

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,*}

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

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

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

4 February 23, 2023

5

Abstract

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

20

Introduction

21 Large-scale networks, including the human brain, may be conceptualized as occupying one or more positions
22 along on a continuum. At one extreme, every node is fully independent of every other node. At the other
23 extreme, all nodes behave identically. Each extreme optimizes key properties of how the network functions.
24 When every node is independent, the network is maximally *expressive*: if we define the network’s “state”
25 as the activity pattern across its nodes, then every state is equally reachable by a network with fully
26 independent nodes. On the other hand, a network of identically behaved nodes optimizes *robustness*: any
27 subset of nodes may be removed from the network without any loss of function or expressive power, as
28 long as any single node remains. Presumably, most natural systems tend to occupy positions between
29 these extremes. We wondered: might the human brain reconfigure itself to be more flexible or more robust

30 according to ongoing demands? In other words, might the brain reconfigure its connections or behaviors
31 under different circumstances to change its position along this continuum?

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

48 We take as a given that brain activity is highly flexible: our brains can exhibit nearly infinite activity
49 patterns. This flexibility implies that our brains activity patterns are highly information rich. However,
50 brain activity patterns are also highly structured. For example, full-brain correlation matrices are stable
51 within (Finn et al., 2015, 2017; Gratton et al., 2018) and across (Yeo et al., 2011; Glerean et al., 2012; Gratton
52 et al., 2018; Cole et al., 2014) individuals. This stability suggests that our brains' activity patterns are at
53 least partially constrained, for example by anatomical, external, or internal factors. Constraints on brain
54 activity that limit its flexibility decrease expressiveness (i.e., its information rate). However, constraints on
55 brain activity also increase its robustness to noise (e.g., “missing” or corrupted signals may be partially
56 recovered). For example, recent work has shown that full-brain activity patterns may be reliably recovered
57 from only a relatively small number of implanted electrodes (Owen et al., 2020; Scangos et al., 2021). This
58 robustness property suggests that the relevant signal (e.g., underlying factors that have some influence over
59 brain activity patterns) are compressible.

60 To the extent that brain activity patterns contain rich task-relevant information, we should be able to
61 use the activity patterns to accurately differentiate between different aspects of the task (e.g., using pattern
62 classifiers; Norman et al., 2006). For example, prior work has shown a direct correspondence between

63 classification accuracy and the information content of a signal (Alvarez, 2002). To the extent that brain
64 activity patterns are compressible, we should be able to generate simplified (e.g., lower dimensional)
65 representations of the data while still preserving the relevant or important aspects of the original signal. In
66 general, information content and compressibility are related but are partially dissociable (Fig. 1).

67 Results

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

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

81 We also wondered how this compression would change across brain regions. We repeated the analysis
82 but limited the brain hubs to 7 networks using the Yeo et al. (2011) network parcellation shown here in the
83 inflated brain (Fig. 3, d.). We found that as complexity of the stimuli increases, decoding accuracy increases
84 with higher cognitive areas. (Fig. 3).

