

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

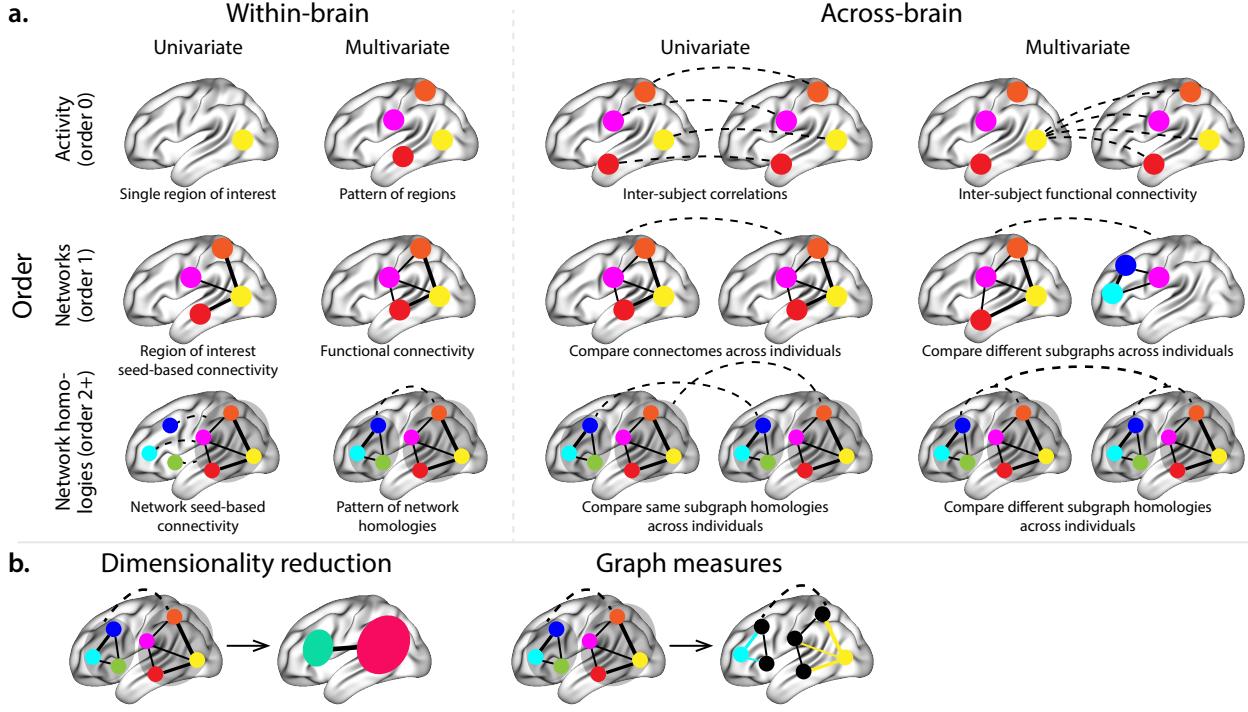
4      November 4, 2020

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

20     **Introduction**

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



**Figure 1: Neural patterns. a. A space of neural 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 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.

multivariate activity patterns versus dynamic correlations between neural activity patterns (e.g., Fong et al., 2019; Manning et al., 2018). An emerging theme in this literature is that cognition is mediated by dynamic interactions between brain structures (Bassett et al., 2006; Demertzi et al., 2019; Lurie et al., 2018; Mack et al., 2017; Preti et al., 2017; Solomon et al., 2019; Sporns & Honey, 2006; Turk-Browne, 2013; Zou et al., 2019). Studies of the neural code to date have primarily focused on univariate or multivariate neural patterns (for review see Norman et al., 2006), or (more recently) on patterns of dynamic first-order correlations (i.e., interactions between pairs of brain structures; Demertzi et al., 2019; Fong et al., 2019; Lurie et al., 2018; Manning et al., 2018; Preti et al., 2017; Zou et al., 2019). What might the future of this line of work hold? For example, is the neural code implemented through higher-order interactions between brain structures (e.g., see Reimann et al., 2017)? Second-order correlations reflect *homologous* patterns of correlation.

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

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

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

65 To gain insights into the above aspects of the neural code, we developed a computational framework  
66 for estimating dynamic high-order correlations in timeseries data. This framework provides an important  
67 advance, in that it enables us to examine patterns of higher-order correlations that are computationally  
68 intractable to estimate via conventional methods. Given a multivariate timeseries, our framework provides  
69 timepoint-by-timepoint estimates of the first-order correlations, second-order correlations, and so on. Our  
70 approach combines a kernel-based method for computing dynamic correlations in timeseries data with a di-  
71 mensionality reduction step (Fig. 1b) that projects the resulting dynamic correlations into a low-dimensional  
72 space. We explored two dimensionality reduction approaches: principle components analysis (PCA; Pear-  
73 son, 1901), which preserves an approximately invertible transformation back to the original data (e.g., this

74 follows related approaches taken by Gonzalez-Castillo et al., 2019; McIntosh & Jirsa, 2019; Toker & Som-  
75 mer, 2019); and a second non-invertible algorithm that explored patterns in eigenvector centrality (Landau,  
76 1895). This latter approach characterizes correlations between each feature dimension's relative *position* in  
77 the network in favor of the specific activity histories of different features (also see Betzel et al., 2019; Reimann  
78 et al., 2017; Sizemore et al., 2018).

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

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

## 101 Results

102 We sought to understand whether high-level cognition is reflected in dynamic patterns of high-order  
103 correlations. To that end, we developed a computational framework for estimating the dynamics of stimulus-  
104 driven high-order correlations in multivariate timeseries data (see *Dynamic inter-subject functional connectivity*

105 (*DISFC*) and *Dynamic higher-order correlations*). We evaluated the efficacy of this framework at recovering  
106 known patterns in several synthetic datasets (see *Synthetic data: simulating dynamic first-order correlations*  
107 and *Synthetic data: simulating dynamic higher-order correlations*). We then applied the framework to a public  
108 fMRI dataset collected as participants listened to an auditorily presented story, listened to a temporally  
109 scrambled version of the story, or underwent a resting state scan (see *Functional neuroimaging data collected*  
110 *during story listening*). We used the relative decoding accuracies of classifiers trained on different sets of  
111 neural features to estimate which types of features reflected ongoing cognitive processing.

## 112 **Recovering known dynamic correlations from synthetic data**

### 113 **Recovering dynamic first-order correlations**

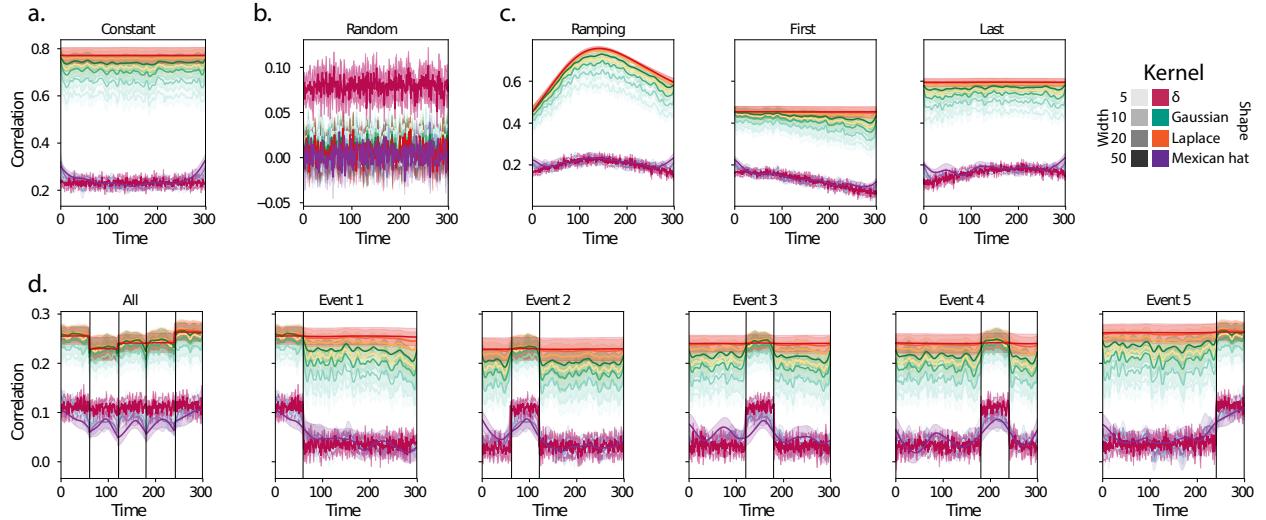
114 We generated synthetic datasets that differed in how the underlying first-order correlations changed over  
115 time. For each dataset, we applied Equation 4 with a variety of kernel shapes and widths. We assessed how  
116 well the true underlying correlations at each timepoint matched the recovered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

