

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>

4            <sup>1</sup>Department of Psychological and Brain Sciences,  
5            Dartmouth College, Hanover, NH

6            <sup>2</sup>Carney Institute for Brain Sciences,  
7            Brown University, Providence, RI

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

9            February 28, 2023

10            **Abstract**

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

27            **Keywords:** information, compression, temporal decoding, dimensionality reduction, neu-  
28            roimaging

29            **Introduction**

30            Large-scale networks, including the human brain, may be conceptualized as occupying one or  
31            more positions along on a continuum. At one extreme, every node is fully independent of every  
32            other node. At the other extreme, all nodes behave identically. Each extreme optimizes key  
33            properties of how the network functions. When every node is independent, the network is

29 maximally *expressive*: if we define the network’s “state” as the activity pattern across its nodes,  
30 then every state is equally reachable by a network with fully independent nodes. On the other  
31 hand, a network of identically behaved nodes optimizes *robustness*: any subset of nodes may  
32 be removed from the network without any loss of function or expressive power, as long as any  
33 single node remains. Presumably, most natural systems tend to occupy positions between these  
34 extremes. We wondered: might the human brain reconfigure itself to be more flexible or more  
35 robust according to ongoing demands? In other words, might the brain reconfigure its connections  
36 or behaviors under different circumstances to change its position along this continuum?

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

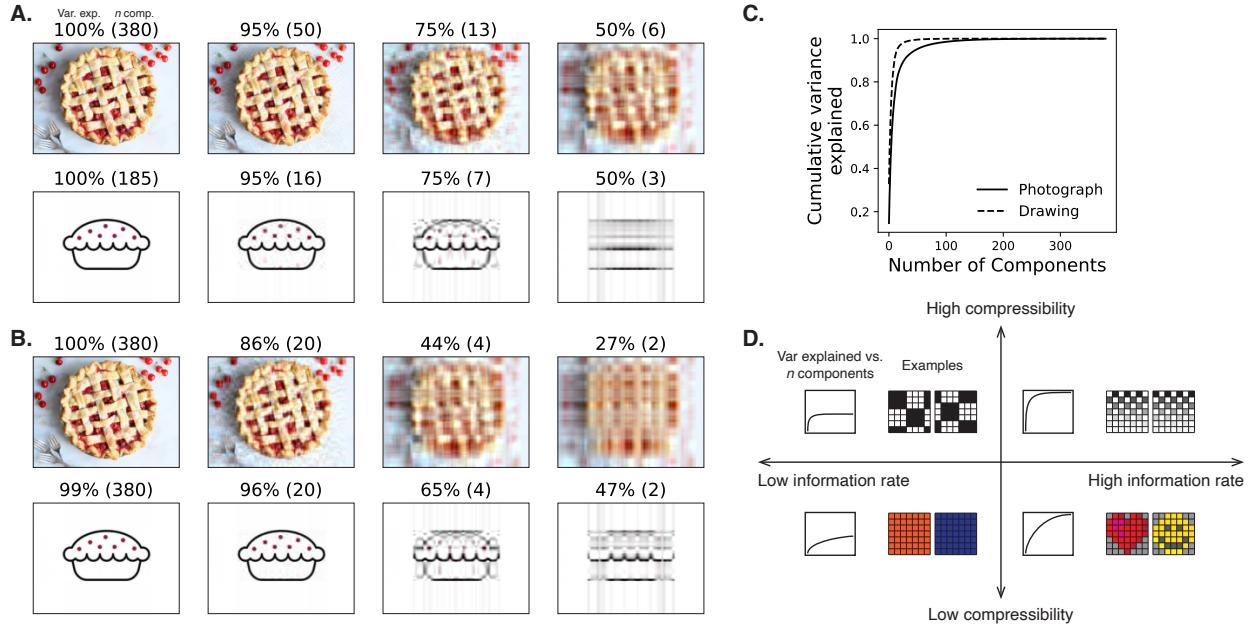
55 We take as a given that brain activity is highly flexible: our brains can exhibit nearly infinite ac-  
56 tivity patterns. This flexibility implies that our brains activity patterns are highly information rich.  
57 However, brain activity patterns are also highly structured. For example, full-brain correlation  
58 matrices are stable within (Finn et al., 2015, 2017; Gratton et al., 2018) and across (Yeo et al., 2011;

59 Gleerean et al., 2012; Gratton et al., 2018; Cole et al., 2014) individuals. This stability suggests that  
60 our brains' activity patterns are at least partially constrained, for example by anatomical, external,  
61 or internal factors. Constraints on brain activity that limit its flexibility decrease expressiveness  
62 (i.e., its information rate). However, constraints on brain activity also increase its robustness to  
63 noise (e.g., “missing” or corrupted signals may be partially recovered). For example, recent work  
64 has shown that full-brain activity patterns may be reliably recovered from only a relatively small  
65 number of implanted electrodes (Owen et al., 2020; Scangos et al., 2021). This robustness property  
66 suggests that the relevant signal (e.g., underlying factors that have some influence over brain  
67 activity patterns) are compressible.