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

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

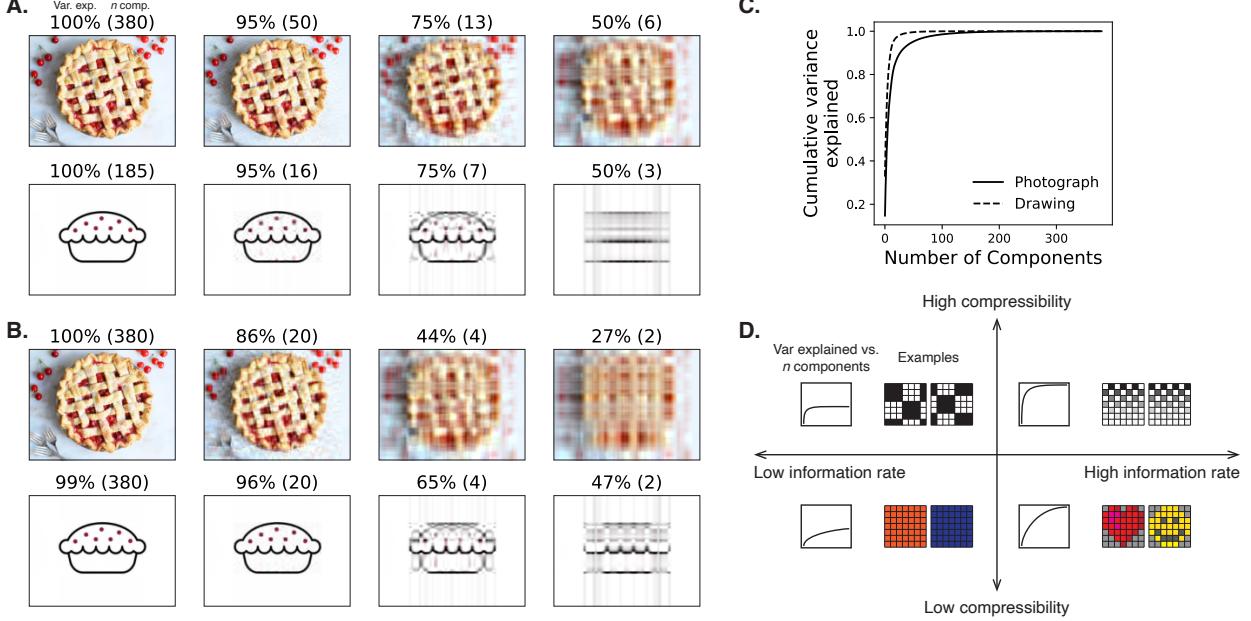


Figure 1: Information content and compressibility. **A. Variance explained for two images.** We applied principal components analysis to a photograph and drawing, treating each row of the images as “observations.” Across columns, we identified the number of components required to explain 100%, 95%, 75%, or 50% of the cumulative variance in each image (the 100% columns denote the original images). The numbers of components are indicated in parentheses, and the resulting “compressed” images are displayed. **B. Representing two images with different numbers of components.** Using the same principal component decompositions as in Panel A, we computed the cumulative proportion of variance explained with 380 (original images), 20, 4, or 2 components. **C. Cumulative variance explained versus number of components.** For the images displayed in Panels A and B, we plot the cumulative proportion of variance explained as a function of the number of components used to represent each image. **D. Information rate and compressibility.** Across multiple images, the information rate (i.e., the amount of information contained in each image; horizontal axis) is high when each individual pixel provides information that cannot be inferred from other pixels. High-information rate images tend to be high-resolution, and low-information rate images tend to be low-resolution. Compressibility is related to the difference between the information required to specify the original versus compressed images (vertical axis). Highly compressible images often contain predictable structure (redundancies) that can be leveraged to represent the images much more efficiently than in their original feature spaces.