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

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

### 128 **Recovering dynamic higher-order correlations**

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



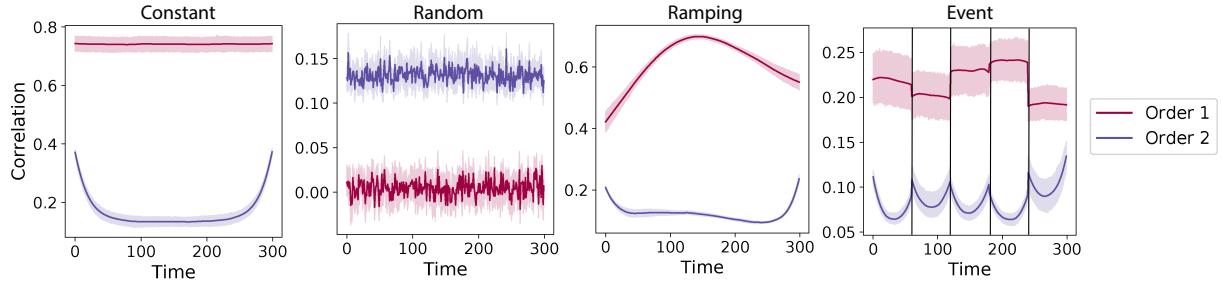
**Figure 2: Recovering known dynamic first-order correlations from synthetic data.** Each panel displays the average correlations between the vectorized upper triangles of the recovered correlation matrix at each timepoint and either the true underlying correlation at each timepoint or a reference correlation 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: simulating dynamic first-order correlations*.

**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.



**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 correlation matrix and the simulated correlation matrix at each timepoint 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 results in this figure used a Laplace kernel (width = 20). **a. Constant correlations.** These first and second order dynamic correlation were simulated from datasets have a stable (unchanging) underlying correlation matrix. **b. Random correlations.** These first and second order dynamic correlation were simulated from datasets are generated using a new independently drawn correlation matrix at each new timepoint. **c. Ramping correlations.** These first and second order dynamic correlation were simulated from datasets are generated by smoothly varying the underlying correlations between the randomly drawn correlation matrices at the first and last timepoints and displays the correlations between the recovered dynamic correlations and the simulalted correlation first and second order correlation matrices. **d. Event-based correlations.** These first and second order dynamic correlation were simulated from datasets are each generated using five randomly drawn correlation matrices that each remain stable for a fifth of the total timecourse.

135 recovered dynamic correlations matched the dynamic correlations we had embedded into the datasets.

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

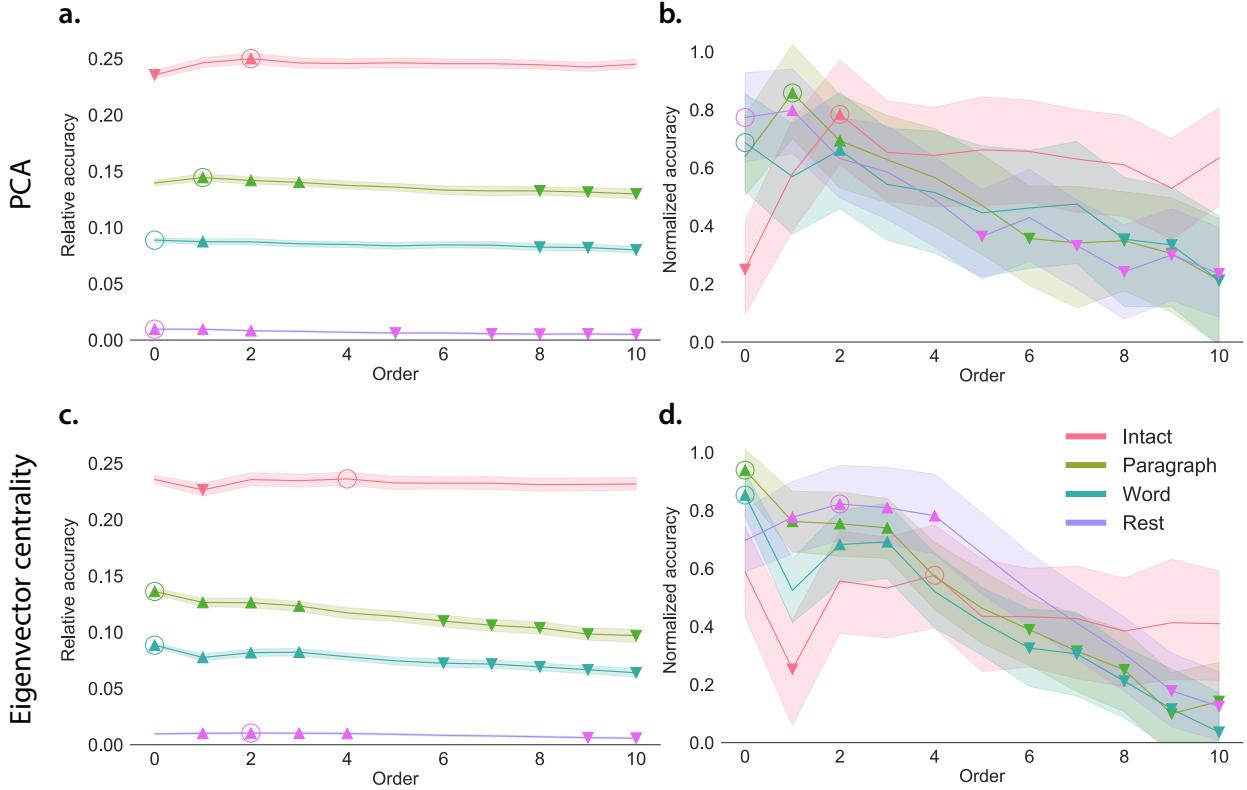
144 Taken together, our explorations using synthetic data indicated that we are able to partially, but not  
 145 perfectly, recover ground truth dynamic first-order and second-order correlations. This suggests that our  
 146 modeling approach provides a meaningful (if noisy) estimate of high-order correlations. We next turned  
 147 to analyses of human fMRI data to examine whether the recovered dynamics might reflect the dynamics of  
 148 human cognition during a naturalistic story-listening task.