68 To the extent that brain activity patterns contain rich task-relevant information, we should  
69 be able to use the activity patterns to accurately differentiate between different aspects of the  
70 task (e.g., using pattern classifiers; Norman et al., 2006). For example, prior work has shown a direct  
71 correspondence between classification accuracy and the information content of a signal (Alvarez,  
72 2002). To the extent that brain activity patterns are compressible, we should be able to generate  
73 simplified (e.g., lower dimensional) representations of the data while still preserving the relevant  
74 or important aspects of the original signal. In general, information content and compressibility are  
75 related but are partially dissociable (Fig. 1). If a given signal (e.g., a representation of brain activity  
76 patterns) contains more information about ongoing cognitive processes, then the peak decoding  
77 accuracy should be high. And if the signal is compressible, a low-dimensional embedding of the  
78 signal will be similarly informative to the original signal (Fig. 1D).

79 Several recent studies suggest that the complexity of brain activity is task-dependent, whereby  
80 simpler tasks with lower cognitive demands are reflected by simpler and more compressible brain  
81 activity patterns, and more complex tasks with higher cognitive demands are reflected by more  
82 complex and less compressible brain activity patterns (Mack et al., 2020; Owen et al., 2021). These  
83 findings complement other work suggesting that functional connectivity (correlation) patterns are  
84 task-dependent (Finn et al., 2017; Owen et al., 2020; Cole et al., 2014), although see Gratton et  
85 al. (2018). Higher-order cognitive processing of a common stimulus also appears to drive more  
86 stereotyped task-related activity and functional connectivity across individuals (Hasson et al.,  
87 2008; Lerner et al., 2011; Simony & Chang, 2020; Simony et al., 2016).

88 The above studies are consistent with two potential descriptions of how cognitive processes are



**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.

89 reflects in brain activity patterns. One possibility is that the information rate of brain activity in-  
90 creases during more complex or higher-level cognitive processing. If so, then the ability to reliably  
91 decode cognitive states from brain activity patterns should improve with task complexity or with  
92 the level (or “depth”) of cognitive processing. A second possibility is that the compressibility of  
93 brain activity patterns increases during more complex or higher-level cognitive processing. If so,  
94 then individual features of brain recordings, or compressed representations of brain recordings,  
95 should carry more information during complex or high-level (versus simple or low-level) cognitive  
96 tasks.

97 We used a previously collected neuroimaging dataset to estimate the extent to which each of  
98 these two possibilities might hold. The dataset we examined comprised functional magnetic reso-  
99 nance imaging (fMRI) data collected as participants listened to an audio recording of a 10-minute  
100 story, temporally scrambled recordings of the story, or underwent a resting state scan (Simony  
101 et al., 2016). Each of these experimental conditions evokes different depths of cognitive process-  
102 ing (Simony et al., 2016; Lerner et al., 2011; Hasson et al., 2008; Owen et al., 2021). We used  
103 across-participant classifiers to decode listening times in each condition, as a proxy for how “in-  
104 formative” the task-specific activity patterns were (Simony & Chang, 2020). We also use principle  
105 components analysis to generate lower-dimensional representations of the activity patterns. We  
106 then repeated the classification analyses after preserving different numbers of components and  
107 examined how classification accuracy changed across the different experimental conditions.