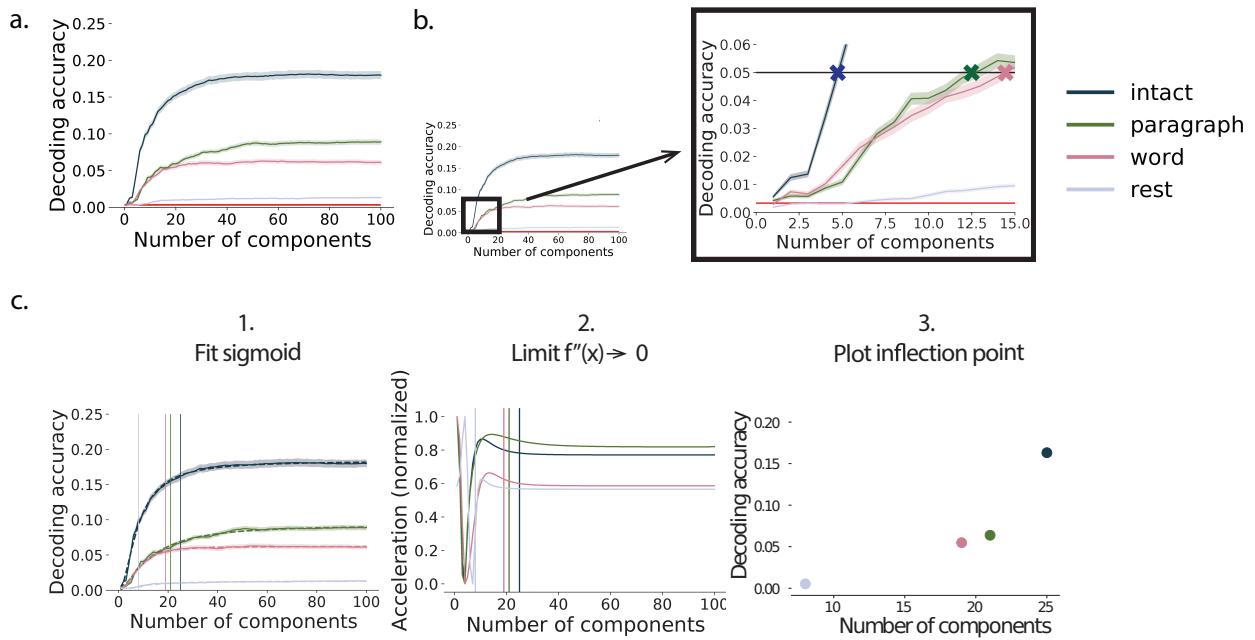


Figure 2: Decoding accuracy. **a. Decoding accuracy by number of components.** Ribbons of each color display cross-validated decoding performance for each condition (intact, paragraph, word, and rest). Decoders were trained using increasingly more principle components and displayed relative to chance (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.