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

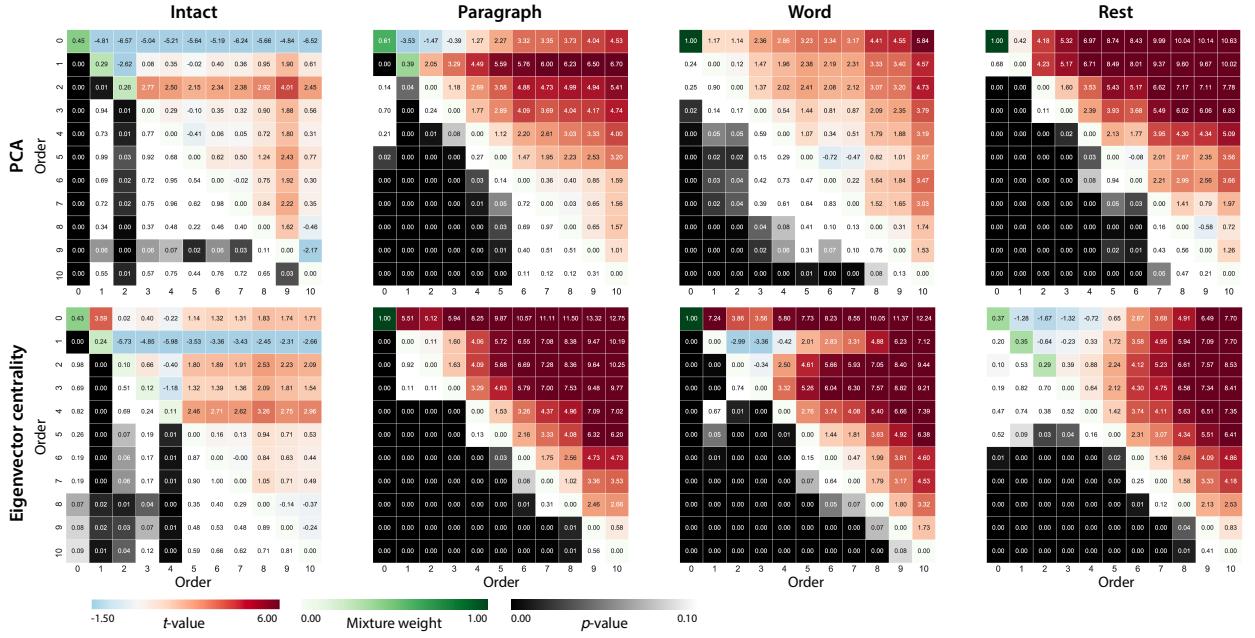
150 We used across-participant temporal decoders to identify cognitively relevant neural patterns in fMRI data  
151 (see *Forward inference and decoding accuracy*). The dataset we examined (collected by Simony et al., 2016)  
152 comprised four experimental conditions that exposed participants to stimuli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

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

175 Our approach to estimating dynamic high-order correlations entails mapping the high-dimensional  
176 feature space of correlations (represented by a  $T$  by  $O(K^2)$  matrix) onto a lower-dimensional feature space  
177 (represented by a  $T$  by  $K$  matrix). We carried out two sets of analyses that differed in how this mapping was  
178 computed. The first set of analyses used PCA to find a low-dimensional embedding of the original dynamic  
179 correlation matrices (Fig. 4a,b). The second set of analyses characterized correlations in dynamics of each  
180 feature's eigenvector centrality, but did not preserve the underlying activity dynamics (Fig. 4c,d).



**Figure 4: 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:  $\frac{1}{400}$ ). 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 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 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. 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 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.

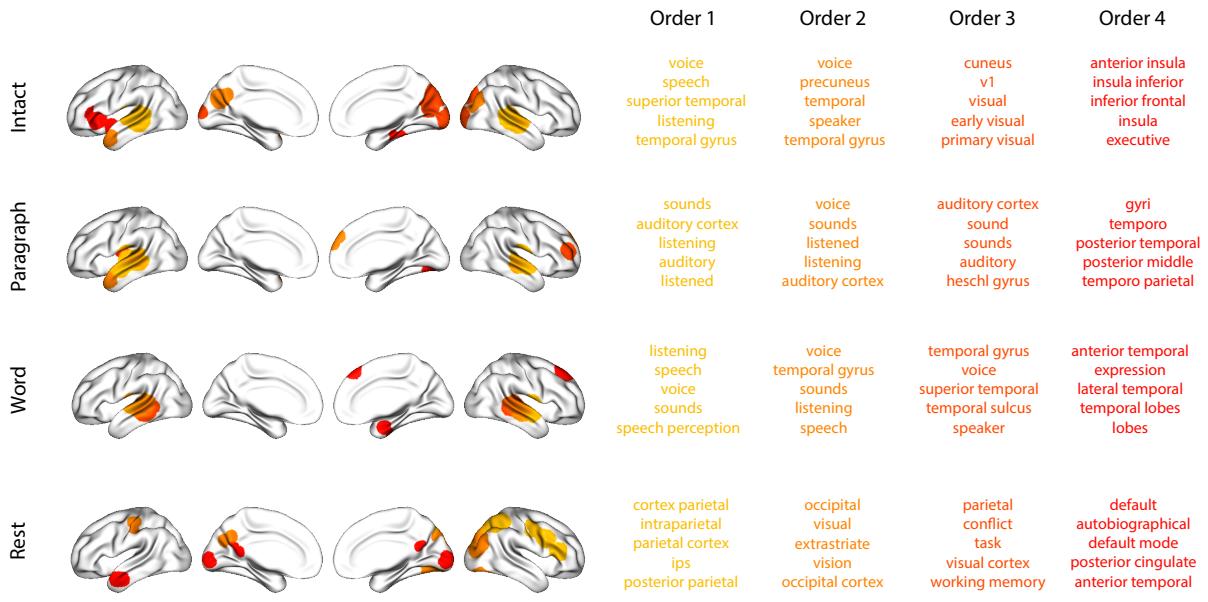


**Figure 5: Statistical summary of decoding accuracies for different neural features.** Each column of matrices displays decoding results for one experimental condition (intact, paragraph, word, and rest). We considered dynamic activity patterns (order 0) and dynamic correlations at different orders (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 matrix, 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, in a decoder that included all feature types (see *Feature weighting and testing*).

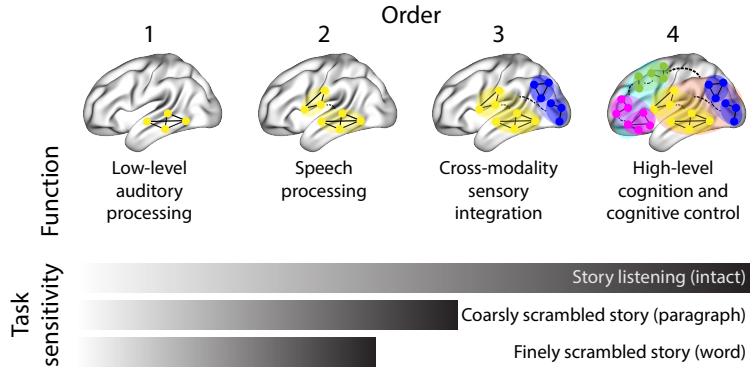
Both sets of temporal decoding analyses yielded qualitatively similar results for the auditory (non-rest) conditions of the experiment (Fig. 4: pink, green, and teal lines; Fig. 5: three leftmost columns). The highest decoding accuracy for participants who listened to the intact (unscrambled) story was 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 slightly above chance. We speculate that the decoders might have picked up on attentional drift, boredom, or tiredness; we hypothesize that these all increased throughout the resting state scan. The decoders might be picking up on aspects of these loosely defined cognitive states that are common across individuals. The PCA-based approach achieved the highest resting state decoding accuracy using order zero features (non-correlational, activation-based), whereas the eigenvector centrality-based approach achieved the highest resting state decoding accuracy using second-order correlations. Taken together, these analyses indicate that high-level cognitive processing (while listening to the intact story) is reflected in the dynamics of high-order correlations in brain activity, whereas lower-level cognitive processing (while listening to scrambled versions of the story that lack rich meaning) is reflected in the dynamics of lower-order correlations and non-correlational activity dynamics. Further, these patterns are associated both with the underlying activity patterns (characterized using PCA) and also with the changing relative positions that different brain areas occupy in their associated networks (characterized using eigenvector centrality).

Having established that patterns of high-order correlations are informative to decoders, we next wondered which specific networks of brain regions contributed most to these patterns. As a representative example, we selected the kernel parameters that yielded decoding accuracies that best matched the average accuracies across all of the kernel parameters we examined. Using Figure 4c as a template, the best-matching kernel was a Laplace kernel with a width of 50 (Fig. 9d). We used this kernel to compute a single  $K$  by  $K$   $n^{\text{th}}$ -order DISFC matrix for each experimental condition. We then used Neurosynth (Rubin et al., 2017) to compute the terms most highly associated with the most strongly correlated pairs of regions in each of these matrices (Fig. 6; see *Reverse inference*).

