

1 **High-order cognition is supported by complex but**
2 **compressible brain activity patterns**

3 Lucy L. W. Owen^{1, 2}, and Jeremy R. Manning^{1,*}

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

6 ²Carney Institute for Brain Sciences,
7 Brown University, Providence, RI

8 *Address correspondence to jeremy.r.manning@dartmouth.edu

9 February 20, 2023

10 **Abstract**

11 We applied dimensionality reduction algorithms and pattern classifiers to functional neuroimaging
12 data collected as participants listened to a story, temporally scrambled versions of the story, or underwent
13 a resting state scanning session. These experimental conditions were intended to require different depths
14 of processing and inspire different levels of engagement. We considered two primary aspects of the data.
15 First, we treated the number of features (components) required to achieve a threshold decoding accuracy
16 as a proxy for the “compressibility” of the neural patterns (where fewer components indicate higher
17 compressibility). Second, we treated the maximum achievable decoding accuracy across participants as
18 an indicator of the “stability” of the recorded patterns. Overall, we found that neural patterns recorded as
19 participants listened to the intact story required fewer features to achieve comparable classification accuracy
20 to the other experimental conditions. However, the peak decoding accuracy (achievable with more features)
21 was also highest during intact story listening. Taken together, our work suggests that our brain networks
22 flexibly reconfigure according to ongoing task demands, and that the activity patterns associated with
23 higher-order cognition and high engagement are both more complex and more compressible than the
24 activity patterns associated with lower-order tasks and lower levels of engagement.

25 **Introduction**

26 Large-scale networks, including the human brain, may be conceptualized as occupying one or more positions
27 along on a continuum. At one extreme, every node is fully independent of every other node. At the other
28 extreme, all nodes are fully coupled and behave identically. Each extreme optimizes key properties of how
29 the network functions. When every node is independent, the network is maximally *expressive*: if we define
30 the network’s “state” as the total set of activity patterns across nodes, then every state is equally reachable by
31 a network with fully independent nodes. On the other hand, a fully coupled network optimizes *robustness*:
32 any subset of nodes, other than the entire network, may be removed from the network without any loss of
33 function or expressive power. Note that a given set of nodes might reconfigure its connections or behaviors
34 under different circumstances to change its position along this continuum according to the needs at hand.

30 Presumably, most systems tend to occupy positions between these extremes. We wondered: might the
31 human brain reconfigure itself to be more flexible or more robust according to ongoing demands?

32 Closely related to the above notions of expressiveness versus robustness are measures of how much
33 *information* (Shannon, 1948) is contained in a given signal or pattern. Formally, information is defined as the
34 amount of uncertainty (i.e., entropy) in a given variables' outcomes, measured in *bits*, or the optimal number
35 of yes/no questions needed to reduce uncertainty about the variable's outcomes to zero. The *redundancy*
36 of a signal denotes the difference how expressive the signal could be (i.e., proportional to the number of
37 unique states or symbols used to transmit the signal) and the actual information rate (i.e., the entropy of
38 each individual state or symbol). If a brain network's nodes are fully independent, then the number of bits
39 required to express a single activity pattern is proportional to the number of nodes. The network would
40 also be minimally redundant, since the status of every node would be needed to fully express a single brain
41 activity pattern. If a brain network's nodes are fully coupled and identical, then the number of bits required
42 to express a single activity pattern is proportional to the number of unique states or values any individual
43 node can take on. Such a network would be highly redundant, since knowing any individual node's state
44 would be sufficient to recover the full-brain activity pattern.

45 We're interested in the complexity of brain patterns that underly different types of thoughts. To explore
46 this question space, we will take brain patterns recorded under different experimental conditions used in
47 Aim 2, and project them into lower dimensional spaces using principle components analysis. We can then
48 ask how well those low-dimensional embeddings of the data retain cognitively relevant information like
49 when in a story someone is listening to.

50 This work has been inspired, in part, by ?. In this paper, they investigated the role of the prefrontal
51 cortex in filtering out irrelevant content. Specifically, they looked at if the vmPFC performs data reduction
52 on incoming information through compression. This was motivated, in part, by orbital frontal cortex (OFC)
53 compression in rats (?). They studied this using a learning paradigm in which participants had to classify
54 insects based on different numbers of feature dimensions. The idea was that participants in some learning
55 blocks, participants could identify the insects based on one feature (low complexity) or several features
56 (high complexity), but importantly the stimuli remained the same across all learning problems. They found
57 that complexity and compression had an inverse relationship; the lower complexity of a conceptual space,
58 the higher the degree of compression. Building on this idea, we wonder if varying degrees of compression
59 is performed throughout the brain. We also want to test this idea, but using varying levels of engagement
60 listening to a naturalistic stimuli.