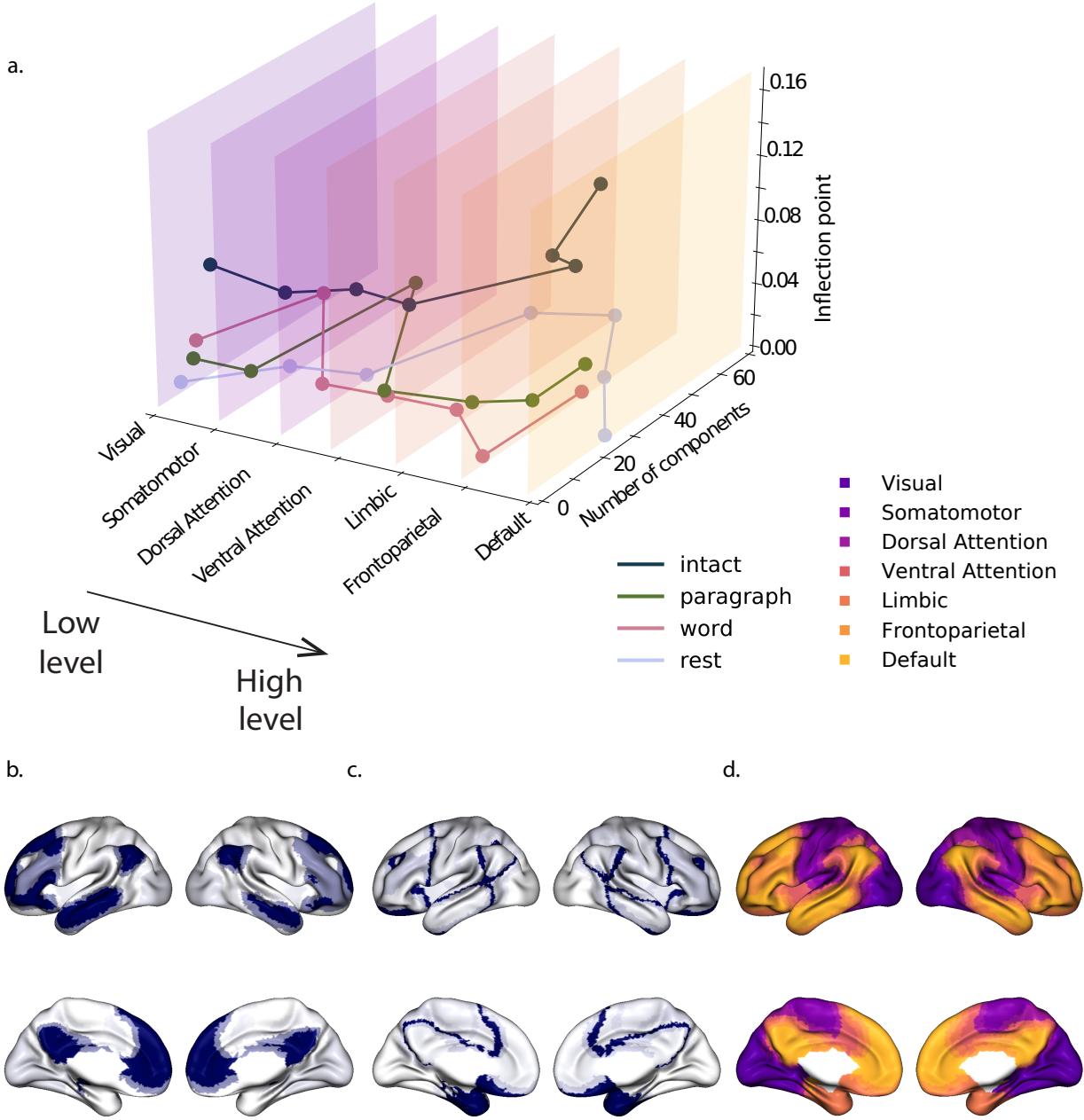


Figure 3: Inflection points by network. a. Inflection point was calculated as explained in Fig. 2, b. Analyses were limited by the brain networks (using the Yeo et al. (2011) network parcellation) and arranged in increasing order relative to the intact condition. b. and c. For the total time in the intact 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.

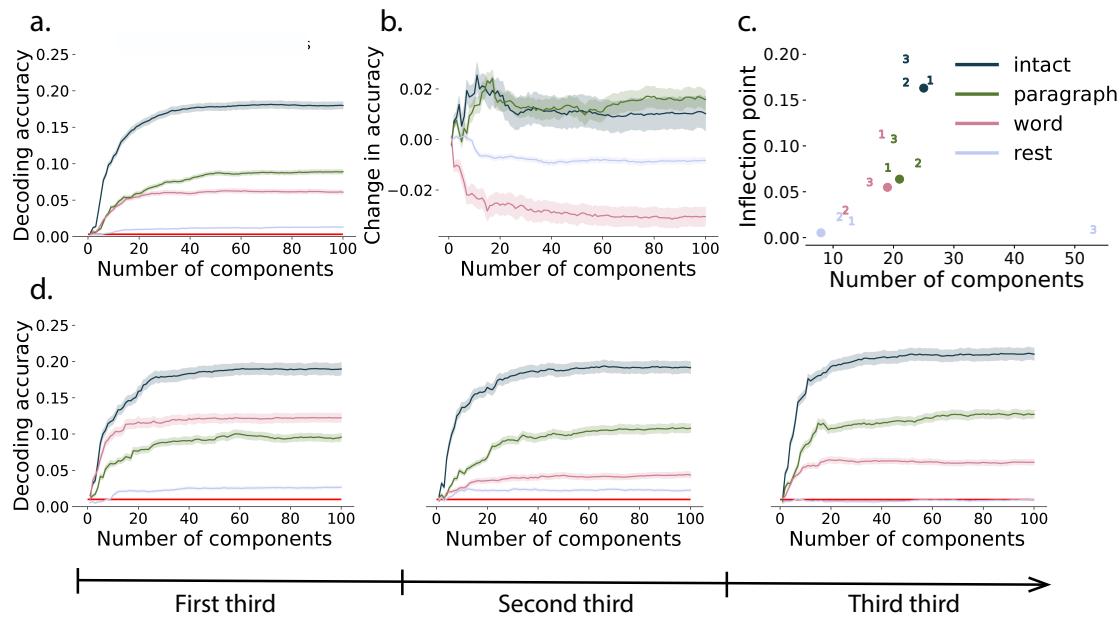


Figure 4: **Inflection points by thirds.** **a.** Decoding accuracy by number of components not broken into thirds (Fig. 2 a.). **b.** and **c.** Quantifying changes in decoding accuracy across time. **b.** Slope of decoding accuracy was calculated by fitting a regression line for each component/condition 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.

