

1                   

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

3                   Lucy L. W. Owen<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>Carney Institute for Brain Sciences,  
Brown University, Providence, RI

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

4                   February 20, 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 are fully coupled and behave identically. Each extreme optimizes key properties of how  
24                  the network functions. When every node is independent, the network is maximally *expressive*: if we define  
25                  the network’s “state” as the total set of activity patterns across nodes, then every state is equally reachable by  
26                  a network with fully independent nodes. On the other hand, a fully coupled network optimizes *robustness*:  
27                  any subset of nodes, other than the entire network, may be removed from the network without any loss of  
28                  function or expressive power. Note that a given set of nodes might reconfigure its connections or behaviors  
29                  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 the above extremes. We wondered: might the  
31 human brain reconfigure itself to be more flexible or more robust according to ongoing demands?

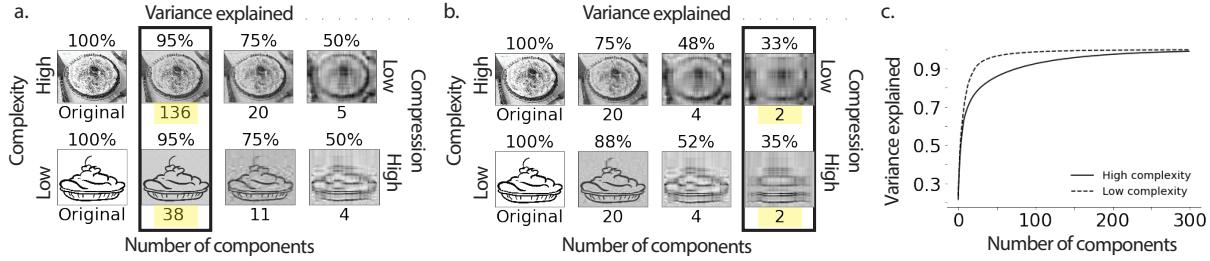
32 We're interested in the complexity of brain patterns that underly different types of thoughts. To explore  
33 this question space, we will take brain patterns recorded under different experimental conditions used in  
34 Aim 2, and project them into lower dimensional spaces using principle components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

37 This work has been inspired, in part, by Mack et al. (2020). In this paper, they investigated the role of  
38 the prefrontal cortex in filtering out irrelevant content. Specifically, they looked at if the vmPFC performs  
39 data reduction on incoming information through compression. This was motivated, in part, by orbital  
40 frontal cortex (OFC) compression in rats (Zhou et al., 2019). They studied this using a learning paradigm in  
41 which participants had to classify insects based on different numbers of feature dimensions. The idea was  
42 that participants in some learning blocks, participants could identify the insects based on one feature (low  
43 complexity) or several features (high complexity), but importantly the stimuli remained the same across all  
44 learning problems. They found that complexity and compression had an inverse relationship; the lower  
45 complexity of a conceptual space, the higher the degree of compression. Building on this idea, we wonder  
46 if varying degrees of compression is performed throughout the brain. We also want to test this idea, but  
47 using varying levels of engagement listening to a naturalistic stimuli.

48 To understand the degree of compression throughout the brain during cognition, we will use the same  
49 fMRI data from Aim 2, collected while participants listened to a story in different scrambling conditions.  
50 We will measure the degree that multivoxel activation patterns are compressed during story listening  
51 using principle components analysis (PCA) a method for low-rank approximation of multidimensional  
52 data (Eckart & Young, 1936). We will explore this using decoding accuracy as a function of the number of  
53 components, or dimensions, in the low-dimensional space under different cognitive conditions.