For all of the story listening conditions (intact, paragraph, and word), we found that first- and second-order correlations were most strongly associated with auditory and speech processing areas. During intact story listening, third-order correlations reflected integration with visual areas, and fourth-order correlations reflected integration with areas associated with high-level cognition and cognitive control, such as the ventrolateral prefrontal cortex. However, during listening to temporally scrambled stories,



**Figure 6: Top terms associated with the endpoints of the strongest correlations.** 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. S1; paragraph: Fig. S2; word: Fig. S3; rest: Fig. S4).



**Figure 7: Proposed high-order network dynamics underlying high-level cognition during story listening.** 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.

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

## Discussion

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

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

An open question not addressed by our study pertains to how different structures integrate incoming information with different time constants. For example, one line of work suggests that the cortical surface comprises a structured map such that nearby brain structures process incoming information at similar timescales. Low-level sensory areas integrate information relatively quickly, whereas higher-level regions integrate information relatively slowly (Baldassano et al., 2017; Chien & Honey, 2019; Hasson et al., 2015, 2008; Honey et al., 2012; Lerner et al., 2014, 2011). A similar hierarchy appears to play a role in predicting future events (C. S. Lee et al., 2020). Other related work in human and mouse brains indicates that the temporal response profile of a given brain structure may relate to how strongly connected that structure is with other brain areas (Fallon et al., 2019). Further study is needed to understand the role of temporal integration at different scales of network interaction, and across different anatomical structures. Importantly, our analyses do not speak to the physiological basis of higher-order dynamics, and could reflect nonlinearities, chaotic patterns, non-stationarities, or multistability, etc. However, our decoding analyses do indicate that higher-order dynamics are consistent across individuals, and therefore unlikely to be driven by non-stimulus-driven dynamics which are unlikely to be similar across individuals.

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

Identifying high-order network dynamics associated with high-level cognition required several important methods advances. First, we used kernel-based dynamic correlations to extended the notion of (static) inter-subject functional connectivity (Simony et al., 2016) to a dynamic measure of inter-subject functional connectivity (DISFC) that does not rely on sliding windows, and that may be computed at individual

267 timepoints. This allowed us to precisely characterize stimulus-evoked network dynamics that were similar  
268 across individuals. Second, we developed a computational framework for efficiently and scalably estimat-  
269 ing high-order dynamic correlations. Our approach uses dimensionality reduction algorithms and graph  
270 measures to obtain low-dimensional embeddings of patterns of network dynamics. Third, we developed  
271 an analysis framework for identifying robust decoding results by carrying out our analyses using a range  
272 of parameter values and then identifying which results were robust to specific parameter choices.

273 **Concluding remarks**

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

279 **Methods**

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

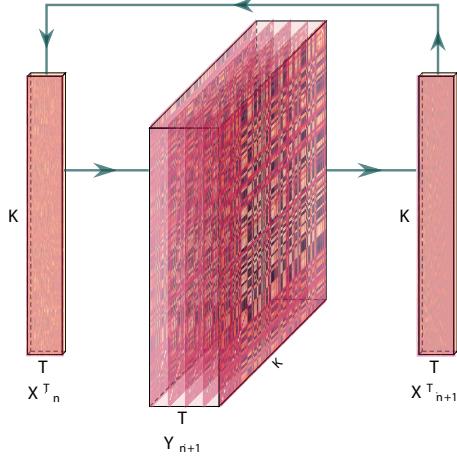


Figure 8: **Estimating dynamic high-order correlations.** Given a  $T$  by  $K$  matrix of multivariate timeseries data,  $\mathbf{X}_n$  (where  $n \in \mathbb{N}, n \geq 0$ ), we use Equation 4 to compute a timeseries of  $K$  by  $K$  correlation matrices,  $\mathbf{Y}_{n+1}$ . We then approximate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

296 patterns of  $n^{\text{th}}$ -order correlations.

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

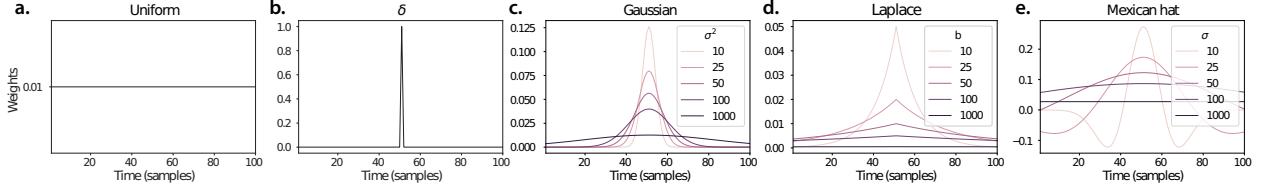
301 **Kernel-based approach for computing dynamic correlations**

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):

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

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

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



**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.

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

305

Given a kernel function  $\kappa_t(\cdot)$  for timepoint  $t$ , evaluated at timepoints  $\tau \in [1, \dots, T]$ , we can update the static correlation formula in Equation 1 to estimate the *instantaneous correlation* at timepoint  $t$ :

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

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

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

306 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  
 307 the kernel  $\kappa_t$ . We evaluate Equation 4 in turn for each pair of columns in  $\mathbf{X}$  and for kernels centered on each  
 308 timepoint in the timeseries, respectively, to obtain a  $T$  by  $K$  by  $K$  timeseries of dynamic correlations,  $\mathbf{Y}$ . For  
 309 convenience, we then reshape the upper triangles and diagonals of each timepoint's symmetric correlation  
 310 matrix into a row vector to obtain an equivalent  $T$  by  $(\frac{K^2-K}{2} + K)$  matrix.

<sup>311</sup> **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,  $\bar{\mathbf{X}}_n^{\setminus p}$ :

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

<sup>312</sup> 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  
<sup>313</sup> matrix  $\bar{\mathbf{C}}$  provides a time-varying extension of the ISFC approach developed by Simony et al. (2016).

<sup>314</sup> **Low-dimensional representations of dynamic correlations**

<sup>315</sup> 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  
<sup>316</sup> to computing a  $T$  by  $K$  low-dimensional representation of those correlations,  $\mathbf{X}_n$ . The first approach uses  
<sup>317</sup> dimensionality reduction algorithms to project  $\mathbf{Y}_n$  onto a  $K$ -dimensional space. The second approach uses

318 graph measures to characterize the relative positions of each feature ( $k \in [1, \dots, K]$ ) in the network defined  
319 by the correlation matrix at each timepoint.

320 **Dimensionality reduction-based approaches to computing  $\mathbf{X}_n$**

321 The modern toolkit of dimensionality reduction algorithms include Principal Components Analysis (PCA;  
322 Pearson, 1901), Probabilistic PCA (PPCA; Tipping & Bishop, 1999), Exploratory Factor Analysis (EFA;  
323 Spearman, 1904), Independent Components Analysis (ICA; Comon et al., 1991; Jutten & Herault, 1991),  
324  $t$ -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  
325 & 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.,  
326 2011), dictionary learning (J. Mairal et al., 2009; J. B. Mairal et al., 2009), and deep auto-encoders (Hinton  
327 & Salakhutdinov, 2006), among others. While complete characterizations of each of these algorithms is  
328 beyond the scope of the present manuscript, the general intuition driving these approaches is to compute  
329 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  
330 approaches place different constraints on what properties  $\mathbf{X}$  must satisfy and which aspects of the data are  
331 compared (and how) in order to optimize how well  $\mathbf{X}$  approximates  $\mathbf{Y}$ .

