

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

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

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

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

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

March 2, 2023

Abstract

We applied dimensionality reduction algorithms and pattern classifiers to functional neuroimaging data collected as participants listened to a story, temporally scrambled versions of the story, or underwent a resting state scanning session. These experimental conditions were intended to require different depths of processing and inspire different levels of cognitive engagement. We considered two primary aspects of the data. First, we treated the maximum achievable decoding accuracy across participants as an indicator of the “informativeness” of the recorded patterns. Second, we treated the number of features (components) required to achieve a threshold decoding accuracy as a proxy for the “compressibility” of the neural patterns (where fewer components indicate greater compression). Overall, we found that the peak decoding accuracy (achievable without restricting the numbers of features) was highest in the intact (unscrambled) story listening condition. However, the number of features required to achieve comparable classification accuracy was also lowest in the intact story listening condition. Taken together, our work suggests that our brain networks flexibly reconfigure according to ongoing task demands, and that the activity patterns associated with higher-order cognition and high engagement are both more informative and more compressible than the activity patterns associated with lower-order tasks and lower levels of engagement.

Keywords: information, compression, temporal decoding, dimensionality reduction, neuroimaging

24 Introduction

25 Large-scale networks, including the human brain, may be conceptualized as occupying one or
26 more positions along on a continuum. At one extreme, every node is fully independent from
27 every other node. At the other extreme, all nodes behave identically. Each extreme optimizes
28 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 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., 2017, 2015; Gratton et al., 2018) and across (Cole et al., 2014;

59 Glerean et al., 2012; Gratton et al., 2018; Yeo et al., 2011) 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 (Cole et al., 2014; Finn et al., 2017; Owen et al., 2020), 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

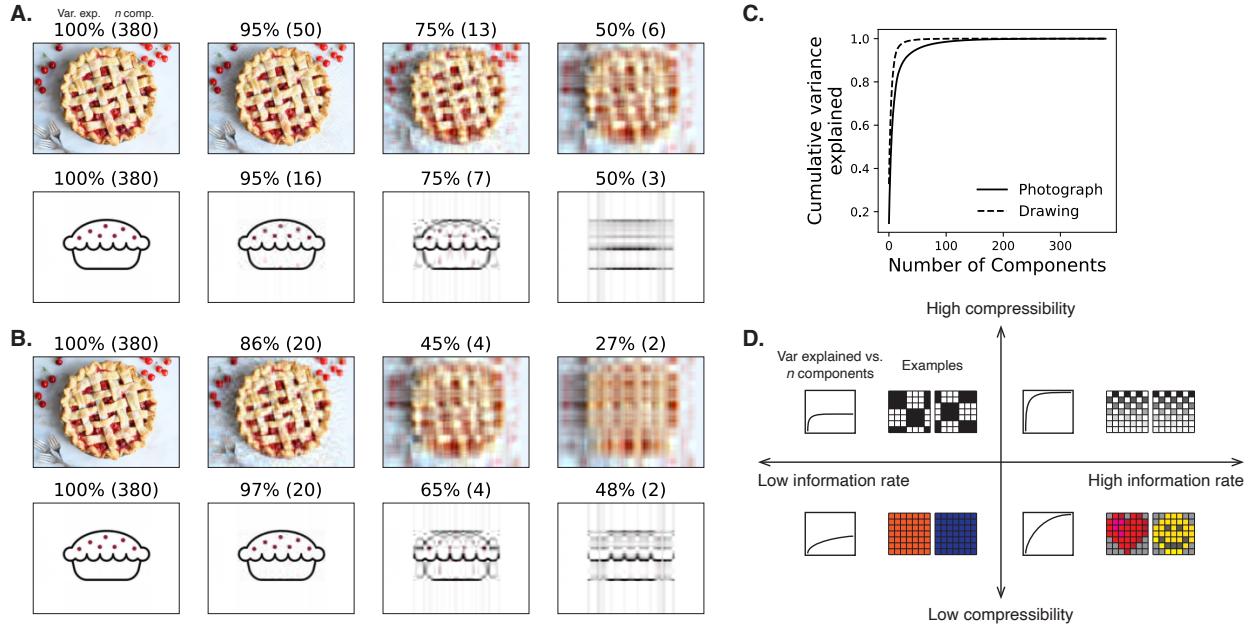


Figure 1: Information content and compressibility. **A. Variance explained for two images.** We applied principal components analysis to a photograph and drawing, treating the rows of the images as “observations.” Across columns, we identified the numbers of components required to explain 100%, 95%, 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. **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 reflected 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 (Hasson et al., 2008; Lerner et al., 2011; Owen et al., 2021; Simony et al., 2016). 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 compressibility for brain activity from
134 each experimental condition, we trained a series of across-participant temporal decoders on com-
135 pressed representations of the data. Figure 2A displays the decoding accuracy as a function of the
136 number of principal components used to represent the data. Several patterns were revealed by the
137 analysis. First, in general (i.e., across experimental conditions), decoding accuracy improves as the
138 number of components increases. However, decoding accuracy peaked at higher levels for exper-
139 imental conditions that exposed participants to cognitively richer stimuli. The peak decoding ac-
140 curacy was highest for the “intact” condition (versus paragraph: $t(XXX) = XXX, p = XXX$; versus
141 word: $t(XXX) = XXX, p = XXX$; versus rest: $t(XXX) = XXX, p = XXX$), next highest for the “para-
142 graph” condition (versus word: $t(XXX) = XXX, p = XXX$; versus rest: $t(XXX) = XXX, p = XXX$),
143 and next highest for the “word” condition (versus rest: $t(XXX) = XXX, p = XXX$). This ordering
144 implies 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