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

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



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

## 67 Evaluation metrics

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

## 71 Timepoint decoding

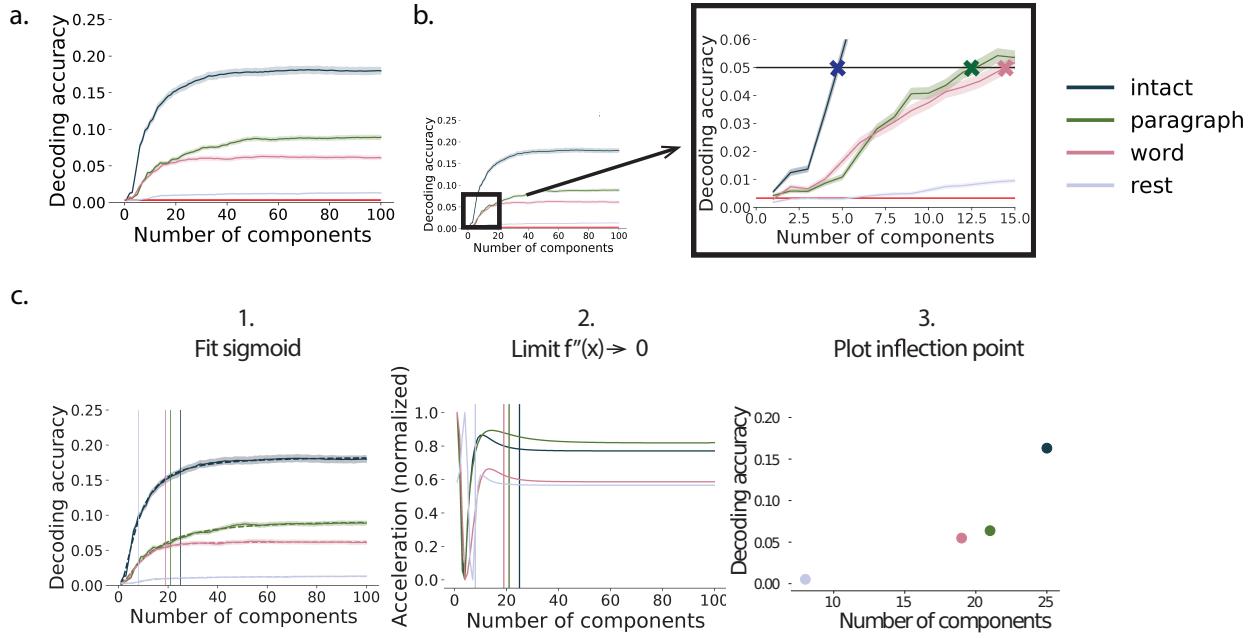
72 To explore how compression varies with complexity, we will use a previous neuroimaging dataset Simony  
 73 et al. (2016) in which participants listened to an audio recording of a story; 36 participants listen to an intact  
 74 version of the story, 17 participants listen to time-scrambled recordings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

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

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

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

81 To assess decoding accuracy, we randomly divide participants for each stimulus into training and testing  
 82 groups. We then compare the groups' activity patterns (using Pearson correlations) to estimate the story  
 83 times each corresponding pattern using more and more principle components (as the data became less



**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.

84 compressed). Specifically, we ask, for each timepoint: what are the correlations between the first group's  
 85 and second group's activity patterns at each order. We note that the decoding test we used is a conservative  
 86 in which we count a timepoint label as incorrect if it is not an exact match.

## 87 Results

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

93 Prior work has shown participants share similar neural responses to richly structured stimuli when  
 94 compared to stimuli with less structure Simony et al. (2016). We replicate this finding, showing as complexity

95 of the stimulus increases, decoding accuracy increases (Fig. 2, a.). Additionally, we found that as complexity  
96 of the stimuli increases, we need fewer components to decode the same amount (Fig. 2, b.). However, we  
97 also found that as complexity of the stimuli increases, more components are required to reach peak decoding  
98 accuracy (Fig. 2, c.). We posit that as the complexity of our thoughts increases, neural compression decreases.  
99 However, as our thoughts become deeper and richer, more reliable information is available at higher neural  
100 compression.