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

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

344 The above dimensionality reduction approaches to approximating a given  $\mathbf{Y}_n$  with a lower-dimensional  
345  $\mathbf{X}_n$  preserve a (potentially recombined and transformed) mapping back to the original data in  $\mathbf{X}_0$ . We also  
346 explore graph measures that instead characterize each feature’s relative *position* in the broader network of  
347 interactions and connections. To illustrate the distinction between the two general approaches we explore,

348 suppose a network comprises nodes  $A$  and  $B$ , along with several other nodes. If  $A$  and  $B$  exhibit uncorrelated  
349 activity patterns, then by definition the functional connection (correlation) between them will be close to  
350 0. However, if  $A$  and  $B$  each interact with *other* nodes in similar ways, we might attempt to capture those  
351 similarities between  $A$ 's and  $B$ 's interactions with those other members of the network.

352 In general, graph measures take as input a matrix of interactions (e.g., using the above notation, a  $K$   
353 by  $K$  correlation matrix or binarized correlation matrix reconstituted from a single timepoint's row of  $\mathbf{Y}$ ),  
354 and return as output a set of  $K$  measures describing how each node (feature) sits within that correlation  
355 matrix with respect to the rest of the population. Widely used measures include betweenness centrality (the  
356 proportion of shortest paths between each pair of nodes in the population that involves the given node  
357 in question; e.g., Barthélémy, 2004; Freeman, 1977; Geisberger et al., 2008; Newman, 2005; Opsahl et al.,  
358 2010); diversity and dissimilarity (characterizations of how differently connected a given node is from others  
359 in the population; e.g., Lin, 2009; Rao, 1982; Ricotta & Szeidl, 2006); eigenvector centrality and pagerank  
360 centrality (measures of how influential a given node is within the broader network; e.g., Bonacich, 2007;  
361 Halu et al., 2013; Lohmann et al., 2010; Newman, 2008); transfer entropy and flow coefficients (a measure of  
362 how much information is flowing from a given node to other nodes in the network; e.g., Honey et al., 2007;  
363 Schreiber, 2000);  $k$ -coreness centrality (a measure of the connectivity of a node within its local subgraph; e.g.,  
364 Alvarez-Hamelin et al., 2005; Christakis & Fowler, 2010); within-module degree (a measure of how many  
365 connections a node has to its close neighbors in the network; e.g., Rubinov & Sporns, 2010); participation  
366 coefficient (a measure of the diversity of a node's connections to different subgraphs in the network; e.g.,  
367 Rubinov & Sporns, 2010); and subgraph centrality (a measure of a node's participation in all of the network's  
368 subgraphs; e.g., Estrada & Rodríguez-Velázquez, 2005); among others.

369 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  
370 characterizes the corresponding graph properties of each column. This results in a new  $T$  by  $K$  matrix,  
371  $\mathbf{X}_n$ , that reflects how the features reflected in the columns of  $\mathbf{X}_{n-1}$  participate in the network during each  
372 timepoint (row).

### 373 **Dynamic higher-order correlations**

374 Because  $\mathbf{X}_n$  has the same shape as the original data  $\mathbf{X}_0$ , approximating  $\mathbf{Y}_n$  with a lower-dimensional  $\mathbf{X}_n$   
375 enables us to estimate high-order dynamic correlations in a scalable way. Given a  $T$  by  $K$  input matrix, the  
376 output of Equation 4 requires  $O(TK^2)$  space to store. Repeated applications of Equation 4 (i.e., computing  
377 dynamic correlations between the columns of the outputted dynamic correlation matrix) each require  
378 exponentially more space; in general the  $n^{\text{th}}$ -order dynamic correlations of a  $T$  by  $K$  timeseries occupies

379  $O(TK^{2^n})$  space. However, when we approximate or summarize the output of Equation 4 with a  $T$  by  $K$  matrix  
380 (as described above), it becomes feasible to compute even very high-order correlations in high-dimensional  
381 data. Specifically, approximating the  $n^{\text{th}}$ -order dynamic correlations of a  $T$  by  $K$  timeseries requires only  
382  $O(TK^2)$  additional space— the same as would be required to compute first-order dynamic correlations. In  
383 other words, the space required to store  $n + 1$  multivariate timeseries reflecting up to  $n^{\text{th}}$  order correlations  
384 in the original data scales linearly with  $n$  using our approach (Fig. 8).

385 **Data**

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

396 **Synthetic data: simulating dynamic first-order correlations**

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

**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,  $\Sigma_m$ :

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

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

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

407 **Random.** We generated a second set of 10 synthetic datasets whose observations at each timepoint were  
408 drawn from a Gaussian distribution with a new randomly constructed (using Eqn. 11) covariance matrix.  
409 Because each timepoint's covariance matrix was drawn independently from the covariance matrices for all  
410 other timepoints, these datasets provided a test of reconstruction accuracy in the absence of any meaningful  
411 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

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

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

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

420 **Synthetic data: simulating dynamic high-order correlations**

421 We developed an iterative procedure for constructing timeseries data that exhibits known dynamic high-  
422 order correlations. The procedure builds on our approach to generating dynamic first-order correlations.  
423 Essentially, once we generate a timeseries with known first-order correlations, we can use the known first-  
424 order correlations as a template to generate a new timeseries of second-order correlations. In turn, we can  
425 generate a timeseries of third-order correlations from the second-order correlations, and so on. In general,  
426 we can generate order  $n$  correlations given a timeseries of order  $n - 1$  correlations, for any  $n > 1$ . Finally,  
427 given the order  $n$  timeseries, we can reverse the preceding process to generate an order  $n - 1$  timeseries, an  
428 order  $n - 2$  order timeseries, and so on, until we obtain an order 0 timeseries of simulated data that reflects  
429 the chosen high-order dynamics.

The central mathematical operations in our procedure are two 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

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

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

430 We can then use repeated applications of Equations 15 and 16 in order to obtain a synthetic dataset.

431 When the template first-order correlations are constructed to exhibit different temporal profiles (e.g.,  
432 using the constant, random, ramping, and event procedures described above), the resulting high-order  
433 correlations and synthetic data will exhibit the same category of temporal profile.

434 Following our approach to generating synthetic data exhibiting known first-order correlations, we  
435 constructed a total of 40 additional multivariate timeseries, collectively reflecting a total of 4 qualitatively  
436 different patterns of dynamic correlations (i.e., 10 datasets reflecting each type of dynamic pattern: constant,  
437 random, ramping, and event).

438 **Functional neuroimaging data collected during story listening**

439 We examined an fMRI dataset collected by Simony et al. (2016) that the authors have made publicly available  
440 at [arks.princeton.edu/ark:/88435/dsp015d86p269k](https://arks.princeton.edu/ark:/88435/dsp015d86p269k). The dataset comprises neuroimaging data collected as  
441 participants listened to an audio recording of a story (intact condition; 36 participants), listened to temporally  
442 scrambled recordings of the same story (17 participants in the paragraph-scrambled condition listened to  
443 the paragraphs in a randomized order and 36 in the word-scrambled condition listened to the words in a  
444 randomized order), or lay resting with their eyes open in the scanner (rest condition; 36 participants). Full  
445 neuroimaging details may be found in the original paper for which the data were collected (Simony et al.,  
446 2016).

447 **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).

457 **Temporal decoding**

458 We sought to identify neural patterns that reflected participants' ongoing cognitive processing of incoming  
459 stimulus information. As reviewed by Simony et al. (2016), one way of homing in on these stimulus-driven  
460 neural patterns is to compare activity patterns across individuals (e.g., using ISFC analyses). In particular,  
461 neural patterns will be similar across individuals to the extent that the neural patterns under consideration  
462 are stimulus-driven, and to the extent that the corresponding cognitive representations are reflected in  
463 similar spatial patterns across people (also see Simony & Chang, 2020). Following this logic, we used an  
464 across-participant temporal decoding test developed by Manning et al. (2018) to assess the degree to which  
465 different neural patterns reflected ongoing stimulus-driven cognitive processing across people (Fig. 10). The  
466 approach entails using a subset of the data to train a classifier to decode stimulus timepoints (i.e., moments  
467 in the story participants listened to) from neural patterns. We use decoding (forward inference) accuracy

468 on held-out data, from held-out participants, as a proxy for the extent to which the inputted neural patterns  
469 reflected stimulus-driven cognitive processing in a similar way across individuals.