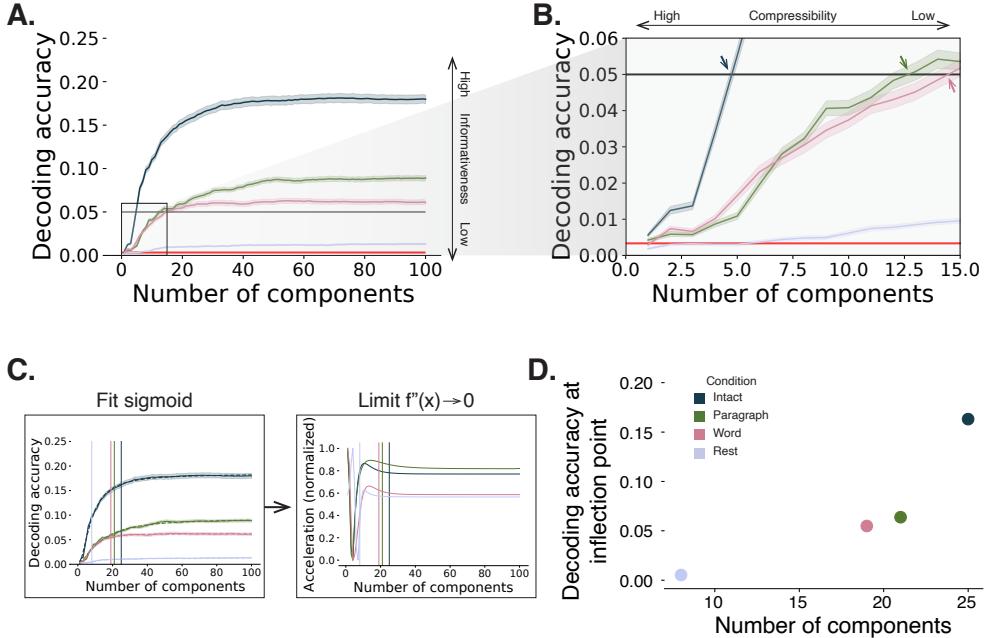


Figure 2: Decoding accuracy and compression. **A. Decoding accuracy by number of components.** Ribbons of each color display cross-validated decoding performance for each condition (intact, paragraph, word, and rest), as a function of the number of components (features) used to represent the data. The 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 for which 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 If informativeness (to the temporal decoders) and compressibility vary with the cognitive
164 richness of the stimulus, might these measures also vary over time *within* a given condition? For
165 example, participants in the intact condition might process the ongoing story more deeply later
166 on in the story (compared with earlier in the story) given the additional narrative background
167 and context they had been exposed to by that point. To examine this possibility, we divided
168 each condition into three successive time segments. We computed decoding curves (Fig. 3A)
169 and inflection points (Fig. 3B) for each segment and condition. We found that, in the two most
170 cognitively rich conditions (intact and paragraph), both decoding accuracy and compressibility,
171 as reflected by the change in decoding curves, increased with listening time (intact: $t(\text{XXX}) =$
172 $\text{XXX}, p = \text{XXX}$; paragraph: $t(\text{XXX}) = \text{XXX}, p = \text{XXX}$). These changes may reflect an increase in
173 comprehension or depth of processing with listening time. In contrast, the decoding accuracy and
174 compressibility *decreased* with listening time in the word condition ($t(\text{XXX}) = \text{XXX}, p = \text{XXX}$) and
175 rest condition ($t(\text{XXX}) = \text{XXX}, p = \text{XXX}$). This might reflect the depletion of attentional resources
176 in the less-engaging word and rest conditions.