61 To understand the degree of compression throughout the brain during cognition, we will use the same
62 fMRI data from Aim 2, collected while participants listened to a story in different scrambling conditions.

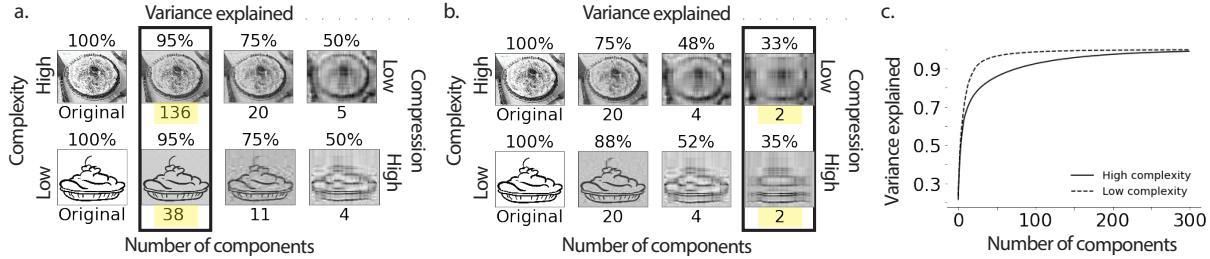


Figure 1: Illustration of compression. Visual analogy for neural compression. Here are 2 images of pies, one more complex than the other. **a.** It takes fewer components to reach the same percent variance explained in the less complex pie, which corresponds to higher compression. **b.** However, with very few components, similar variance is explained in both pies. **c.** Plots the cumulative explained variance for more and more components.

63 We will measure the degree that multivoxel activation patterns are compressed during story listening
 64 using principle components analysis (PCA) a method for low-rank approximation of multidimensional
 65 data (Eckart & Young, 1936). We will explore this using decoding accuracy as a function of the number of
 66 components, or dimensions, in the low-dimensional space under different cognitive conditions.

67 You can imagine two reasonable predictions of how cognition is reflected in brain patterns. The first is
 68 as our thoughts become more complex, they are supported by more complex brain patterns, and require
 69 more components to decode. The second is that when thoughts are deeper and more complicated, the units
 70 of neural activity would carry more information, and would require therefore fewer components to decode.

71 This idea can be explored in this visual analogy (Fig. ??) for neural compression. Here there are two
 72 images of pies, the top pie is more complex than the bottom. On the left we're illustrating that it takes fewer
 73 components to reach the same 95 percent variance explained in the less complex pie, which corresponds to
 74 higher compression. However, on the right with very few components similar variance is explaining both
 75 pies.

76 We investigated the dimensionality of neural patterns by training classifiers using more and more
 77 principle components. Or, in other words, we used less and less compression to decode. We applied the
 78 approach to a neuroimaging dataset comprising data collected as participants listened to a story varying in
 79 cognitive richness (Simony et al., 2016).

80 Evaluation metrics

81 We will evaluate the degree of compression of held-out neuroimaging data by assessing the time at which
 82 it was collected. We will use this evaluation (timepoint decoding) as a proxy for gauging how much
 83 explanatory power the compressed data held with respect to the observed data.

84 **Timepoint decoding**

85 To explore how compression varies with complexity, we will use a previous neuroimaging dataset Simony
86 et al. (2016) in which participants listened to an audio recording of a story; 36 participants listen to an intact
87 version of the story, 17 participants listen to time-scrambled recordings of the same story where paragraphs
88 were scrambled, 36 participants listen to word-scrambled version and 36 participants lay in rest condition.

89 Following the analyses conducted by (HTFA) Manning et al. (2018), we first apply *hierarchical topographic*
90 *factor analysis* (HTFA) to the fMRI datasets to obtain a time series of 700 node activities for every participant.

91 We then apply dimensionality reduction (Incremental PCA) for each group.

92 We then compare the groups' activity patterns (using Pearson correlations) to estimate the story times
93 each corresponding pattern using more and more principle components.