## 108 Results

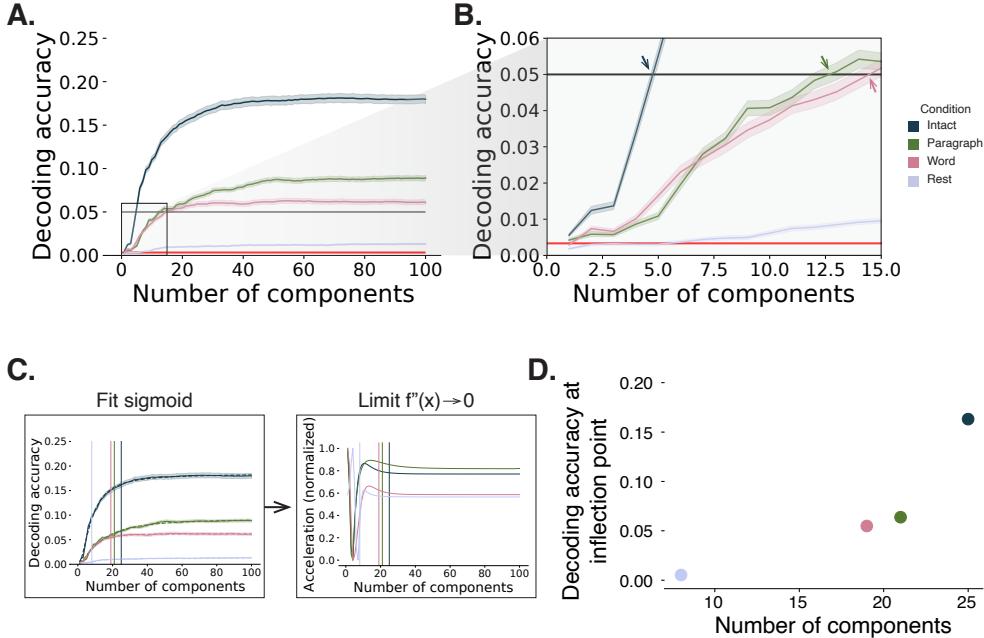
109 We sought to understand whether higher-level cognition is reflected by more reliable and in-  
110 formative brain activity patterns, and how compressibility of brain activity patterns relates to  
111 cognitive complexity. We developed a computational framework for systematically assessing the  
112 informativeness and compressibility of brain activity patterns recorded under different cognitive  
113 circumstances. We used across-participant decoding accuracy (see *Forward inference and decoding*  
114 *accuracy*) as a proxy for informativeness. To estimate the compressibility of the brain patterns, we  
115 used group principal components analysis (PCA) to project the brain patterns into  $k$ -dimensional  
116 spaces, for different values of  $k$  (see *Hierarchical Topographic Factor Analysis (HTFA)* and *Principal*

117 components analysis (PCA)). For more compressible brain patterns, decoding accuracy should be  
118 more robust to small values of  $k$ .

119 We analyzed a dataset collected by Simony et al. (2016) that comprised four experimental  
120 conditions. These conditions exposed participants to stimuli that systematically varied in cognitive  
121 engagement. In the *intact* experimental condition, participants listened to an audio recording of  
122 a 10-minute story. In the *paragraph*-scrambled experimental condition, participants listened to a  
123 temporally scrambled version of the story, where the paragraphs occurred out of order, but where  
124 the same set of paragraphs was presented over the entire listening interval. All participants in  
125 this condition experienced the scrambled paragraphs in the same order. In the *word*-scrambled  
126 experimental condition, participants listened to a temporally scrambled version of the story, where  
127 the words occurred in a random order. Again, all participants in this condition experienced the  
128 scrambled words in the same order. Finally, in the *rest* experimental condition, participants lay  
129 in the scanner with no overt stimulus, while keeping their eyes open and blinking as needed.  
130 This public dataset provided a convenient means for testing our hypothesis that different levels  
131 of cognitive processing and engagement affect how informative and compressible the associated  
132 brain patterns are.

133 To evaluate the relation between informativeness and compression for brain activity from each  
134 experimental condition, we trained a series of across-participant temporal decoders on compressed  
135 representations of the data. Figure 2A displays the decoding accuracy as a function of the number  
136 of principal components used to represent the data. Several patterns were revealed by the analysis.  
137 First, in general (i.e., across experimental conditions), decoding accuracy improves as the number  
138 of components increases. However, decoding accuracy peaked at a higher level for experimental  
139 conditions that exposed participants to cognitively richer stimuli. The peak decoding accuracy  
140 was highest for the “intact” condition (versus paragraph:  $t(\text{XXX}) = \text{XXX}, p = \text{XXX}$ ; versus word:  
141  $t(\text{XXX}) = \text{XXX}, p = \text{XXX}$ ; versus rest:  $t(\text{XXX}) = \text{XXX}, p = \text{XXX}$ ), next highest for the “paragraph”  
142 condition (versus word:  $t(\text{XXX}) = \text{XXX}, p = \text{XXX}$ ; versus rest:  $t(\text{XXX}) = \text{XXX}, p = \text{XXX}$ ), and next  
143 highest for the “word” condition (versus rest:  $t(\text{XXX}) = \text{XXX}, p = \text{XXX}$ ). This ordering implies  
144 that cognitively richer conditions evoke more stable brain activity patterns across people.