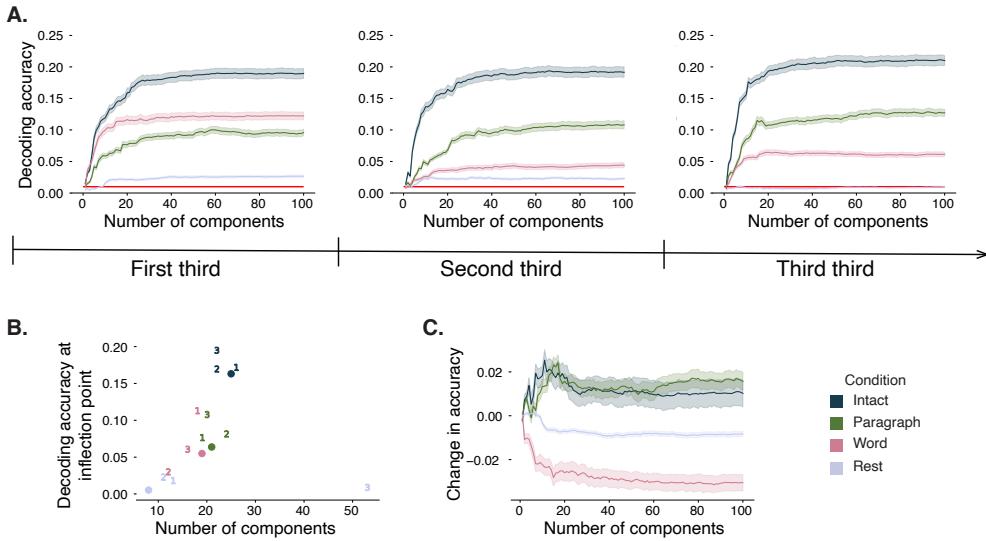


Figure 3: Changes in decoding accuracy and compression over time. **A. Decoding accuracy by number of components, by story segment.** Each family of curves is plotted in the same format as Figure 2A but reflects data only from one third of the dataset. **B. Inflection points by condition and segment.** 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.

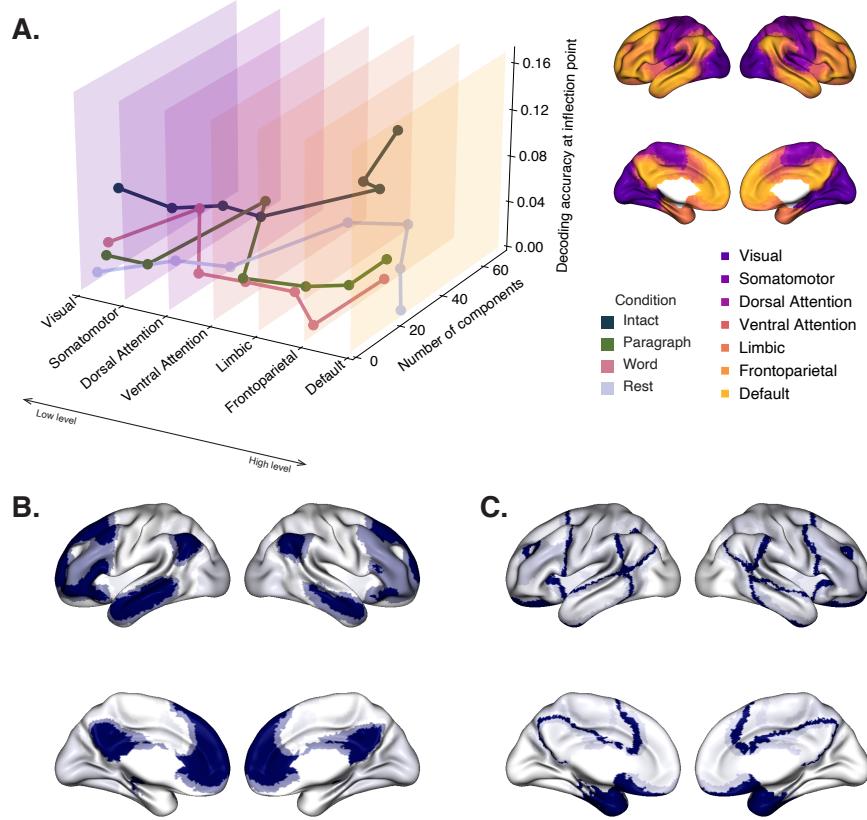


Figure 4: Network-specific decoding accuracy and compression. **A. Decoding accuracy and number of components for network-specific inflection points.** We considered the seven networks identified by Yeo et al. (2011). We computed each network’s inflection point, for each experimental condition, using 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.

