bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507854; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Perception and Memory Retrieval States are Reflected in Distributed Patterns of Background Functional Connectivity

Y. Peeta Li[1*], Yida Wang[2], Nicholas B. Turk-Browne[3,4], Brice A. Kuhl[1], J. Benjamin Hutchinson[1]

[1] Department of Psychology, University of Oregon, Eugene, OR

[2] Amazon Web Services, Palo Alto, CA

[3] Department of Psychology, Yale University, New Haven, CT

[4] Wu Tsai Institute, Yale University, New Haven, CT

∗ Corresponding Author.
E-mail address: peetal@uoregon.edu (Y. Peeta Li).

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507854; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

ABSTRACT

The same visual input can serve as the target of perception or as a trigger for memory retrieval depending on whether cognitive processing is externally oriented (perception) or internally oriented (memory retrieval). While numerous human neuroimaging studies have characterized how visual stimuli are differentially processed during perception versus memory retrieval, perception and memory retrieval may also be associated with distinct neural states that are independent of stimulus-evoked neural activity. Here, we combined human fMRI with full correlation matrix analysis (FCMA) to reveal potential differences in "background" functional connectivity across perception and memory retrieval states. We found that perception and retrieval states could be discriminated with high accuracy based on patterns of connectivity across (1) the control network, (2) the default mode network (DMN), and (3) retrosplenial cortex (RSC). Clusters in the control network increased connectivity with each other during the perception state, whereas clusters in the DMN were more strongly coupled during the retrieval state. Interestingly, RSC switched its coupling between networks as the cognitive state shifted from retrieval to perception. Finally, we show that background connectivity captured distinct cognitive aspects compared to traditional classification of stimulus-evoked responses. Together, our results reveal that perception and memory retrieval are associated with sustained cognitive states that manifest as distinct patterns of connectivity among large-scale brain networks.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507854; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

INTRODUCTION

When a familiar visual stimulus (e.g., a colleague) is encountered, it might serve as the target of perceptual scrutiny (e.g., their current facial expression) or the trigger for episodic memory retrieval (e.g., something they said in the past). Numerous cognitive theories and neuroimaging studies have contrasted the cognitive processes and brain regions engaged during perception versus retrieval (Bosch et al., 2014; Chun & Johnson, 2011; Kosslyn et al., 1995; McClelland et al., 1995; O'Reilly & McClelland, 1994; Polyn et al., 2005; Wheeler et al., 2000). The majority of this work has focused on investigating the operations performed upon the *stimulus*, such as how the stimulus is encoded into long-term memory or how the stimulus is used to drive memory recollection (Favila et al., 2020; Fernandez et al., 2022; Kim, 2013). However, evidence also suggests that perception and memory retrieval are characterized by distinct 'states' that are sustained over time and independent from external stimuli (Duncan et al., 2012; Duncan & Shohamy, 2016; Ezzyat et al., 2017; Guderian et al., 2009; Hasselmo et al., 1996). Characterizing how perception and retrieval states are implemented in the brain represents an important challenge, but one that is complicated by methodological and analytical factors.

One potentially powerful way to characterize distinct cognitive states is by measuring patterns of functional connectivity within the brain (Cohen & D'Esposito, 2016; Fritch et al., 2021; Shirer et al., 2012; Song & Rosenberg, 2021). Functional connectivity is typically computed over extended windows of time and is therefore well-suited to measure states that putatively persist across, or in the absence of, external stimuli. However, a critical issue in applying functional connectivity during cognitive tasks is that any observed correlations in neural activity may be largely or entirely driven by stimulus-evoked neural responses. That is, if two brain regions consistently respond to external stimuli, this will induce apparent "connectivity" between these regions. Several fMRI studies have addressed this concern by using "background" connectivity—an approach in which stimulus-evoked responses are explicitly modeled and removed, with connectivity then computed using the residual (background) activity (Al-Aidroos et al., 2012; Cole et al., 2019; Duncan et al., 2014; Norman-Haignere et al., 2012; Turk-Browne, 2013). Conceptually, this approach can isolate interactions between brain regions that reflect sustained, endogenous processing as opposed to transient, exogenous responses to stimuli (Summerfield et al., 2006; Turk-Browne, 2013).

To date, only a limited number of fMRI studies have used background connectivity to test for differences between perception and memory retrieval states (Cooper & Ritchey, 2019; Duncan et al., 2014). Moreover, these studies have only considered potential interactions (connectivity) between a relatively limited number of brain regions. For example, Cooper and Ritchey (2019) examined background FC patterns exclusively between *a priori*, memory-related brain regions in the posterior-medial (PM) and anterior-temporal (AT) networks. They found that regions in these networks exhibited stronger background connectivity during retrieval compared to perception, providing important evidence that perception versus retrieval states can be linked to stimulus-independent processes. While this type of targeted, seed-based analysis is highly

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507854; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

valuable for testing specific hypotheses about regions of a priori interest, it is inherently blind to interactions involving non-seed regions (Turk-Browne, 2013). In other words, discoveries are systematically biased to come from the regions that are tested (Wang et al., 2015). Thus, there is value to unbiased approaches that allow for background FC to be more comprehensively measured—ideally, across the whole brain. That said, while conducting whole-brain functional connectivity may be theoretically appealing, it can be computationally intractable, particularly if applied without down-sampling data. For example, brain volumes consisting of 50,000 voxels would yield 1.25B voxel pairs ('connections') to analyze. For this reason, previous attempts to measure whole-brain functional connectivity have tended to substantially reduce the dimensionality of data by grouping voxels into regions or parcels (e.g., Pantazatos et al., 2012; Shirer et al., 2012; Watanabe et al., 2012).

Here, we sought to identify patterns of background connectivity associated with perception versus retrieval states using full correlation matrix analysis (FCMA) applied to human fMRI data. In contrast, to typical seed-based connectivity measures, FCMA comprehensively considers connectivity between every possible pair of voxels in the brain using sophisticated approaches to overcome historical computational limitations (Kumar et al., 2022; Turk-Browne, 2013; Wang et al., 2015). Specifically, FCMA uses a combination of machine learning (support vector machine; SVM) and parallel computing to efficiently map patterns of connectivity to stimulus or task information (Turk-Browne, 2013; Wang et al., 2015). The computational efficiency afforded by this approach is substantial in that it can reduce computation time from weeks to hours and has the potential to reveal information in fine-grained connectivity patterns that would be missed by conventional seed-based functional connectivity (Wang et al., 2015). We applied FCMA using an experiment that manipulated perception versus retrieval states while carefully controlling for three potentially confounding variables: (1) subjects performed the same judgments on the same images across both perception and retrieval tasks, minimizing differences in task demands and sensory content; (2) subjects were familiarized with the images in all conditions, minimizing differences in novelty/encoding across perception and retrieval; and (3) behavioral accuracy was matched across key conditions, minimizing differences in task difficulty.

To preview, we report four main findings: (1) background connectivity, which is orthogonal to evoked responses, allowed for classification of perception versus retrieval states with remarkably high accuracy; (2) perception and retrieval states can be discriminated parsimoniously based on background connectivity patterns in a relatively small number of clusters that span three functional communities: the control network, default mode network (DMN), and retrosplenial cortices (RSC); (3) connections within the control network were relatively stronger during perception states whereas connections within the DMN were relatively stronger during retrieval states; and (4) RSC shifted its coupling with the control network and DMN as a function of cognitive state, suggesting that it acts as a hub for transitioning between perception and retrieval.

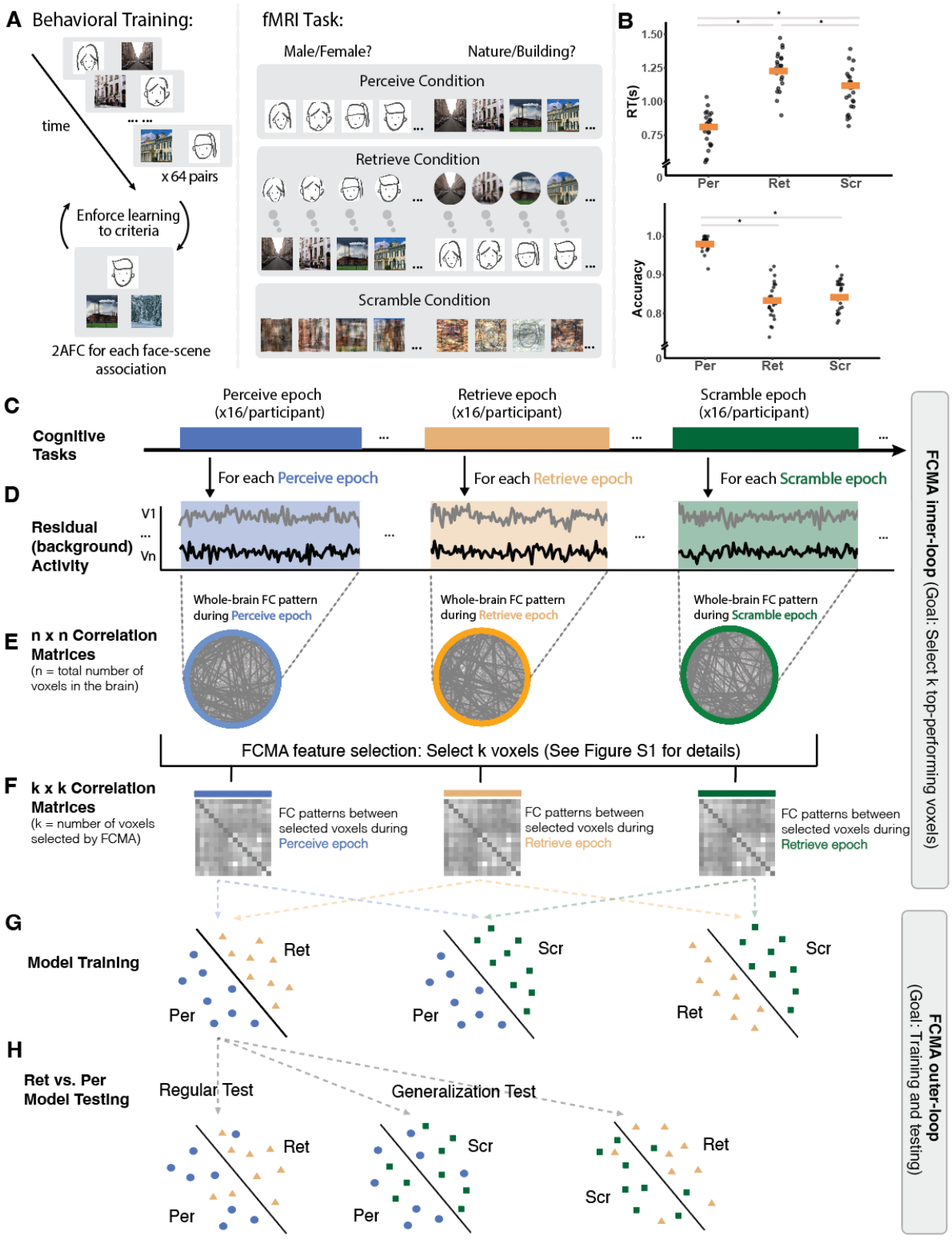
bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507854; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

RESULTS

Perception and Retrieval Tasks

We designed an fMRI task that required subjects to perform judgments on information that was either available in the perceptual environment (perception) or had to be retrieved from memory (retrieval), in both cases matched in terms of visual content and task difficulty (**Figure 1A**). During an initial behavioral training session outside of the scanner, 24 subjects were extensively trained to remember pairs of face-scene associations, with the faces being either male or female and the scenes being either manmade or natural (See **Methods: *Procedure*** for details). During the scanning session, subjects were presented with blocks of either eight faces or eight scenes, one at a time. In the *Perceive* condition, they were asked to make either a male/female or manmade/natural decision for the visually presented image. In the *Retrieve* condition, they were asked to retrieve from memory the image associated with the visually presented image and to make a male/female or manmade/natural decision for that retrieved image (of the opposite category); for example, if the subject saw a face image, they would need to retrieve the specific associated scene image and perform a manmade/natural judgement. Notably, the *Perceive* and *Retrieve* conditions differed in cognitive demands, with the latter being more difficult as captured by judgement accuracy (**Figure 1B**). To account for potential neural differences introduced by task difficulty and not retrieval versus perceptual states per se, we introduced a third condition, the *Scramble* condition. Specifically, subjects performed the same perception task as in the *Perceive* condition, but with visually scrambled images to increase the task difficulty. The amount of scrambling was determined through behavioral piloting in a separate sample of subjects to match accuracy between the *Retrieve* and *Scramble* conditions. Subjects completed 16 epochs of each of the three task conditions, evenly divided into two fMRI runs, resulting in a total of six fMRI runs; the order of the task conditions was randomized across subjects. **Figure 1B** shows the behavioral performance for each condition. One-way (condition) repeated measures ANOVAs revealed a main effect of condition for both reaction time ($F_{(2, 46)} = 288.8$, $p < 0.001$, $\eta^2 = 0.97$) and accuracy ($F_{(2, 46)} = 91.21$, $p < 0.001$, $\eta^2 = 0.90$). Importantly, although accuracy in the *Perceive* condition was significantly greater than that in the *Retrieve* condition ($t_{(23)} = 13.4$, $p < 0.001$, 95% CI = [0.16, 0.22], Cohen's d = 3.55), the *Retrieve* and *Scramble* conditions did not significantly differ, as intended ($t_{(23)} = -1.65$, $p = 0.11$, 95% CI = [-0.07, 0.01], Cohen's d = 0.45).

