

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 22, 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

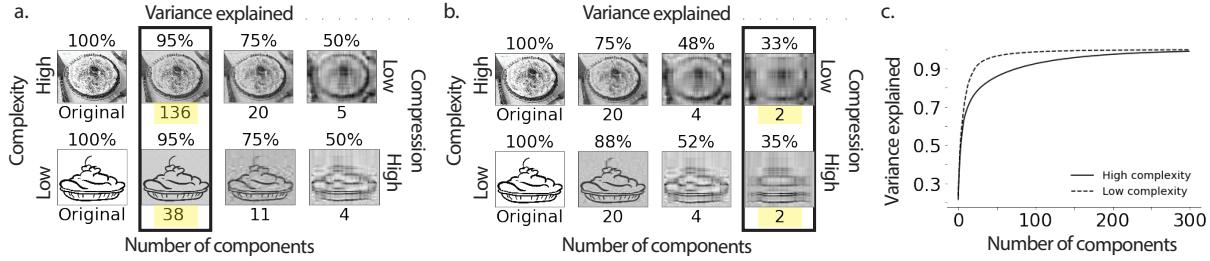


Figure 1: **Illustration of compression.** Visual analogy for neural compression. Here are 2 images of pies, one more complex than the other. **a.** It takes fewer components to reach the same percent variance explained in the less complex pie, which corresponds to higher compression. **b.** However,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

63 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. ??).

67 Specifically, if a given signal (e.g., a representation of brain activity patterns) contains more information
 68 about ongoing cognitive processes, then the peak decoding accuracy should be high. And if the signal
 69 is compressible, a low-dimensional embedding of the signal will be similarly informative to the original
 70 signal. A signal can be informative but not compressible, informative but

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

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

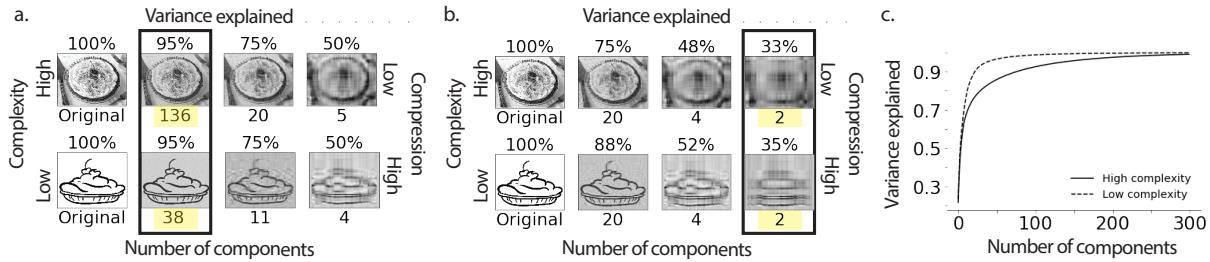


Figure 2: **Illustration of compression.** Visual analogy for neural compression. Here are 2 images of pies, one more complex than the other. **a.** It takes fewer components to reach the same percent variance explained in the less complex pie, which corresponds to higher compression. **b.** However,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

86 listening to a naturalistic stimuli.

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

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

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

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

106 **Evaluation metrics**

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

110 **Timepoint decoding**

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

115 Following the analyses conducted by (HTFA) Manning et al. (2018), we first apply *hierarchical topographic*
116 *factor analysis* (HTFA) to the fMRI datasets to obtain a time series of 700 node activities for every participant.
117 We then apply dimensionality reduction (Incremental PCA) for each group.

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

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

126 **Results**

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

132 Prior work has shown participants share similar neural responses to richly structured stimuli when
133 compared to stimuli with less structure Simony et al. (2016). We replicate this finding, showing as complexity
134 of the stimulus increases, decoding accuracy increases (Fig. 3, a.). Additionally, we found that as complexity