94 To assess decoding accuracy, we randomly divide participants for each stimulus into training and testing
95 groups. We then compare the groups' activity patterns (using Pearson correlations) to estimate the story
96 times each corresponding pattern using more and more principle components (as the data became less
97 compressed). Specifically, we ask, for each timepoint: what are the correlations between the first group's
98 and second group's activity patterns at each order. We note that the decoding test we used is a conservative
99 in which we count a timepoint label as incorrect if it is not an exact match.

100 **Results**

101 By training classifiers using more and more principle components to decode, and comparing across conditions
102 with varying degrees of cognitive richness, we can assess the explanatory power of the compressed
103 data held with respect to the observed data (see *Methods*). We note that our primary goal was not to achieve
104 perfect decoding accuracy, but rather to use decoding accuracy as a benchmark for assessing whether
105 different neural features specifically capture cognitively relevant brain patterns.

106 Prior work has shown participants share similar neural responses to richly structured stimuli when
107 compared to stimuli with less structure Simony et al. (2016). We replicate this finding, showing as complexity
108 of the stimulus increases, decoding accuracy increases (Fig. 2, a.). Additionally, we found that as complexity
109 of the stimuli increases, we need fewer components to decode the same amount (Fig. 2, b.). However, we
110 also found that as complexity of the stimuli increases, more components are required to reach peak decoding
111 accuracy (Fig. 2, c.). We posit that as the complexity of our thoughts increases, neural compression decreases.
112 However, as our thoughts become deeper and richer, more reliable information is available at higher neural
113 compression.

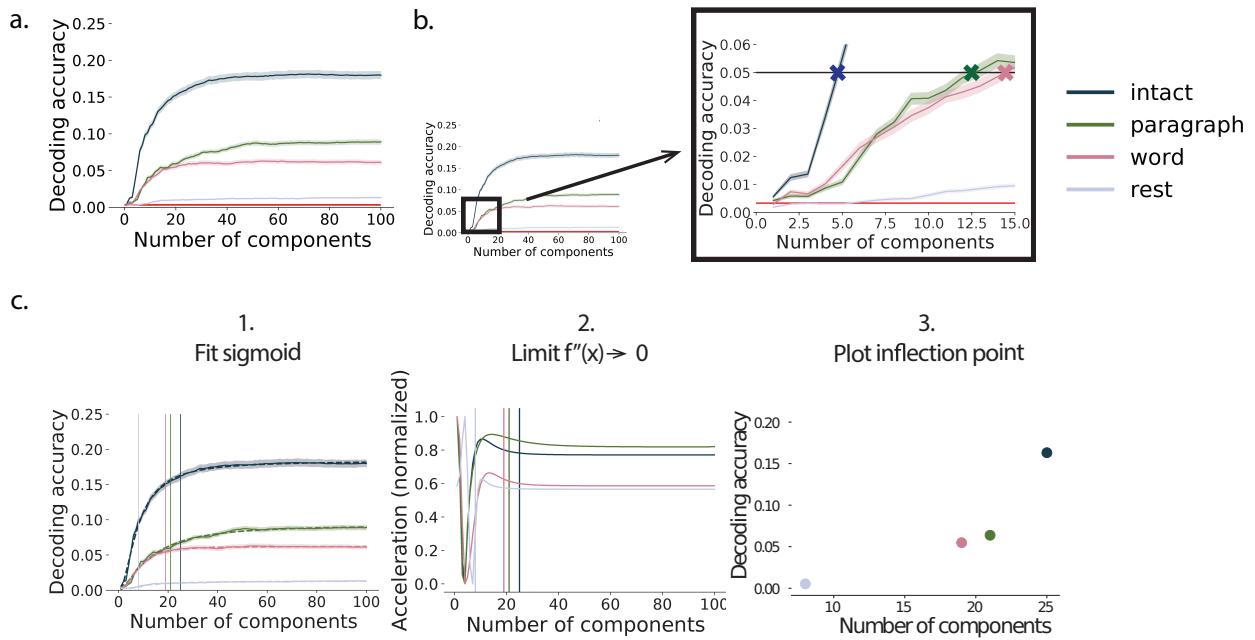


Figure 2: **Decoding accuracy.** **a. Decoding accuracy by number of components.** Ribbons of each color display cross-validated decoding performance for each condition (intact, paragraph, word, and rest). Decoders were trained using increasingly more principle components and displayed relative to chance (red line). **b. Fixed decoding accuracy by number of components.** We zoom in on the plot shown in **a.** and add a line denoting fixed decoding accuracy (.05). We plot where the intact, paragraph, and word conditions intersect. **c. Explanation of inflection metric.** First we fit a sigmoid function to the decoding accuracy by number of components. Second, we found where the second derivative is both positive and less than .0001. Last, we then plot that inflection point as a single metric to capture the slope and asymptote of the curve.