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507854; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507854; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 1. Task paradigm and analysis flowchart. (**A**) Behavioral training and in-scanner tasks. Subjects formed mental associations between face[1] and scene images outside of the scanner and then made judgments on either perceptual or mnemonic information in the scanner. (**B**) In-scanner behavioral performance, with the orange bar indicating the sample mean and asterisks indicating $p < 0.001$. (**C**) Each condition had 16 epochs per subject, with each epoch lasting 40 s. (**D**) The residual activity for each task epoch was computed by regressing out the stimulus-evoked component using a finite impulse response general linear model. **(E**) Whole-brain voxel-wise background FC matrices were computed for each task epoch. These matrices were *n*-by-*n* shape, where *n* is the total number of voxels in the brain (*n* = 92,745). (**F**) Background FC matrices of FCMA-selected voxels. FCMA was implemented to select a certain number of voxels whose connectivity patterns were most accurate for separating perception and retrieval states based on two-way classifications of epochs from different conditions. The details of the feature selection process are shown in **Figure S1**. In essence, FCMA reduced the whole-brain correlation matrix to a *k*-by-*k* correlation matrix for each epoch, where *k* is the number of voxels selected by FCMA. The current study tested multiple values of *k*, ranging from 100 to 15,000. (**G**) Model training. Using a cross-validation framework, SVM classifiers with precomputed linear kernel were trained to separate all task condition comparisons. (**H**) Model testing. Trained models were tested using the left-out subject for each fold. During the regular test, the model was tested on the same task condition comparison as it was trained on. During the generalization test, the model was tested on a different task condition comparison than it was trained on.

Perception and Retrieval Involve Distinct Intrinsic Connectivity Patterns

We first examined whether perception and retrieval states involve different "state-related" whole-brain FC patterns. Using background connectivity analysis, the stimulus-evoked component of the preprocessed blood oxygen level dependent (BOLD) time series was regressed out using a general linear model (Al-Aidroos et al., 2012; Cole et al., 2019; Córdova et al., 2016). The residuals were then used for computing voxel-wise whole-brain correlation matrix for each epoch, yielding 16 subject-specific whole-brain background FC matrices for each of the three conditions (each FC matrix was 92745 voxels by 92745 voxels; **Figure 1E**). To examine the potential differences between background FC patterns associated with perception and retrieval states, we applied full correlation matrix analysis (FCMA) to test whether a trained support vector machine (SVM) could successfully separate perception epochs (*Perceive* and *Scramble*) from retrieval epochs (*Retrieve;* **Figure 1G-H**).

Notably, the whole-brain correlation matrix is not easily interpretable to the human eye given its high dimensionality. However, FCMA allows identification of the most important/diagnostic regions of the brain involved in discriminating cognitive states (**Figure 1F**). Specifically, FCMA selected the *k* best performing voxels comprising smaller, clearer FC patterns which could then be used for follow-up analyses. To evaluate binary classifier performance within the top-performing voxels, we computed the receiver operating characteristic (ROC) curves and area under the ROC curve (AUC) for each classifier per subject, with larger AUC indicating better model performance (Hanley & McNeil, 1982). We started by setting *k* at a low value of 100 voxels and gradually increased it until classifier performance asymptoted. As seen in **Figure 2A,** accuracy plateaued when the mask reached 3,000 voxels, and the overall pattern of results did not dramatically change as a function of the mask size used. Accordingly, we performed follow-up analyses on the top *k* = 3,000 voxel masks. These analyses suggested that classifiers trained on background FC patterns successfully differentiated epochs of perception from retrieval states

[1] Due to the preprint policy of biorxiv, we are unable to show the actual face stimuli used in the experiment. For added details on the stimuli, see **Methods: *Material***.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507854; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(*Perceive* vs. *Retrieve*: M_{AUC} = 0.87 ± 0.06, $t_{(23)}$ = 28.62, p < .001, 95% CI = [0.84, 0.89], Cohen's d = 5.84; *Retrieve* vs. *Scramble*: M_{AUC} = 0.83 ± 0.08; $t_{(23)}$ 19.39, p < .001, 95% CI = [0.79, 0.86], Cohen's d = 3.96; one-sample t-test against chance-level performance of $\bar{\mu}$ = 0.5; see also **Figure S2A** for results in terms of percent accuracy). Control analyses showed that state-related differences were selectively captured by background FC measures and not by left-over stimuli-evoked components in the residual timeseries. In particular, models using residual activity patterns (i.e., simple MVPA on residual timeseries) in the same voxel masks failed to discriminate any task condition comparisons (all model performances ≈ 50%; **Figure S2B**).

One potential concern with the FCMA results above is that the above-chance classification performance was not strictly related to differences in perception versus retrieval states *per se* and might reflect the contribution of confounding factors. That is, in addition to capturing differences in perception and retrieval, the *Perceive* vs. *Retrieve* comparison also varied in terms of task difficulty (i.e., accuracy) and the *Scramble vs. Retrieve* comparison also varied in terms of sensory input (i.e., scrambled vs. intact stimuli). To ensure that the variations in background FC patterns (among FCMA-selected voxels) were most strongly induced by state-related differences, we performed two extra sets of analyses to rule out these potential confounds. First, we tested the background FC pattern separability of *Perceive* vs. *Scramble* conditions—a task condition comparison that differed in terms of both task difficulty and visual content *but not* in perception/retrieval cognitive states. If the background FC patterns contain mostly state-related information, the consistent perception state for both *Perceive* and *Scramble* should worsen classifier performance. Consistent with this hypothesis, AUC significantly differed between the non-state-related and state-related background FC classifiers ($F_{(2,46)}$ = 20.30, p < 0.001, η^2 = 0.47), with post-hoc tests revealing that *Perceive-Scramble* AUC was significantly lower than *Retrieve-Perceive* ($t_{(23)}$ = -5.62, p < 0.001, 95% CI = [-0.19, -0.09], Cohen's d = 1.45; **Figure 2A**) and *Retrieve-Scramble* ($t_{(46)}$ = -4.10, p < 0.001, 95% CI = [-0.15, -0.05], Cohen's d = 0.97). No significant difference was observed between the other two FC classifiers, although there was a trend for higher AUC in the *Retrieve-Perceive* comparison ($t_{(46)}$ = 1.82, p = 0.08, 95% CI = [-0.01, 0.08], Cohen's d = 0.52).

In the second set of analyses, we leveraged classification generalization to assess the degree to which background FC patterns capture differences within each of the three dimensions: (1) cognitive state (i.e., perception vs. retrieval); (2) task difficulty; and (3) visual content. This was operationalized as how well a classifier trained on one pair of conditions generalized to another pair of conditions that differed on the same dimension (**Figure 1F**). We predict that generalizability should be greatest for task condition comparisons that involve state-related differences (i.e., *Perceive* vs. *Retrieve* to/from *Scramble* vs. *Retrieve*; **Figure 2B**). Indeed, we found that background FC classifiers differed in their sensitivities to the three dimensions (averaged bidirectional generalization AUC: cognitive task state: M_{AUC} = 0.78 ± 0.07; task difficulty: M_{AUC} = 0.68 ± 0.07; visual content: M_{AUC} = 0.53 ± 0.11; $F_{(2,69)}$ = 52.80, p < 0.001, η^2 = 0.60,

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507854; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 2C). Followed up t-tests suggested that background FC classifier had the greatest sensitivity to state-related differences than other differences (*p*s < 0.001). Together, these two analyses provide additional support to the initial finding that background FC patterns (among FCMA-selected voxels) contain information about perception and retrieval states above and beyond other differences between conditions

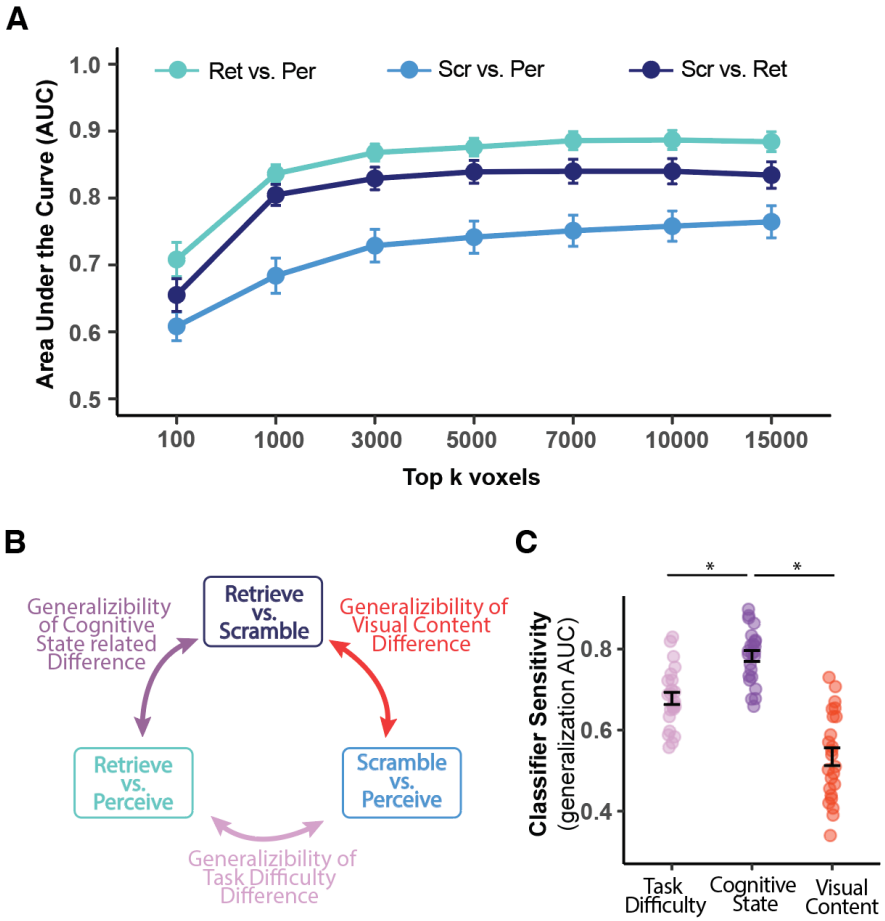


Figure 2. Classification of task conditions based on different features. **A**) Performance of background FC classifiers as a function of the number of voxels selected by FCMA. Area under the receiver operating characteristic curve (AUC) was computed for each leave-one-subject-out testing fold to examine the performance of the binary classifiers. The error bars indicate the standard error of the mean AUC across all subjects being tested. The results suggest that the top 3,000 voxels were able to differentiate task conditions as well as larger sets of voxels and so were used for follow-up analyses. **B**) Schematic diagram for measuring sensitivity to different aspects of cognitive processes using generalization tests. For example, given that both *Retrieve* vs. *Scramble* and *Retrieve* vs. *Perceive* comparisons involve state-related differences, success in classifying *Retrieve* vs. *Perceive* with the background FC classifier trained on *Retrieve* vs. *Scramble* (and vice versa) would provide strong evidence that the classifier is sensitive to state-related differences. **C)** Sensitivity of background FC classifiers to three distinctions: cognitive state (i.e., perception vs. retrieval), visual content (i.e., scrambled vs. intact), and task difficulty (i.e., low vs. high accuracy). Sensitivity was quantified as the average AUC from classifiers trained on one set of conditions and tested on another set along the same task-related dimension (generalization test). Error bars indicate the standard error of the mean across all testing folds. Asterisks indicate $p < 0.05$.