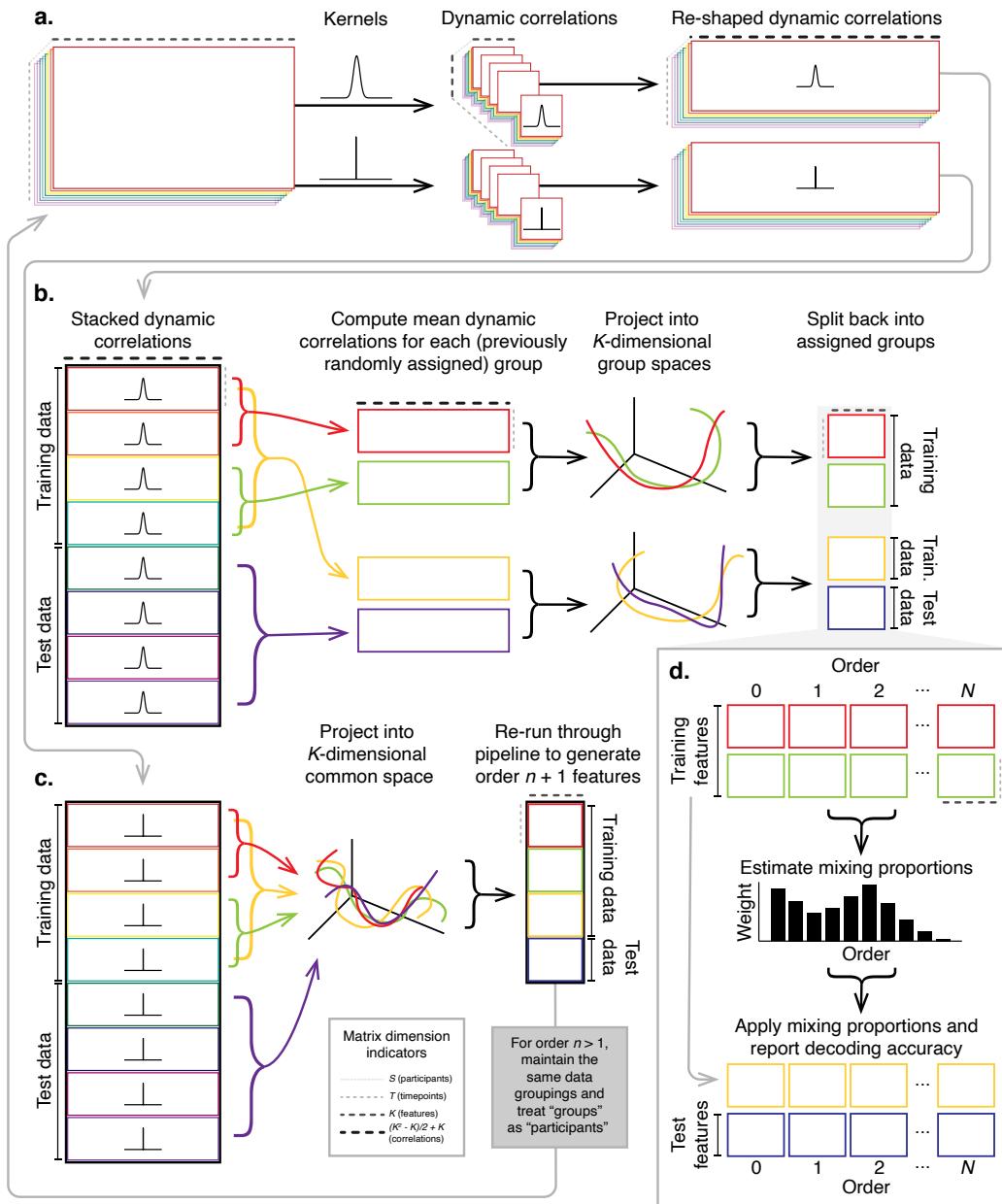
470 **Forward inference and decoding accuracy**

471 We used an across-participant correlation-based classifier to decode which stimulus timepoint matched  
472 each timepoint's neural pattern (Fig. 10). We first divided the participants into two groups: a template group,  
473  $\mathcal{G}_{\text{template}}$  (i.e., training data), and a to-be-decoded group,  $\mathcal{G}_{\text{decode}}$  (i.e., test data). We used Equation 7 to  
474 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  
475  $\bar{\mathbf{C}}_{\text{template}}$  and  $\bar{\mathbf{C}}_{\text{decode}}$  to form a number-of-timepoints by number-of-timepoints decoding matrix,  $\Lambda$ . In this  
476 way, the rows of  $\Lambda$  reflected timepoints from the template group, while the columns reflected timepoints  
477 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  
478  $\bar{\mathbf{C}}_{\text{template}}$  with which it was most highly correlated. We then repeated this decoding procedure, but using  
479  $\mathcal{G}_{\text{decode}}$  as the template group and  $\mathcal{G}_{\text{template}}$  as the to-be-decoded group. Given the true timepoint labels (for  
480 each group), we defined the *decoding accuracy* as the average proportion of correctly decoded timepoints,  
481 across both groups. We defined the *relative decoding accuracy* as the difference between the decoding accuracy  
482 and chance accuracy (i.e.,  $\frac{1}{T}$ ).

483 **Feature weighting and testing**

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

487 One challenge to fairly evaluating high-order correlations is that if the kernel used in Equation 4 is  
488 wider than a single timepoint, each repeated application of the equation will result in further temporal  
489 blur. Because our primary assessment metric is temporal decoding accuracy, this unfairly biases against  
490 detecting meaningful signal in higher-order correlations (relative to lower-order correlations). We attempted  
491 to mitigate temporal blur in estimating each  $\mathbf{X}_n$  by using a Dirac  $\delta$  function kernel (which places all of its  
492 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  
493 then used a new (potentially wider, as described below) kernel to compute  $\mathbf{X}_n$  from  $\mathbf{X}_{n-1}$ . In this way,  
494 temporal blurring was applied only in the last step of computing  $\mathbf{X}_n$ . We note that, because each  $\mathbf{X}_n$  is a  
495 low-dimensional representation of the corresponding  $\mathbf{Y}_n$ , the higher-order correlations we estimated reflect  
496 true correlations in the data with lower-fidelity than estimates of lower-order correlations. Therefore, even  
497 after correcting for temporal blurring, our approach is still biased against finding meaningful signal in



**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.

498 higher-order correlations.

499 After computing each  $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_{n-1}$  for each participant, we divided participants into two equally sized  
500 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  $\mathcal{G}_{\text{train}_1}$   
501 and  $\mathcal{G}_{\text{train}_2}$ . We then computed  $\Lambda$  (temporal correlation) matrices for each type of neural feature, using  $\mathcal{G}_{\text{train}_1}$   
502 and  $\mathcal{G}_{\text{train}_2}$ . This resulted in  $n + 1$   $\Lambda$  matrices (one for the original timeseries of neural activations, and one  
503 for each of  $n$  orders of dynamic correlations). Our objective was to find a set of weights for each of these  
504  $\Lambda$  matrices such that the weighted average of the  $n + 1$  matrices yielded the highest decoding accuracy.  
505 We used quasi-Newton gradient ascent (Nocedal & Wright, 2006), using decoding accuracy (for  $\mathcal{G}_{\text{train}_1}$  and  
506  $\mathcal{G}_{\text{train}_2}$ ) as the objective function to be maximized, to find an optimal set of training data-derived weights,  
507  $\phi_{0,1,\dots,n}$ , where  $\sum_{i=0}^n \phi_i = 1$  and where  $\phi_i \geq 0 \forall i \in [0, 1, \dots, n]$ .