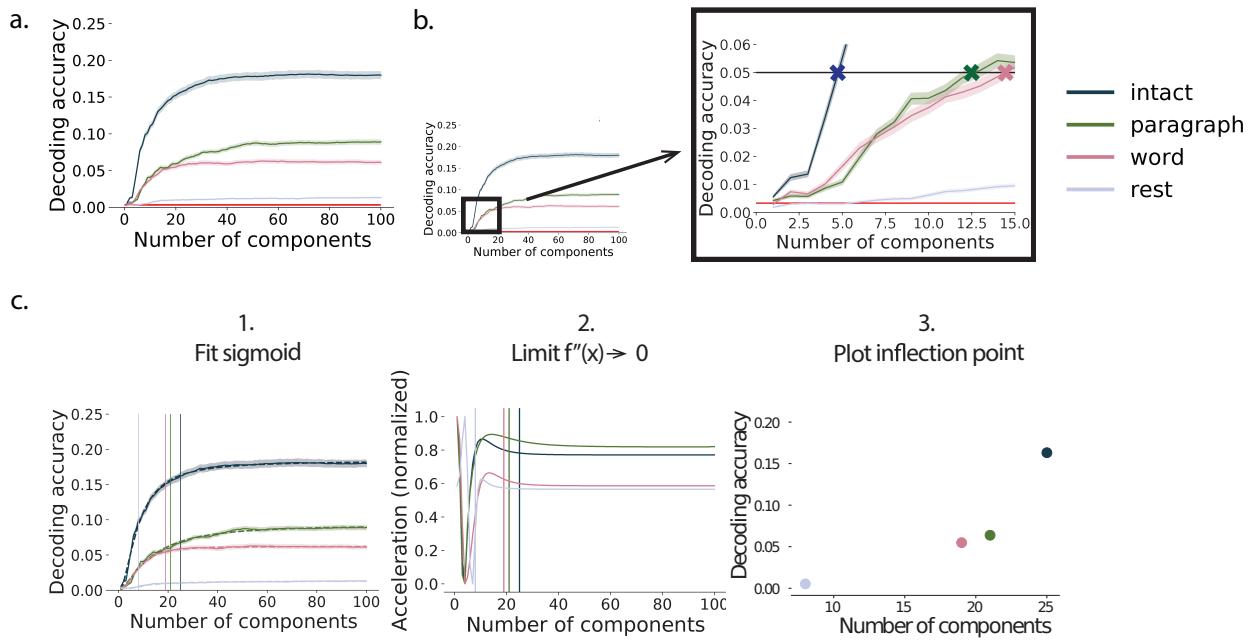


Figure 3: **Decoding accuracy.** **a. Decoding accuracy by number of components.** Ribbons of each color display cross-validated decoding performance for each condition (intact, paragraph, word, and rest). Decoders were trained using increasingly more principle components and displayed relative to chance (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.

135 of the stimuli increases, we need fewer components to decode the same amount (Fig. 3, b.). However, we
136 also found that as complexity of the stimuli increases, more components are required to reach peak decoding
137 accuracy (Fig. 3, c.). We posit that as the complexity of our thoughts increases, neural compression decreases.
138 However, as our thoughts become deeper and richer, more reliable information is available at higher neural
139 compression.

140 We also wondered how this compression would change across brain regions. We repeated the analysis
141 but limited the brain hubs to 7 networks using the Yeo et al. (2011) network parcellation shown here in the
142 inflated brain (Fig. 4, d.). We found that as complexity of the stimuli increases, decoding accuracy increases
143 with higher cognitive areas. (Fig. 4).

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

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

153 Discussion

154 - We trained classifiers using more and more principle components to decode, and compared across condi-
155 tions with varying degrees of cognitive richness. -We found that as listening conditions become more
156 cognitively rich, decoding accuracy increased. -Also, decoding accuracy increased as understanding of the
157 narrative accumulated over time, in more complex listening conditions. - Decoding accuracy also increased
158 in higher cognitive areas, in more complex listening conditions. -We found that as story listening conditions
159 become more complex, more components are required to decode. -We also found we could decode better
160 with more impoverished data when there is the underlying structure of the narrative providing more
161 cognitive richness. -We posit that as the complexity of our thoughts increases, neural compression decreases.
162 However, as our thoughts become deeper and richer, more reliable information is available at higher neural
163 compression.