145 The cognitively richer conditions also displayed steeper initial slopes. For example, the intact  
146 condition decoders reached an arbitrarily chosen threshold of 5% accuracy using fewer components



**Figure 2: Decoding accuracy and compression.** **A. Decoding accuracy by number of components.** Ribbons of each color display cross-validated decoding performance for each condition (intact, paragraph, word, and rest), as a function of the number of components (features) used to represent the data. The horizontal red line denotes chance performance, and the horizontal black line denotes 5% decoding accuracy (used as a reference in Panel B). **B. Numbers of components required to reach a fixed decoding accuracy threshold, by condition.** The panel displays a zoomed-in view of the inset in Panel A. Intersections between each condition’s decoding accuracy curve and the 5% decoding accuracy reference line are marked by arrows. **C. Estimating inflection points.** We sought to identify an “inflection point” for each decoding curve, denoting the number of components at which the decoding curve asymptotes. We fit sigmoid functions to each decoding curve (left sub-panel) and then computed the minimum number of components where the second derivative of the sigmoid was both positive and less than a threshold value of 0.0001. **D. Inflection points by condition.** Each dot displays the number of components ( $x$ -axis) and decoding accuracy ( $y$ -axis) at one condition’s inflection point. All error ribbons denote bootstrap-estimated 95% confidence intervals.

147 than the paragraph condition decoders ( $t(\text{XXX}) = \text{XXX}, p = \text{XXX}$ ) or word condition decoders  
148 ( $t(\text{XXX}) = \text{XXX}, p = \text{XXX}$ ), and decoding accuracy never exceeded 5% for the rest condition. This  
149 suggests that brain activity patterns evoked by cognitively richer conditions are more compressible,  
150 such that representing the data using the same number of principal components provides more  
151 information to the temporal decoders (Fig. 2B).

152 In every experimental condition, decoding accuracy appeared to asymptote (i.e., hit an upper  
153 limit) beyond some characteristic number of components that differed across conditions. To  
154 quantify the “inflection points” at which the decoding curves in Figure 2A flattened out, we fit a  
155 sigmoid function to the average decoding curve for each condition. We defined the inflection point  
156 for each condition as the point on the fitted sigmoid where the second derivative was both positive  
157 and less than a threshold value of 0.0001 (i.e., approaching 0 from the right). These inflection  
158 points reflect a “balance” between higher decoding accuracy (which tends to be better when  
159 more components are used) and compression (which is better for fewer components). Plotting  
160 each condition’s inflection point (Fig. 2D) reveals that both the number of components and the  
161 decoding accuracy at each inflection point increase systematically across conditions in proportion  
162 to cognitive richness.

