

1 **High-level cognition is supported by information-rich but**
2 **compressible brain activity patterns**

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

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

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

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

9 December 19, 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 maximum
16 achievable decoding accuracy across participants as an indicator of the “informativeness” of
17 the recorded patterns. Second, we treated the number of features (components) required to
18 achieve a threshold decoding accuracy as a proxy for the “compressibility” of the neural pat-
19 terns (where fewer components indicate greater compression). Overall, we found that the peak
20 decoding accuracy (achievable without restricting the numbers of features) was highest in the
21 intact (unscrambled) story listening condition. However, the number of features required to
22 achieve comparable classification accuracy was also lowest in the intact story listening con-
23 dition. Taken together, our work suggests that our brain networks flexibly reconfigure according
24 to ongoing task demands, and that the activity patterns associated with higher-order cogni-
25 tion and high engagement are both more informative and more compressible than the activity
26 patterns associated 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 from
32 every other node. At the other extreme, all nodes behave identically. Each extreme optimizes
33 key 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, then
30 every state is equally reachable by a network with fully independent nodes. On the other hand, a
31 network of identically behaved nodes optimizes *robustness*: any subset of nodes may be removed
32 from the network without any loss of function or expressive power, as long as any single node
33 remains. Presumably, most natural systems tend to occupy positions between these extremes. We
34 wondered: might the human brain reconfigure itself to be more flexible or more robust according
35 to ongoing demands? In other words, might the brain reconfigure its connections or behaviors
36 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 between
44 how expressive the signal *could* be (i.e., proportional to the number of unique states or symbols
45 used to transmit the signal) and the actual information rate (i.e., the entropy of each individual
46 state or symbol). If a brain network’s nodes are fully independent, then the number of bits required
47 to express a single activity pattern is proportional to the number of nodes. The network would
48 also be minimally redundant, since the status of every node would be needed to fully express a
49 single brain activity pattern. If a brain network’s nodes are fully coupled and identical, then the
50 number of bits required to express a single activity pattern is proportional to the number of unique
51 states or values any individual node can take on. Such a network would be highly redundant,
52 since knowing any individual node’s state would be sufficient to recover the full-brain activity
53 pattern. Highly redundant systems are also robust, since there is little total information loss due
54 to removing any given observation.

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

59 across (Cole et al., 2014; Glerean et al., 2012; Gratton et al., 2018; Yeo et al., 2011) individuals. This
60 stability suggests that our brains' activity patterns are at least partially constrained, for example
61 by anatomical, external, or internal factors. Constraints on brain activity that limit its flexibility
62 decrease expressiveness (i.e., its information rate). However, constraints on brain activity also
63 increase its robustness to noise (e.g., "missing" or corrupted signals may be partially recovered).
64 For example, recent work has shown that full-brain activity patterns may be reliably recovered
65 from only a relatively small number of implanted electrodes (Owen et al., 2020; Scangos et al.,
66 2021). This robustness property suggests that the relevant signal (e.g., underlying factors that
67 have some influence over brain activity patterns) are compressible.

68 To the extent that brain activity patterns contain rich task-relevant information, we should be
69 able to use the activity patterns to accurately differentiate between different aspects of a task (e.g.,
70 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
75 are often related but are also dissociable (Fig. 1). If a given signal (e.g., a representation of brain
76 activity patterns) contains more information about ongoing cognitive processes, then the peak
77 decoding accuracy should be high. In the simulations shown in Figure 1C we construct synthetic
78 datasets that have high or low levels of informativeness by varying temporal autocorrelations
79 in the data (see *Synthetic data*). If a signal is compressible, then a low-dimensional embedding
80 of the signal will be similarly informative as the original signal. In the simulations shown in
81 Figure 1C we construct synthetic datasets that have high or low levels of compressibility by
82 varying the covariance structure across features (see *Synthetic data*). As shown in Figure 1D, highly
83 informative datasets yield higher decoding accuracies than less informative datasets (i.e., the peaks
84 of the curves in the top panels of Fig. 1D are higher than the peaks of the curves in the bottom
85 panels). Highly compressible datasets show steeper slopes when we plot decoding accuracy as a
86 function of the number of components used to represent the data (i.e., the slopes of the curves in
87 the left panels of Fig. 1D are steeper than the slopes of the curves in the right panels). Whereas
88 characterizing the informativeness and compressibility of synthetic data can be instructive, we

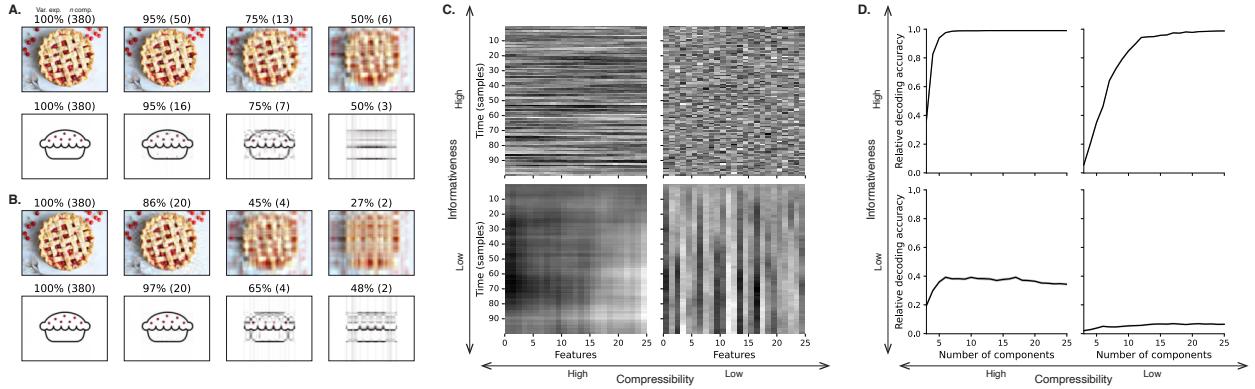


Figure 1: Information content and compressibility. **A. Variance explained for two images.** Consider two example images: a photograph and simple drawing. The images have the same resolutions, but their content is very different. The photograph is more information rich: when we compare the original photograph and drawing (leftmost column), we can see that the photograph captures much more detail. We can also apply principal components analysis to the images, treating the rows of the images as “observations.” Across columns, we identified the numbers 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. This analysis reveals that the drawing is more compressible: just 16 components can explain 95% of the variance in the drawing, whereas 50 components are required to explain 95% of the variance in the photograph. **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. This analysis provides another way of characterizing the compressibility of each image by showing that the same number of components can explain more variance in the drawing than in the photograph. **C. Template data from four synthetic datasets.** We constructed four synthetic datasets, each comprising 25 features (columns) observed across 100 samples (rows) from each of 10 simulated “participants.” The datasets were constructed to contain different levels of informativeness and compressibility (see *Synthetic data*). **D. Decoding accuracy by number of components.** For each synthetic dataset, we trained across-participant classifiers to decode timepoint labels. Each panel displays the decoding accuracy as a function of the number of components used to represent the data. Error ribbons denote bootstrap-estimated 95% confidence intervals.

89 are ultimately interested in understanding how these properties relate to brain activity patterns
90 recorded under different cognitive circumstances.

91 Several recent studies suggest that the complexity of brain activity is task-dependent, whereby
92 simpler tasks with lower cognitive demands are reflected by simpler and more compressible brain
93 activity patterns, and more complex tasks with higher cognitive demands are reflected by more
94 complex and less compressible brain activity patterns (Mack et al., 2020; Owen et al., 2021). These
95 findings complement other work suggesting that functional connectivity (correlation) patterns are
96 task-dependent (Cole et al., 2014; Finn et al., 2017; Owen et al., 2020), although see Gratton et
97 al. (2018). Higher-order cognitive processing of a common stimulus also appears to drive more
98 stereotyped task-related activity and functional connectivity across individuals (Hasson et al.,
99 2008; Lerner et al., 2011; Simony & Chang, 2020; Simony et al., 2016).