Background Connectivity and Evoked Activity Patterns Capture Distinct Cognitive Processes

Previous research has shown that stimulus-evoked, multi-voxel activity patterns reflect ongoing cognitive processes and can be used to train classifiers for separating task conditions (Norman et al., 2006). We

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507854; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

investigated whether patterns of stimulus-evoked activity reflect similar or distinct aspects of perception and retrieval states compared with patterns of background FC. Specifically, we trained MVPA classifiers on activity patterns for each pair of task conditions and compared performance to classifiers trained using background FC patterns (**Figure 3 left and middle**). Stimulus-evoked activity patterns led to reliable classification of task conditions (*p*s < 0.001). Interestingly, and different from background FC classifiers, activity classifiers did not perform better when discriminating between perception and retrieval conditions (*Retrieve* vs. *Perceive*: M_{auc} = 0.86 $\pm$ 0.06) than between the two perception conditions (*Scramble* vs. *Perceive*: M_{auc}= 0.90 $\pm$ 0.12; $t_{(46)}$ = 0.42, p = 0.67, 95% CI = [-0.05, 0.07], Cohen's d = 0.12). Indeed, a two-way repeated measures ANOVA confirmed an interaction between classifier condition pair (*Retrieve* vs. *Perceive*, *Scramble* vs. *Perceive*) and neural measure (stimulus-evoked activity, background FC): $F_{(2,46)}$ = 9.70, p < 0.001, η^2 = 0.30).

Based on these findings, stimulus-evoked activity and background FC may capture distinct component processes of perception and retrieval. To test this hypothesis more directly, we trained hybrid classifiers that combined both stimulus-evoked activity and background FC. The decision confidences of the hybrid classifier were computed as the averaged decision function outputs from both FC and MVPA classifiers. The rationale is that the hybrid classifiers should achieve better performance than either classifier on its own if the two neural measures capture distinct state-related processes (Manning et al., 2018). As expected, a two-way repeated measures ANOVA revealed a significant main effect of classifier types (i.e., FC vs. MVPA vs. hybrid; $F_{(2,46)}$ = 68.32, p < 0.001, η^2 = 0.75; **Figure 3**). Follow-up analyses showed that the average AUC of hybrid classifiers was significantly greater than background FC classifiers ($t_{(23)}$ = 9.95, p < 0.001, 95% CI = [0.11, 0.17], Cohen's d = 2.58) and evoked-activity MVPA classifiers ($t_{(23)}$ = 3.06, p = 0.006, 95% CI = [0.01, 0.04], Cohen's d = 0.88). Together, these results suggest that evoked activity and background FC reflect distinct neural signatures of cognitive processes, with the latter capturing more state-related (e.g., perception vs. retrieval states) information.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507854; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

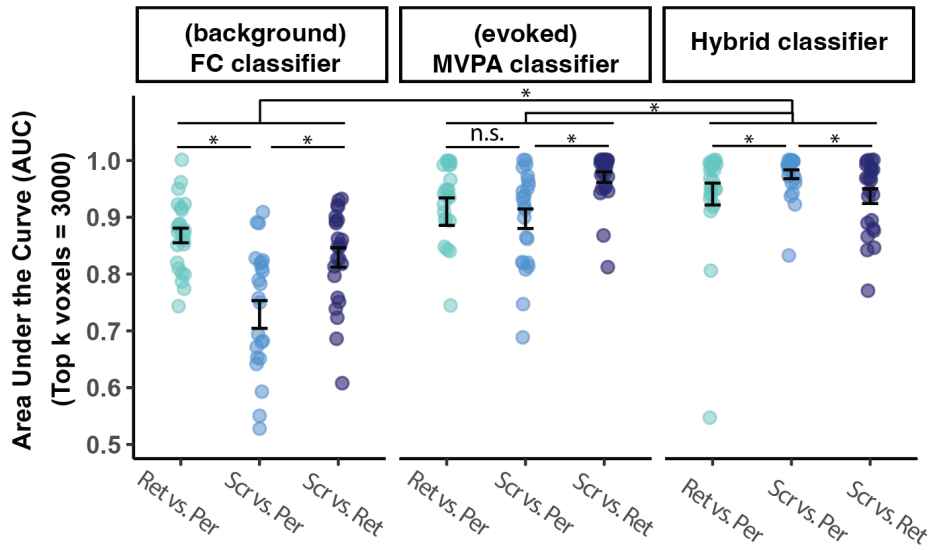


Figure 3. Classification of task conditions based on different neural measures. Comparing performance of binary classifiers trained with different neural measures: background FC (left), evoked activity (middle), hybrid FC + activity (right). Each dot represents a subject. The error bars indicate the standard error of the mean AUC across all subjects being tested.

Regions and Functional Communities Underlying Perception and Retrieval States

The results so far suggest that background patterns of FC among voxels selected with FCMA capture differences between perception and retrieval cognitive states. However, interpretation of these results is complicated by two key issues: 1) the top voxels were not necessarily identical across cross-validation folds and 2) the number of connections between voxels remains quite large. In terms of the first issue, because of leave-one-subject-out cross-validation, different sets of voxels were selected by FCMA for each left-out testing subject/fold. Accordingly, to allow further characterization and interpretation of state-related background FC patterns, we combined the FCMA inner-loop (**Figure S1**) with permutation-based statistical inference tests to obtain a shared set of voxels across all testing folds (yielding roughly 4,000 voxels) whose background FC patterns captured state-related differences (for details see **Methods**: ***Information Mapping***). In terms of the second issue, even with a shared voxel mask, it may still be intractable to interpret FC patterns consisting of millions of connections (i.e., $\sim 8 \times 10^6$ unique connections among 4,000 voxels). For this reason, we further reduced the dimensions of FC patterns using clustering. We first identified spatially contiguous clusters of voxels in the shared mask. In order to determine an effective extent threshold for cluster size, we only included clusters whose FC contributed to task condition classification performance above and beyond larger clusters. That is, we trained classifiers to differentiate each task condition comparison starting with all voxels in the largest cluster and sequentially adding voxels from the next largest cluster until the benefit of adding in clusters on performance plateaued (**Figure S3A**). This FCMA-then-clustering pipeline revealed 16 clusters of interests, whose cluster-level background FC patterns provided a parsimonious summary of the neural sources distinguishing between perception and retrieval states (**Figure 4A**; **Table S1**).

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507854; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Notably, in addition to its greater interpretability, this FCMA-then-clustering approach also produced superior classification accuracy to the use of *a priori* parcellation-based clusters (Feilong et al., 2021; Schaefer et al., 2018)). That is, by repeating the analysis pipeline with predefined brain parcels (instead of fine-grained voxel-wise analysis), the 16 clusters derived from voxel-level analyses had significantly greater discrimination performance compared to the best performing 16 predefined parcels in the parcel-level analysis ($F_{(1,138)} = 60.25$, $p < 0.001$, $\eta^2 = 0.30$; **Figure S3B right**). This result held across different parcellation granularities (400 and 1000 parcels; Schaefer et al., 2018). Moreover, cluster-level background FC patterns retained their preference for state-related vs. non-state-related differences (**Figure S3A**, **B**; $F_{(2,69)} = 19.31$, $p < 0.001$, $\eta^2 = 0.36$; *Scramble* vs. *Perceive* classifier had the worst model performance; *p*s < 0.001).

Although the analyses above inform *which* brain regions might be most involved in differentiating perception and retrieval cognitive states, they do not provide information about *how* the regions are differentially connected. Indeed, previous research has suggested that functionally coupled brain regions form large-scale functional communities (Yeo et al., 2011) and that a cluster may be associated with different functional communities across different cognitive states (Braun et al., 2015). With this in mind, we next examined the functional community structures of the 16 clusters of interest during each task condition (i.e., *Perceive*, *Retrieve*, and *Scramble*). Specifically, we applied the Louvain community detection algorithm (Blondel et al., 2008) to the group-averaged background FC matrices for each condition. The algorithm was repeated 1,000 times on each weighted graph and the partition that maximized modularity for each condition was selected (Rubinov & Sporns, 2011; see **Methods: *Community Detection*** for details). Interestingly, the community structures underlying all three task conditions were consistent in that the 16 clusters were consistently partitioned into 3 functional communities (**Figure 4A**). The first functional community consisted of regions from the conventional default mode network (Buckner et al., 2008), including the bilateral inferior parietal lobule, precuneus, medial prefrontal cortex, posterior cingulate cortex, and the middle temporal gyrus, which hereafter we refer to as the DMN. The second functional community consisted of the bilateral prefrontal cortex, bilateral intraparietal sulcus, superior frontal gyrus, and temporal gyrus, most of which are part of the frontoparietal control network (Marek & Dosenbach, 2018); we refer to this as the Control network[2]. The last community consisted of ventral and dorsal retrosplenial cortices (RSC; Gilmore et al., 2016).

[2] Note that some of the clusters (right intraparietal sulcus and the inferior temporal gyrus) had peak coordinates in the dorsal-attention network (according to Schaefer et al., 2018). Given that the remainder of the clusters (5 out of 7 clusters) fell into the Control Network we use that as the shorthand label.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507854; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

As an initial step to characterize the contribution of these different functional communities, we examined their evoked activation profiles during each task condition, using the averaged beta values estimated by an FIR model (**Figure S4**; see **Method: *Stimulus-Evoked and Residual Activity***). Consistent with network-level activity reported in past work, we found that DMN clusters deactivated (task-negative) during both perception and retrieval states, whereas Control network clusters activated (task-positive; Kim et al., 2015). Interestingly, the activation profile of RSC clusters differed between perception and retrieval states: task-positive during the retrieval state but task-negative during the two conditions in the perception state. Moreover, force-directed plots (**Figure 4B**) suggested that RSC was more closely connected with the Control network during retrieval (*Retrieve* condition) and with the DMN during perception (*Perceive* and *Scramble* conditions). Given the apparent state-specific traits of RSC, we performed a series of follow-up analyses.

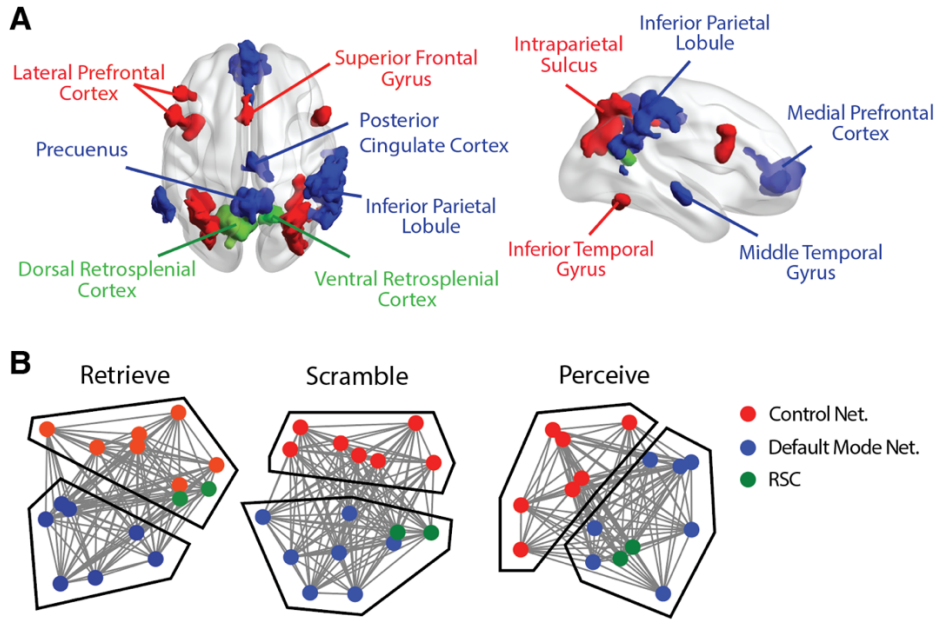


Figure 4. **A**) The 16 clusters of interest identified through the information mapping pipeline, partitioned into 3 functional communities. The color indicates the functional community assignment of each cluster. **B**) Force-directed graphs generated using the Fruchterman-Reingold algorithm implemented in NetworkX (V2.7.1; Fruchterman & Reingold, 1991). The projected physical distance in the graph indicates the degree to which nodes are being functionally connected. The color indicates the functional community allegiance of each cluster. Solid black lines represent manual delineations of the coupling pattern of RSC nodes.