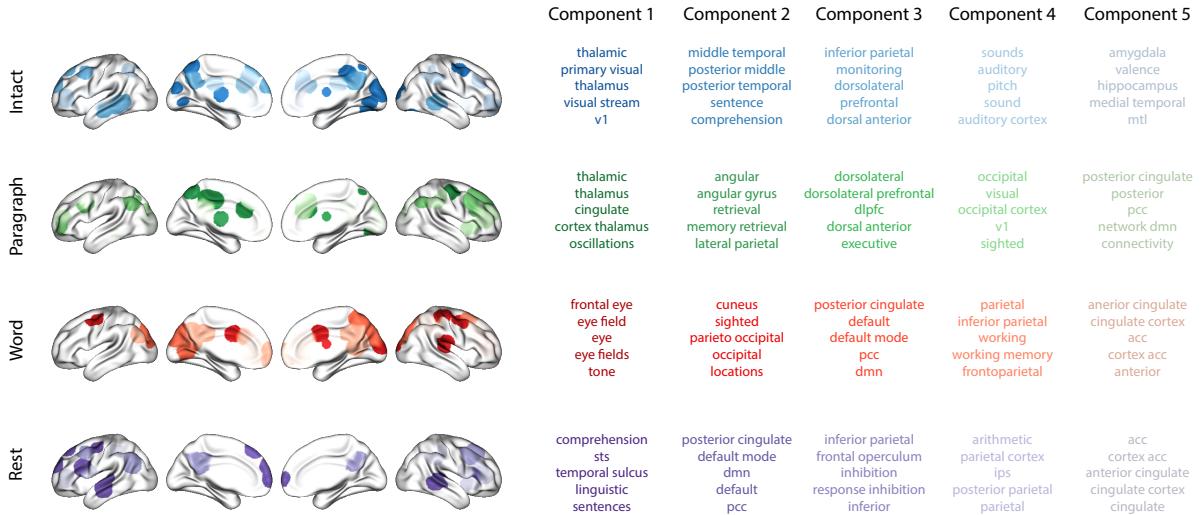


Figure 5: Top terms associated with the highest-weighted components by condition. Each row corresponds to an experimental condition, and the colors correspond to the component number (ranked by proportion of variance explained). The inflated brain plots display the top 20 highest-weighted hubs (see *Topographic 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.

177 We also wondered how informativeness and compressibility in the different experimental
 178 conditions might vary across brain networks. We used a network parcellation identified by
 179 Yeo et al. (2011) to segment the brain into seven distinct networks. The networks can be sorted
 180 (roughly) in order from lower-level to higher-level cortex as follows (Fig. 4A): visual, somatomotor,
 181 dorsal attention, ventral attention, limbic, frontoparietal, and default mode. Next, we computed
 182 decoding curves separately for the activity patterns within each network and identified each
 183 network's inflection point, for each experimental condition. Moving from low-order networks
 184 to higher-order networks, we found that decoding accuracy (for the intact condition) tended to
 185 increase (Fig. 4B). This suggests that higher-order networks may carry more content-relevant or
 186 stimulus-driven "information." We found no clear trends in the numbers of components at each
 187 network's inflection point across networks or conditions (Fig. 4C).

188 In addition to examining different networks in isolation, we wondered about the general
 189 structure of the full-brain (i.e., potentially multi-network) activity patterns reflected by different
 190 principal components across different experimental conditions. Figure 5 displays inflated brain
 191 maps of the top five highest-weighted components, for each experimental condition. We also

192 used Neurosynth (Rubin et al., 2017) to identify, for each component, the top five terms associated
193 with each map (see *Reverse inference*). We noticed (by inspection) several common themes across
194 the sets of terms associated with each component and condition. Memory-related components
195 included terms like “middle temporal,” “memory retrieval,” and “working memory.” Sensory
196 processing related components included terms like “primary visual,” “auditory cortex,” “v1,” and
197 so on. Other components were associated with sensory integration (e.g., “thalamic,” “cingulate,”
198 etc.), sentence comprehension (e.g., “sentence,” “comprehension”), emotion and valence (e.g.,
199 “amygdala,” “valance”), or the default mode network (e.g., “default mode”). The components we
200 identified were relatively stable across story segments (Fig. S1).

201 Discussion

202 We examined fMRI data collected as participants listened to an auditory recording of a story, scram-
203 bled recordings of the story, or underwent a resting state scan. We found that cognitively richer
204 stimuli evoked more reliable (i.e., consistent across people) and information rich brain activity pat-
205 terns. The brain patterns evoked by cognitively richer stimuli were also more compressible, in that
206 each individual component provided more “signal” to temporal decoders relative to components
207 of data from less cognitively rich conditions (Fig. 2). Over time (e.g., as the experiment progressed),
208 these phenomena were strengthened. Specifically, across story segments, data from more cogni-
209 tively rich conditions became more informative and compressible, and data from less cognitively
210 rich conditions became *less* informative and compressible (Fig. 3). We also repeated these analyses
211 separately for different brain networks. We found that networks traditionally associated with
212 higher-level cognitive functions tended to provide more informative brain patterns than networks
213 traditionally associated with lower level cognitive functions (Fig. 4). Finally, we examined the
214 most dominant components of the brain activity patterns from each experimental condition. We
215 used a reverse inference approach (Rubin et al., 2017) to identify the terms in the neuroimaging
216 literature most commonly associated with the corresponding maps. As summarized in Figure 6,
217 we found that terms associated with memory and sensory processing were associated with the
218 strongest components in all three story listening conditions. Terms associated with sensory in-
219 tegration were associated with the strongest components in the intact and paragraph-scrambled