164 Based on prior work (?) and following the direction of the field (Turk-Browne, 2013) we think our
165 thoughts might be encoded in dynamic network patterns, and possibly higher order network patterns

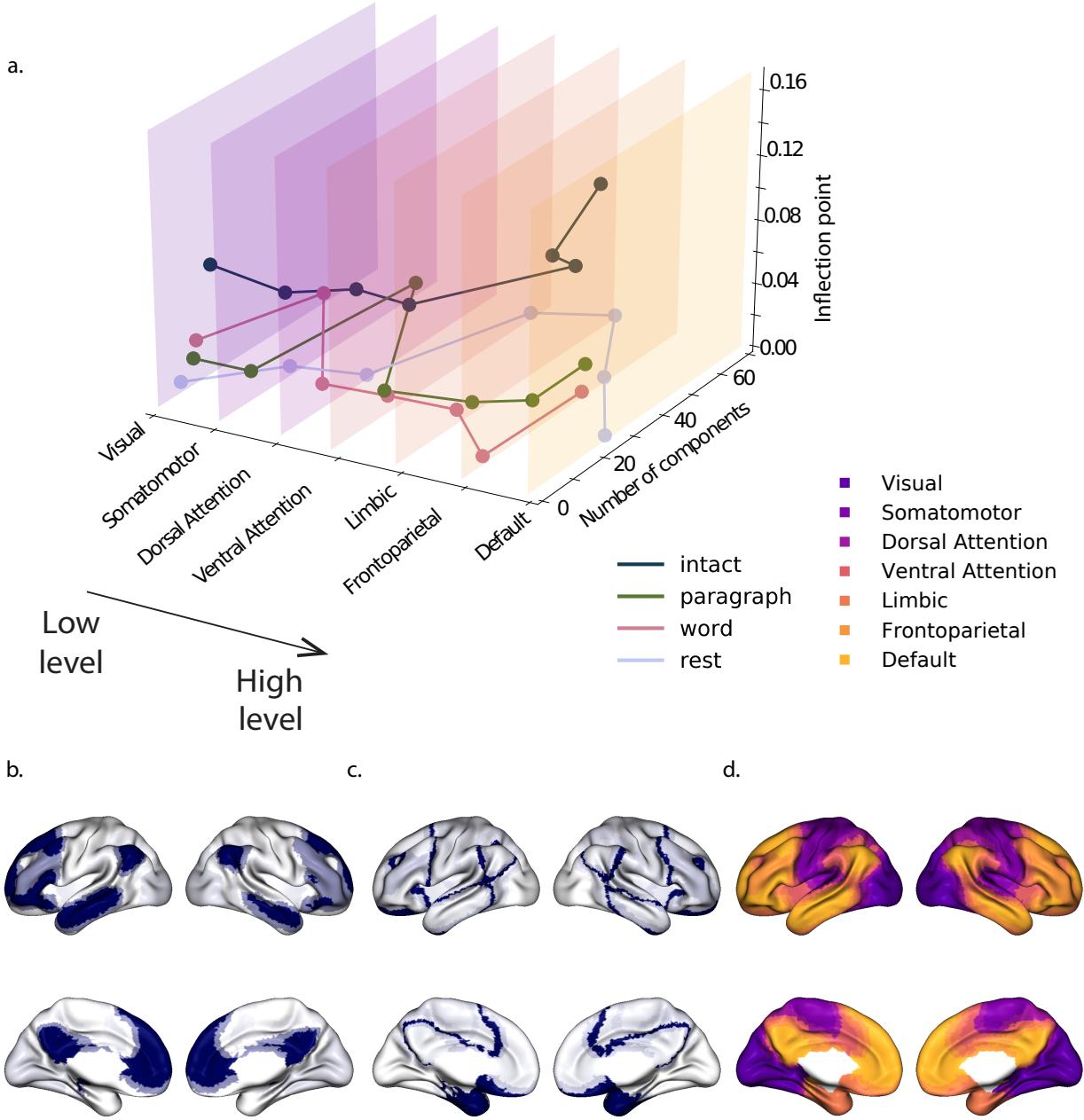


Figure 4: Inflection points by network. a. Inflection point was calculated as explained in Fig. 3, b. Analyses were limited by the brain networks (using the Yeo et al. (2011) network parcellation) and arranged in increasing order relative to the intact condition. b. and c. For the total time in the intact 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.