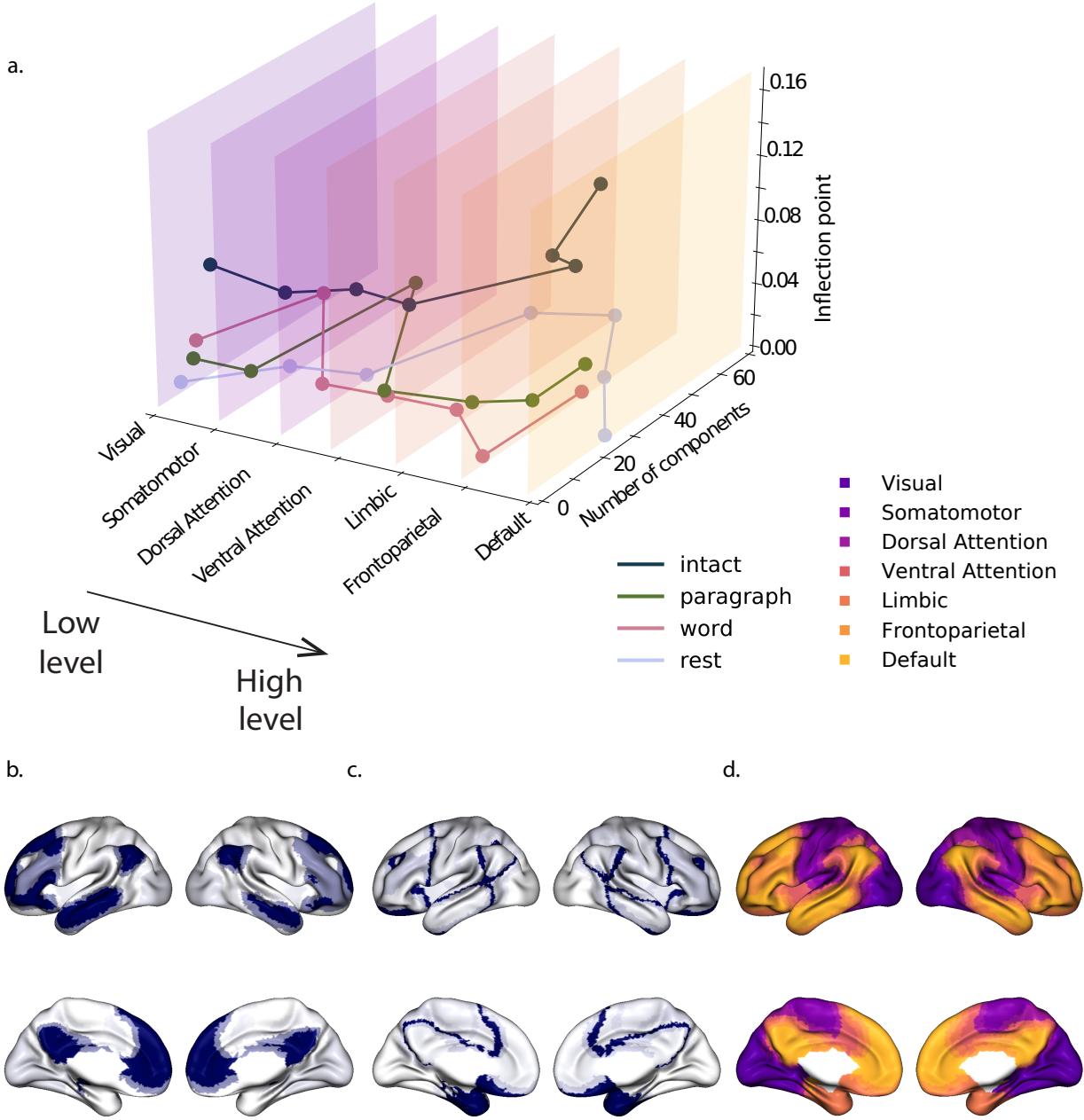


Figure 3: Inflection points by network. a. Inflection point was calculated as explained in Fig. 2, b. Analyses were limited by the brain networks (using the Yeo et al. (2011) network parcellation) and arranged in increasing order relative to the intact condition. b. and c. For the total time in the intact condition, we are plotting the relative inflection points (b.) and corresponding number of components (c.) by network. d. The network parcellation defined by Yeo et al. (2011) is displayed on the inflated brain maps. The colors and network labels serve as a legend for a. and d.