Within- and Between-Community Background FC Discriminates Perception from Retrieval States

Our FCMA-then-clustering pipeline successfully identified a tractable number of connections among selected brain clusters that robustly differentiated perception from retrieval states. In the next set of analyses, we aimed to characterize the nature of the difference in background FC by comparing within- and between-community FC strength across these states. For each subject, we averaged across background FC matrices for each epoch within a condition, obtaining a 16 by 16 background FC matrix for each condition. We then averaged the *Perceive* and *Scramble* matrices to obtain a single measure of

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507854; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

background FC for perception state for each subject. Note that we expected to observe background FC differences because the clusters were selected because of their sensitivity to state-related changes; the goal of this analysis is to interpret this difference.

We first examined the background FC strength for all cluster pairs within the same functional community during perception and retrieval states. **Figure 5A** shows the group-level differences in cluster-to-cluster background FC strength across the two states (perception minus retrieval). Visual inspection revealed that most intra-Control network connections (17 out of 21 connections) were stronger during the perception state, whereas the majority of intra-DMN connections (17 out of 21 connections) were stronger during the retrieval state. Quantitively, a repeated-measures ANOVA with factors of cognitive state (perception vs. retrieval state) and functional community (DMN, Control network, and RSC) revealed a significant interaction ($F_{(2, 46)}$ = 19.05, p < 0.001, η^2 = 0.45; **Figure 5B**). This interaction was driven by the fact that the averaged background FC strength among RSC clusters ($t_{(23)}$ = 2.94, p = 0.007, 95% CI = [0.02, 0.1], Cohen's d = 0.49) and cluster pairs within the Control network ($t_{(23)}$ = 2.26, p = 0.03, 95% CI = [0.01, 0.06], Cohen's d = 0.42) was stronger during perception compared to retrieval state (**Figure 5A left**), whereas it was numerically stronger among clusters in DMN during retrieval than perception state ($t_{(23)}$ = 1.87, p = 0.07, 95% CI = [-0.06, 0.01], Cohen's d = 0.32; **Figure 5A right**). Additionally, the ANOVA revealed a significant main effect of functional community ($F_{(2, 46)}$ = 228.56, p < 0.001, η^2 = 0.91). Clusters within the Control network had overall stronger connectivity density compared to those within in the DMN ($t_{(23)}$ = 5.54, p < 0.001, 95% CI = [0.05, 0.1], Cohen's d = 0.93). Note that this result holds even after accounting for the anatomical distances between clusters (**Figure S5**). Greater coupling between RSC regions may have possibly been driven by the fact that the two RSC clusters were close to each other anatomically. Lastly, the ANOVA did not show a significant main effect of cognitive state ($F_{(1, 23)}$ = 2.06, p = 0.16, η^2 = 0.08), suggesting that the overall background FC densities were comparable across perception and retrieval.

To better understand across network connectivity, we next examined the background FC strength for cluster pairs in different functional communities across perception and retrieval states. Specifically, as a function of the two states, we assessed each set of between-network connections separately, both in an individual, cluster-wise manner as well as averaged across all between-network connections. Based on our results in **Figure 4B** suggesting relatively stable connectivity between Control and DMN and more flexible connectivity between RSC and the other networks, we present the results from Control and DMN separate from RSC. Certain clusters had significantly different connections with DMN as a function of the cognitive state (**Figure 5C**). For example, rPFC tended to couple with DMN clusters more strongly during the retrieval state, whereas SFG had stronger coupling with DMN clusters during the perception state (**Figure 5D**). Nevertheless, the averaged background FC strength of cluster pairs in Control network and DMN were comparable across the two cognitive states ($t_{(23)}$ = 1.38, p = 0.18, 95% CI = [-0.1, -0.02],

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507854; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Cohen's d = 0.39). RSC shifted from coupling with the Control network to DMN, as cognitive states shifted from retrieval to perception (**Figure 5E**). A repeated-measures ANOVA with factors of cognitive state (perception and retrieval) and functional community pair (i.e., the averaged connectivity measure between RSC nodes and regions in either Control or DMN regions) revealed a significant interaction ($F_{(1, 23)}$ = 100.94, $p < 0.0001$, $\eta^2 = 0.81$; **Figure 5F**). Specifically, RSC nodes had stronger averaged background connectivity with DMN nodes during the perception state ($t_{(23)} = 4.83$, $p < 0.001$, 95% CI = [0.05, 0.12], Cohen's d = 0.76), but stronger background connectivity with Control nodes during the retrieval state. ($t_{(23)}$ = 2.35, $p = 0.03$, 95% CI = [0.01, 0.09], Cohen's d = 0.43). The ANOVA did not show a main effect of cognitive state ($F_{(1, 23)}$ 1.22, $p = 0.28$, $\eta^2 = 0.05$) or network pair ($F_{(1, 23)} = 0.01$, $p = 0.91$, $\eta^2 < 0.01$).

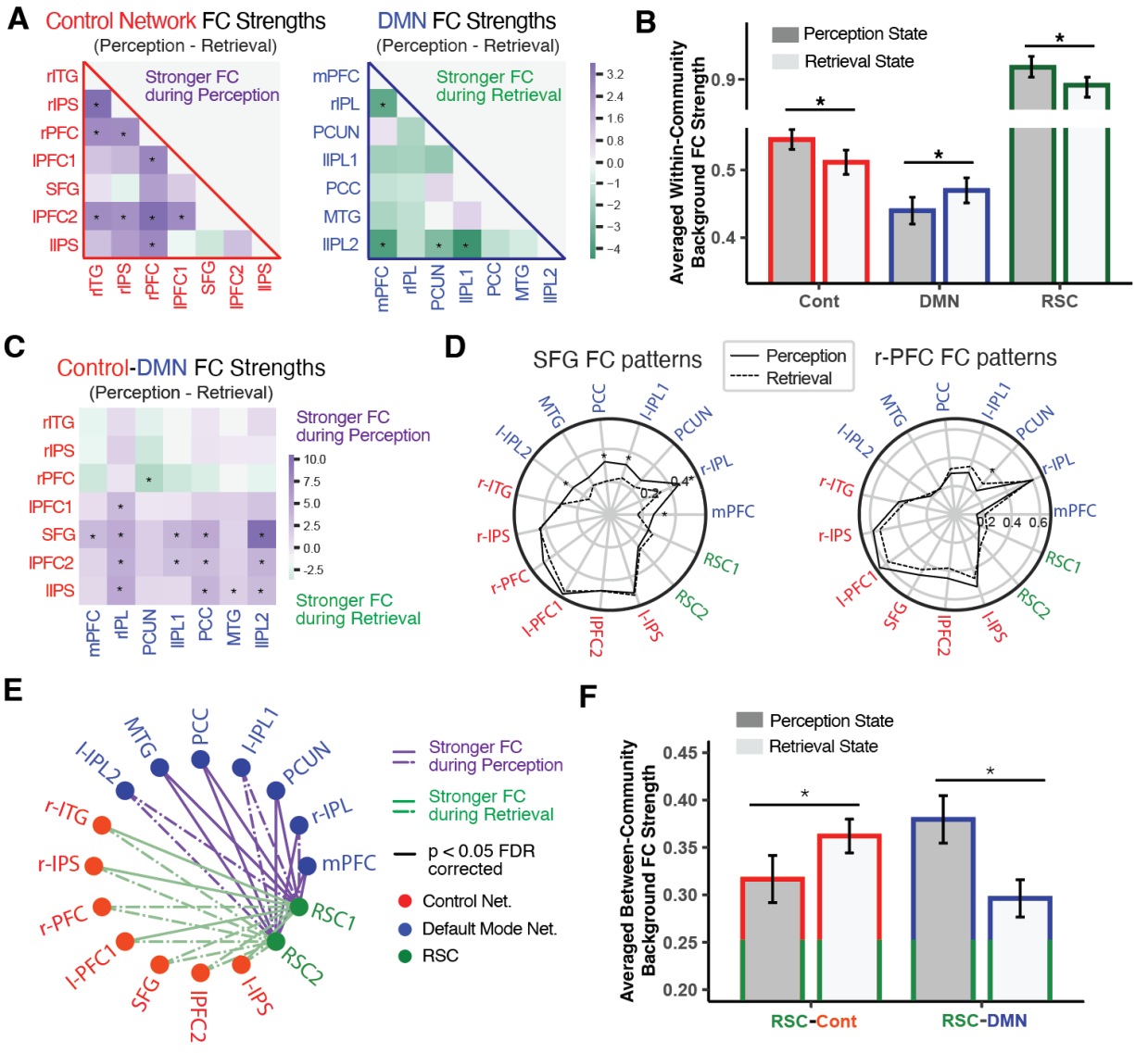


Figure 5. Connectivity configurations during perception and retrieval cognitive states. **A**) Group-level differences in background FC strength between each pair of clusters (within the same functional community) across perception and retrieval. The color of each cell represents the sign and magnitude of the *t* values, with a positive value indicating

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507854; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

stronger coupling during perception state and a negative value indicating stronger coupling during retrieval state. Asterisks indicate $p < 0.05$ after FDR correction. **B**) Background FC averaged across all pairwise connections within the same functional community across perception and retrieval states. The error bars indicate the standard error of the mean across all subjects. Asterisks indicate $p<0.05$ **C**) Group-level differences in background FC between pairs of clusters in different functional communities (DMN and Control network) across perception and retrieval states. The color of each cell represents the sign and magnitude of the *t* values and asterisks indicate $p < 0.05$ after FDR correction. **D**) Background FC patterns of superior frontal gyrus (SFG) and right prefrontal cortex (rPFC) during perception and retrieval states. Asterisks imply $p < 0.05$ after FDR correction. **E**) Group-level differences in background FC between RSC clusters and DMN/Control clusters across perception and retrieval states. The color of each line represents the sign of the *t* values, with purple indicating stronger coupling during the perception state and green suggesting stronger coupling during the retrieval state. Solid lines indicate $p < 0.05$ after FDR correction. **F**) Background FC strength averaged across all pairwise connections between RSC clusters and DMN/Control clusters. The error bars indicate the standard error of the mean across all subjects. Asterisks indicate $p<0.05$

Retrosplenial Cortex Plays Unique Role across Perception and Retrieval States

Although we have focused until now on differences between conditions related to cognitive state (i.e., perception vs. retrieval, **Figure 6A left**), performing the task required tracking two other forms of information: (1) the visual category[3] currently being presented (i.e., face vs. scene, **Figure 6A middle**) and (2) the behavioral judgement to perform (i.e., male/female vs. natural/manmade, **Figure 6A right**). Here, we performed pattern similarity analyses to assess how strongly these task components were represented in each cluster and functional community. Further, given the divergence in sensitivity between background FC and evoked activity patterns reported above, we performed these analyses using both neural metrics. For each cluster, we calculated pattern similarity between pairs of epochs (**Figure 6B**) and compared within-class similarity (e.g., for visual category: face-face/scene-scene) and between-class similarity (e.g., face-scene) to index the cluster's sensitivity to a given component. We then averaged across clusters within each functional community, yielding 3 sensitivity indices for each functional community per subject.