94 Discussion

95 - We trained classifiers using more and more principle components to decode, and compared across condi-
96 tions with varying degrees of cognitive richness. -We found that as listening conditions become more
97 cognitively rich, decoding accuracy increased. -Also, decoding accuracy increased as understanding of the
98 narrative accumulated over time, in more complex listening conditions. - Decoding accuracy also increased
99 in higher cognitive areas, in more complex listening conditions. -We found that as story listening conditions
100 become more complex, more components are required to decode. -We also found we could decode better
101 with more impoverished data when there is the underlying structure of the narrative providing more
102 cognitive richness. -We posit that as the complexity of our thoughts increases, neural compression decreases.
103 However, as our thoughts become deeper and richer, more reliable information is available at higher neural
104 compression.

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

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

114 We first validated our model using synthetic data, and explored how recovery varied with different
115 underlying data structures and kernels. We then applied the approach to an fMRI dataset (Simony et al.,
116 2016) in which participants listened to an audio recording of a story, as well as scrambled versions of the
117 same story (where the scrambling was applied at different temporal scales). We trained classifiers to take
118 the output of the model and decode the timepoint in the story (or scrambled story) that the participants
119 were listening to. We found that, during the intact listening condition in the experiment, classifiers that
120 incorporated higher-order correlations yielded consistently higher accuracy than classifiers trained only
121 on lower-order patterns (Fig. ??, a.&d.). By contrast, these higher-order correlations were not necessary
122 to support decoding the other listening conditions and (minimally above chance) during a control rest
123 condition. This suggests that the cognitive processing that supported the most cognitively rich listening
124 conditions involved second-order (or higher) network dynamics.

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

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

129 **Concluding remarks**

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

140 **Acknowledgements**

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

147 **Author contributions**

148 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.

149 **References**

150 Alvarez, S. A. (2002). *An exact analytical relation among recall, precision, and classification accuracy in information*
151 *retrieval* (Technical Report No. BCCS-02-01). Boston College.

- 152 Cole, M. W., Bassett, D. S., Power, J. D., Braver, T. S., & Petersen, S. E. (2014). Intrinsic and task-evoked
153 network architectures of the human brain. *Neuron*, 83(1), 238–251.
- 154 Finn, E. S., Scheinost, D., Finn, D. M., Shen, X., Papademetris, X., & Constable, R. T. (2017). Can brain state
155 be manipulated to emphasize individual differences in functional connectivity. *NeuroImage*, 160, 140–151.
- 156 Finn, E. S., Shen, X., Scheinost, D., Rosenberg, M. D., Huang, J., Chun, M. M., ... Constable, R. T. (2015).
157 Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. *Nature
158 Neuroscience*, 18, 1664–1671.
- 159 Glerean, E., Salmi, J., Lahnakoski, J. M., Jääskeläinen, I. P., & Sams, M. (2012). Functional magnetic resonance
160 imaging phase synchronization as a measure of dynamic functional connectivity. *Brain Connectivity*, 2(2),
161 91–101.
- 162 Gratton, C., Laumann, T. O., Nielsen, A. N., Greene, D. J., Gordon, E. M., Gilmore, A. W., ... Petersen, S. E.
163 (2018). Functional brain networks are dominated by stable group and individual factors, not cognitive or
164 daily variation. *Neuron*, 98(2), 439–452.
- 165 Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: multi-voxel pattern
166 analysis of fMRI data. *Trends in Cognitive Sciences*, 10(9), 424–430.
- 167 Owen, L. L. W., Muntianu, T. A., Heusser, A. C., Daly, P., Scangos, K. W., & Manning, J. R. (2020). A Gaussian
168 process model of human electrocorticographic data. *Cerebral Cortex*, 30(10), 5333–5345.
- 169 Scangos, K. W., Khambhati, A. N., Daly, P. M., Owen, L. L. W., Manning, J. R., Ambrose, J. B., ... Chang,
170 E. F. (2021). Biomarkers of depression symptoms defined by direct intracranial neurophysiology. *Frontiers
171 in Human Neuroscience*, In press.
- 172 Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27(3),
173 379–423.
- 174 Simony, E., Honey, C. J., Chen, J., & Hasson, U. (2016). Dynamic reconfiguration of the default mode
175 network during narrative comprehension. *Nature Communications*, 7(12141), 1–13.
- 176 Turk-Browne, N. B. (2013). Functional interactions as big data in the human brain. *Science*, 342, 580–584.
- 177 Yeo, B. T. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., ... Buckner, R. L.
178 (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity.
179 *Journal of Neurophysiology*, 106(3), 1125–1165.