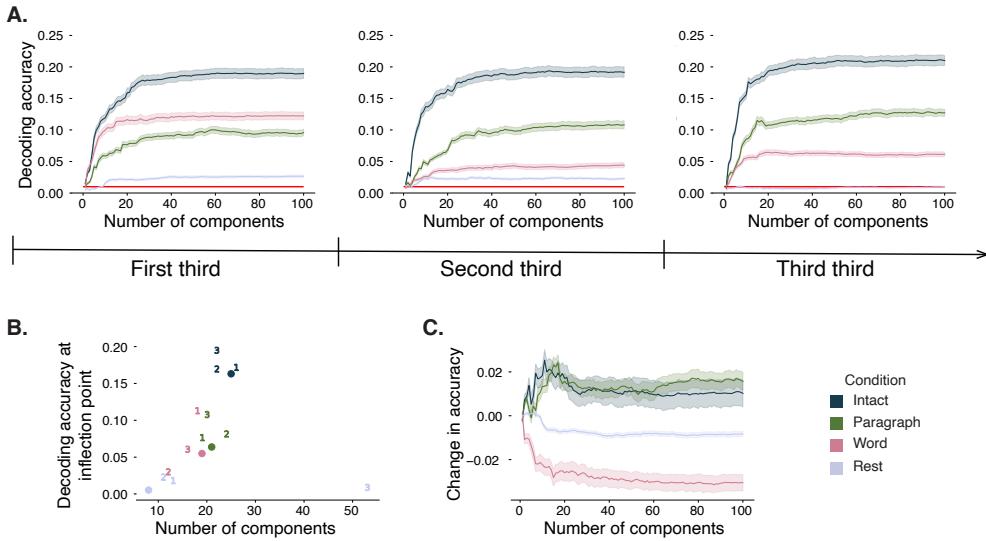
## 163 Discussion

## 164 Methods

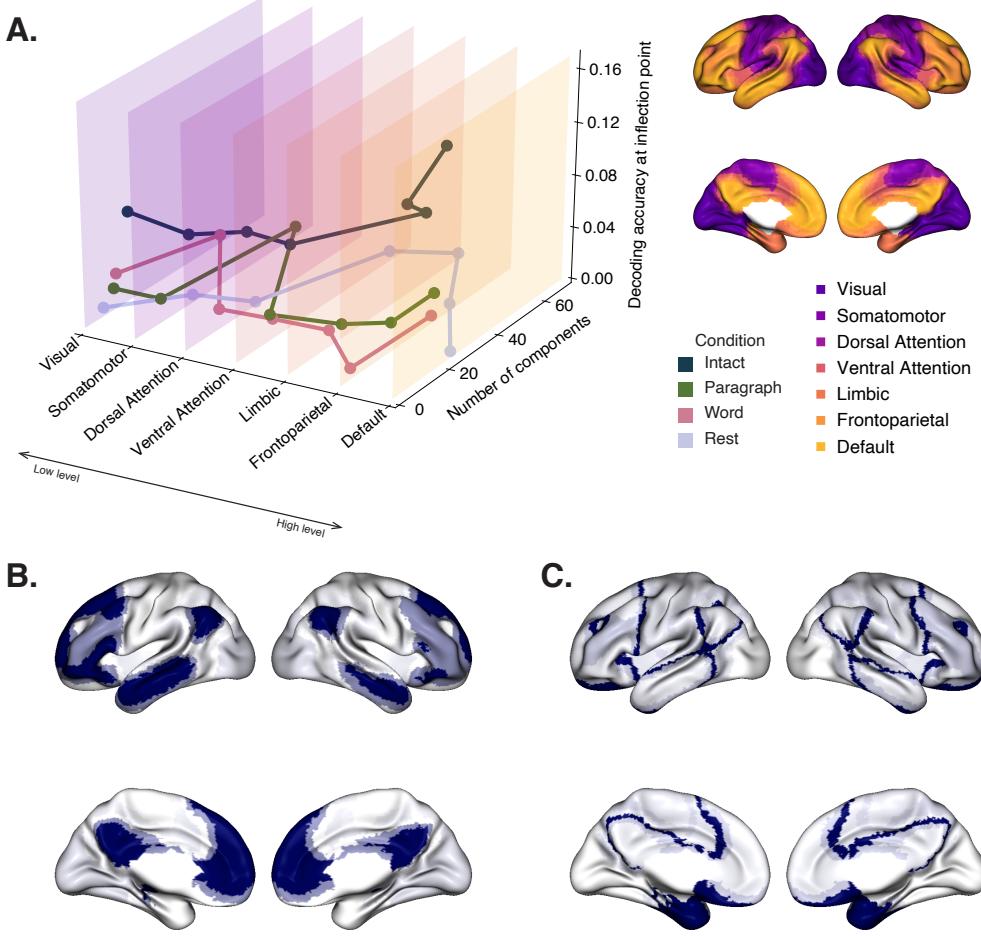
165 We measured properties of recorded neuroimaging data under different task conditions that varied  
166 systematically in cognitive engagement and depth of processing. We were especially interested in  
167 how *informative* and *compressible* the activity patterns were under these different conditions (Fig. 1).

### 168 Functional neuroimaging data collected during story listening

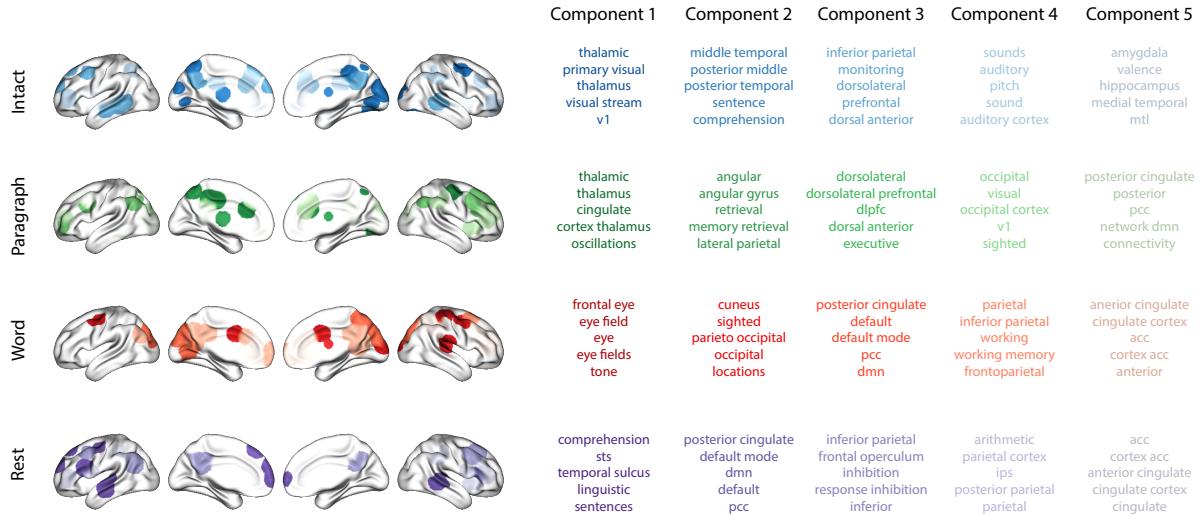
169 We examined an fMRI dataset collected by Simony et al. (2016) that the authors have made publicly  
170 available at [arks.princeton.edu/ark:/88435/dsp015d86p269k](http://arks.princeton.edu/ark:/88435/dsp015d86p269k). The dataset comprises neuroimaging  
171 data collected as participants listened to an audio recording of a story (intact condition; 36 par-  
172 ticipants), listened to temporally scrambled recordings of the same story (17 participants in the