Pattern similarity measures of background FC suggested that clusters in different functional communities had different levels of sensitivity to cognitive state (**Figure 6C**; $F_{(2,46)} = 6.32$, $p = 0.003$, $\eta^2 = 0.22$). In particular, RSC sensitivity was significantly greater than the Control network ($t_{(23)} = 2.98$, $p = 0.006$, 95% CI = [0.01, 0.05], Cohen's d = 0.78) and numerically greater than the DMN ($t_{(23)} = 1.95$, $p = 0.06$, 95% CI = [-0.01, 0.04], Cohen's d = 0.49). Note that although clusters were initially selected because their background FC patterns differentiated between cognitive states, their sensitivity is biased to be greater. However, this bias should not necessarily extend to differences in sensitivity between communities. Pattern similarity of evoked activity indicated that clusters in different functional communities also showed different level of sensitivity to visual category (**Figure 6D**; $F_{(2,46)} = 48.35$, $p < 0.001$, $\eta^2 = 0.68$) and behavioral judgment (**Figure 6E**; $F_{(2,46)} = 26.92$, $p < 0.001$, $\eta^2 = 0.54$). Importantly, the evoked activity patterns of RSC were significantly more sensitive to these two components than the DMN (visual-

[3] Here we mean "visual category" to be the underlying stimuli used regardless of how scrambled they were. Although images were slightly scrambled, they still afforded well-off-chance M/F N/M judgements, making them dissimilar to traditional completely "scrambled" stimuli. Note that we also quantified visual-category sensitivity using only non-scrambled stimuli and the result patterns did not change.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507854; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

category: $t_{(23)}$ = 9.02, p < 0.001, 95% CI = [0.12, 0.19], Cohen's d = 2.26; behavioral-judgment: $t_{(23)}$= 5.70, p < 0.001, 95% CI = [0.04, 0.08], Cohen's d = 1.38) and the Control network (visual-category: $t_{(23)}$= 4.52, p < 0.001, 95% CI = [0.05, 0.13], Cohen's d = 1.21; behavioral-judgment: $t_{(23)}$= 5.00, p < 0.001, 95% CI = [0.03, 0.07], Cohen's d = 1.25). The one-way ANOVAs for all other communities across measures (activity or connectivity) failed to reach significance (ps > 0.25). Together, these pattern similarity results highlight how RSC might be involved in many critical aspects of retrieving versus perceiving information (cognitive state, the visual content, and the behavioral judgment). However, importantly, depending on whether the patterns are defined based on connectivity or activity, RSC can differentially capture the cognitive states of retrieval and perception relative to other regions.

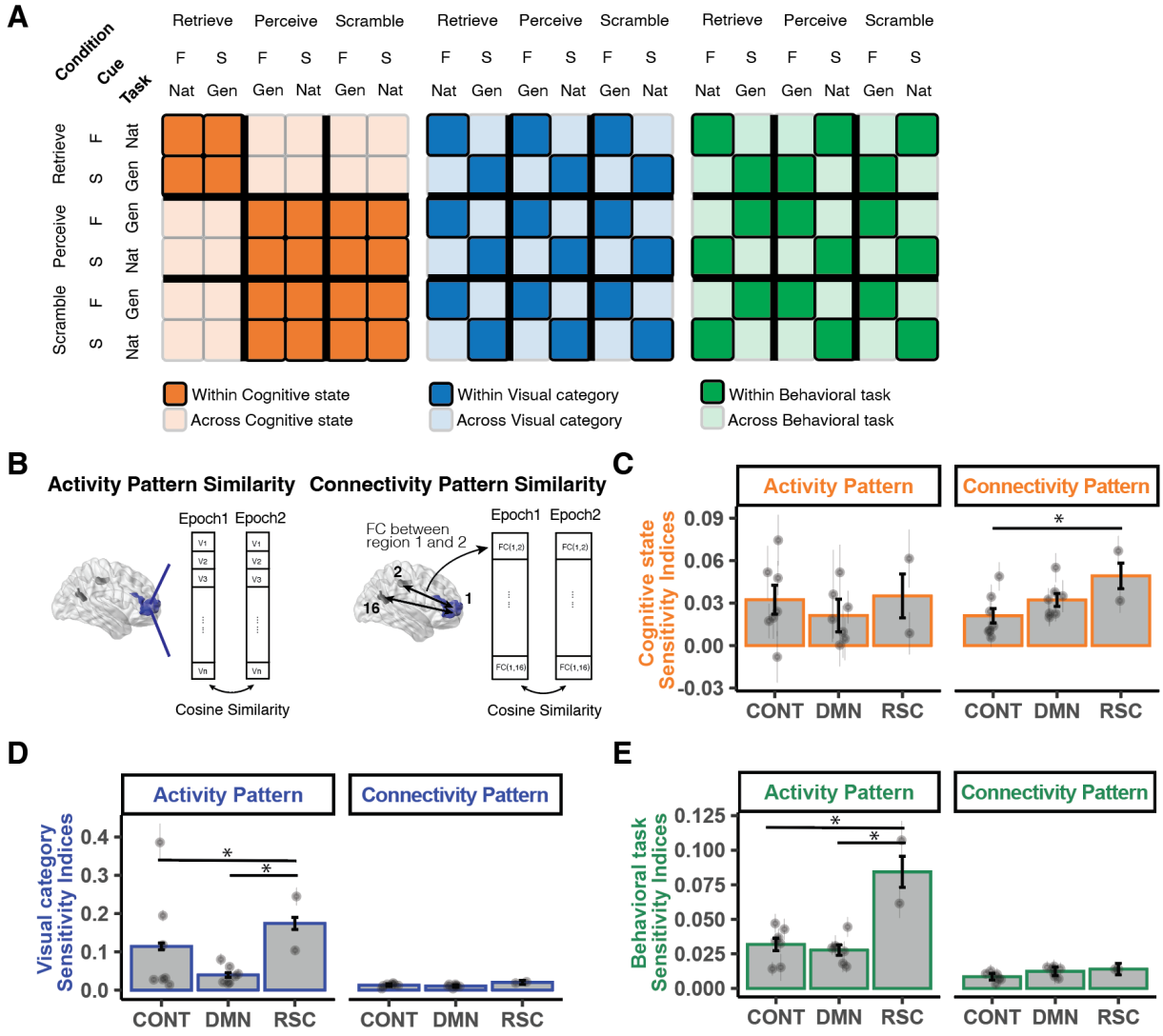


Figure 6. Pattern similarity analyses using both background FC patterns and stimulus-evoked activity patterns. **A**) Diagram for computing pattern similarity measures between pairs of epochs for each of the three task components (cognitive state, visual category, behavioral task). Darker colors indicate epochs of the same class (e.g., visual categories being face-face) whereas light colors indicate epochs of different classes (e.g., visual categories being face-scene). **B**) Diagram for computing pattern similarity measures with both FC and activity patterns. When computing activity pattern similarity measures, each epoch is represented by a *n*-dimensional vector, where *n* is the total number

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507854; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

of voxels in the respective cluster. When computing connectivity pattern similarity, each epoch is represented by a 15-dimensional vector, representing the background FC measure between this cluster and the other 15 clusters. **C**-**E**) Sensitivity indices of each functional community with regard to each of the three task component. Sensitivity indices were quantified as the average difference of within- vs. between-class epoch similarity. Each individual dot indicates the average across all subjects for a cluster and the error bars indicate the standard error of the mean across all subjects. Each bar represents the average across all subjects for a functional community and the error bars indicate the standard error of the mean across all clusters within the respective functional community. Asterisks indicate $p < 0.05$.

DISCUSSION

The goal of the current study was to characterize and differentiate perception versus retrieval states in a way that captures the complexity of whole-brain functional connectivity (FC). First, we found that patterns of background FC across perception versus retrieval states were systematically different from one another (**Figure 2A**). Moreover, the differences were best captured by background FC patterns between 16 clusters across 3 hypothesized functional communities (**Figure 4; Table S1**). Our whole-brain analysis pipeline allowed us to extend findings from previous research (Cooper & Ritchey, 2019) by identifying important brain clusters and coupling patterns beyond memory-related brain regions (**Figure 5**). Second, our results showed that background FC and evoked activity tend to capture distinct component processes (**Figure 3**), with the former being more “state-related” and the latter being more “stimulus-related” (Summerfield et al., 2006). Third, we demonstrated the utility of full correlation matrix analysis (FCMA; Kumar et al., 2022; Wang et al., 2015) and showed how the feature selection process of FCMA, paired with cluster-based dimensionality reduction, can be used to improve the interpretability of high-dimensional FC results (**Figure S1, S3**). We conclude by highlighting how the above findings are consistent with a framework of selective attention where, given the same perceptual input, directing attention externally would promote perception of the input-related sensory features, whereas directing attention internally would promote retrieval of the input-associated episodes (Chun et al., 2011).

Background Functional Connectivity Configurations Underlying Perception and Retrieval States

The current study had subjects use the same visual input as either the target of a perceptual judgement or the trigger for episodic memory. Our results suggest that in order to successfully perform tasks in these different conditions the brain produces distinct background FC configurations to maintain a sustained cognitive state for either perception or retrieval (**Figure 2A**). Similar state-related shifts in background FC configurations have been reported in previous work. Cooper and Ritchey (2019) found that pre-defined regions of interest in memory-related brain systems (anterior temporal and posterior medial networks; Ranganath & Ritchey, 2012) showed stronger background FC during retrieval over perception. Using a purely data-driven approach, a subset of our findings are largely consistent with Cooper and Ritchey (2019), as we found that the retrieval (vs. perception) state is characterized by stronger background coupling between clusters in the default mode network (DMN), which is important for internal cognitive processes (Buckner et al., 2008; Yeshurun et al., 2021). Specifically, our pipeline highlighted seven clusters in the DMN (blue clusters in **Figure 4A**; **Figure 5A, B**), including the bilateral inferior parietal lobule (IPL), posterior cingulate cortex (PCC), precuneus (PCUN), medial prefrontal cortex (mPFC) and

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507854; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

middle temporal gyrus (MTG). Note that IPL, PCUN, and PCC are considered parts of the posterior medial (PM) network for memory-guided behaviors (Ranganath & Ritchey, 2012), and IPL, PCC and mPFC have been shown to form a "core recollection network", supporting memory retrieval success (Rugg & Vilberg, 2013). Indeed, several studies have showed that functional interactions between these regions contribute to different aspects of episodic memory (Cooper & Ritchey, 2019; Geib et al., 2017; King et al., 2015). In particular, the interaction between the IPL and PCUN may connect episodic features to form an integrated neural representation, while conceptual knowledge and existing schemas are integrated by PCC and mPFC (Ranganath & Ritchey, 2012; Ritchey & Cooper, 2020).

Our findings further extend previous work by showing that the perception (vs. retrieval) state is characterized by increased background FC within a functional community with clusters in both the Control Network and the Dorsal Attention network (red clusters in **Figure 4A**; **Figure 5A, B**). Although there are situations where these networks might be engaged during memory-related processes (Hutchinson et al., 2014; Rosen et al., 2016, 2018), much evidence suggests their consistent involvement in processing the external world. That is, previous research has consistently implicated these regions as being a part of a larger task-positive network that is typically engaged when the brain processes external stimuli (Fornito et al., 2012; Golland et al., 2008). Together with the results concerning the DMN, our findings are consistent with the idea that the brain has two anatomically separable systems that primarily correspond to "externally oriented" versus "internally oriented" processing (Golland et al., 2008). Although our analyses did not implicate the hippocampus (see below), an interesting possibility is that the hippocampus may switch 'modes' of processing (externally- vs. internally-oriented) and consequently drive the network-level differences observed in cortex here (Duncan et al., 2012; Hasselmo et al., 1996; Honey et al., 2017).

Our whole-brain data-driven approach also revealed the importance of the background FC patterns in retrosplenial cortex (RSC; green clusters in **Figure 4A**) for accurately characterizing perception versus retrieval states. Specifically, we showed that RSC-DMN coupling was substantially greater during perception whereas RSC-Control coupling was greater during retrieval (**Figure 5E**, **F**). This connectivity pattern may seem counterintuitive at first, but interestingly it replicates findings of previous research. RSC is hypothesized to be part of the PM network , and its background FC patterns were also examined in Cooper and Ritchey (2019). Their results suggested that, although most ROIs in the PM network showed stronger coupling with each other during retrieval compared to perception, the background FC patterns of RSC did not demonstrate observable enhancement during retrieval state with any other regions in the PM network (c.f., Figure 3C right in Cooper & Ritchey, 2019). Instead of showing stronger coupling with PM regions during the retrieval state, they found that RSC had numerically stronger background FC with task-positive regions, such as the inferior temporal cortex, which is consistent with what we observed (**Figure 5E**). Benefiting from the whole-brain data-driven approach, the current study revealed a more complete

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507854; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

picture of RSC background FC during perception and retrieval states, identifying regions in the DMN and Control networks beyond classical memory-related systems (Cooper & Ritchey, 2019).