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

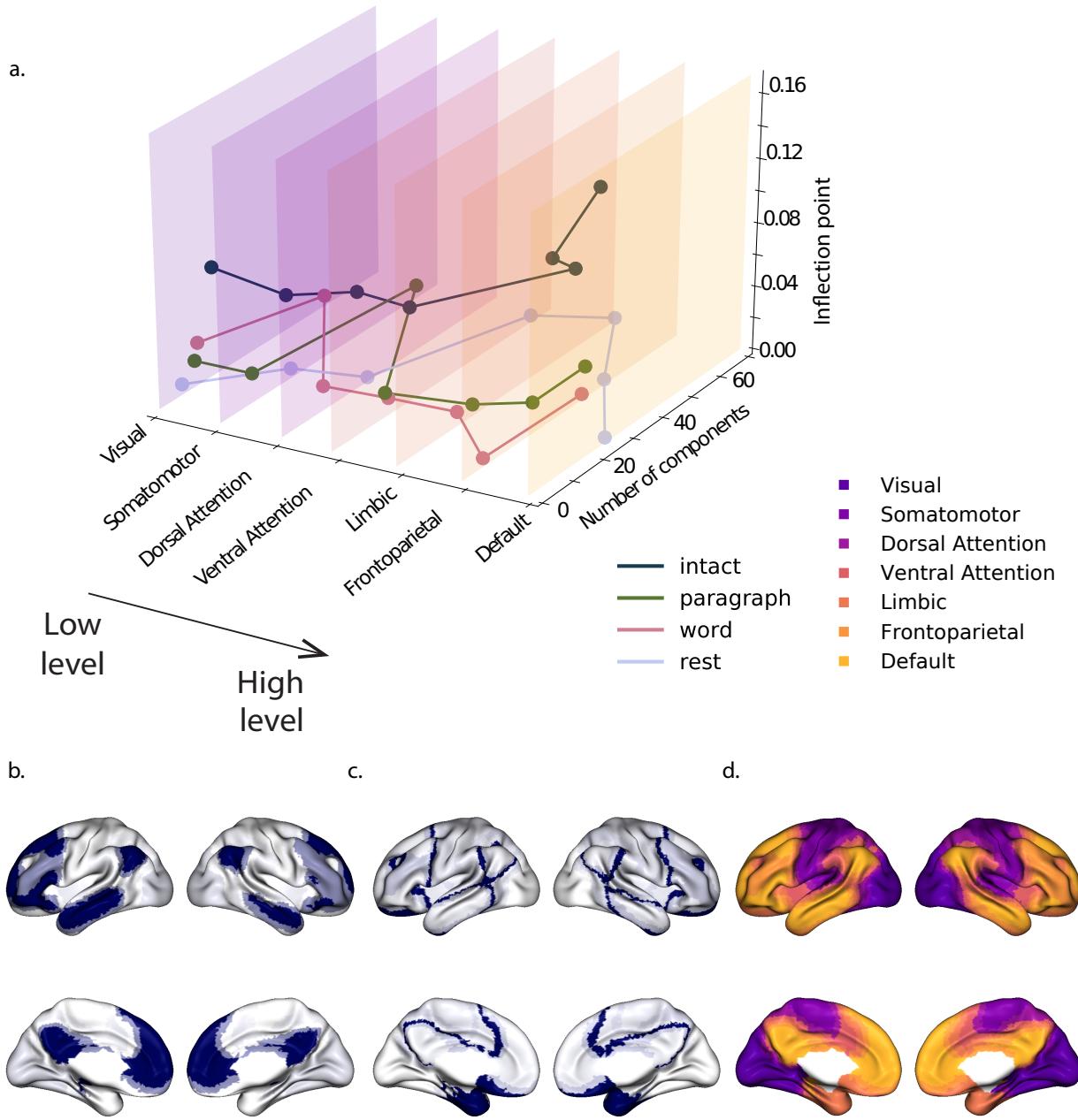
105 We were also curious how compression would change across time. If, there is some understanding of  
106 the narrative that accumulates over time, we should be able to see that difference. We found increases  
107 in decoding accuracy with the same number or fewer components for more complex, cognitively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