100 The above studies are consistent with two potential descriptions of how cognitive processes are
101 reflected in brain activity patterns. One possibility is that the information rate of brain activity in-
102 creases during more complex or higher-level cognitive processing. If so, then the ability to reliably
103 decode cognitive states from brain activity patterns should improve with task complexity or with
104 the level (or “depth”) of cognitive processing. A second possibility is that the compressibility of
105 brain activity patterns increases during more complex or higher-level cognitive processing. If so,
106 then individual features of brain recordings, or compressed representations of brain recordings,
107 should carry more information during complex or high-level (versus simple or low-level) cognitive
108 tasks.

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

119 examined how classification accuracy changed across the different experimental conditions.

120 **Results**

121 We sought to understand whether higher-level cognition is reflected by more reliable and in-
122 formative brain activity patterns, and how compressibility of brain activity patterns relates to
123 cognitive complexity. We developed a computational framework for systematically assessing the
124 informativeness and compressibility of brain activity patterns recorded under different cognitive
125 circumstances. We used across-participant decoding accuracy (see *Forward inference and decoding*
126 *accuracy*) as a proxy for informativeness. To estimate the compressibility of the brain patterns, we
127 used group principal components analysis (PCA) to project the brain patterns into k -dimensional
128 spaces, for different values of k (see *Hierarchical Topographic Factor Analysis (HTFA)* and *Principal*
129 *components analysis (PCA)*). For more compressible brain patterns, decoding accuracy should be
130 more robust to small values of k .

131 We analyzed a dataset collected by Simony et al. (2016) that comprised four experimental con-
132 ditions. These conditions exposed participants to stimuli that systematically varied in cognitive
133 engagement. In the *intact* experimental condition, participants listened to an audio recording of a
134 10-minute Moth Radio Hour story, *Pie Man*, by Jim O’Grady. In the *paragraph*-scrambled experi-
135 mental condition, participants listened to a temporally scrambled version of the story, where the
136 paragraphs occurred out of order, but where the same set of paragraphs was presented over the
137 entire listening interval. All participants in this condition experienced the scrambled paragraphs
138 in the same order. In the *word*-scrambled experimental condition, participants listened to a tem-
139 porally scrambled version of the story, where the words occurred in a random order. Again, all
140 participants in this condition experienced the scrambled words in the same order. Finally, in the
141 *rest* experimental condition, participants lay in the scanner with no overt stimulus, while keeping
142 their eyes open and blinking as needed. This public dataset provided a convenient means for
143 testing our hypothesis that different levels of cognitive processing and engagement affect how
144 informative and compressible the associated brain patterns are.

145 To evaluate the relation between informativeness and compressibility for brain activity from
146 each experimental condition, we trained a series of across-participant temporal decoders on com-

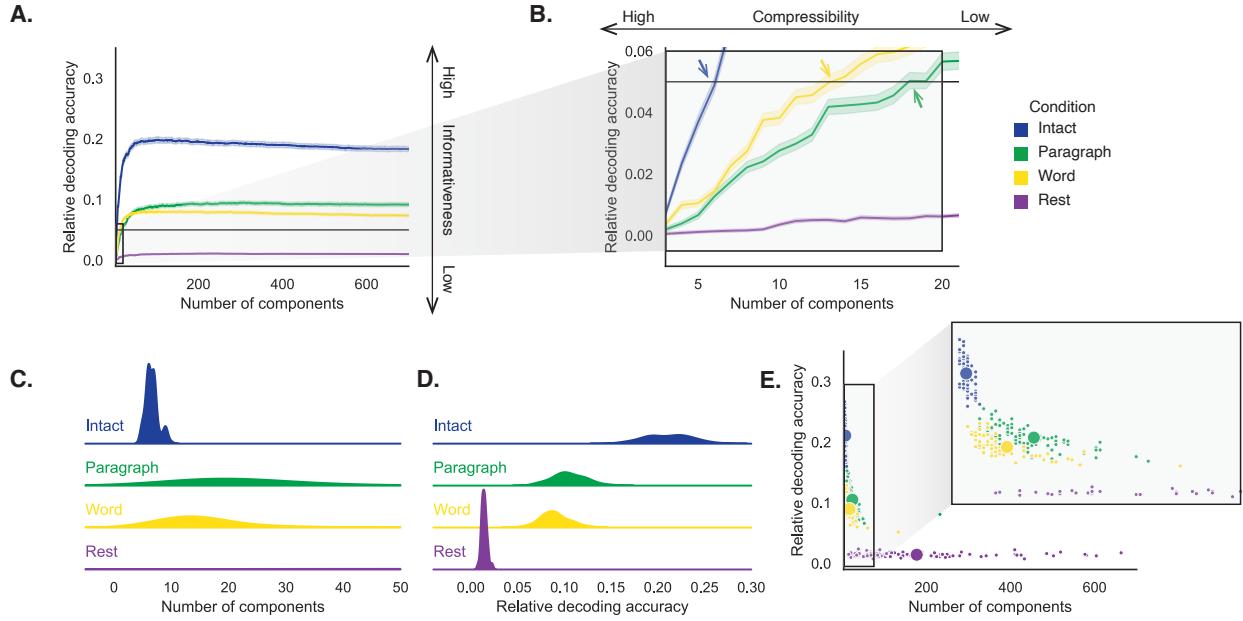


Figure 2: Decoding accuracy and compression. A. Decoding accuracy by number of components. Ribbons of each color display cross-validated decoding performance for each condition (intact, paragraph, word, and rest), as a function of the number of components (features) used to represent the data. For each condition, decoding accuracy has been normalized by subtracting expected “chance” decoding performance (estimated as $\frac{1}{T}$, where T is the number of timepoints in the given condition). This normalization was intended to enable fairer comparisons across conditions; a relative decoding accuracy of 0.0 denotes “chance” performance. 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. All error ribbons in Panels A and B denote bootstrap-estimated 95% confidence intervals. **C. Estimating “compressibility” for each condition.** The density plots display the numbers of components required to reach at least 5% (corrected) decoding accuracy, across all cross validation runs. For the “rest” condition, where decoding accuracy never reached 5% accuracy, we display the distribution of the numbers of components required to achieve the peak decoding accuracy in each fold. (This distribution is nearly flat, as also illustrated in Panel E.) **D. Estimating “informativeness” for each condition.** The density plots display the peak (corrected) decoding accuracies achieved in each condition, across all cross validation runs. **E. Informativeness versus compressibility.** Each dot displays the minimum number of components required to achieve at least 5% corrected decoding accuracy (x -coordinate; for the rest condition the numbers of components are estimated using the peak decoding accuracies as in Panel C) and the peak (corrected) decoding accuracy (y -coordinate). The smaller dots correspond to individual cross validation runs and the larger dots display the averages across runs. The inset displays a zoomed-in view of the indicated region in the main panel.

147 pressed representations of the data. Figure 2A displays the decoding accuracy as a function
148 of the number of principal components used to represent the data (also see Fig. S1). Several
149 patterns were revealed by the analysis. First, in general (i.e., across experimental conditions),
150 decoding accuracy tends to improve as the number of components are increased. However, de-
151 coding accuracy peaked at higher levels for experimental conditions that exposed participants
152 to cognitively richer stimuli (Fig. 2D). The peak decoding accuracy was highest for the “intact”
153 condition (versus paragraph: $t(99) = 35.205, p < 0.001$; versus word: $t(99) = 43.172, p < 0.001$);
154 versus rest: $t(99) = 81.361, p < 0.001$), next highest for the “paragraph” condition (versus word:
155 $t(99) = 6.243, p < 0.001$; versus rest: $t(99) = 50.748, p < 0.001$), and next highest for the “word”
156 condition (versus rest: $t(99) = 48.791, p < 0.001$). This ordering implies that cognitively richer
157 conditions evoke more stable brain activity patterns across people.