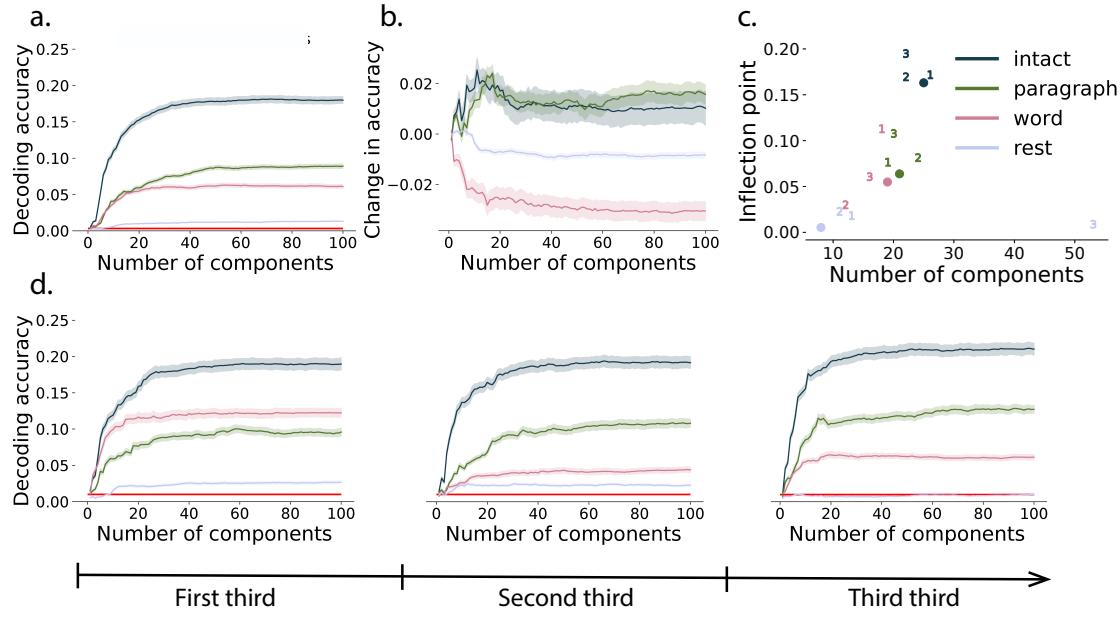


Figure 4: Inflection points by thirds. **a.** Decoding accuracy by number of components not broken into thirds (Fig. 2 a.). **b. and c.** Quantifying changes in decoding accuracy across time. **b.** Slope of decoding accuracy was calculated by fitting a regression line for each component/condition for each third. **c.** We also repeated the analysis (Fig. 2, b.) to obtain the inflection point for each condition and for each third. **d.** Decoding accuracy by number of components for each third of the scan time. We repeated the same analysis in Fig. 2 a. but breaking the scan time for each condition into 3 intervals.

114 We also wondered how this compression would change across brain regions. We repeated the analysis

115 but limited the brain hubs to 7 networks using the Yeo et al. (2011) network parcellation shown here in the
116 inflated brain (Fig. 3, d.). We found that as complexity of the stimuli increases, decoding accuracy increases
117 with higher cognitive areas. (Fig. 3).

118 We were also curious how compression would change across time. If, there is some understanding of
119 the narrative that accumulates over time, we should be able to see that difference. We found increases
120 in decoding accuracy with the same number or fewer components for more complex, cognitively rich,
121 conditions. We also found decreases in decoding accuracy for the word-scrambled and rest condition.

122 Overall, we found that as story listening conditions become more complex, more components are
123 required to decode. We also found we could decode better with more impoverished data when there is the
124 underlying structure of the narrative providing more cognitive richness. We posit that as the complexity
125 of our thoughts increases, neural compression decreases. However, as our thoughts become deeper and
126 richer, more reliable information is available at higher neural compression.