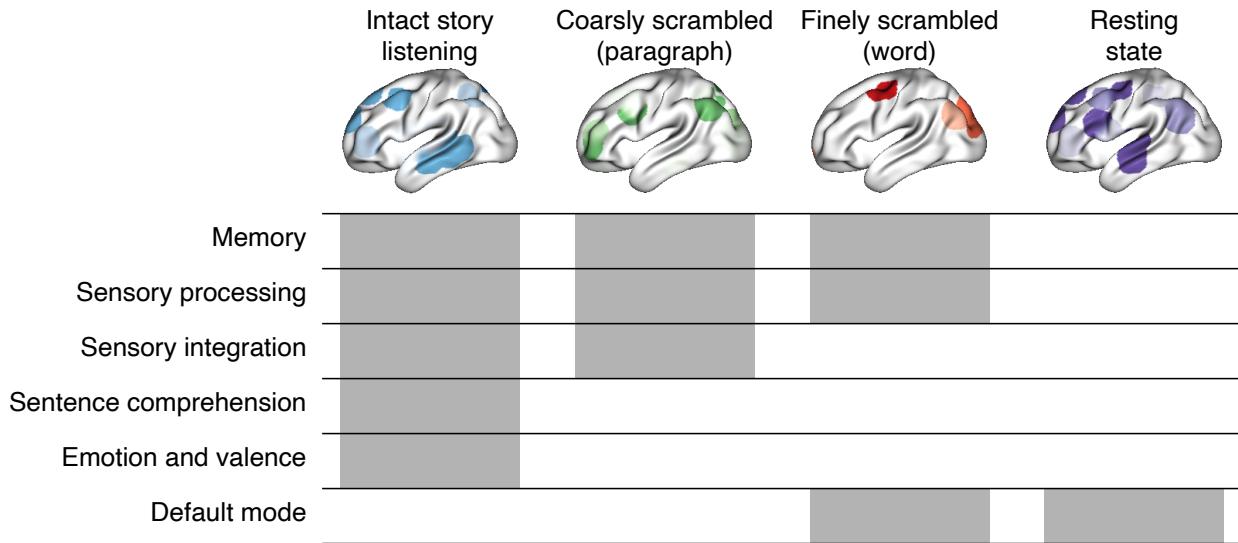


Figure 6: Summary of functions associated with top-weighted components by condition. Each column corresponds to an experimental condition. Brain maps in the top row are reproduced from Figure 5, for reference. Cognitive functions summarized from the top Neurosynth-derived terms in Figure 5 are listed in the rows on the left. Shaded cells denote which experimental conditions were associated with one or more top-weighted principal components associated with the given function.

220 conditions. Terms associated with sentence comprehension, emotion, and valence were associated
 221 with the strongest components in the intact condition. Finally, terms associated with the default
 222 mode network were associated with the strongest components in the word-scrambled and resting
 223 state conditions. Taken together, our findings indicate that the informativeness and compressibil-
 224 ity of our brain activity patterns are task-dependent, and these properties change systematically
 225 with factors like cognitive richness and depth of processing.

226 Our explorations of informativeness and compressibility are related to a much broader litera-
 227 ture on the correlational and causal structure of brain activity patterns (Adachi et al., 2012; Bassett
 228 & Sporns, 2017; Brovelli et al., 2004; Bullmore & Sporns, 2009; Dhamala et al., 2008; Korzeniewska
 229 et al., 2008; Owen et al., 2021; Preti et al., 2017; Rogers et al., 2007; Rubinov & Sporns, 2010; Size-
 230 more et al., 2018; Smith, Beckmann, et al., 2013; Smith, Vidaurre, et al., 2013; Sporns & Betzel, 2016;
 231 Sporns & Honey, 2006; Sporns & Zwi, 2004; Srinivasan et al., 2007; Tomasi & Volkow, 2011; Yeo
 232 et al., 2011). Correlations or causal associations between different brain regions simultaneously
 233 imply that full-brain activity patterns will be compressible and also that those activity patterns
 234 will contain redundancies. For example, the extent of which activity patterns at one brain area can

235 be inferred or predicted from activity patterns at other areas (e.g., Owen et al., 2020; Scangos et al.,
236 reflects overlap in the information available in or represented by those brain areas. If brain
237 patterns in one area are recoverable using brain patterns in another area, then the “signal” used to
238 convey the activity patterns could be compressed by removing the recoverable activity. Predictable
239 (and therefore redundant) brain activity patterns are also more robust to signal corruption. For
240 example, even if the activity patterns at one region are unreadable or unreliable at a given moment,
241 that unreliability could be compensated for by other regions’ activity patterns that were predictive
242 of the unreliable region.

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

254 In the information theory sense (Shannon, 1948), when a signal is transmitted using a fixed
255 alphabet of “symbols,” the information rate decreases as the signal is compressed (e.g., fewer sym-
256 bols transmitted per unit time, using an alphabet with fewer symbols, etc.). Our finding that each
257 individual brain component (symbol) becomes more informative as cognitive richness increases
258 suggests that the “alphabet” of brain activity patterns is also task-dependent. Other work suggests
259 that the representations that are *reflected* by brain activity patterns may also change with task de-
260 mands. For example, our brains may represent the same perceptual stimulus differently depending
261 on which aspects of the stimulus or which combinations of features are task-relevant (Mack et al.,
262 2020).