Although the findings in RSC are consistent with past observations, conceptual interpretation of its role here remains speculative. We do note that previous human fMRI and rodent studies suggest that RSC is involved in connecting external and internal states (Bicanski & Burgess, 2018; Yeshurun et al., 2021). For example, a study in mice found that RSC integrates both allocentric mapping (the animal's location in the external world) and egocentric frame (the animal's internal representation of the location) to navigate through a maze (Alexander & Nitz, 2015) by combining sensory inputs and mnemonic information from the medial temporal network (Bicanski & Burgess, 2018). Similarly, human RSC has been proposed to be a hub for connecting external and internal worlds (Yeshurun et al., 2021), such that it integrates external cues with self-generated information to guide behavior (Ranganath & Ritchey, 2012). In this respect, background FC in RSC during perception and retrieval may capture the role of RSC in bridging perceptual and mnemonic information; however, it is unclear why RSC would express higher coupling with the functional community putatively less involved with the task at hand (e.g., with DMN during the perception state). Accordingly, we believe an important direction of future research will be to fully characterize how the RSC assists in establishing perception and retrieval states.

Whereas some prior studies have identified changes in background connectivity within the hippocampus (Bein et al., 2020; Duncan et al., 2012; Honey et al., 2017) the hippocampus did not pop up using our whole brain approach. In particular, these studies have found that coupling between CA1 and CA3/DG is stronger during retrieval compared to perception states. By contrast, clusters identified in the current study did not highlight the importance of hippocampal subregions. We believe that this could be due to the fact that FCMA was configured to generate group-level voxel masks, ignoring individual differences in hippocampal subfield anatomy. The hippocampus is a relatively small brain region, and its subfields often include on the order of dozens of voxels (under typical scanning parameters) that may differ in their precise anatomical location across individuals. As a result, analyses of hippocampal subregions require subject-level hippocampal subfield segmentation, with each subject obtaining a slightly different segmentation map. In this respect, a shared group-level voxel mask may not be able to capture the important but subtle interactions among hippocampal subfields.

Background FC Captures "State-Related" Signals and Evoked Activity Reflect "Stimulus-Related" Signals

Different mental states can be reflected in distributed and overlapping patterns of evoked activity in the brain. An influential line of work has aimed to decode this information using the technique referred to as multivariate pattern analysis (MVPA) (Norman et al., 2006). At the same time, another line of complementary work has investigated the inter-regional connectivity structure of the brain by examining

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507854; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

the patterns of BOLD correlations (functional connectivity; FC) between multiple brain regions (Smith, 2012). Both approaches have been fruitful, resulting in tremendous insights into our understandings of human cognitive processes (Haxby, 2012; Song & Rosenberg, 2021). Importantly, some previous research has suggested that these two neural measures are likely to capture and reflect distinct, or at least non-overlapping aspects of cognitive processes. For example, Song et al. (2021) found that when viewing or listening to narratives, ongoing attentional engagement can only be predicted by FC-based models, whereas models trained with regional activity patterns failed to capture this information. Additionally, Manning et al. (2018) showed that an ensemble model that relied on both FC- and activity-based neural measures outperformed models that utilized either measure on its own. This result suggests that FC and activity patterns can capture partially non-overlapping variance in cognitive processes.

Extending these findings, the current study suggests that background FC-based measures are more sensitive to differences in cognitive states whereas activity-based measures are more likely reflect differences in stimulus-related features of task. Specifically, we found that background FC-based classifiers better separated task conditions that involved state-related (i.e., perception vs. retrieval) than those only involved stimulus-related differences (e.g., visual content). On the contrary, MVPA classifiers did not demonstrate a clear preference for state-related comparisons over stimulus-related distinctions (**Figure 3**). Our findings are in line with the theory that the fMRI data acquired at each voxel are composed of both state-related and stimulus-related activity, and that FC-based measures could be better suited to capture "state-related" signals whereas activity patterns better capture the "event-related" component (Summerfield et al., 2006). It is also worth noting that there are different types of FC-based measures; the current study primarily tested the background FC measures by regressing out the stimulus-evoked component using a general linear model (Al-Aidroos et al., 2012; Bejjanki et al., 2017; Norman-Haignere et al., 2012; Tompary et al., 2018). However, some studies computed their FC measures without regressing out the stimulus-evoked component (e.g., Song et al., 2021) or rely on stimulus-evoked parametric measures (e.g., beta series correlation; Bein et al., 2020). Future work might further examine whether these types of FC-based measures also preferentially capture state-related aspects of the cognitive process.

The Utility of Feature Selection in Whole-brain Voxel-wise FC Analyses

The current study used full correlation matrix analysis (FCMA; Kumar et al., 2022; Wang et al., 2015) to explore whether cognitive states are encoded in whole-brain voxel-wise background FC patterns. Our approach is systematically different from those used in previous studies, such as connectome-based predictive modeling (Shen et al., 2017) in two ways. First, FCMA operates on voxel-level connectivity matrices instead of the commonly-used, lower-dimensional parcel-averaged time series. Second, FCMA uses a nested leave-one-subject-out cross validation framework that enables more efficient feature selection to reduce large connectivity matrix to a tractable size. Here we discuss the potential trade-off

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507854; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

regarding these two distinctions. First, our results suggest that FCMA identified regions that, even after clustering over contiguous voxels, retained a higher sensitivity for differences in cognitive states compared to parcel-level analyses (**Figure S3B right**). This finding is consistent with previous research suggesting that voxel-/vertex-level FC patterns are more sensitive to other cognitive measures (concerning intelligence) compared to relatively coarse-grained parcel-level analysis (Feilong et al., 2021). The differences in sensitivity could be due to the fact that many parcellation schemes were defined from whole-brain resting-state functional connectivity profiles rather than task-based connectivity and previous work has suggested that the functional architecture of the brain might change across resting and task states (Cole et al., 2014) and may also vary across different task states (Krienen et al., 2014). It is worth mentioning that the use of predefined parcellation schemes can significantly improve sensitivity (as measured by effect size) in other types of analyses (e.g., univariate analyses) compared to voxel-/vertex-level analysis (Li et al., 2021). Thus, future studies need to further investigate the tradeoffs for predefined parcellation schemes in connectivity analyses.

The second way the current approach differs from prior approaches is in terms of interpretability. That is, machine learning models often face a trade-off between prediction accuracy and model interpretability (Feilong et al., 2021). In the context of FC-based models, previous work has primarily focused on constructing a model for making the most accurate predictions on trait-like demographic variation (Finn et al., 2015), behavioral performance (Rosenberg et al., 2016, 2020), or task conditions (Gonzalez-Castillo et al., 2015; Shirer et al., 2012). Despite high prediction accuracies in these models, it is often hard to interpret the FC configuration given the vast number of connections. For example, Rosenberg et al. (2016) identified a “high-attention network”, consisted of 757 edges across the entire brain, whose pattern of connectivity reliably predicted better performance on a sustained attention task. Indeed, such an approach likely captures many nuances in the neuronal distributed process of attention and has high predictive accuracy. However, it is possible that such high accuracy comes at the expense of interpretability. That is, constraining the number of edges and nodes in the descriptive network might sacrifice some degree of prediction accuracy, but the resulting FC configuration might be more easily interpreted. The current study attempted to do this by combining a whole-brain voxel-level FC model with feature selection using a nested cross-validation framework. Specifically, we quantified the “utility” of each connection using the machine learning training and testing framework (See **Method**: ***Full Correlation Matrix Analysis on Residual Activity***; **Figure S1**). As a result, we were able to select the most useful connections in an unbiased, automated fashion, reducing large correlation matrices to a tractable size. Remarkably, the model based on connections among only a set of 16 regions retained comparable AUC scores compared to models based on connections among 3000 voxels (Figure 2a; 0.83 vs. 0.87)[4]. Thus,

[4] Note that there was overlap between the feature selection process and the final testing process (both used all subjects) in the 16 cluster analysis. Thus, the performance of the 16 clusters is slightly biased towards higher performance.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507854; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

the current study demonstrates the value of using the FCMA voxel-to-cluster pipeline in order to yield the most relevant and interpretable FC configuration profile while largely maintaining prediction performance.

Ideas and Speculation: Perception Versus Retrieval as Outcomes of Selective Attention

Reconfiguration of background FC patterns has been examined as a neural mechanism underlying, or resulting from, selective attention (Desimone & Duncan, 1995). Specifically, the "switching-train-track" framework for selective attention argues that selective attention prioritizes goal-relevant information by strengthening the interaction (as measured by the changes in background FC) between goal-relevant brain regions (Al-Aidroos et al., 2012; Miller & Cohen, 2001; Turk-Browne, 2013). For example, previous work has shown that, when presented with face-scene composite stimuli, early visual cortex shows stronger intrinsic FC coupling with face-specialized regions (e.g., fusiform face area) when subjects attended to the face component but with scene-specialized regions (e.g., parahippocampal place area) when subjects attended to the scene component (Al-Aidroos et al., 2012; Córdova et al., 2016; Norman-Haignere et al., 2012; Tompary et al., 2018). Here we speculate that the difference between perception and retrieval could be understood and investigated as an key dimension of selective attention. Specifically, a recent taxonomy dichotomized attention into external attention, with the target being perceptual information and internal attention, with the target being self-generated trains of thoughts (Chun et al., 2011). Under this dichotomy, perception versus retrieval can be thought of as a form of external vs. internal attention states for information processing. When a given input is being attended externally, the sensory information of the input is being preferentially processed while the mnemonic information is suppressed. On the other hand, when the same input cues internally-directed attention, the mnemonic information will be enhanced while the sensory information is suppressed (Chun et al., 2011; Gazzaley & Nobre, 2012). The switching-train-track account of selective attention has been mostly studied by examining FC changes within the ventral visual stream when subjects were attending externally to different visual categories (e.g., face vs. scene). The current study has provided evidence that this account of selective attention may generalize to the whole brain and to more complex cognitive states (e.g., perception vs. retrieval). In particular, although previous research studying perception vs. retrieval has focused primarily on decoding the "train of thoughts" (i.e., stimulus-related features; Rissman & Wagner, 2012), the selective attention perspective highlights the importance of investigating state-specific "train tracks"—i.e., background FC patterns—and understanding how such train tracks establish and sustain particular cognitive states.

METHODS

Subjects

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507854; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Twenty-seven adults with normal or corrected-to-normal vision were recruited to participate for monetary compensation at Princeton University. Three subjects were excluded due to excessive head motion for a total of 24 subjects in the current sample (eleven reported male, mean age = 23.3 years). The Princeton University Institutional Review Board approved the study protocol, and all subjects provided informed consent. The sample size is on par with previous studies that examined functional connectivity changes with memory (Cooper & Ritchey, 2019) and used FC patterns to differentiate task states (Shirer et al., 2012).

Materials

Stimuli consisted of 64 scene and 64 face images. The scene images were collected from the "Massive Memory" dataset (Konkle et al., 2010; http://konklab.fas.harvard.edu/#). The face images were obtained from the FEI face database (Thomaz & Giraldi, 2010; https://fei.edu.br/~cet/facedatabase.html) and contained emotionally neural expressions. Scrambled images were generated as the weighted average between the actual image and its phase-scrambled version (Oppenheim & Lim, 1981; Stojanoski & Cusack, 2014). The script for creating the phase-scrambled images was adopted from code by Nicolaas Prins (e.g., as described here: https://github.com/rordenlab/spmScripts/blob/master/bmp_scramble.m).

Procedure

During the training phase (~20-30 minutes prior to scanning), subjects viewed randomly assigned pairs of images (one face and one scene) and indicated how successfully they were in forming a mental association between the images (form-association-task). After viewing all stimuli pairs, they were given a 2-alternative force choice task (AFC) in which a face would be presented above two scenes (or vice versa) and the subjects had to indicate which of the bottom images had been paired with the top image before. The cycle of association-forming and 2-AFC task continued until subjects got each association correct two times in a row. While undergoing fMRI scanning, subjects completed 3 task conditions across 6 functional runs using a block design (2 runs for each condition). Each run started with a 6-s blank lead-in period, followed by 8 task epochs and ended with a 6-s lead-out period. Each epoch consisted of a 4-s presentation of instructions, followed by 8 2-s presentations of single images presented at central fixation separated by a 1-s interstimulus interval. Each sequence of stimuli was followed by a 12-s inter-block-interval. Together, the duration of each trial, epoch, and run was 3 s, 40 s and 332 s, respectively. Note that all epochs within a functional run are of the same condition, and that all visual stimuli of an epoch are of the same category (i.e., either all faces or all scenes).