158 The cognitively richer conditions also displayed steeper initial slopes. For example, the intact
159 condition decoders reached an arbitrarily chosen threshold of 5% accuracy using fewer components
160 than the paragraph condition decoders ($t(99) = -7.429, p < 0.001$) or word condition decoders
161 ($t(99) = -7.300, p < 0.001$), and decoding accuracy never exceeded 5% for the rest condition. This
162 suggests that brain activity patterns evoked by cognitively richer conditions are more compressible,
163 such that representing the data using the same number of principal components provides more
164 information to the temporal decoders (Figs. 2B, C). Taken together, as shown in Figure 2E, we
165 found that brain activity patterns evoked by cognitively richer conditions tended to be both more
166 informative (i.e., associated with higher peak decoding accuracies) *and* more compressible (i.e.,
167 requiring fewer components to achieve the 5% accuracy threshold).

168 If informativeness (to the temporal decoders) and compressibility vary with the cognitive
169 richness of the stimulus, might these measures also vary over time *within* a given condition? For
170 example, participants in the intact condition might process the ongoing story more deeply later
171 on in the story (compared with earlier in the story) given the additional narrative background
172 and context they had been exposed to by that point. To examine this possibility, we divided each
173 condition into four successive time segments. We computed decoding curves (Fig. 3A) and the
174 numbers of components required to achieve 5% decoding accuracy (Fig. 3B) for each segment and
175 condition. We found that, in the two most cognitively rich conditions (intact and paragraph), both
176 decoding accuracy and compressibility, as reflected by the change in decoding curves, increased

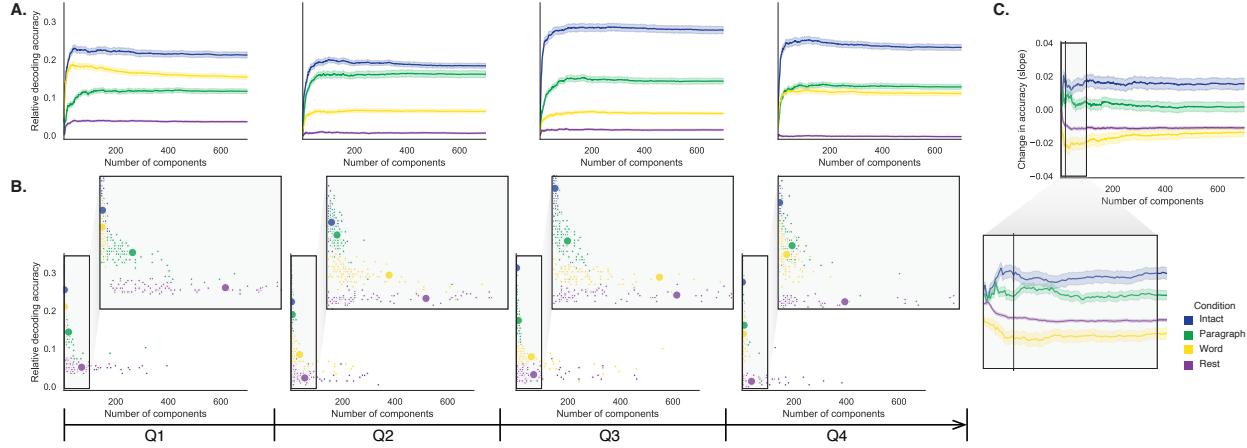


Figure 3: Changes in decoding accuracy and compression over time. **A. Decoding accuracy by number of components, by story segment.** Each family of curves is plotted in the same format as Figure 2A but reflects data only from one quarter of the dataset. **B. Informativeness versus compressibility by condition and segment.** Each scatter plot is in the same format as Figure 2E, but reflects data only from one quarter of the dataset. **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, third, and fourth quarters of the dataset (corresponding to the columns of Panels A and B). The y -axis denotes the slopes of the regression lines. The black vertical line marks $k = 20$ components, as referenced in the main text. All error ribbons denote bootstrap-estimated 95% confidence intervals.

177 with listening time (e.g., at the annotated reference point of $k = 20$ components in Fig. 3C: intact:
 178 $t(99) = 7.915, p < 0.001$; paragraph: $t(99) = 2.354, p = 0.021$). These changes may reflect an increase
 179 in comprehension or depth of processing with listening time. In contrast, the decoding accuracy
 180 and compressibility *decreased* with listening time in the word condition ($t(99) = -10.747, p < 0.001$)
 181 and rest condition ($t(99) = -22.081, p < 0.001$). This might reflect the depletion of attentional
 182 resources in the less-engaging word and rest conditions.

183 We also wondered how informativeness and compressibility in the different experimental
 184 conditions might vary across brain networks. We used a network parcellation identified by Yeo et
 185 al. (2011) to segment the brain into seven distinct networks. The networks can be sorted (roughly)
 186 in order from lower-level to higher-level cortex as follows (Figs. 4A–C): visual, somatomotor,
 187 dorsal attention, ventral attention, limbic, frontoparietal, and default mode. Next, we computed
 188 decoding curves separately for the activity patterns within each network and identified each
 189 network's inflection point, for each experimental condition. Moving from low-order networks
 190 to higher-order networks, we found that decoding accuracy tended to increase, particularly in