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

## 114 Discussion

115 - We trained classifiers using more and more principle components to decode, and compared across condi-  
116 tions with varying degrees of cognitive richness. -We found that as listening conditions become more  
117 cognitively rich, decoding accuracy increased. -Also, decoding accuracy increased as understanding of the  
118 narrative accumulated over time, in more complex listening conditions. - Decoding accuracy also increased  
119 in higher cognitive areas, in more complex listening conditions. -We found that as story listening conditions  
120 become more complex, more components are required to decode. -We also found we could decode better  
121 with more impoverished data when there is the underlying structure of the narrative providing more  
122 cognitive richness. -We posit that as the complexity of our thoughts increases, neural compression decreases.  
123 However, as our thoughts become deeper and richer, more reliable information is available at higher neural  
124 compression.

125 Based on prior work (?) and following the direction of the field (Turk-Browne, 2013) we think our



**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.

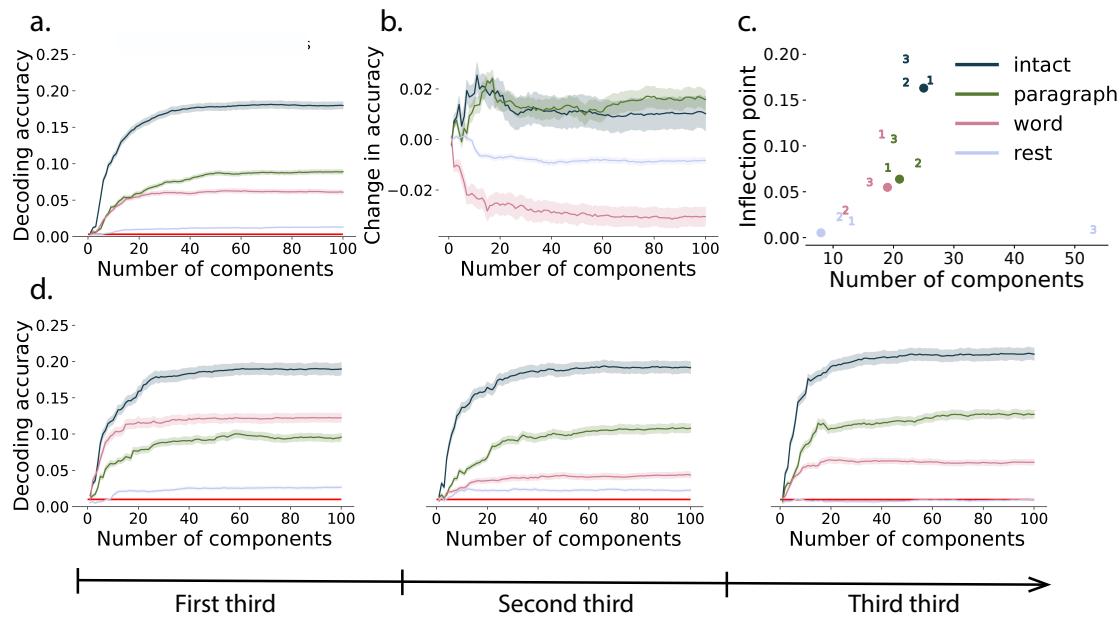


Figure 4: **Inflection points by thirds.** **a.** Decoding accuracy by number of components not broken into thirds (Fig. 2 a.). **b.** and **c.** Quantifying changes in decoding accuracy across time. **b.** Slope of decoding accuracy was calculated by fitting a regression line for each component/condition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

126 thoughts might be encoded in dynamic network patterns, and possibly higher order network patterns  
127 (Fig. ??). We sought to test this hypothesis by developing an approach to inferring high-order network  
128 dynamics from timeseries data.