**Figure 3: Changes in decoding accuracy and compression over time.** **A. Decoding accuracy by number of components, by story segment.** Each family of curves is plotted in the same format as Figure 2A but reflects data only from one third of the dataset. **B. Inflection points by condition and segment.** The dots re-plot the inflection points from Figure 2D for reference. The numbers denote the inflection points for each third of the dataset (1: first third; 2: second third; 3: third third; colors denote experimental conditions). **C. Change in decoding accuracy over time, by number of components.** For each number of components ( $x$ -axis) and condition (color), we fit a regression line to the decoding accuracies obtained for the first, second, and third thirds of the dataset (corresponding to the left, middle, and right columns of Panel A, respectively). The  $y$ -axis denotes the slopes of the regression lines. All error ribbons denote bootstrap-estimated 95% confidence intervals.



**Figure 4: Network-specific decoding accuracy and compression. A. Decoding accuracy and number of components for network-specific inflection points.** We considered the seven networks identified by Yeo et al. (2011). We computed each network's inflection point, for each experimental condition, using the procedure described in Figure 2C. **B. Network-specific decoding accuracy.** Each of the seven networks are colored according to the decoding accuracy at the network's inflection point for the "intact" experimental condition (corresponding to the dark blue curve in Panel A). **C. Network-specific compression.** Each of the seven networks are colored according to the number of components at the network's inflection point for the intact experimental condition. Larger numbers of components reflect lower compressibility.



**Figure 5: Top terms associated with the highest-weighted components by condition.** Each row corresponds to an experimental condition, and the colors correspond to the component number (ranked by proportion of variance explained). The inflated brain plots display the top 20 highest-weighted hubs (see *Topographic Factor Analysis*) for each components'. The lists on the right display the top five Neurosynth terms (Rubin et al., 2017) decoded from each components' brain map. Analogous maps computed separately for each story segment may be found in Figure S1.

173 paragraph-scrambled condition listened to the paragraphs in a randomized order and 36 in the  
 174 word-scrambled condition listened to the words in a randomized order), or lay resting with their  
 175 eyes open in the scanner (rest condition; 36 participants). Full neuroimaging details may be found  
 176 in the original paper for which the data were collected Simony et al. (2016). Procedures were  
 177 approved by the Princeton University Committee on Activities Involving Human Subjects, and by  
 178 the Western Institutional Review Board (Puyallup, WA). All subjects were native English speakers  
 179 with normal hearing and provided written informed consent.

## 180 Hierarchical topographic factor analysis (HTFA)

181 Following our prior related work, we used HTFA Manning et al. (2018) to derive a compact  
 182 representation of the neuroimaging data. In brief, this approach approximates the timeseries  
 183 of voxel activations (44,415 voxels) using a much smaller number of radial basis function (RBF)  
 184 nodes (in this case, 700 nodes, as determined by an optimization procedure; Manning et al., 2018)).  
 185 This provides a convenient representation for examining full-brain activity patterns and network  
 186 dynamics. All of the analyses we carried out on the neuroimaging dataset were performed in this

187 lower-dimensional space. In other words, each participant's data matrix,  $\mathbf{X}$ , was a number-of-  
188 timepoints ( $T$ ) by 700 matrix of HTFA-derived factor weights (where the row and column labels  
189 were matched across participants). Code for carrying out HTFA on fMRI data may be found as  
190 part of the BrainIAK toolbox Capota et al. (2017); Kumar et al. (2021), which may be downloaded  
191 at [brainiak.org](http://brainiak.org).

192 **Principal components analysis (PCA)**

193 We applied group PCA (Smith et al., 2014) separately to the HTFA-derived representations of the  
194 data (i.e., factor loadings) from each experimental condition. Specifically, for each condition, we  
195 considered the set of all participants'  $T$  by 700 factor weight matrices. We used group PCA to  
196 project these 700-dimensional matrices into a series of  $k$ -dimensional spaces, for  $k \in \{3, 4, 5, \dots, 700\}$ .  
197 This yielded a set of number-of-participants matrices, each with  $T$  rows and  $k$  columns.