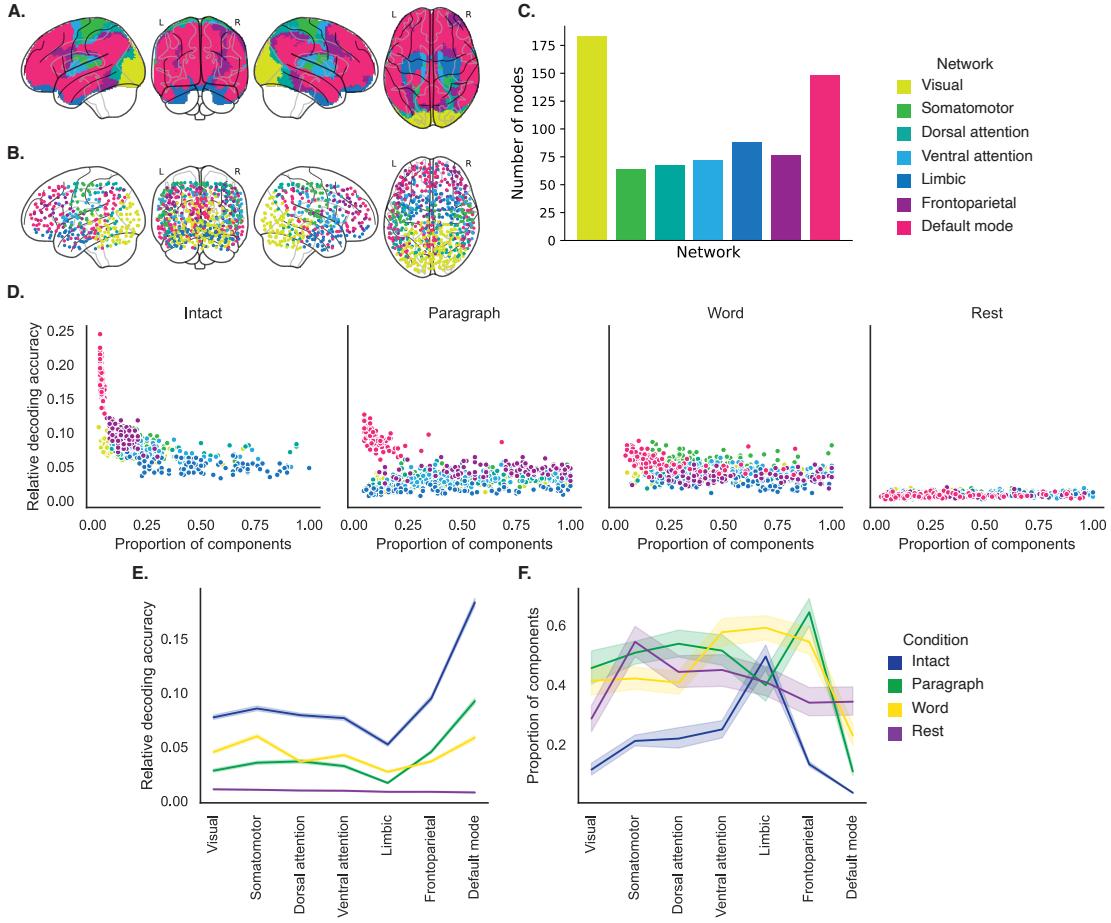


Figure 4: Network-specific decoding accuracy and compression. **A. Network parcellation map.** The glass brain displays the location of each of seven networks identified by Yeo et al. (2011). **B. Node labels.** We assigned network labels to each of 700 Hierarchical Topographic Factor Analysis-derived nodes (see *Hierarchical topographic factor analysis (HTFA)*). For each node, we evaluated its radial basis function (RBF) at the locations of every voxel in the parcellation map displayed in Panel A. We summed up these RBF weights separately for the voxels in each network. We defined each node's network label as its highest-weighted network across all voxels. **C. Per-network node counts.** The bar plot displays the total number of HTFA-derived nodes assigned to each network. **D. Informativeness versus compressibility by network and condition.** Each dot corresponds to a single cross validation run. The dots' coordinates denote the minimum proportion of components (relative to the number of nodes in the given network) required to achieve at least 5% corrected decoding accuracy (x -coordinate; for the rest condition these proportions are estimated using the peak decoding accuracies) and peak (corrected) decoding accuracy (y -coordinate). We repeated the analysis separately for each network (color) and experimental condition (column). **E. Informativeness by network.** For each network, sorted (roughly) from lower-order networks on the left to higher-order networks on the right, the curves display the peak (corrected) decoding accuracy for each experimental condition (color). **F. Compressibility by network.** For each network, the curves display the proportion of components (relative to the total possible number of components displayed in Panel C) required to achieve at least 5% decoding accuracy (or, for the rest condition, peak decoding accuracy, as described above). Error ribbons in Panels E and F denote bootstrap-estimated 95% confidence intervals.

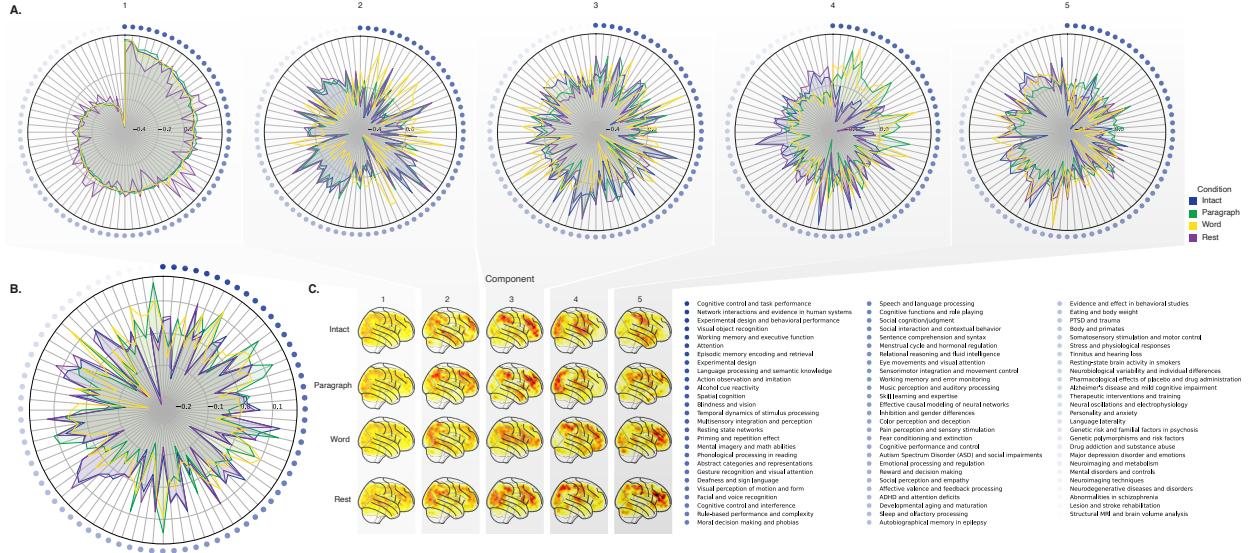


Figure 5: Neurosynth topic weightings by component. We used a reverse inference procedure (see *Reverse inference*) to estimate the correspondences between brain images and a set of 80 topics derived from the Neurosynth database of neuroimaging articles (Rubin et al., 2017). **A. Topic correlations by component.** For each of the top five highest-weighted principal components (columns) derived from each experimental condition (colors), the radar plots display the correlations between the component images and the per-topic images derived for each topic (topic identities are denoted by the blue dots; a legend for the topic labels is in the lower right of the figure). An annotated list of the top-weighted topics for each component and condition may be found in Figure S2 and the top-weighted terms for each topic may be found in Table S1. **B. Topic correlations averaged across components.** The radar plot is in the same format as the plots in Panel A, but here we display the per-condition average correlations across the top five components (for each condition) reflected in Panel A. **C. Component images.** Each plot displays a right sagittal view of a glass brain projection of the top five principal components (columns) for each experimental condition (rows). Additional projections for each component may be found in Figure S3.

191 the higher-level experimental conditions (Fig. 4D, E). This suggests that higher-order networks
 192 may carry more content-relevant or stimulus-driven “information.” We found no clear trends
 193 in the proportions of components required to achieve 5% decoding accuracy across networks or
 194 conditions (Fig. 4F).

195 In addition to examining different networks in isolation, we wondered about the general
 196 structure of the full-brain (i.e., potentially multi-network) activity patterns reflected by different
 197 principal components across different experimental conditions. As shown in Figure 5, we used
 198 Neurosynth (Rubin et al., 2017) to identify, for each component, the associations with each of
 199 80 themes (see *Reverse inference*). In general, the first principal components across all of the
 200 experimental conditions tended to weight most heavily on themes related to cognitive control,

201 memory, language processing, attention, and perception. Other components appeared to vary
202 more across conditions.

203 To gain further insights into which brain functions might be most closely associated with
204 the top-weighted components from each experimental condition, we manually grouped each
205 Neurosynth-derived topic into a smaller set of core cognitive functions. Separately for each com-
206 ponent, we computed the average weightings across all topics that were tagged as being associated
207 with each of these cognitive functions (Figs. 6A, S5A). To help visualize these associations, we used
208 the patterns of associations for each component to construct graphs whose nodes were experimen-
209 tal conditions and cognitive functions (Figs. 6B, S5B). We also computed correlations between the
210 sets of per-topic weightings from each of the top-weighted components from each experimental
211 condition (Fig. 6C, D) and between the brain maps for each condition's components (Fig. S5C, D).
212 Taken together, we found that each component appeared to weight on a fundamental set of cogni-
213 tive functions that varied by experimental condition. For example, the top principal components
214 from every condition weighted similarly (across conditions) on the full set of Neurosynth topics
215 (Fig. 5A) and cognitive functions (Figs. 6A, B and S5 A, B), suggesting that these components
216 might reflect a set of functions or activity patterns that are common across all conditions. The
217 second components' weightings were similar across the intact, paragraph, and rest conditions
218 (highest-weighted functions: cognitive control, memory, social cognition, and resting state), but
219 different for the word condition (highest-weighted functions: sensory perception and cognitive
220 control). The fourth components' weighting grouped the paragraph and word conditions (highest-
221 weighted functions: memory, language processing, and cognitive control) and the intact and rest
222 conditions (highest-weighted functions: emotion, social cognition). We also used ChatGPT (Ope-
223 nAI, 2023) to sort the list of manually tagged cognitive functions from lowest-level to highest-level
224 (Tab. S2, Fig. 6E; also see *Ranking cognitive processes*). We found that higher-level functions tended
225 to be weighted on more heavily by top components from the intact and paragraph conditions
226 than lower-level functions. The top components from the word condition showed the opposite
227 tendency, whereby *lower*-level functions tended to be weighted on more heavily than higher-level
228 functions. The components from the rest condition showed almost no differences in the weights
229 associated with high-level versus low-level functions.