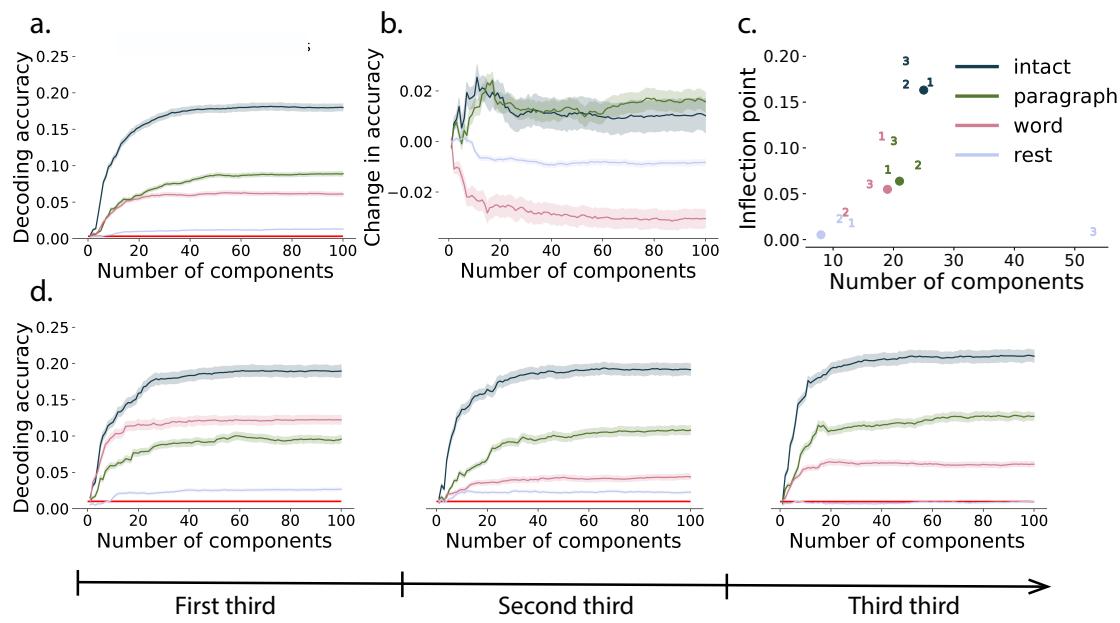


Figure 5: **Inflection points by thirds.** **a.** Decoding accuracy by number of components not broken into thirds (Fig. 3 a.). **b.** and **c.** Quantifying changes in decoding accuracy across time. **b.** Slope of decoding accuracy was calculated by fitting a regression line for each component/condition for each third. **c.** We also repeated the analysis (Fig. 3, 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. 3 a. but breaking the scan time for each condition into 3 intervals.

166 (Fig. ??). We sought to test this hypothesis by developing an approach to inferring high-order network
167 dynamics from timeseries data.

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

173 We first validated our model using synthetic data, and explored how recovery varied with different
174 underlying data structures and kernels. We then applied the approach to an fMRI dataset (Simony et al.,
175 2016) in which participants listened to an audio recording of a story, as well as scrambled versions of the
176 same story (where the scrambling was applied at different temporal scales). We trained classifiers to take
177 the output of the model and decode the timepoint in the story (or scrambled story) that the participants
178 were listening to. We found that, during the intact listening condition in the experiment, classifiers that
179 incorporated higher-order correlations yielded consistently higher accuracy than classifiers trained only
180 on lower-order patterns (Fig. ??, a.&d.). By contrast, these higher-order correlations were not necessary
181 to support decoding the other listening conditions and (minimally above chance) during a control rest
182 condition. This suggests that the cognitive processing that supported the most cognitively rich listening
183 conditions involved second-order (or higher) network dynamics.

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

188 **Concluding remarks**

189 How can we better understand how brain patterns change over time? How can we quantify the potential
190 network dynamics that might be driving these changes? One way to judge the techniques of the future is
191 to look at the trajectory of the fMRI field so far has taken so far (Fig. ??). The field started with univariate
192 activation, measuring the average activity for each voxel. Analyses of multivariate activation followed,
193 looking at spatial patterns of activity over voxels. Next, correlations of activity were explored, first with
194 measures like resting connectivity that take temporal correlation between a seed voxel and all other voxels
195 then with full connectivity that measure all pairwise correlations. Additionally, this path of increasing
196 complexity also moved from static to dynamic measurements. One logical next step in this trajectory would

¹⁹⁷ be dynamic higher-order correlations. We have created a method to support these calculations by scalably
¹⁹⁸ approximating dynamic higher-order correlations.