198 **Temporal decoding**

199 We sought to identify neural patterns that reflected participants' ongoing cognitive processing of  
200 incoming stimulus information. As reviewed by Simony et al. (2016), one way of homing in on  
201 these stimulus-driven neural patterns is to compare activity patterns across individuals. In partic-  
202 ular, neural patterns will be similar across individuals to the extent that the neural patterns under  
203 consideration are stimulus-driven, and to the extent that the corresponding cognitive representa-  
204 tions are reflected in similar spatial patterns across people Simony & Chang (2020). Following this  
205 logic, we used an across-participant temporal decoding test developed by Manning et al. (2018) to  
206 assess the degree to which different neural patterns reflected ongoing stimulus-driven cognitive  
207 processing across people. The approach entails using a subset of the data to train a classifier to  
208 decode stimulus timepoints (i.e., moments in the story participants listened to) from neural pat-  
209 terns. We use decoding (forward inference) accuracy on held-out data, from held-out participants,  
210 as a proxy for the extent to which the inputted neural patterns reflected stimulus-driven cognitive  
211 processing in a similar way across individuals.

212 **Forward inference and decoding accuracy**

213 We used an across-participant correlation-based classifier to decode which stimulus timepoint  
214 matched each timepoint's neural pattern. For a given value of  $k$  (i.e., number of principal com-  
215 ponents), we first used group PCA to project the data from each condition into a  $k$ -dimensional  
216 space. Next, we divided the participants into two groups: a template group,  $\mathcal{G}_{\text{template}}$  (i.e., training  
217 data), and a to-be-decoded group,  $\mathcal{G}_{\text{decode}}$  (i.e., test data). We averaged the projected data within  
218 each group to obtain a single  $T$  by  $k$  matrix for each group. Next, we correlated the rows of the two  
219 averaged matrices to form a  $T$  by  $T$  decoding matrix,  $\Lambda$ . In this way, the rows of  $\Lambda$  reflected time-  
220 points from the template group, while the columns reflected timepoints from the to-be-decoded  
221 group. We used  $\Lambda$  to assign temporal labels to each timepoint (row) from the test group's ma-  
222 trix, using the row of the training group's matrix with which it was most highly correlated. We  
223 repeated this decoding procedure, but using  $\mathcal{G}_{\text{decode}}$  as the template group and  $\mathcal{G}_{\text{template}}$  as the  
224 to-be-decoded group. Given the true timepoint labels (for each group), we defined the decoding  
225 accuracy as the average proportion of correctly decoded timepoints, across both groups (where  
226 chance performance is  $\frac{1}{T}$ ). In Figures 2 and 3 we report the decoding accuracy for each condition  
227 and value of  $k$ , averaged across  $n = 100$  cross validation folds.

228 **Reverse inference**

229 To help interpret the brain activity patterns we found within the contexts of other studies, we  
230 created summary maps of each principal component, for each experimental condition, by summing  
231 together the 20 HTFA-derived RBF nodes (see *Hierarchical Topographic Factor Analysis*) with the  
232 highest absolute value weights for each of the top 5 components (Figs. 5, S1). We then carried  
233 out a meta analysis using Neurosynth Rubin et al. (2017) to identify the 5 terms most commonly  
234 associated with the given map.

235 **Data and code availability**

236 All of the code used to produce the figures and results in this manuscript, along with links to the  
237 corresponding datasets, may be found at [github.com/ContextLab/pca\\_paper](https://github.com/ContextLab/pca_paper).

<sup>238</sup> **Acknowledgements**

<sup>239</sup> We acknowledge discussions with Rick Betzel, Emily Finn, and Jim Haxby. Our work was sup-  
<sup>240</sup> ported in part by NSF CAREER Award Number 2145172 to J.R.M. The content is solely the re-  
<sup>241</sup> sponsibility of the authors and does not necessarily represent the official views of our supporting  
<sup>242</sup> organizations. The funders had no role in study design, data collection and analysis, decision to  
<sup>243</sup> publish, or preparation of the manuscript.

<sup>244</sup> **Author contributions**

<sup>245</sup> Conceptualization: J.R.M. and L.L.W.O. Methodology: J.R.M. and L.L.W.O. Implementation: L.L.W.O.  
<sup>246</sup> Analysis: L.L.W.O. Writing, Reviewing, and Editing: J.R.M. and L.L.W.O. Supervision: J.R.M.

<sup>247</sup> **References**

- <sup>248</sup> Alvarez, S. A. (2002). *An exact analytical relation among recall, precision, and classification accuracy in*  
<sup>249</sup> *information retrieval* (Technical Report No. BCCS-02-01). Boston College.
- <sup>250</sup> Capota, M., Turek, J., Chen, P.-H., Zhu, X., Manning, J. R., Sundaram, N., ... Shin, Y. S. (2017).  
<sup>251</sup> *Brain imaging analysis kit*.
- <sup>252</sup> Cole, M. W., Bassett, D. S., Power, J. D., Braver, T. S., & Petersen, S. E. (2014). Intrinsic and  
<sup>253</sup> task-evoked network architectures of the human brain. *Neuron*, 83(1), 238–251.
- <sup>254</sup> Finn, E. S., Scheinost, D., Finn, D. M., Shen, X., Papademetris, X., & Constable, R. T. (2017).  
<sup>255</sup> Can brain state be manipulated to emphasize individual differences in functional connectivity.  
<sup>256</sup> *NeuroImage*, 160, 140–151.
- <sup>257</sup> Finn, E. S., Shen, X., Scheinost, D., Rosenberg, M. D., Huang, J., Chun, M. M., ... Constable, R. T.  
<sup>258</sup> (2015). Functional connectome fingerprinting: identifying individuals using patterns of brain  
<sup>259</sup> connectivity. *Nature Neuroscience*, 18, 1664–1671.