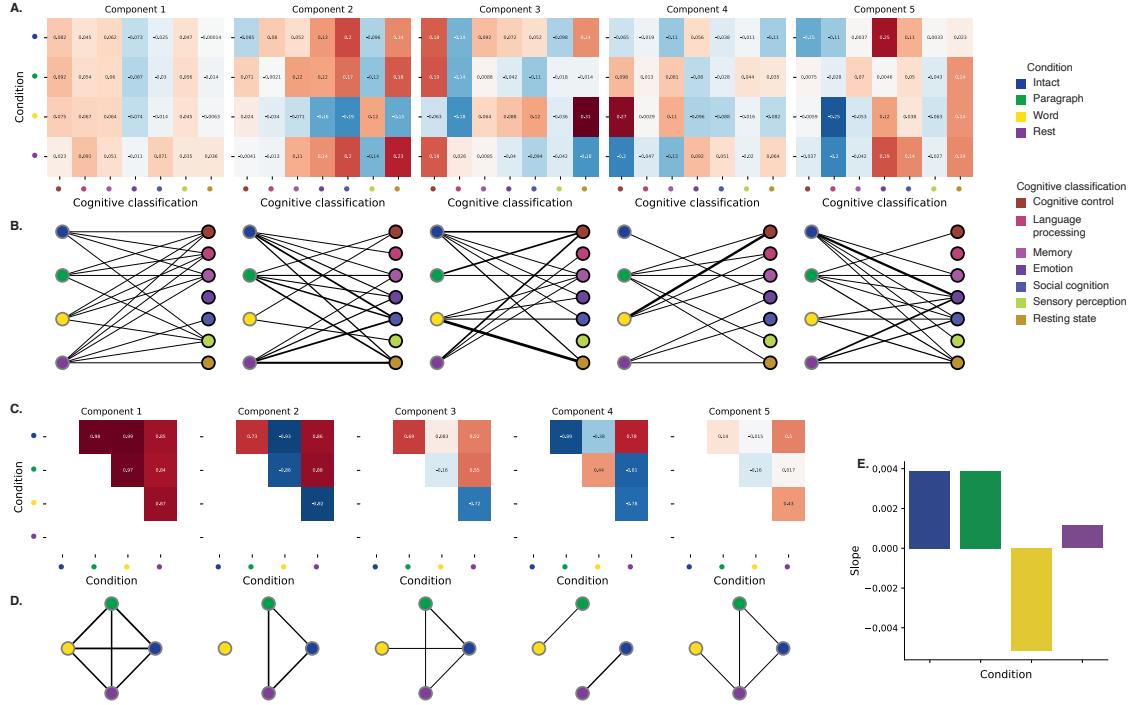


Figure 6: Summary of functions associated with top-weighted components by condition. A. Top-weighted topics by condition. Here we display per-condition (rows, indicated by colored dots) topic correlations, averaged across topics that pertain to each of several broad cognitive functions (columns within each sub-panel, indicated by colored dots). Each sub-panel reflects correlations for the components indicated in the panel titles. A legend for the condition and several core cognitive function classifications is displayed in the lower right of the figure. Table S1 provides a list of each topic's top-weighted terms, along with each topic's manually labeled cognitive classification. A full list of the topics most highly associated with each component may be found in Figure S2. **B. Associations between per-condition components and cognitive functions.** The network plots denote positive average correlations between the component images for each condition (gray-outlined dots on the left sides of each network; colors denote conditions) and topic-specific brain maps associated with each indicated cognitive function (black-outlined dots on the right sides of each network; colors denote cognitive functions). The line thicknesses are proportional to the correlation values (correlation coefficients are noted in the heat maps in Panel A). **C. Correlations between each principal component, by condition.** The heat maps display the correlations between the brain maps (Fig. S3) for each principal component (sub-panel), across each pair of conditions (rows and columns of each sub-panel's matrix, indicated by colored dots). **D. Associations between per-condition components, by component.** Each sub-panel's network plot summarizes the pattern of correlations between the n^{th} top-weighted principal components (sub-panel) for each experimental condition (gray-outlined dots). The line thicknesses are proportional to the correlation values (correlation coefficients are noted in the heat maps in Panel C). **E. Change in weights by condition.** The bar heights reflect the slopes of regression lines, fit separately to the top five components from each condition, between the ChatGPT-derived "rank" of each cognitive classification and the correlations between the component and topic maps associated with cognitive processes at the given rank (see *Ranking cognitive processes*). Also see Fig. S5 for additional information.

230 **Discussion**

231 We examined fMRI data collected as participants listened to an auditory recording of a story, scram-
232 bled recordings of the story, or underwent a resting state scan. We found that cognitively richer
233 stimuli evoked more reliable (i.e., consistent across people) and information rich brain activity pat-
234 terns. The brain patterns evoked by cognitively richer stimuli were also more compressible, in that
235 each individual component provided more “signal” to temporal decoders relative to components
236 of data from less cognitively rich conditions (Fig. 2). Over time (e.g., as the experiment progressed),
237 these phenomena were strengthened. Specifically, across story segments, data from more cogni-
238 tively rich conditions became more informative and compressible, and data from less cognitively
239 rich conditions became *less* informative and compressible (Fig. 3). We also repeated these analyses
240 separately for different brain networks. We found that networks traditionally associated with
241 higher-level cognitive functions tended to provide more informative brain patterns than networks
242 traditionally associated with lower level cognitive functions (Fig. 4). Finally, we examined the most
243 dominant components of the brain activity patterns from each experimental condition. We used a
244 reverse inference approach (Rubin et al., 2017) to identify the terms in the neuroimaging literature
245 most commonly associated with the corresponding maps. As summarized in Figure 6E, we found
246 that the intact and paragraph conditions tended to weight on higher-level cognitive processes
247 more than lower-level cognitive processes, whereas the word condition weighted on lower-level
248 processes more than higher-level processes and the rest condition showed no difference in high-
249 level versus low-level weighting. Taken together, our findings indicate that the informativeness
250 and compressibility of our brain activity patterns are task-dependent, and these properties change
251 systematically with factors like cognitive richness and depth of processing.

252 Our explorations of informativeness and compressibility are related to a much broader litera-
253 ture on the correlational and causal structure of brain activity patterns and networks (Adachi et
254 al., 2012; Bassett & Sporns, 2017; Brovelli et al., 2004; Bullmore & Sporns, 2009; Dhamala et al.,
255 2008; Korzeniewska et al., 2008; Lynn & Bassett, 2021; Owen et al., 2021; Preti et al., 2017; Rogers
256 et al., 2007; Rubinov & Sporns, 2010; Sizemore et al., 2018; Smith, Beckmann, et al., 2013; Smith,
257 Vidaurre, et al., 2013; Sporns & Betzel, 2016; Sporns & Honey, 2006; Sporns & Zwi, 2004; Srinivasan
258 et al., 2007; Tomasi & Volkow, 2011; Yeo et al., 2011). Correlations or causal associations between

259 different brain regions simultaneously imply that full-brain activity patterns will be compressible
260 and also that those activity patterns will contain redundancies. For example, the extent to which
261 activity patterns at one brain area can be inferred or predicted from activity patterns at other
262 areas (e.g., Owen et al., 2020; Scangos et al., 2021), reflects overlap in the information available
263 in or represented by those brain areas. If brain patterns in one area are recoverable using brain
264 patterns in another area, then a “signal” used to convey the activity patterns could be compressed
265 by removing the recoverable activity. Predictable (and therefore redundant) brain activity patterns
266 are also more robust to signal corruption. For example, even if the activity patterns at one region
267 are unreadable or unreliable at a given moment, that unreliability could be compensated for by
268 other regions’ activity patterns that were predictive of the unreliable region.

269 Our findings that informativeness and compressibility change with task demands may follow
270 from task-dependent changes in full-brain correlation patterns. A number of prior studies have
271 found that so-called “functional connectivity” (i.e., correlational) patterns vary reliably across
272 tasks, events, and situations (Cole et al., 2014; Owen et al., 2021; Simony et al., 2016; Smith et
273 al., 2009). By examining how these task-dependent changes in correlations affect informativeness
274 and compressibility, our work suggests a potential reason why the statistical structure of brain
275 activity patterns might vary with cognitive task or with cognitive demands. For lower-level tasks,
276 or for tasks that require relatively little “deep” cognitive processing, our brains may optimize
277 activity patterns for robustness and redundancy over expressiveness, for example to maximize
278 reliability. For higher-level tasks, or for tasks that require deeper cognitive processing, our brains
279 may sacrifice some redundancy in favor of greater expressiveness.