263 Different brain networks also varied in how informative and compressible their activity pat-
264 terns were across experimental conditions (e.g., Fig. 4). This might follow from evolutionary

optimizations that reflect the relevant constraints or demands placed on those networks. One possibility is that cortex is organized in a hierarchy of networks “concerned with” or selective to different levels of processing or function. To the extent that different levels of processing (e.g., low-level sensory processing versus “deeper” higher-level processing) reflect different stimulus timescales (e.g., Manning, 2020), the network differences we observed might also relate to the timescales at which each network is maximally sensitive (Baldassano et al., 2017; Hasson et al., 2008; Lerner et al., 2011; Regev et al., 2018).

Concluding remarks

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

Methods

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

Functional neuroimaging data collected during story listening

We examined an fMRI dataset collected by Simony et al. (2016) that the authors have made publicly available at arks.princeton.edu/ark:/88435/dsp015d86p269k. The dataset comprises neuroimaging data collected as participants listened to an audio recording of a story (intact condition; 36 participants), listened to temporally scrambled recordings of the same story (17 participants in the paragraph-scrambled condition listened to the paragraphs in a randomized order and 36 in the word-scrambled condition listened to the words in a randomized order), or lay resting with their eyes open in the scanner (rest condition; 36 participants). Full neuroimaging details may be found

291 in the original paper for which the data were collected (Simony et al., 2016). Procedures were
292 approved by the Princeton University Committee on Activities Involving Human Subjects, and by
293 the Western Institutional Review Board (Puyallup, WA). All subjects were native English speakers
294 with normal hearing and provided written informed consent.

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

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

307 **Principal components analysis (PCA)**

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

313 **Temporal decoding**

314 We sought to identify neural patterns that reflected participants’ ongoing cognitive processing of
315 incoming stimulus information. As reviewed by Simony et al. (2016), one way of homing in on
316 these stimulus-driven neural patterns is to compare activity patterns across individuals. In partic-

317 ular, neural patterns will be similar across individuals to the extent that the neural patterns under
318 consideration are stimulus-driven, and to the extent that the corresponding cognitive representa-
319 tions are reflected in similar spatial patterns across people (Simony & Chang, 2020). Following this
320 logic, we used an across-participant temporal decoding test developed by Manning et al. (2018) to
321 assess the degree to which different neural patterns reflected ongoing stimulus-driven cognitive
322 processing across people. The approach entails using a subset of the data to train a classifier to
323 decode stimulus timepoints (i.e., moments in the story participants listened to) from neural pat-
324 terns. We use decoding (forward inference) accuracy on held-out data, from held-out participants,
325 as a proxy for the extent to which the inputted neural patterns reflected stimulus-driven cognitive
326 processing in a similar way across individuals.

327 **Forward inference and decoding accuracy**

328 We used an across-participant correlation-based classifier to decode which stimulus timepoint
329 matched each timepoint's neural pattern. For a given value of k (i.e., number of principal compo-
330 nents), we first used group PCA to project the data from each condition into a shared k -dimensional
331 space. Next, we divided the participants into two groups: a template group, $\mathcal{G}_{\text{template}}$ (i.e., training
332 data), and a to-be-decoded group, $\mathcal{G}_{\text{decode}}$ (i.e., test data). We averaged the projected data within
333 each group to obtain a single T by k matrix for each group. Next, we correlated the rows of the two
334 averaged matrices to form a T by T decoding matrix, Λ . In this way, the rows of Λ reflected time-
335 points from the template group, while the columns reflected timepoints from the to-be-decoded
336 group. We used Λ to assign temporal labels to each timepoint (row) from the test group's ma-
337 trix, using the row of the training group's matrix with which it was most highly correlated. We
338 repeated this decoding procedure, but using $\mathcal{G}_{\text{decode}}$ as the template group and $\mathcal{G}_{\text{template}}$ as the
339 to-be-decoded group. Given the true timepoint labels (for each group), we defined the decoding
340 accuracy as the average proportion of correctly decoded timepoints, across both groups (where
341 chance performance is $\frac{1}{T}$). In Figures 2 and 3 we report the decoding accuracy for each condition
342 and value of k , averaged across $n = 100$ cross validation folds.

343 **Reverse inference**

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

350 **Data and code availability**

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

353 **Acknowledgements**

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

359 **Author contributions**

360 Conceptualization: J.R.M. and L.L.W.O. Methodology: J.R.M. and L.L.W.O. Implementation: L.L.W.O.
361 Analysis: L.L.W.O. Writing, Reviewing, and Editing: J.R.M. and L.L.W.O. Supervision: J.R.M.