¹⁹⁹ Acknowledgements

²⁰⁰ We acknowledge discussions with Luke Chang, Hany Farid, Paxton Fitzpatrick, Andrew Heusser, Eshin
²⁰¹ Jolly, Qiang Liu, Matthijs van der Meer, Judith Mildner, Gina Notaro, Stephen Satterthwaite, Emily Whitaker,
²⁰² Weizhen Xie, and Kirsten Ziman. Our work was supported in part by NSF EPSCoR Award Number 1632738
²⁰³ to J.R.M. and by a sub-award of DARPA RAM Cooperative Agreement N66001-14-2-4-032 to J.R.M. The
²⁰⁴ content is solely the responsibility of the authors and does not necessarily represent the official views of our
²⁰⁵ supporting organizations.

²⁰⁶ Author contributions

²⁰⁷ 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.

²⁰⁸ References

- ²⁰⁹ Alvarez, S. A. (2002). *An exact analytical relation among recall, precision, and classification accuracy in information retrieval* (Technical Report No. BCCS-02-01). Boston College.
- ²¹¹ Cole, M. W., Bassett, D. S., Power, J. D., Braver, T. S., & Petersen, S. E. (2014). Intrinsic and task-evoked
²¹² network architectures of the human brain. *Neuron*, 83(1), 238–251.
- ²¹³ Eckart, C., & Young, G. (1936). The approximation of one matrix by another of lower rank. *Psychometrika*,
²¹⁴ 1, 211–218.
- ²¹⁵ Finn, E. S., Scheinost, D., Finn, D. M., Shen, X., Papademetris, X., & Constable, R. T. (2017). Can brain state
²¹⁶ be manipulated to emphasize individual differences in functional connectivity. *NeuroImage*, 160, 140–151.
- ²¹⁷ Finn, E. S., Shen, X., Scheinost, D., Rosenberg, M. D., Huang, J., Chun, M. M., ... Constable, R. T. (2015).
²¹⁸ Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. *Nature Neuroscience*, 18, 1664–1671.
- ²²⁰ Glerean, E., Salmi, J., Lahnakoski, J. M., Jääskeläinen, I. P., & Sams, M. (2012). Functional magnetic resonance
²²¹ imaging phase synchronization as a measure of dynamic functional connectivity. *Brain Connectivity*, 2(2),
²²² 91–101.

- 223 Gratton, C., Laumann, T. O., Nielsen, A. N., Greene, D. J., Gordon, E. M., Gilmore, A. W., ... Petersen, S. E.
224 (2018). Functional brain networks are dominated by stable group and individual factors, not cognitive or
225 daily variation. *Neuron*, 98(2), 439–452.
- 226 Manning, J. R., Zhu, X., Willke, T. L., Ranganath, R., Stachenfeld, K., Hasson, U., ... Norman, K. A. (2018).
227 A probabilistic approach to discovering dynamic full-brain functional connectivity patterns. *NeuroImage*,
228 180, 243–252.
- 229 Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: multi-voxel pattern
230 analysis of fMRI data. *Trends in Cognitive Sciences*, 10(9), 424–430.
- 231 Owen, L. L. W., Muntianu, T. A., Heusser, A. C., Daly, P., Scangos, K. W., & Manning, J. R. (2020). A Gaussian
232 process model of human electrocorticographic data. *Cerebral Cortex*, 30(10), 5333–5345.
- 233 Scangos, K. W., Khambhati, A. N., Daly, P. M., Owen, L. L. W., Manning, J. R., Ambrose, J. B., ... Chang,
234 E. F. (2021). Biomarkers of depression symptoms defined by direct intracranial neurophysiology. *Frontiers*
235 in *Human Neuroscience*, In press.
- 236 Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27(3),
237 379–423.
- 238 Simony, E., Honey, C. J., Chen, J., & Hasson, U. (2016). Dynamic reconfiguration of the default mode
239 network during narrative comprehension. *Nature Communications*, 7(12141), 1–13.
- 240 Turk-Browne, N. B. (2013). Functional interactions as big data in the human brain. *Science*, 342, 580–584.
- 241 Yeo, B. T. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., ... Buckner, R. L.
242 (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity.
243 *Journal of Neurophysiology*, 106(3), 1125–1165.