280 In the information theory sense (Shannon, 1948), when a signal is transmitted using a fixed
281 alphabet of “symbols,” the information rate decreases as the signal is compressed (e.g., fewer sym-
282 bols transmitted per unit time, using an alphabet with fewer symbols, etc.). Our finding that each
283 individual brain component (symbol) becomes more informative as cognitive richness increases
284 suggests that the “alphabet” of brain activity patterns is also task-dependent. Other work suggests
285 that the representations that are *reflected* by brain activity patterns may also change with task de-
286 mands. For example, our brains may represent the same perceptual stimulus differently depending
287 on which aspects of the stimulus or which combinations of features are task-relevant (Mack et al.,
288 2020).

289 Different brain networks also varied in how informative and compressible their activity pat-
290 terns were across experimental conditions (e.g., Fig. 4). This might follow from evolutionary
291 optimizations that reflect the relevant constraints or demands placed on those networks. One
292 possibility is that cortex is organized in a hierarchy of networks “concerned with” or selective to
293 different levels of processing or function. To the extent that different levels of processing (e.g.,
294 low-level sensory processing versus “deeper” higher-level processing) reflect different stimulus
295 timescales (e.g., Manning, 2020), the network differences we observed might also relate to the
296 timescales at which each network is maximally sensitive (Baldassano et al., 2017; Hasson et al.,
297 2008; Lerner et al., 2011; Regev et al., 2018).

298 **Concluding remarks**

299 Cognitive neuroscientists are still grappling with basic questions about the fundamental “rules”
300 describing how our brains respond, and about how brain activity patterns and the associated
301 underlying cognitive representations and computations are linked. We identified two aspects of
302 brain activity patterns, informativeness and compressibility, that appear to change systematically
303 with task demands and across brain networks. Our work helps to clarify how the “neural code”
304 might be structured, and how the code might vary across tasks and brain areas.

305 **Methods**

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

309 Unless otherwise noted, all statistical tests are two-sided, all error bars and error ribbons
310 denote bootstrap-estimated 95% confidence intervals across participants, and all reported *p*-values
311 are corrected for multiple comparisons using the Benjamini-Hochberg procedure (Benjamini &
312 Hochberg, 1995).

313 **Functional neuroimaging data collected during story listening**

314 We examined an fMRI dataset collected by Simony et al. (2016) that the authors have made publicly
315 available at arks.princeton.edu/ark:/88435/dsp015d86p269k. The dataset comprises neuroimaging
316 data collected as participants listened to an audio recording of a story (intact condition; 36 par-
317 ticipants), listened to temporally scrambled recordings of the same story (17 participants in the
318 paragraph-scrambled condition listened to the paragraphs in a randomized order and 36 in the
319 word-scrambled condition listened to the words in a randomized order), or lay resting with their
320 eyes open in the scanner (rest condition; 36 participants). Full neuroimaging details may be found
321 in the original paper for which the data were collected (Simony et al., 2016). Procedures were
322 approved by the Princeton University Committee on Activities Involving Human Subjects, and by
323 the Western Institutional Review Board (Puyallup, WA). All subjects were native English speakers
324 with normal hearing and provided written informed consent.

325 **Hierarchical topographic factor analysis (HTFA)**

326 Following our prior related work, we used HTFA (Manning et al., 2018) to derive a compact
327 representation of the neuroimaging data. In brief, this approach approximates the timeseries
328 of voxel activations (44,415 voxels) using a much smaller number of radial basis function (RBF)
329 nodes (in this case, 700 nodes, as determined by an optimization procedure; Manning et al., 2018).
330 This provides a convenient representation for examining full-brain activity patterns and network
331 dynamics. All of the analyses we carried out on the neuroimaging dataset were performed in
332 this lower-dimensional space. In other words, each participant's data matrix was a number-of-
333 timepoints (T) by 700 matrix of HTFA-derived factor weights (where the row and column labels
334 were matched across participants). Code for carrying out HTFA on fMRI data may be found as
335 part of the BrainIAK toolbox (Capota et al., 2017; Kumar et al., 2021), which may be downloaded
336 at brainiak.org.

337 **Principal components analysis (PCA)**

338 We applied group PCA (Smith et al., 2014) separately to the HTFA-derived representations of the
339 data (i.e., factor loadings) from each experimental condition. Specifically, for each condition, we

340 considered the set of all participants' T by 700 factor weight matrices. We used group PCA to project
341 these 700-dimensional matrices into a series of shared k -dimensional spaces, for $k \in \{3, 4, 5, \dots, 700\}$.
342 This yielded a set of number-of-participants matrices, each with T rows and k columns.

343 **Temporal decoding**

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

357 **Forward inference and decoding accuracy**

358 We used an across-participant correlation-based classifier to decode which stimulus timepoint
359 matched each timepoint's neural pattern. For a given value of k (i.e., number of principal compo-
360 nents), we first used group PCA to project the data from each condition into a shared k -dimensional
361 space. Next, we divided the participants into two groups: a template group, $\mathcal{G}_{\text{template}}$ (i.e., training
362 data), and a to-be-decoded group, $\mathcal{G}_{\text{decode}}$ (i.e., test data). We averaged the projected data within
363 each group to obtain a single T by k matrix for each group. Next, we correlated the rows of the two
364 averaged matrices to form a T by T decoding matrix, Λ . In this way, the rows of Λ reflected time-
365 points from the template group, while the columns reflected timepoints from the to-be-decoded
366 group. We used Λ to assign temporal labels to each timepoint (row) from the test group's ma-

367 trix, using the row of the training group’s matrix with which it was most highly correlated. We
368 repeated this decoding procedure, but using $\mathcal{G}_{\text{decode}}$ as the template group and $\mathcal{G}_{\text{template}}$ as the
369 to-be-decoded group. Given the true timepoint labels (for each group), we defined the decoding
370 accuracy as the average proportion of correctly decoded timepoints, across both groups (where
371 chance performance is $\frac{1}{T}$). In Figures 2 and 3 we report the decoding accuracy for each condition
372 and value of k , averaged across $n = 100$ cross validation folds.

373 **Reverse inference**

374 To help interpret the brain activity patterns we found within the contexts of other studies, we
375 created summary maps of each principal component, for each experimental condition. Each
376 principal component comprises 700 “weights” on each of the HTFA-derived RBF nodes (see
377 *Hierarchical Topographic Factor Analysis*). For each node, we evaluated its RBF at the locations of
378 every voxel in the standard 2 mm MNI152 template brain and multiplied the RBF by the node’s
379 weight. The sum of these weighted RBF activation maps provides a full-brain image, in MNI152
380 space, of the given principal component (Fig. S3).

381 Next, we considered 80 topics estimated using Latent Dirichlet Allocation (Blei et al., 2003)
382 applied to 9,204 functional neuroimaging articles in the Neurosynth database (Rubin et al., 2017).
383 The topics, as well as associated brain maps identified using Neurosynth, were identified and
384 reported in several prior studies (Chen et al., 2020; Fox et al., 2014; Sul et al., 2017). The topic
385 labels for each topic were generated automatically with the following ChatGPT (OpenAI, 2023)
386 prompt: “Please help me come up with intuitive labels for topics I found by fitting a topic model
387 to thousands of neuroscience and psychology articles. I’ll paste in the top 10 highest-weighted
388 words for each topic, and I’d like you to respond with a suggested label. For each topic, please
389 respond with just the topic label and no other formatting or text. Here are the next topic’s top
390 words:” followed by a comma-separated list of the given topic’s top-weighted words reflected
391 in the Table S1. For some topics, ChatGPT responded with a longer-form response rather than a
392 concise topic label. In these instances, on a case-by-case basis, we used a second follow-up prompt
393 to achieve the given topic’s label: “Could you please come up with a more concise label for that
394 topic?”. We then manually identified a set of 11 cognitive labels that were intended to encapsulate

395 a representative range of widely studied low-level and high-level cognitive functions. In choosing
396 the set of cognitive labels, we jointly considered each topic's ChatGPT-derived topic label, along
397 with the top-weighted words for the topic. We attempted to generate a concise set of labels that
398 still spanned the full set of cognitive functions reflected across the 80 topics. Topics that appeared
399 unrelated to specific cognitive functions (e.g., topics related to specific methods or clinical themes)
400 are designated with dashes in Table S1.