In the *Perceive* task condition, subjects were asked to identify the visual features of each cue on the screen. That is, if a face cue was presented, subjects were instructed to make a male/female judgment of the face via a button box, whereas if a scene cue was shown, subjects were instructed to make natural/man-made judgment of the scene. On the other hand, during the *Retrieve* task condition, subjects

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507854; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

were asked to judge the gender or naturalness of the cue-associated image (i.e., the pair mate of the cue from the training phase). For example, if a face cue was shown, subjects needed to retrieve the specific associated scene image (not presented) and make a natural/man-made decision on that remembered information. Likewise, if a scene cue was shown, subjects were supposed to retrieve the specific associated face image and make a male/female decision on that remembered information. To ensure that the neural correlates we later identified were not driven solely by differences in task difficulty (as measured by accuracy) between *Perceive* and *Retrieve* task conditions, we included a *Scramble* condition. During the *Scramble* condition, subjects completed the same task as they did in the *Perceive* condition, but the visual cues were scrambled using a weighted average of the original image and its phase-scrambled version. The weight of this average was established based on a behavioral pilot study using the same design which sought to vary the weight until accuracy between the *Scramble* and *Retrieve* task conditions was matched (**Figure 1**).

Image Acquisition and Preprocessing

The fMRI data were acquired with a 3T scanner (Siemens Prisma) at the Princeton Neuroscience Institute. Functional data were acquired using a T2*-weighted multiband EPI sequence (repetition time = 1 s, echo time = 26 ms, flip angle = 50°, FOV = 260 x 260, resolution = 2.5 x 2.5 x 2.5 mm, multiband acceleration factor = 4) with 44 axial slices aligned to the anterior commissure/posterior commissure. A whole-brain T1-weighted MPRAGE 3D anatomical volume (1 x 1 x 1 mm voxels) was collected to improve registration. One phase and two magnitude field maps were collected to correct field inhomogeneities.

The first 6 lead-in volumes of each functional run were manually discarded before entering the preprocessing pipeline. Image preprocessing was performed using fMRIPrep 20.1.0rc1 (Esteban et al., 2019). All functional images were corrected for slice-acquisition time, head motion, and susceptibility distortion, and were normalized to a standard template, yielding preprocessed BOLD runs in MNI152NLin2009cAsym space. Following use of fMRIPrep, the minimally preprocessed functional runs were further processed using FSL (Woolrich et al., 2001) with a Nipype implementation (Gorgolewski et al., 2011). All functional images were smoothed with a 5.0 mm FWHM Gaussian kernel and high-pass filtered at 0.01 Hz. For each subject, the intensity values in each voxel in each of the 6 functional runs were then normalized using the mean and standard deviation of the resting period. The resting period for each functional run was defined as the 6 lead-out volumes plus all 12-s inter-block intervals of that run, which were shifted for 4 TRs to account for the hemodynamic delay. This normalization is intended to remove the BOLD signal differences across runs and thus all 6 runs were then concatenated to one time series and used for further modeling.

Stimulus-Evoked and Residual Activity

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507854; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

We computed residual activity by regressing out the stimulus-evoked component from the preprocessed data in order mitigate stimulus-evoked coactivation confounds (Al-Aidroos et al., 2012; Cole et al., 2019). First, we constructed a confound regression model using FSL (implemented in NiPype) in order to minimize the effect of the following confound variables (obtained from fMRIPrep): six head motion parameters and the mean time series from white matter and cerebrospinal fluid. The resulting timeseries from regressing out the confound regressors are referred to as the stimulus-evoked activity timeseries in all subsequent analyses, as they are fully preprocessed, yet contain the stimulus-evoked components. Second, we estimated and removed the stimulus-driven components from the stimulus-evoked timeseries using a finite impulse response (FIR) model, which modeled the first 36 TRs for every epoch separately for face and scene epochs, resulting in 36 (TR) x 2 (epoch category) x 3 (condition) = 216 regressors. FIR is believed to be the optimal GLM for removing stimulus-evoked response because it does not assume the shape of the hemodynamic response function (Cole et al., 2019; Norman-Haignere et al., 2012). The residual timeseries data are referred to as the residual activity and used for computing background FC for all subsequent analyses.

Full Correlation Matrix Analysis on Residual Activity

We utilized full correlation matrix analysis (FCMA) as implemented in the Brain Imaging Analysis Kit (BrainIAK; version 0.11; http://brainiak.org) to conduct an unbiased, whole-brain voxel-wise FC analysis that systematically considers all pairwise correlations in the brain to explore the differences in connectivity configurations between perception and retrieval states. All FCMA jobs were executed on Talapas, the HPC cluster at University of Oregon (https://racs.uoregon.edu/talapas). Each FCMA inner loop job was spread across 4 nodes, with each node supporting 28 threads and a job took around 4 hours to complete. Each node was equipped with two intel E5-2690v4 processors, with 128 GB of memory. FCMA took in the residual activity and computed a full correlation matrix (i.e., whole-brain voxel-wise correlation matrix; 92745 voxels x 92745 voxels) for each task epoch. Therefore, for each subject, 8 (epoch) x 2 (run) = 16 full correlation matrices were computed per task condition (i.e., *Perceive*, *Retrieve* and *Scramble*; **Figure 1 c-e**). Using these full correlation matrices, FCMA aimed to (i) examine whether perception (*Perceive* and *Scramble*) versus retrieval (*Retrieve*) states can be successfully decoded from background connectivity patterns, and (ii) identify the connectivity configuration that characterizes each cognitive task state. FCMA achieved these goals by implementing a nested leave-one-subject-out cross-validation (LOOCV) framework. Specifically, the outer loop contains 23 training subject and 1 left-out test subject for each outer-loop iteration; and the inner-loop uses 22 training subjects and 1 left-out test subject within the outer training set for each inner-loop iteration. Importantly, the inner-loop intends to select the top k most useful voxels (based on their connectivity patterns) from the training data (**Figure S1**) and the outer loop aims to train classifiers on connectivity patterns of the selected voxels and test their ability to predict left-out data (**Figure 1G, H**). Note that we also performed parcel-level analysis to demonstrate the sensitivity advantage of our more fine-grained approach. We utilized the MNI version of the Schaefer parcellation

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507854; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

scheme (https://github.com/ThomasYeoLab/CBIG/tree/master/stable_projects/brain_parcellation/Schaefer2018_LocalGlobal/Parcellations/MNI; Schaefer et al., 2018), and performed analyses on parcellations of different granularity (400 and 1000 parcels), which yielded comparable patterns of results as one another.

For each iteration, the inner-loop worked with the training data from 23 subjects (N-1 subjects) and performed a separate, nested LOOCV on the training set. In an unbiased manner, the inner-loop tested the accuracy of using each voxel's seed maps (i.e., how much the voxel is connected to all other voxels in the brain) to differentiate perception (i.e., *Perceive* and *Scramble*) from retrieval (i.e., *Retrieve*) state (**Figure S1**). Specifically, each voxel would get an accuracy score for separating *Retrieve* and *Perceive* FC patterns, and another score for separating *Retrieve* and *Scramble* FC patterns. The minimum score of the two was assigned to the voxel and all assigned scores were averaged across the inner-loop LOOCV iterations (i.e., training using 22 subjects and testing on a left out subject). The minimum score was used here to make sure one of the two comparisons does not drive the overall effect and thus voxel selection prioritized discovery of regions which were sensitive to both comparisons, making the influence of single confounds less likely. For example, a voxel's seed map might be able to separate *Retrieve* from *Scramble* well by picking up the visual content information (i.e., scramble vs. intact), but not separate *Retrieve* from *Perceive* well when the visual content difference was absent. Computing the minimum accuracy scores allowed us to more accurately measure how well each voxel's seed map encoded cognitive task state differences.

The resulting scores indicated the ability of each voxel to differentiate perception from retrieval states independent of the testing data (i.e. the left out subject in the outer loop). Based on these scores, masks of the top *k* most useful voxels were created from each inner-loop and the whole-brain full correlation matrices were reduced to *k* x *k* correlation matrices. The current study tested all results with *k* = 100, 1000, 3000, 5000, 7000, 10000 and 15000 voxels, and reported the results on *k* = 3000. We reported our main results using the top 3000 voxels for two reasons. First, model performances dramatically improved as the voxel masks were enlarged from *k* = 100 to *k* = 3000. However, model performance seemed to plateau when the top 3000 voxels were selected. Second, our information mapping pipeline (See **Methods: *Information Mapping***) identified a shared mask that included about 3500 voxels. The results from k = 3000 therefore provides the most direct comparison and is the most representative for subsequent analyses. Using these *k* x *k* connectivity matrices, the FCMA outer-loop trained 3 classifiers (one for each pair of task conditions, e.g., *Perceive* vs. *Retrieve*) and tested the model performance using the left-out testing data.

Regular and Generalization Tests

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507854; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Classifiers were first trained and tested on the same task condition comparison. To quantify and compare model performance for each task comparison, we computed the area under the receiver operating characteristic curve (AUC) for each classifier. One way ANOVA was used to compare the AUCs for the three classifiers. The goals of the regular tests were to examine (i) whether task conditions could be successfully decoded from background connectivity patterns and (ii) whether task condition comparisons that involve cognitive task state differences (e.g., *Retrieve* vs. *Perceive*) could be decoded with higher accuracies relative to the comparison that did not involve such differences (e.g., *Scramble* vs. *Perceive*).

During the generalization test, a classifier was trained on one task condition comparison (e.g., *Perceive* vs. *Retrieve*) but tested on a different comparison (e.g., *Perceive* vs. *Scramble*). The generalization test can be conducted between each pair of task condition comparisons; we argue that the bidirectionally averaged generalization AUCs should indicate the classifiers' sensitivity to a certain dimension of task differences (**Figure 2D**). For example, both the *Retrieve* vs. *Perceive* and *Retrieve* vs. *Scramble* comparisons involve cognitive task state differences. As a result, high bidirectionally averaged AUC scores across classifiers trained on *Retrieve* vs. *Perceive* and tested on *Retrieve* vs. *Scramble* and vice versa would indicate that the classifiers were indeed trained to pick up on cognitive task state information. On the other hand, high bidirectionally averaged AUC score across *Retrieve* vs. *Perceive* and *Scramble* vs. *Perceive* comparisons would suggest that the classifiers were trained to pick up on task difficulty differences. Likewise, high bidirectionally averaged AUC scores across *Scramble* vs. *Perceive* and *Scramble* vs. *Retrieve* comparisons would suggest that the classifiers were trained to detect visual features. Thus, the averaged bidirectional generalization accuracies served to measure the degree to which differences in cognitive task states, task accuracies, and visual content drove classifier performance. The primary goal of the generalization test, then, was to examine whether classifiers were able to selectively detect connectivity patterns underlying perception versus retrieval task state differences.

Multivoxel Pattern Classification Analyses

Pattern classification analyses were performed using a support vector classifier (C = 1) implemented in the Scikit-learn module in Python (Pedregosa et al., 2011). We performed pattern classification using both the residual and the stimulus-evoked activities (See section: ***Task Evoked and Residual Activity***). First, using the *k* voxel masks generated by FCMA inner-loop (See section: ***Full Correlation Matrix Analysis on Residual Activity***), we trained pattern classifiers using residual activity to perform the regular tests (i.e, train and test a model with the same task condition comparison; **Figure S2**). This analysis aims to test whether FCMA classifier performances were driven by any coactivation confounds left in the residual activity. Second, with each set of *k* voxel masks, we also trained pattern classifiers using stimulus-evoked activity to perform the regular test (**Figure 2B middle**). This analysis aimed to examine whether simple MVPA and background FC patterns rely on the same type of cognitive processes.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507854; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Information Mapping

Due to the leave-one-out cross-validation scheme, each fold of the outer-loop produced non-identical sets of top voxels. To obtain a common mask, we combined the FCMA inner-loop with non-parametric permutation tests to generate a group-level mask that included all voxels that differentiated perception from retrieval states better than chance. We obtained this group-level mask in three steps. First, we averaged the classification accuracy composite score (i.e., the minimum of the accuracy scores for separating *Retrieve* vs. *Perceive* and *Retrieve* vs. *Scramble* FC patterns) for each voxel across the 24 voxel maps (**Figure S1**; one for each cross-validation fold) and used this as the observed score for each voxel. Second, we shuffled the labels of all three task conditions for each subject and repeated the outer-loop training-testing pipeline using the shuffled labels for 100 iterations. As a result, we generated a null distribution of the classification accuracy composite scores for each voxel. Using the mean and standard deviation of the classification accuracy score null distribution for each voxel, we z-transformed the observed feature selection score computed above (i.e., the group-averaged composite accuracy score). We chose a relatively stringent voxel-wise primary threshold ($P < 0.0001$) to control for false positive rates and to avoid finding large voxel clusters that span the brain (Woo et al., 2014). Voxels that passed the threshold were divided into clusters using 3dClusterize and the corresponding cluster-extent threshold was computed using 3dClustSim with AFNI (whole-volume alpha-values: -athr < 0.01). In total, 62 clusters passed the cluster-extent threshold, with size ranging from 620 to 2 voxels. In the last phase of the information mapping, we selected clusters according to their size and whether they improved model performance when combined with all larger clusters. Specifically, we began with the largest cluster (620 voxels), then we sequentially added voxels from the next largest cluster to our mask and used the FCMA outer loop to test how well this set of voxels differentiated perception from retrieval states. In other words, we iteratively built a set of models, and measured the change in classification performance as the mask accumulated each new cluster of voxels (**Figure S3A**). The rationale for performing cluster selection is to identify the smallest number of "sufficient" nodes whose network dynamics capture the differences between perception and retrieval states and ultimately reduce the full correlation matrix of the brain to a tractable set of representative ROIs.