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

512 We used a permutation-based procedure to form stable estimates of decoding accuracy for each set of  
513 neural features. In particular, we computed the decoding accuracy for each of 10 random group assignments  
514 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  
515 features.

## 516 Identifying robust decoding results

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

527 To identify decoding results that were robust to specific classifier parameter choices, we repeated our  
528 decoding analyses after substituting into Equation 4 each of a variety of kernel shapes and widths. We

529 examined Gaussian (Fig. 9c), Laplace (Fig. 9d), and Mexican Hat (Fig. 9e) kernels, each with widths of 5, 10,  
530 20, and 50 samples. We then report the average decoding accuracies across all of these parameter choices.  
531 This enabled us to (partially) factor out performance characteristics that were parameter-dependent, within  
532 the set of parameters we examined.

### 533 Reverse inference

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

## 545 Acknowledgements

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

## 552 Author contributions

553 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:  
554 L.L.W.O. and J.R.M.

## 555 References

- 556 Allen, E. A., Damaraju, E., Plis, S. M., Erhardt, E. B., Eichele, T., & Calhoun, V. D. (2012). Tracking  
557 whole-brain connectivity dynamics in the resting state. *Cerebral Cortex*, 24(3), 663–676.
- 558 Alvarez-Hamelin, I., Dall'Asta, L., Barrat, A., & Vespignani, A. (2005). *k*-corr decomposition: a tool for the  
559 visualiztion of large scale networks. *arXiv*, cs/0504107v2.
- 560 Baldassano, C., Chen, J., Zadbood, A., Pillow, J. W., Hasson, U., & Norman, K. A. (2017). Discovering event  
561 structure in continuous narrative perception and memory. *Neuron*, 95(3), 709–721.
- 562 Barthélémy, M. (2004). Betweenness centrality in large complex networks. *European Physical Journal B*, 38,  
563 163–168.
- 564 Bassett, D., Meyer-Lindenberg, A., Achard, S., Duke, T., & Bullmore, E. (2006, December). Adaptive  
565 reconfiguration of fractal small-world human brain functional networks. *Proceedings of the National  
566 Academy of Sciences, USA*, 103(51), 19518–23.
- 567 Beaty, R. E., Benedek, M., Silvia, P. J., & Schacter, D. L. (2016). Creative cognition and brain network  
568 dynamics. *Trends in Cognitive Sciences*, 20(2), 87–95.
- 569 Betzel, R. F., Byrge, L., Esfahlani, F. Z., & Kennedy, D. P. (2019). Temporal fluctuations in the brain's modular  
570 architecture during movie-watching. *bioRxiv*, doi.org/10.1101/750919.
- 571 Bonacich, P. (2007). Some unique properties of eigenvector centrality. *Social Networks*, 29(4), 555–564.
- 572 Bullmore, E., & Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and  
573 functional systems. *Nature Reviews Neuroscience*, 10(3), 186–198.
- 574 Capota, M., Turek, J., Chen, P.-H., Zhu, X., Manning, J. R., Sundaram, N., . . . Shin, Y. S. (2017). *Brain imaging  
575 analysis kit*. Retrieved from <https://doi.org/10.5281/zenodo.59780>
- 576 Chang, C., & Glover, G. H. (2010). Time-frequency dynamics of resting-state brain connectivity measured  
577 with fMRI. *NeuroImage*, 50, 81–98.
- 578 Chien, H.-Y. S., & Honey, C. J. (2019). Constructing and forgetting temporal context in the human cerebral  
579 cortex. *bioRxiv*, doi.org/10.1101/761593.
- 580 Christakis, N. A., & Fowler, J. H. (2010). Social network sensors for early detection of contagious outbreaks.  
581 *PLoS One*, 5(9), e12948.

- 582 Combrisson, E., Vallat, R., O'Reilly, C., Jas, M., Pascarella, A., l Saive, A., ... Jerbi, K. (2019). Visbrain: a  
583 multi-purpose GPU-accelerated open-source suite for multimodal brain data visualization. *Frontiers in*  
584 *Neuroinformatics*, 13(14), 1–14.
- 585 Comon, P., Jutten, C., & Herault, J. (1991). Blind separation of sources, part II: Problems statement. *Signal*  
586 *Processing*, 24(1), 11 - 20.
- 587 Demertzi, A., Tagliazucchi, E., Dehaene, S., Deco, G., Barttfeld, P., Raimondo, F., ... Sitt, J. D. (2019). Human  
588 consciousness is supported by dynamic complex patterns of brain signal coordination. *Science Advances*,  
589 5(2), eaat7603.
- 590 Estrada, E., & Rodríguez-Velázquez, J. A. (2005). Subgraph centrality in complex networks. *Physical Review*  
591 *E*, 71(5), 056103.
- 592 Etzel, J. A., Gazzola, V., & Keysers, C. (2009). An introduction to anatomical ROI-based fMRI classification.  
593 *Brain Research*, 1281, 114–125.
- 594 Fallon, J., Ward, P., Parkes, L., Oldham, S., Arnatkevičiūtė, A., Fornito, A., & Fulcher, B. D. (2019).  
595 Timescales of spontaneous activity fluctuations relate to structural connectivity in the brain. *bioRxiv*,  
596 doi.org/10.1101/655050.
- 597 Fong, A. H. C., Yoo, K., Rosenberg, M. D., Zhang, S., Li, C.-S. R., Scheinost, D., ... Chun, M. M. (2019).  
598 Dynamic functional connectivity during task performance and rest predicts individual differences in  
599 attention across studies. *NeuroImage*, 188, 14–25.
- 600 Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry*, 40(1), 35–41.
- 601 Geisberger, R., Sanders, P., & Schultes, D. (2008). Better approximation of betweenness centrality. *Proceedings*  
602 *of the meeting on Algorithm Engineering and Experiments*, 90–100.
- 603 Gershman, S., Blei, D., Pereira, F., & Norman, K. (2011). A topographic latent source model for fMRI data.  
604 *NeuroImage*, 57, 89–100.
- 605 Gonzalez-Castillo, J., Caballero-Gaudes, C., Topolski, N., Handwerker, D. A., Pereira, F., & Bandettini, P. A.  
606 (2019). Imaging the spontaneous flow of thought: distinct periods of cognition contribute to dynamic  
607 functional connectivity during rest. *NeuroImage*, 202(116129).
- 608 Halu, A., Mondragón, R. J., Panzarasa, P., & Bianconi, G. (2013). Multiplex PageRank. *PLoS One*, 8(10),  
609 e78293.

- 610 Hasson, U., Chen, J., & Honey, C. J. (2015). Hierarchical process memory: memory as an integral component  
611 of information processing. *Trends in Cognitive Science*, 19(6), 304-315.
- 612 Hasson, U., Yang, E., Vallines, I., Heeger, D. J., & Rubin, N. (2008). A hierarchy of temporal receptive windows  
613 in human cortex. *Journal of Neuroscience*, 28(10), 2539-2550. doi: 10.1523/JNEUROSCI.5487-07.2008
- 614 Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and  
615 overlapping representations of faces and objects in ventral temporal cortex. *Science*, 293, 2425–2430.
- 616 Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks.  
617 *Science*, 313(5786), 504–507.
- 618 Honey, C. J., Kötter, R., Breakspear, M., & Sporns, O. (2007). Network structure of cerebral cortex shapes  
619 functional connectivity on multiple time scales. *Proceedings of the National Academy of Science USA*, 104(24),  
620 10240–10245.
- 621 Honey, C. J., Thesen, T., Donner, T. H., Silbert, L. J., Carlson, C. E., Devinsky, O., . . . Hasson, U. (2012). Slow  
622 cortical dynamics and the accumulation of information over long timescales. *Neuron*, 76, 423-434. doi:  
623 10.1016/j.neuron.2012.08.011
- 624 Huth, A. G., de Heer, W. A., Griffiths, T. L., Theunissen, F. E., & Gallant, J. L. (2016). Natural speech reveals  
625 the semantic maps that tile human cerebral cortex. *Nature*, 532, 453–458.
- 626 Huth, A. G., Nisimoto, S., Vu, A. T., & Gallant, J. L. (2012). A continuous semantic space describes  
627 the representation of thousands of object and action categories across the human brain. *Neuron*, 76(6),  
628 1210-1224.
- 629 Jutten, C., & Herault, J. (1991). Blind separation of sources, part I: An adaptive algorithm based on  
630 neuromimetic architecture. *Signal Processing*, 24(1), 1–10.
- 631 Kamitani, Y., & Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nature  
632 Neuroscience*, 8, 679–685.
- 633 Landau, E. (1895). Zur relativen Wertbemessung der Turnierresultate. *Deutsches Wochenschach*, 11, 366–369.
- 634 Lee, C. S., Aly, M., & Baldassano, C. (2020). Anticipation of temporally structured events in the brain.  
635 *bioRxiv*, 10.1101/2020.10.14.338145.
- 636 Lee, D. D., & Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature*,  
637 401, 788–791.

