Razi, A. (2018). Dynamic effective connectivity in resting state  
696 fMRI. *NeuroImage*, 180, 594–608.
- 697 Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. *The London, Edinburgh,*  
698 *and Dublin Philosophical Magazine and Journal of Science*, 2, 559-572.
- 699 Pereira, F., Lou, B., Pritchett, B., Ritter, S., Gershman, S. J., Kanwisher, N., . . . Fedorenko, E. (2018). Toward  
700 a universal decoder of linguistic meaning from brain activation. *Nature Communications*, 9(963), 1–13.
- 701 Preti, M. G., Bolton, T. A. W., & Van De Ville, D. (2017). The dynamic functional connectome: state-of-the-art  
702 and perspectives. *NeuroImage*, 160, 41–54.
- 703 Rao, C. R. (1982). Diversity and dissimilarity coefficients: a unified approach. *Theoretical Population Biology*,  
704 21(1), 24–43.
- 705 Reimann, M. W., Nolte, M., Scolamiero, M., Turner, K., Perin, R., Chindemi, G., . . . Markram, H. (2017).  
706 Cliques of neurons bound into cavities provide a missing link between structure and function. *Frontiers*  
707 *in Computational Neuroscience*, 11(48), 1–16.

- 708 Ricotta, C., & Szeidl, L. (2006). Towards a unifying approach to diversity measures: Bridging the gap  
709 between the Shannon entropy and Rao's quadratic index. *Theoretical Population Biology*, 70(3), 237–243.
- 710 Roy, D. S., Park, Y.-G., Ogawa, S. K., Cho, J. H., Choi, H., Kamensky, L., ... Tonegawa, S. (2019). Brain-  
711 wide mapping of contextual fear memory engram ensembles supports the dispersed engram complex  
712 hypothesis. *bioRxiv*, doi.org/10.1101/668483.
- 713 Rubin, T. N., Kyoejo, O., Gorgolewski, K. J., Jones, M. N., Poldrack, R. A., & Yarkoni, T. (2017). Decoding  
714 brain activity using a large-scale probabilistic functional-anatomical atlas of human cognition. *PLoS  
715 Computational Biology*, 13(10), e1005649.
- 716 Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpreta-  
717 tions. *NeuroImage*, 52, 1059–1069.
- 718 Schreiber, T. (2000). Measuring information transfer. *Physical Review Letters*, 85(2), 461–464.
- 719 Shappell, H., Caffo, B. S., Pekar, J. J., & Lindquist, M. A. (2019). Improved state change estimation in  
720 dynamic functional connectivity using hidden semi-Markov models. *NeuroImage*, 191, 243–257.
- 721 Simony, E., & Chang, C. (2020). Analysis of stimulus-induced brain dynamics during naturalistic paradigms.  
722 *NeuroImage*, 216, 116461.
- 723 Simony, E., Honey, C. J., Chen, J., & Hasson, U. (2016). Dynamic reconfiguration of the default mode  
724 network during narrative comprehension. *Nature Communications*, 7(12141), 1–13.
- 725 Sizemore, A. E., Giusti, C., Kahn, A., Vettel, J. M., Betzel, R. F., & Bassett, D. S. (2018). Cliques and cavities  
726 in the human connectome. *Journal of Computational Neuroscience*, 44(1), 115–145.
- 727 Solomon, S. H., Medaglia, J. D., & Thompson-Schill, S. L. (2019). Implementing a concept network model.  
728 *Behavior Research Methods*, 51, 1717–1736.
- 729 Spearman, C. (1904). General intelligence, objectively determined and measured. *American Journal of  
730 Psychology*, 15, 201–292.
- 731 Sporns, O., & Honey, C. J. (2006). Small worlds inside big brains. *Proceedings of the National Academy of  
732 Science USA*, 103(51), 19219–19220.
- 733 Thompson, W. H., Richter, C. G., Plavén-Sigray, P., & Fransson, P. (2018). Simulations to benchmark  
734 time-varying connectivity methods for fMRI. *PLoS Computational Biology*, 14(5), e1006196.

- 735 Tipping, M. E., & Bishop, C. M. (1999). Probabilistic principal component analysis. *Journal of Royal Statistical Society, Series B*, 61(3), 611–622.
- 736
- 737 Toker, D., & Sommer, F. T. (2019). Information integration in large brain networks. *PLoS Computational Biology*, 15(2), e1006807.
- 738
- 739 Tong, F., & Pratte, M. S. (2012). Decoding patterns of human brain activity. *Annual Review of Psychology*, 63, 483–509.
- 740
- 741 Turk-Browne, N. B. (2013). Functional interactions as big data in the human brain. *Science*, 342, 580–584.
- 742 van der Maaten, L. J. P., & Hinton, G. E. (2008). Visualizing high-dimensional data using t-SNE. *Journal of Machine Learning Research*, 9, 2579-2605.
- 743
- 744 Vidaurre, D., Abeysuriya, R., Becker, R., Quinn, A. J., Alfaro-Almagro, F., Smith, S. M., & Woolrich, M. W. (2018). Discovering dynamic brain networks from big data in rest and task. *NeuroImage*, 180, 646–656.
- 745
- 746 Zar, J. H. (2010). *Biostatistical analysis*. Prentice-Hall/Pearson.
- 747 Zheng, M., Allard, A., Hagmann, P., & Serrano, M. A. (2019). Geometric renormalization unravels self-similarity of the multiscale human connectome. *arXiv*, 1904.11793.
- 748
- 749 Zou, Y., Donner, R. V., Marwan, N., Donges, J. F., & Kurths, J. (2019). Complex network approaches to nonlinear time series analysis. *Physics Reports*, 787, 1–97.
- 750