- 260 Glerean, E., Salmi, J., Lahnakoski, J. M., Jääskeläinen, I. P., & Sams, M. (2012). Functional magnetic  
261 resonance imaging phase synchronization as a measure of dynamic functional connectivity.  
262 *Brain Connectivity*, 2(2), 91–101.
- 263 Gratton, C., Laumann, T. O., Nielsen, A. N., Greene, D. J., Gordon, E. M., Gilmore, A. W., ...  
264 Petersen, S. E. (2018). Functional brain networks are dominated by stable group and individual  
265 factors, not cognitive or daily variation. *Neuron*, 98(2), 439–452.
- 266 Hasson, U., Yang, E., Vallines, I., Heeger, D. J., & Rubin, N. (2008). A hierarchy of temporal  
267 receptive windows in human cortex. *The Journal of Neuroscience*, 28(10), 2539–2550.
- 268 Kumar, M., Anderson, M. J., Antony, J. W., Baldassano, C., Brooks, P. P., Cai, M. B., ... Norman,  
269 K. A. (2021). BrainIAK: the brain image analysis kit. *Aperture*, In press.
- 270 Lerner, Y., Honey, C. J., Silbert, L. J., & Hasson, U. (2011). Topographic mapping of a hierarchy of  
271 temporal receptive windows using a narrated story. *The Journal of Neuroscience*, 31(8), 2906–2915.
- 272 Mack, M. L., Preston, A. R., & Love, B. C. (2020). Ventromedial prefrontal cortex compressesion  
273 during concept learning. *Nature Communications*, 11(46), doi.org/10.1038/s41467-019-13930-8.
- 274 Manning, J. R., Zhu, X., Willke, T. L., Ranganath, R., Stachenfeld, K., Hasson, U., ... Norman,  
275 K. A. (2018). A probabilistic approach to discovering dynamic full-brain functional connectivity  
276 patterns. *NeuroImage*, 180, 243–252.
- 277 Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: multi-voxel  
278 pattern analysis of fMRI data. *Trends in Cognitive Sciences*, 10(9), 424–430.
- 279 Owen, L. L. W., Chang, T. H., & Manning, J. R. (2021). High-level cognition during story listening is  
280 reflected in high-order dynamic correlations in neural activity patterns. *Nature Communications*,  
281 12(5728), doi.org/10.1038/s41467-021-25876-x.
- 282 Owen, L. L. W., Muntianu, T. A., Heusser, A. C., Daly, P., Scangos, K. W., & Manning, J. R. (2020). A  
283 Gaussian process model of human electrocorticographic data. *Cerebral Cortex*, 30(10), 5333–5345.

- 284 Rubin, T. N., Kyojo, O., Gorgolewski, K. J., Jones, M. N., Poldrack, R. A., & Yarkoni, T. (2017).  
285 Decoding brain activity using a large-scale probabilistic functional-anatomical atlas of human  
286 cognition. *PLoS Computational Biology*, 13(10), e1005649.
- 287 Scangos, K. W., Khambhati, A. N., Daly, P. M., Owen, L. L. W., Manning, J. R., Ambrose, J. B.,  
288 ... Chang, E. F. (2021). Biomarkers of depression symptoms defined by direct intracranial  
289 neurophysiology. *Frontiers in Human Neuroscience*, In press.
- 290 Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*,  
291 27(3), 379–423.
- 292 Simony, E., & Chang, C. (2020). Analysis of stimulus-induced brain dynamics during naturalistic  
293 paradigms. *NeuroImage*, 216, 116461.
- 294 Simony, E., Honey, C. J., Chen, J., & Hasson, U. (2016). Dynamic reconfiguration of the default  
295 mode network during narrative comprehension. *Nature Communications*, 7(12141), 1–13.
- 296 Smith, S. M., Hyvärinen, A., Varoquaux, G., Miller, K. L., & Beckmann, C. F. (2014). Group-PCA  
297 for very large fMRI datasets. *NeuroImage*, 101, 738–749.
- 298 Yeo, B. T. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., ... Buckner,  
299 R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional  
300 connectivity. *Journal of Neurophysiology*, 106(3), 1125–1165.