- 638 Lerner, Y., Honey, C. J., Katkov, M., & Hasson, U. (2014). Temporal scaling of neural responses to compressed  
639 and dilated natural speech. *Journal of Neurophysiology*, 111, 2433–2444.
- 640 Lerner, Y., Honey, C. J., Silbert, L. J., & Hasson, U. (2011). Topographic mapping of a hierarchy of  
641 temporal receptive windows using a narrated story. *Journal of Neuroscience*, 31(8), 2906-2915. doi: 10.1523/  
642 JNEUROSCI.3684-10.2011
- 643 Liégeois, R., Li, J., Kong, R., Orban, C., De Ville, D. V., Ge, T., ... Yeo, B. T. T. (2019). Resting brain  
644 dynamics at different timescales capture distinct aspects of human behavior. *Nature Communications*,  
645 10(2317), doi.org/10.1038/s41467-019-10317-7.
- 646 Lin, J. (2009). Divergence measures based on the Shannon entropy. *IEEE Transactions on Information Theory*,  
647 37(1), 145–151.
- 648 Lohmann, G., Margulies, D. S., Horstmann, A., Pleger, B., Lepsien, J., Goldhahn, D., ... Turner, R. (2010).  
649 Eigenvector centrality mapping for analyzing connectivity patterns in fMRI data of the human brain.  
650 *PLoS One*, 5(4), e10232.
- 651 Lurie, D., Kessler, D., Bassett, D., Betzel, R., Breakspear, M., Keilholz, S., ... Calhoun, V. (2018). On the  
652 nature of time-varying functional connectivity in resting fMRI. *PsyArXiv*, doi.org/10.31234/osf.io/xtzre.
- 653 Mack, M. L., Preston, A. R., & Love, B. C. (2017). Medial prefrontal cortex compresses concept representations  
654 through learning. *bioRxiv*, doi.org/10.1101/178145.
- 655 Mairal, J., Ponce, J., Sapiro, G., Zisserman, A., & Bach, F. R. (2009). Supervised dictionary learning. *Advances  
656 in Neural Information Processing Systems*, 1033–1040.
- 657 Mairal, J. B., Bach, F., Ponce, J., & Sapiro, G. (2009). Online dictionary learning for sparse coding. *Proceedings  
658 of the 26th annual international conference on machine learning*, 689–696.
- 659 Manning, J. R., Ranganath, R., Norman, K. A., & Blei, D. M. (2014). Topographic factor analysis: a Bayesian  
660 model for inferring brain networks from neural data. *PLoS One*, 9(5), e94914.
- 661 Manning, J. R., Zhu, X., Willke, T. L., Ranganath, R., Stachenfeld, K., Hasson, U., ... Norman, K. A. (2018).  
662 A probabilistic approach to discovering dynamic full-brain functional connectivity patterns. *NeuroImage*,  
663 180, 243–252.
- 664 McInnes, L., Healy, J., & Melville, J. (2018). UMAP: Uniform manifold approximation and projection for  
665 dimension reduction. *arXiv*, 1802(03426).

- 666 McIntosh, A. R., & Jirsa, V. K. (2019). The hidden repertoire of brain dynamics and dysfunction. *Network*  
667 *Neuroscience*, doi.org/10.1162/netn\_a\_00107.
- 668 Mitchell, T., Shinkareva, S., Carlson, A., Chang, K., Malave, V., Mason, R., & Just, M. (2008). Predicting  
669 human brain activity associated with the meanings of nouns. *Science*, 320(5880), 1191.
- 670 Newman, M. E. J. (2005). A measure of betweenness centrality based on random walks. *Social Networks*, 27,  
671 39–54.
- 672 Newman, M. E. J. (2008). The mathematics of networks. *The New Palgrave Encyclopedia of Economics*, 2, 1–12.
- 673 Nishimoto, S., Vu, A., Naselaris, T., Benjamini, Y., Yu, B., & Gallant, J. (2011). Reconstructing visual  
674 experience from brain activity evoked by natural movies. *Current Biology*, 21, 1 - 6.
- 675 Nocedal, J., & Wright, S. J. (2006). *Numerical optimization*. New York: Springer.
- 676 Norman, K. A., Newman, E., Detre, G., & Polyn, S. M. (2006). How inhibitory oscillations can train neural  
677 networks and punish competitors. *Neural Computation*, 18, 1577–1610.
- 678 Opsahl, T., Agneessens, F., & Skvoretz, J. (2010). Node centrality in weighted networks: generalizing degree  
679 and shortest paths. *Social Networks*, 32, 245–251.
- 680 Park, H.-J., Friston, K. J., Pae, C., Park, B., & Razi, A. (2018). Dynamic effective connectivity in resting state  
681 fMRI. *NeuroImage*, 180, 594–608.
- 682 Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. *The London, Edinburgh,*  
683 *and Dublin Philosophical Magazine and Journal of Science*, 2, 559-572.
- 684 Pereira, F., Lou, B., Pritchett, B., Ritter, S., Gershman, S. J., Kanwisher, N., . . . Fedorenko, E. (2018). Toward  
685 a universal decoder of linguistic meaning from brain activation. *Nature Communications*, 9(963), 1–13.
- 686 Preti, M. G., Bolton, T. A. W., & Van De Ville, D. (2017). The dynamic functional connectome: state-of-the-art  
687 and perspectives. *NeuroImage*, 160, 41–54.
- 688 Rao, C. R. (1982). Diversity and dissimilarity coefficients: a unified approach. *Theoretical Population Biology*,  
689 21(1), 24–43.
- 690 Reimann, M. W., Nolte, M., Scolamiero, M., Turner, K., Perin, R., Chindemi, G., . . . Markram, H. (2017).  
691 Cliques of neurons bound into cavities provide a missing link between structure and function. *Frontiers*  
692 *in Computational Neuroscience*, 11(48), 1–16.

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

- 720 Tipping, M. E., & Bishop, C. M. (1999). Probabilistic principal component analysis. *Journal of Royal Statistical Society, Series B*, 61(3), 611–622.
- 722 Toker, D., & Sommer, F. T. (2019). Information integration in large brain networks. *PLoS Computational Biology*, 15(2), e1006807.
- 724 Tong, F., & Pratte, M. S. (2012). Decoding patterns of human brain activity. *Annual Review of Psychology*, 63, 483–509.
- 726 Turk-Browne, N. B. (2013). Functional interactions as big data in the human brain. *Science*, 342, 580–584.
- 727 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.
- 729 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.
- 731 Zar, J. H. (2010). *Biostatistical analysis*. Prentice-Hall/Pearson.
- 732 Zheng, M., Allard, A., Hagmann, P., & Serrano, M. A. (2019). Geometric renormalization unravels self-similarity of the multiscale human connectome. *arXiv*, 1904.11793.
- 734 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.