362 **References**

363 Adachi, Y., Osada, T., Sporns, O., Watanabe, T., Matsui, T., Miyamoto, K., & Miyashita, Y.
364 (2012). Functional connectivity between anatomically unconnected areas is shaped by collective
365 network-level effects in the macaque cortex. *Cerebral Cortex*, 22(7), 1586–1592.

- 366 Alvarez, S. A. (2002). *An exact analytical relation among recall, precision, and classification accuracy in*
367 *information retrieval* (Technical Report No. BCCS-02-01). Boston College.
- 368 Baldassano, C., Chen, J., Zadbood, A., Pillow, J. W., Hasson, U., & Norman, K. A. (2017). Discovering
369 event structure in continuous narrative perception and memory. *Neuron*, 95(3), 709–721.
- 370 Bassett, D. S., & Sporns, O. (2017). Network neuroscience. *Nature Neuroscience*, 20(3), 353–364.
- 371 Brovelli, A., Ding, M., Ledberg, A., Chen, Y., Nakamura, R., & Bressler, S. (2004). Beta oscillations in
372 a large-scale sensorimotor cortical network: directional influences revealed by Granger causality.
373 *Proceedings of the National Academy of Sciences, USA*, 101(26), 9849–9854.
- 374 Bullmore, E., & Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural
375 and functional systems. *Nature Reviews Neuroscience*, 10(3), 186–198.
- 376 Capota, M., Turek, J., Chen, P.-H., Zhu, X., Manning, J. R., Sundaram, N., ... Shin, Y. S. (2017).
377 *Brain imaging analysis kit*.
- 378 Cole, M. W., Bassett, D. S., Power, J. D., Braver, T. S., & Petersen, S. E. (2014). Intrinsic and
379 task-evoked network architectures of the human brain. *Neuron*, 83(1), 238–251.
- 380 Dhamala, M., Rangarajan, G., & Ding, M. (2008). Analyzing information flow in brain networks
381 with nonparametric Granger causality. *NeuroImage*, 41(2), 354–362.
- 382 Finn, E. S., Scheinost, D., Finn, D. M., Shen, X., Papademetris, X., & Constable, R. T. (2017).
383 Can brain state be manipulated to emphasize individual differences in functional connectivity.
384 *NeuroImage*, 160, 140–151.
- 385 Finn, E. S., Shen, X., Scheinost, D., Rosenberg, M. D., Huang, J., Chun, M. M., ... Constable, R. T.
386 (2015). Functional connectome fingerprinting: identifying individuals using patterns of brain
387 connectivity. *Nature Neuroscience*, 18, 1664–1671.
- 388 Glerean, E., Salmi, J., Lahnakoski, J. M., Jääskeläinen, I. P., & Sams, M. (2012). Functional magnetic
389 resonance imaging phase synchronization as a measure of dynamic functional connectivity.
390 *Brain Connectivity*, 2(2), 91–101.

- 391 Gratton, C., Laumann, T. O., Nielsen, A. N., Greene, D. J., Gordon, E. M., Gilmore, A. W., ...
392 Petersen, S. E. (2018). Functional brain networks are dominated by stable group and individual
393 factors, not cognitive or daily variation. *Neuron*, 98(2), 439–452.
- 394 Hasson, U., Yang, E., Vallines, I., Heeger, D. J., & Rubin, N. (2008). A hierarchy of temporal
395 receptive windows in human cortex. *The Journal of Neuroscience*, 28(10), 2539–2550.
- 396 Korzeniewska, A., Crainiceanu, C. M., Kus, R., Franaszczuk, P. J., & Crone, N. E. (2008). Dynamics
397 of event-related causality in brain electrical activity. *Human Brain Mapping*, 29(10).
- 398 Kumar, M., Anderson, M. J., Antony, J. W., Baldassano, C., Brooks, P. P., Cai, M. B., ... Norman,
399 K. A. (2021). BrainIAK: the brain image analysis kit. *Aperture*, In press.
- 400 Lerner, Y., Honey, C. J., Silbert, L. J., & Hasson, U. (2011). Topographic mapping of a hierarchy of
401 temporal receptive windows using a narrated story. *The Journal of Neuroscience*, 31(8), 2906–2915.
- 402 Mack, M. L., Preston, A. R., & Love, B. C. (2020). Ventromedial prefrontal cortex compressesion
403 during concept learning. *Nature Communications*, 11(46), doi.org/10.1038/s41467-019-13930-8.
- 404 Manning, J. R. (2020). Context reinstatement. In M. J. Kahana & A. D. Wagner (Eds.), *Handbook of
405 human memory*. Oxford University Press.
- 406 Manning, J. R., Zhu, X., Willke, T. L., Ranganath, R., Stachenfeld, K., Hasson, U., ... Norman,
407 K. A. (2018). A probabilistic approach to discovering dynamic full-brain functional connectivity
408 patterns. *NeuroImage*, 180, 243–252.
- 409 Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: multi-voxel
410 pattern analysis of fMRI data. *Trends in Cognitive Sciences*, 10(9), 424–430.
- 411 Owen, L. L. W., Chang, T. H., & Manning, J. R. (2021). High-level cognition during story listening is
412 reflected in high-order dynamic correlations in neural activity patterns. *Nature Communications*,
413 12(5728), doi.org/10.1038/s41467-021-25876-x.
- 414 Owen, L. L. W., Muntianu, T. A., Heusser, A. C., Daly, P., Scangos, K. W., & Manning, J. R. (2020). A
415 Gaussian process model of human electrocorticographic data. *Cerebral Cortex*, 30(10), 5333–5345.