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

134 We first validated our model using synthetic data, and explored how recovery varied with different  
135 underlying data structures and kernels. We then applied the approach to an fMRI dataset (Simony et al.,  
136 2016) in which participants listened to an audio recording of a story, as well as scrambled versions of the  
137 same story (where the scrambling was applied at different temporal scales). We trained classifiers to take  
138 the output of the model and decode the timepoint in the story (or scrambled story) that the participants  
139 were listening to. We found that, during the intact listening condition in the experiment, classifiers that  
140 incorporated higher-order correlations yielded consistently higher accuracy than classifiers trained only  
141 on lower-order patterns (Fig. ??, a.&d.). By contrast, these higher-order correlations were not necessary  
142 to support decoding the other listening conditions and (minimally above chance) during a control rest  
143 condition. This suggests that the cognitive processing that supported the most cognitively rich listening  
144 conditions involved second-order (or higher) network dynamics.

145 Although we found decoding accuracy was best when incorporating higher-order network dynamics  
146 for all but rest condition, it is unclear if this is a product of the brain or the data collection technique. It could  
147 be that the brain is second-order or it could be that fMRI can only reliably give second-order interactions.  
148 Exploring this method with other data collection technique will be important to disentangle this question.

## 149 **Concluding remarks**

150 How can we better understand how brain patterns change over time? How can we quantify the potential  
151 network dynamics that might be driving these changes? One way to judge the techniques of the future is  
152 to look at the trajectory of the fMRI field so far has taken so far (Fig. ??). The field started with univariate  
153 activation, measuring the average activity for each voxel. Analyses of multivariate activation followed,  
154 looking at spatial patterns of activity over voxels. Next, correlations of activity were explored, first with  
155 measures like resting connectivity that take temporal correlation between a seed voxel and all other voxels  
156 then with full connectivity that measure all pairwise correlations. Additionally, this path of increasing

157 complexity also moved from static to dynamic measurements. One logical next step in this trajectory would  
158 be dynamic higher-order correlations. We have created a method to support these calculations by scalably  
159 approximating dynamic higher-order correlations.

160 **Acknowledgements**

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

167 **Author contributions**

168 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.

169 **References**

- 170 Eckart, C., & Young, G. (1936). The approximation of one matrix by another of lower rank. *Psychometrika*,  
171 1, 211–218.
- 172 Mack, M. L., Preston, A. R., & Love, B. C. (2020). Ventromedial prefrontal cortex compression during  
173 concept learning. *Nature Communications*, 11(1), 46. Retrieved from <https://doi.org/10.1038/s41467-019-13930-8> doi: 10.1038/s41467-019-13930-8
- 174
- 175 Manning, J. R., Zhu, X., Willke, T. L., Ranganath, R., Stachenfeld, K., Hasson, U., . . . Norman, K. A. (2018).  
176 A probabilistic approach to discovering dynamic full-brain functional connectivity patterns. *NeuroImage*,  
177 180, 243–252.
- 178 Simony, E., Honey, C. J., Chen, J., & Hasson, U. (2016). Dynamic reconfiguration of the default mode  
179 network during narrative comprehension. *Nature Communications*, 7(12141), 1–13.
- 180 Turk-Browne, N. B. (2013). Functional interactions as big data in the human brain. *Science*, 342, 580–584.

- <sup>181</sup> Yeo, B. T. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., . . . Buckner, R. L.  
<sup>182</sup> (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity.  
<sup>183</sup> *Journal of Neurophysiology*, 106(3), 1125–1165.
- <sup>184</sup> Zhou, J., Gardner, M. P. H., Stalnaker, T. A., Ramus, S. J., Wikenheiser, A. M., Niv, Y., & Schoenbaum,  
<sup>185</sup> G. (2019). Rat orbitofrontal ensemble activity contains multiplexed but dissociable representations of  
<sup>186</sup> value and task structure in an odor sequence task. *Current Biology*, 29(6), 897–907.e3. Retrieved from  
<sup>187</sup> <https://www.sciencedirect.com/science/article/pii/S0960982219300776> doi: <https://doi.org/10.1016/j.cub.2019.01.048>