127 **Discussion**

128 - We trained classifiers using more and more principle components to decode, and compared across condi-
129 tions with varying degrees of cognitive richness. -We found that as listening conditions become more
130 cognitively rich, decoding accuracy increased. -Also, decoding accuracy increased as understanding of the
131 narrative accumulated over time, in more complex listening conditions. - Decoding accuracy also increased
132 in higher cognitive areas, in more complex listening conditions. -We found that as story listening conditions
133 become more complex, more components are required to decode. -We also found we could decode better
134 with more impoverished data when there is the underlying structure of the narrative providing more
135 cognitive richness. -We posit that as the complexity of our thoughts increases, neural compression decreases.
136 However, as our thoughts become deeper and richer, more reliable information is available at higher neural
137 compression.

138 Based on prior work (?) and following the direction of the field (Turk-Browne, 2013) we think our
139 thoughts might be encoded in dynamic network patterns, and possibly higher order network patterns
140 (Fig. ??). We sought to test this hypothesis by developing an approach to inferring high-order network
141 dynamics from timeseries data.

142 One challenge in studying dynamic interactions is the computational resources required to calculate
143 higher-order correlations. We developed a computationally tractable model of network dynamics (Fig. ??)
144 that takes in a feature timeseries and outputs approximated first-order dynamics (i.e., dynamic functional
145 correlations), second-order dynamics (reflecting homologous networks that dynamically form and disperse),
146 and higher-order network dynamics (up to tenth-order dynamic correlations).

147 We first validated our model using synthetic data, and explored how recovery varied with different
148 underlying data structures and kernels. We then applied the approach to an fMRI dataset (Simony et al.,
149 2016) in which participants listened to an audio recording of a story, as well as scrambled versions of the
150 same story (where the scrambling was applied at different temporal scales). We trained classifiers to take
151 the output of the model and decode the timepoint in the story (or scrambled story) that the participants
152 were listening to. We found that, during the intact listening condition in the experiment, classifiers that
153 incorporated higher-order correlations yielded consistently higher accuracy than classifiers trained only
154 on lower-order patterns (Fig. ??, a.&d.). By contrast, these higher-order correlations were not necessary
155 to support decoding the other listening conditions and (minimally above chance) during a control rest
156 condition. This suggests that the cognitive processing that supported the most cognitively rich listening
157 conditions involved second-order (or higher) network dynamics.

158 Although we found decoding accuracy was best when incorporating higher-order network dynamics

159 for all but rest condition, it is unclear if this is a product of the brain or the data collection technique. It could
160 be that the brain is second-order or it could be that fMRI can only reliably give second-order interactions.
161 Exploring this method with other data collection technique will be important to disentangle this question.

162 **Concluding remarks**

163 How can we better understand how brain patterns change over time? How can we quantify the potential
164 network dynamics that might be driving these changes? One way to judge the techniques of the future is
165 to look at the trajectory of the fMRI field so far has taken so far (Fig. ??). The field started with univariate
166 activation, measuring the average activity for each voxel. Analyses of multivariate activation followed,
167 looking at spatial patterns of activity over voxels. Next, correlations of activity were explored, first with
168 measures like resting connectivity that take temporal correlation between a seed voxel and all other voxels
169 then with full connectivity that measure all pairwise correlations. Additionally, this path of increasing
170 complexity also moved from static to dynamic measurements. One logical next step in this trajectory would
171 be dynamic higher-order correlations. We have created a method to support these calculations by scalably
172 approximating dynamic higher-order correlations.

173 **Acknowledgements**

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

180 **Author contributions**

181 Concept: J.R.M. and L.L.W.O. Implementation: L.L.W.O., and J.R.M. Analyses: L.L.W.O and J.R.M.

182 **References**

183 Eckart, C., & Young, G. (1936). The approximation of one matrix by another of lower rank. *Psychometrika*,
184 1, 211–218.

- 185 Manning, J. R., Zhu, X., Willke, T. L., Ranganath, R., Stachenfeld, K., Hasson, U., ... Norman, K. A. (2018).
- 186 A probabilistic approach to discovering dynamic full-brain functional connectivity patterns. *NeuroImage*,
- 187 180, 243–252.
- 188 Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27(3),
- 189 379–423.
- 190 Simony, E., Honey, C. J., Chen, J., & Hasson, U. (2016). Dynamic reconfiguration of the default mode
- 191 network during narrative comprehension. *Nature Communications*, 7(12141), 1–13.
- 192 Turk-Browne, N. B. (2013). Functional interactions as big data in the human brain. *Science*, 342, 580–584.
- 193 Yeo, B. T. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., ... Buckner, R. L.
- 194 (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity.
- 195 *Journal of Neurophysiology*, 106(3), 1125–1165.