- 416 Preti, M. G., Bolton, T. A. W., & Van De Ville, D. (2017). The dynamic functional connectome:
417 state-of-the-art and perspectives. *NeuroImage*, 160, 41–54.
- 418 Regev, M., Simony, E., Lee, K., Tan, K. M., Chen, J., & Hasson, U. (2018). Propagation of information
419 along the cortical hierarchy as a function of attention while reading and listening to stories.
420 *Cerebral Cortex*.
- 421 Rogers, B. P., Morgan, V. L., Newton, A. T., & Gore, J. C. (2007). Assessing functional connectivity
422 in the human brain by fMRI. *Magnetic Resonance Imaging*, 25(11), 1347–1357.
- 423 Rubin, T. N., Kyojo, O., Gorgolewski, K. J., Jones, M. N., Poldrack, R. A., & Yarkoni, T. (2017).
424 Decoding brain activity using a large-scale probabilistic functional-anatomical atlas of human
425 cognition. *PLoS Computational Biology*, 13(10), e1005649.
- 426 Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: uses and
427 interpretations. *NeuroImage*, 52, 1059–1069.
- 428 Scangos, K. W., Khambhati, A. N., Daly, P. M., Owen, L. L. W., Manning, J. R., Ambrose, J. B.,
429 ... Chang, E. F. (2021). Biomarkers of depression symptoms defined by direct intracranial
430 neurophysiology. *Frontiers in Human Neuroscience*, In press.
- 431 Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*,
432 27(3), 379–423.
- 433 Simony, E., & Chang, C. (2020). Analysis of stimulus-induced brain dynamics during naturalistic
434 paradigms. *NeuroImage*, 216, 116461.
- 435 Simony, E., Honey, C. J., Chen, J., & Hasson, U. (2016). Dynamic reconfiguration of the default
436 mode network during narrative comprehension. *Nature Communications*, 7(12141), 1–13.
- 437 Sizemore, A. E., Giusti, C., Kahn, A., Vettel, J. M., Betzel, R. F., & Bassett, D. S. (2018). Cliques and
438 cavities in the human connectome. *Journal of Computational Neuroscience*, 44(1), 115–145.
- 439 Smith, S. M., Beckmann, C. F., Andersson, J., Auerbach, E. J., Bijsterbosch, J., Douaud, G., ...
440 Glasser, M. F. (2013). Resting-state fMRI in the Human Connectome Project. *NeuroImage*, 80,
441 144–168.

- 442 Smith, S. M., Fox, P. T., Miller, K. L., Glahn, D. C., Fox, P. M., Mackay, C. E., ... Beckmann,
443 C. F. (2009). Correspondence of the brain's functional architecture during activation and rest.
444 *Proceedings of the National Academy of Sciences, USA*, 106(31), 13040–13045.
- 445 Smith, S. M., Hyvärinen, A., Varoquaux, G., Miller, K. L., & Beckmann, C. F. (2014). Group-PCA
446 for very large fMRI datasets. *NeuroImage*, 101, 738–749.
- 447 Smith, S. M., Vidaurre, D., Beckmann, C. F., Glasser, M. F., Jenkinson, M., Miller, K. L., ... Van
448 Essen, D. C. (2013). Functional connectomics from resting-state fMRI. *Trends in Cognitive Sciences*,
449 17(12), 666–682.
- 450 Sporns, O., & Betzel, R. F. (2016). Modular brain networks. *Annual Review of Psychology*, 67,
451 613–640.
- 452 Sporns, O., & Honey, C. J. (2006). Small worlds inside big brains. *Proceedings of the National Academy
of Sciences, USA*, 103(51), 19219–19220.
- 454 Sporns, O., & Zwi, J. D. (2004). The small world of the cerebral cortex. *Neuroinformatics*, 2(2),
455 145–162.
- 456 Srinivasan, R., Winter, W. R., Ding, J., & Nunez, P. L. (2007). EEG and MEG coherence: Measures of
457 functional connectivity at distinct spatial scales of neocortical dynamics. *Journal of Neuroscience
Methods*, 166, 41–52.
- 459 Tomasi, D., & Volkow, N. D. (2011). Association between functional connectivity hubs and brain
460 networks. *Cerebral Cortex*, 21, 2003–2013.
- 461 Yeo, B. T. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., ... Buckner,
462 R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional
463 connectivity. *Journal of Neurophysiology*, 106(3), 1125–1165.