401 Finally, following an approach used in several prior studies (Chen et al., 2020; Fox et al., 2014;
402 Sul et al., 2017) we treated the correlation between a given component's brain map and each topic's
403 brain map as an approximate measure of how much the component was reflective of the given
404 topic. This resulted in a set of 80 "weights" (correlation coefficients) for each component's brain
405 map, with one weight per Neurosynth-derived topic.

406 **Ranking cognitive processes**

407 We manually identified 11 cognitive labels spanning the set of 80 Neurosynth-derived topics:
408 cognitive control, language processing, memory, emotion, social cognition, spatial cognition, at-
409 tention, reward, sensory perception, motor control, and resting state. We then used ChatGPT
410 to automatically "rank" the processes from high-level to low-level using the following prompt:
411 "Please rank these cognitive processes from highest-level to lowest-level, where higher values
412 indicate higher-order or higher-level processes. Return the result as a csv file with a header row
413 and two columns: 'Cognitive label' and 'Rank'. Here are the processes: cognitive control, lan-
414 guage processing, memory, emotion, social cognition, spatial cognition, attention, reward, sensory
415 perception, motor control, resting state". Table S2 displays the output. We used these labels
416 in the analysis presented in Figure 6E to help summarize difference in topic weightings across
417 experimental conditions.

418 **Synthetic data**

419 To help illustrate the relationship between informativeness and compressibility (Fig. 1), we gen-
420 erated four synthetic datasets, varying in informativeness and compressibility. Each dataset com-
421 prised simulated observations of $K = 25$ features across $N = 100$ timepoints, from each of $S = 10$

422 participants. To create each dataset, we first constructed a “template” matrix of N timepoints by K
423 features. We then generated participant-specific data by adding independent noise to each entry
424 in template matrix, drawn from the unit normal distribution (i.e., with a mean of 0 and a variance
425 of 1). We repeated this process for each participant, yielding S participant-specific matrices for
426 each dataset.

427 Since we estimate informativeness using the temporal decoding accuracy across participants,
428 highly informative data will tend to have observations that are highly timepoint specific. Relatively
429 uninformative data, in contrast, will tend to have more similar observations across timepoints. To
430 generate data with “high informativeness,” we constructed template matrices whose rows (ob-
431 servations) were drawn independently from zero-mean multivariate normal distributions. The
432 covariances of these distributions were determined according to the desired compressibility of
433 the data, as described below. We used a multi-step process to generate data with “low informa-
434 tiveness.” First we generated new template matrices using the same procedure as for the “high
435 informativeness” datasets. We then multiplied each matrix by a constant ($\rho = 0.1$) and computed
436 the cumulative sum of each matrix’s rows. This yielded matrices whose rows were highly similar
437 across observations.

438 Compressibility reflects the extent to which decoding accuracy is affected by reducing the num-
439 ber of components used to represent the data. Highly compressible data will tend to exhibit more
440 similarities across features, whereas less compressible data will tend to show greater independence
441 across features. To generate data with “high compressibility,” we set the covariance matrix of the
442 multivariate normal distribution to a toeplitz matrix whose first row was given by $[K, K - 1, \dots, 1]$.
443 To generate data with “low compressibility,” we set the covariance matrix to the identity matrix.

444 Template matrices for datasets with high informativeness and high compressibility, high in-
445 formativeness and low compressibility, low informativeness and high compressibility, and low
446 informativeness and low compressibility are displayed in Figure 1C. The corresponding decoding
447 curves are displayed in Figure 1D.

448 Data and code availability

449 All of the code used to produce the figures and results in this manuscript, along with links to the
450 corresponding data, may be found at github.com/ContextLab/pca_paper.

451 Acknowledgements

452 We acknowledge discussions with Rick Betzel, Luke Chang, Emily Finn, and Jim Haxby. Our
453 work was supported in part by NSF CAREER Award Number 2145172 to J.R.M. The content is
454 solely the responsibility of the authors and does not necessarily represent the official views of our
455 supporting organizations. The funders had no role in study design, data collection and analysis,
456 decision to publish, or preparation of the manuscript.

457 Author contributions

458 Conceptualization: J.R.M. and L.L.W.O. Methodology: J.R.M. and L.L.W.O. Implementation:
459 J.R.M. and L.L.W.O. Analysis: J.R.M. and L.L.W.O. Writing, Reviewing, and Editing: J.R.M. and
460 L.L.W.O. Funding acquisition: J.R.M. Supervision: J.R.M.

461 References

- 462 Adachi, Y., Osada, T., Sporns, O., Watanabe, T., Matsui, T., Miyamoto, K., & Miyashita, Y.
463 (2012). Functional connectivity between anatomically unconnected areas is shaped by collective
464 network-level effects in the macaque cortex. *Cerebral Cortex*, 22(7), 1586–1592.
- 465 Alvarez, S. A. (2002). *An exact analytical relation among recall, precision, and classification accuracy in*
466 *information retrieval* (Technical Report No. BCCS-02-01). Boston College.
- 467 Baldassano, C., Chen, J., Zadbood, A., Pillow, J. W., Hasson, U., & Norman, K. A. (2017). Discov-
468 ering event structure in continuous narrative perception and memory. *Neuron*, 95(3), 709–721.
- 469 Bassett, D. S., & Sporns, O. (2017). Network neuroscience. *Nature Neuroscience*, 20(3), 353–364.

- 470 Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: a practical and
471 powerful approach to multiple testing. *Journal of Royal Statistical Society, Series B*, 57, 289–300.
- 472 Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning
473 Research*, 3, 993–1022.
- 474 Brovelli, A., Ding, M., Ledberg, A., Chen, Y., Nakamura, R., & Bressler, S. (2004). Beta oscillations in
475 a large-scale sensorimotor cortical network: directional influences revealed by Granger causality.
476 *Proceedings of the National Academy of Sciences, USA*, 101(26), 9849–9854.
- 477 Bullmore, E., & Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural
478 and functional systems. *Nature Reviews Neuroscience*, 10(3), 186–198.
- 479 Capota, M., Turek, J., Chen, P.-H., Zhu, X., Manning, J. R., Sundaram, N., ... Shin, Y. S. (2017).
480 *Brain imaging analysis kit*.
- 481 Chen, P.-H. A., Jolly, E., Cheong, J. H., & Chang, L. J. (2020). Intersubject representational
482 similarity analysis reveals individual variations in affective experience when watching erotic
483 movies. *NeuroImage*, 216(116851), doi.org/10.1016/j.neuroimage.2020.116851.
- 484 Cole, M. W., Bassett, D. S., Power, J. D., Braver, T. S., & Petersen, S. E. (2014). Intrinsic and
485 task-evoked network architectures of the human brain. *Neuron*, 83(1), 238–251.
- 486 Dhamala, M., Rangarajan, G., & Ding, M. (2008). Analyzing information flow in brain networks
487 with nonparametric Granger causality. *NeuroImage*, 41(2), 354–362.
- 488 Finn, E. S., Scheinost, D., Finn, D. M., Shen, X., Papademetris, X., & Constable, R. T. (2017).
489 Can brain state be manipulated to emphasize individual differences in functional connectivity.
490 *NeuroImage*, 160, 140–151.
- 491 Finn, E. S., Shen, X., Scheinost, D., Rosenberg, M. D., Huang, J., Chun, M. M., ... Constable, R. T.
492 (2015). Functional connectome fingerprinting: identifying individuals using patterns of brain
493 connectivity. *Nature Neuroscience*, 18, 1664–1671.

- 494 Fox, A. S., Chang, L. J., Gorgolewski, K. J., & Yarkoni, T. (2014). Bridging psychology and
495 genetics using large-scale spatial analysis of neuroimaging and neurogenetic data. *bioRxiv*,
496 doi.org/10.1101/012310.
- 497 Glerean, E., Salmi, J., Lahnakoski, J. M., Jääskeläinen, I. P., & Sams, M. (2012). Functional magnetic
498 resonance imaging phase synchronization as a measure of dynamic functional connectivity.
499 *Brain Connectivity*, 2(2), 91–101.
- 500 Gratton, C., Laumann, T. O., Nielsen, A. N., Greene, D. J., Gordon, E. M., Gilmore, A. W., ...
501 Petersen, S. E. (2018). Functional brain networks are dominated by stable group and individual
502 factors, not cognitive or daily variation. *Neuron*, 98(2), 439–452.
- 503 Hasson, U., Yang, E., Vallines, I., Heeger, D. J., & Rubin, N. (2008). A hierarchy of temporal
504 receptive windows in human cortex. *The Journal of Neuroscience*, 28(10), 2539–2550.
- 505 Korzeniewska, A., Crainiceanu, C. M., Kus, R., Franaszczuk, P. J., & Crone, N. E. (2008). Dynamics
506 of event-related causality in brain electrical activity. *Human Brain Mapping*, 29(10).
- 507 Kumar, M., Anderson, M. J., Antony, J. W., Baldassano, C., Brooks, P. P., Cai, M. B., ... Norman,
508 K. A. (2021). BrainIAK: the brain image analysis kit. *Aperture*, In press.
- 509 Lerner, Y., Honey, C. J., Silbert, L. J., & Hasson, U. (2011). Topographic mapping of a hierarchy of
510 temporal receptive windows using a narrated story. *The Journal of Neuroscience*, 31(8), 2906–2915.
- 511 Lynn, C. W., & Bassett, D. S. (2021). Quantifying the compressibility of complex networks.
512 *Proceedings of the National Academy of Sciences, USA*, 118(32), e2023473118.
- 513 Mack, M. L., Preston, A. R., & Love, B. C. (2020). Ventromedial prefrontal cortex compression
514 during concept learning. *Nature Communications*, 11(46), doi.org/10.1038/s41467-019-13930-8.
- 515 Manning, J. R. (2020). Context reinstatement. In M. J. Kahana & A. D. Wagner (Eds.), *Handbook of
516 human memory*. Oxford University Press.
- 517 Manning, J. R., Zhu, X., Willke, T. L., Ranganath, R., Stachenfeld, K., Hasson, U., ... Norman,
518 K. A. (2018). A probabilistic approach to discovering dynamic full-brain functional connectivity
519 patterns. *NeuroImage*, 180, 243–252.

- 520 Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: multi-voxel
521 pattern analysis of fMRI data. *Trends in Cognitive Sciences*, 10(9), 424–430.
- 522 OpenAI. (2023, March). *ChatGPT*. Personal communication.
- 523 Owen, L. L. W., Chang, T. H., & Manning, J. R. (2021). High-level cognition during story listening is
524 reflected in high-order dynamic correlations in neural activity patterns. *Nature Communications*,
525 12(5728), doi.org/10.1038/s41467-021-25876-x.
- 526 Owen, L. L. W., Muntianu, T. A., Heusser, A. C., Daly, P., Scangos, K. W., & Manning, J. R. (2020). A
527 Gaussian process model of human electrocorticographic data. *Cerebral Cortex*, 30(10), 5333–5345.
- 528 Preti, M. G., Bolton, T. A. W., & Van De Ville, D. (2017). The dynamic functional connectome:
529 state-of-the-art and perspectives. *NeuroImage*, 160, 41–54.
- 530 Regev, M., Simony, E., Lee, K., Tan, K. M., Chen, J., & Hasson, U. (2018). Propagation of information
531 along the cortical hierarchy as a function of attention while reading and listening to stories.
532 *Cerebral Cortex*.
- 533 Rogers, B. P., Morgan, V. L., Newton, A. T., & Gore, J. C. (2007). Assessing functional connectivity
534 in the human brain by fMRI. *Magnetic Resonance Imaging*, 25(11), 1347–1357.
- 535 Rubin, T. N., Kyoejo, O., Gorgolewski, K. J., Jones, M. N., Poldrack, R. A., & Yarkoni, T. (2017).
536 Decoding brain activity using a large-scale probabilistic functional-anatomical atlas of human
537 cognition. *PLoS Computational Biology*, 13(10), e1005649.
- 538 Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: uses and
539 interpretations. *NeuroImage*, 52, 1059–1069.
- 540 Scangos, K. W., Khambhati, A. N., Daly, P. M., Owen, L. L. W., Manning, J. R., Ambrose, J. B.,
541 ... Chang, E. F. (2021). Biomarkers of depression symptoms defined by direct intracranial
542 neurophysiology. *Frontiers in Human Neuroscience*, In press.
- 543 Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*,
544 27(3), 379–423.

- 545 Simony, E., & Chang, C. (2020). Analysis of stimulus-induced brain dynamics during naturalistic
546 paradigms. *NeuroImage*, 216, 116461.
- 547 Simony, E., Honey, C. J., Chen, J., & Hasson, U. (2016). Dynamic reconfiguration of the default
548 mode network during narrative comprehension. *Nature Communications*, 7(12141), 1–13.
- 549 Sizemore, A. E., Giusti, C., Kahn, A., Vettel, J. M., Betzel, R. F., & Bassett, D. S. (2018). Cliques and
550 cavities in the human connectome. *Journal of Computational Neuroscience*, 44(1), 115–145.
- 551 Smith, S. M., Beckmann, C. F., Andersson, J., Auerbach, E. J., Bijsterbosch, J., Douaud, G., ...
552 Glasser, M. F. (2013). Resting-state fMRI in the Human Connectome Project. *NeuroImage*, 80,
553 144–168.
- 554 Smith, S. M., Fox, P. T., Miller, K. L., Glahn, D. C., Fox, P. M., Mackay, C. E., ... Beckmann,
555 C. F. (2009). Correspondence of the brain's functional architecture during activation and rest.
556 *Proceedings of the National Academy of Sciences, USA*, 106(31), 13040–13045.
- 557 Smith, S. M., Hyvärinen, A., Varoquaux, G., Miller, K. L., & Beckmann, C. F. (2014). Group-PCA
558 for very large fMRI datasets. *NeuroImage*, 101, 738–749.
- 559 Smith, S. M., Vidaurre, D., Beckmann, C. F., Glasser, M. F., Jenkinson, M., Miller, K. L., ... Van
560 Essen, D. C. (2013). Functional connectomics from resting-state fMRI. *Trends in Cognitive Sciences*,
561 17(12), 666–682.
- 562 Sporns, O., & Betzel, R. F. (2016). Modular brain networks. *Annual Review of Psychology*, 67,
563 613–640.
- 564 Sporns, O., & Honey, C. J. (2006). Small worlds inside big brains. *Proceedings of the National Academy
565 of Sciences, USA*, 103(51), 19219–19220.
- 566 Sporns, O., & Zwi, J. D. (2004). The small world of the cerebral cortex. *Neuroinformatics*, 2(2),
567 145–162.
- 568 Srinivasan, R., Winter, W. R., Ding, J., & Nunez, P. L. (2007). EEG and MEG coherence: Measures of
569 functional connectivity at distinct spatial scales of neocortical dynamics. *Journal of Neuroscience
570 Methods*, 166, 41–52.

- 571 Sul, S., Güroğlu, B., Crone, E. A., & Chang, L. J. (2017). Medial prefrontal cortical thinning
572 mediates shifts in other-regarding preferences during adolescence. *Scientific Reports*, 7(8510),
573 doi.org/10.1038/s41598-017-08692-6.
- 574 Tomasi, D., & Volkow, N. D. (2011). Association between functional connectivity hubs and brain
575 networks. *Cerebral Cortex*, 21, 2003–2013.
- 576 Yeo, B. T. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., . . . Buckner,
577 R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional
578 connectivity. *Journal of Neurophysiology*, 106(3), 1125–1165.