Community Detection

We averaged across all voxels within each cluster to obtain cluster-level time series data for each epoch. Cluster-level background functional connectivity (FC) matrices were then computed for each epoch for every subject, and then averaged within each task condition, resulting in one group-averaged connectivity matrix for the *Perceive*, *Scramble,* and *Retrieve* conditions respectively. Community detection on weighted graphs (i.e., the connectivity matrices calculated above) was performed using the Louvain algorithm via NetworkX in Python. Following the approach used in prior work (Barnett et al., 2021; Ji et al., 2019), we ran the algorithm 1000 times on the weighted graphs to tune the resolution parameter in

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507854; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

order to maximize modularity, which in turn captures how well a network can be subdivided into non-overlapping groups (Rubinov & Sporns, 2011). The tuned resolution parameters (gamma) were 1.10, 1.13, 1.19, for Perceive, Retrieve and Scramble task conditions, respectively. To visualize intra- and intramodular connectivity, we used the Fruchterman-Reingold force-directed projection implemented in NetworkX to project the graphs onto 2D spaces.

Within- and Between-Communities Connectivity

We computed a subject-level FC matrix for each task condition. We averaged the FC matrices for *Perceive* and *Scramble* conditions to create composite matrices that reflect FC dynamics underlying the perception state. To examine the coupling pattern across communities, we averaged across the connectivity strength between nodes within the same functional community using both perception and retrieval state FC matrices. A paired sample t-test was used to examine whether DMN-Control connectivity strength differ between perception and retrieval states; and a two-way repeated measures ANOVA was then used to examine whether RSC nodes changed their coupling pattern with respect to the DMN and Control network nodes. To examine within-community connectivity strength, we averaged the connectivity strength between all nodes within the DMN, Control, and RSC regions for both perception and retrieval states. Similarly, a two-way repeated measures ANOVA was used to assess whether a functional community was biased toward a certain cognitive task state. All statistical analyses were performed using Pingouin 0.5.1 with Python3.

Pattern Similarity Analyses on Stimulus-Evoked Activity

The goal of this set of analyses was to identify any potential differences between the three functional communities in their roles in performing cognitive tasks. We focused on three different cognitive aspects and tested the degree to which each functional community was sensitive to (i) the current visual categories (i.e., face vs. scene), (ii) the current behavioral tasks (i.e., gender vs. naturalness judgement), and (iii) current cognitive state (i.e., perception or retrieval state). To do that, we performed pattern similarity analyses using both stimulus-evoked activity patterns and background connectivity patterns.

To obtain the stimulus-evoked activity patterns, we extracted the 24 task TRs for each epoch (after being shifted 4 s to account for hemodynamic delay) from the post-confound regression time-series. These stimulus-evoked estimates were then averaged along the time dimension and reshaped into a vector for each cluster (length of the vector is the number of voxels in that cluster). On the other hand, background connectivity patterns for each cluster were computed as the temporal correlation between this cluster and other 15 clusters, reshaped into a 15-dimensional vector for each ROI (**Figure 6B**). To compute pattern similarity measures, the Fisher's Z transformed correlations between each pair of vectors from different functional runs were calculated (Kriegeskorte et al., 2009; **Figure 6A**). Sensitivity was quantified as the difference between within-state epoch pattern similarity and between-state epoch pattern similarity. For

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507854; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

example, when examining sensitivities for the current task judgement, pattern similarities were computed among all epochs with the same judgement (i.e., gender to gender and naturalness to naturalness) and compared to those with different judgments (i.e., gender to naturalness). Sensitivity was calculated by using the average within-state pattern similarity score minus the average between-states pattern similarity score. Thus, a significantly positive sensitivity index would suggest that the given ROI produced relatively distinct activity and/or connectivity patterns for the two aspects of the task (state, content, or judgement). We averaged the sensitivity scores across ROIs within the same functional communities. A one-way ANOVA was used to compare sensitivity scores of each cognitive process across the three functional communities.

Data and Code Availability

Processed fMRI data including both the stimulus-evoked and residual time series supporting the primary findings of this study are available on the Open Science Framework (OSF) at https://osf.io/yfwc7/. Scripts for performing and reproducing the specific analyses described in this paper can be found through Github at https://github.com/peetal/Decode_AttenStates.

Funding and Acknowledgement

This work was supported by the Intel Corporation, the John Templeton Foundation, and NIH R01 EY021755. We express our deepest thanks for the feedback we received from Dr. Ken Norman and Dr. Sam Nastase. We are also grateful for helpful discussions and support from members of the Hutchinson Lab of Cognitive Neuroscience, the Dubrow Lab, and the Kuhl Lab.

Credit Authorship Contribution Statement

Study conception and design: Nicholas B. Turk-Browne, J. Benjamin Hutchinson

Data Collection: J. Benjamin Hutchinson

Analysis and interpretation of results: Y. Peeta Li, Yida Wang, Brice A. Kuhl, J. Benjamin Hutchinson

Draft manuscript preparation: Y. Peeta Li, Brice A. Kuhl, J. Benjamin Hutchinson

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507854; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

REFERENCE

Al-Aidroos, N., Said, C. P., & Turk-Browne, N. B. (2012). Top-down attention switches coupling between low-level and high-level areas of human visual cortex. *Proceedings of the National Academy of Sciences*, *109*(36), 14675–14680. https://doi.org/10.1073/pnas.1202095109

Alexander, A. S., & Nitz, D. A. (2015). Retrosplenial cortex maps the conjunction of internal and external spaces. *Nature Neuroscience*, *18*(8), 1143–1151. https://doi.org/10.1038/nn.4058

Barnett, A. J., Reilly, W., Dimsdale-Zucker, H. R., Mizrak, E., Reagh, Z., & Ranganath, C. (2021). Intrinsic connectivity reveals functionally distinct cortico-hippocampal networks in the human brain. *PLoS Biology*, *19*(6), e3001275–e3001275. https://doi.org/10.1371/journal.pbio.3001275

Bein, O., Duncan, K., & Davachi, L. (2020). Mnemonic prediction errors bias hippocampal states. *Nature Communications*, *11*(1), 3451. https://doi.org/10.1038/s41467-020-17287-1

Bejjanki, V. R., Silveira, R. A. da, Cohen, J. D., & Turk-Browne, N. B. (2017). Noise correlations in the human brain and their impact on pattern classification. *PLOS Computational Biology*, *13*(8), e1005674. https://doi.org/10.1371/journal.pcbi.1005674

Bicanski, A., & Burgess, N. (2018). A neural-level model of spatial memory and imagery. *ELife*, *7*, e33752. https://doi.org/10.7554/eLife.33752

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, *2008*(10), P10008. https://doi.org/10.1088/1742-5468/2008/10/P10008

Bosch, S. E., Jehee, J. F. M., Fernández, G., & Doeller, C. F. (2014). Reinstatement of associative memories in early visual cortex is signaled by the hippocampus. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *34*(22), 7493–7500. https://doi.org/10.1523/JNEUROSCI.0805-14.2014

Braun, U., Schäfer, A., Walter, H., Erk, S., Romanczuk-Seiferth, N., Haddad, L., Schweiger, J. I., Grimm, O., Heinz, A., Tost, H., Meyer-Lindenberg, A., & Bassett, D. S. (2015). Dynamic reconfiguration of frontal brain networks during executive cognition in humans. *Proceedings of the National Academy of Sciences*, *112*(37), 11678–11683.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507854; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network. *Annals of the New York Academy of Sciences*, *1124*(1), 1–38. https://doi.org/10.1196/annals.1440.011

Chun, M. M., Golomb, J. D., & Turk-Browne, N. B. (2011). A taxonomy of external and internal attention. *Annual Review of Psychology*, *62*(1), 73–101. https://doi.org/10.1146/annurev.psych.093008.100427

Chun, M. M., & Johnson, M. K. (2011). Memory: Enduring traces of perceptual and reflective attention. *Neuron*, *72*(4), 520–535. https://doi.org/10.1016/j.neuron.2011.10.026

Cohen, J. R., & D'Esposito, M. (2016). The Segregation and Integration of Distinct Brain Networks and Their Relationship to Cognition. *Journal of Neuroscience*, *36*(48), 12083–12094. https://doi.org/10.1523/JNEUROSCI.2965-15.2016

Cole, M. W., Bassett, D. S., Power, J. D., Braver, T. S., & Petersen, S. E. (2014). Intrinsic and task-evoked network architectures of the human brain. *Neuron*, *83*(1), 238–251. https://doi.org/10.1016/j.neuron.2014.05.014

Cole, M. W., Ito, T., Schultz, D., Mill, R., Chen, R., & Cocuzza, C. (2019). Task activations produce spurious but systematic inflation of task functional connectivity estimates. *NeuroImage*, *189*, 1–18. https://doi.org/10.1016/j.neuroimage.2018.12.054

Cooper, R. A., & Ritchey, M. (2019). Cortico-hippocampal network connections support the multidimensional quality of episodic memory. *ELife*, *8*, e45591. https://doi.org/10.7554/eLife.45591

Córdova, N. I., Tompary, A., & Turk-Browne, N. B. (2016). Attentional modulation of background connectivity between ventral visual cortex and the medial temporal lobe. *Neurobiology of Learning and Memory*, *134*, 115–122. https://doi.org/10.1016/j.nlm.2016.06.011

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193–222. https://doi.org/10.1146/annurev.ne.18.030195.001205

Duncan, K. D., & Shohamy, D. (2016). Memory states influence value-based decisions. *Journal of Experimental Psychology: General*, *145*(11), 1420–1426. https://doi.org/10.1037/xge0000231

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507854; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Duncan, K., Sadanand, A., & Davachi, L. (2012). Memory's penumbra: Episodic memory decisions induce lingering mnemonic biases. *Science*, *337*(6093), 485–487. https://doi.org/10.1126/science.1221936

Duncan, K., Tompary, A., & Davachi, L. (2014). Associative encoding and retrieval are predicted by functional connectivity in distinct hippocampal area ca1 pathways. *Journal of Neuroscience*, *34*(34), 11188–11198. https://doi.org/10.1523/JNEUROSCI.0521-14.2014

Esteban, O., Markiewicz, C. J., Blair, R. W., Moodie, C. A., Isik, A. I., Erramuzpe, A., Kent, J. D., Goncalves, M., DuPre, E., Snyder, M., Oya, H., Ghosh, S. S., Wright, J., Durnez, J., Poldrack, R. A., & Gorgolewski, K. J. (2019). fMRIPrep: A robust preprocessing pipeline for functional MRI. *Nature Methods*, *16*(1), 111–116. https://doi.org/10.1038/s41592-018-0235-4

Ezzyat, Y., Kragel, J. E., Burke, J. F., Levy, D. F., Lyalenko, A., Wanda, P., O'Sullivan, L., Hurley, K. B., Busygin, S., Pedisich, I., Sperling, M. R., Worrell, G. A., Kucewicz, M. T., Davis, K. A., Lucas, T. H., Inman, C. S., Lega, B. C., Jobst, B. C., Sheth, S. A., … Kahana, M. J. (2017). Direct brain stimulation modulates encoding states and memory performance in humans. *Current Biology*, *27*(9), 1251–1258. https://doi.org/10.1016/j.cub.2017.03.028

Favila, S. E., Lee, H., & Kuhl, B. A. (2020). Transforming the Concept of Memory Reactivation. *Trends in Neurosciences*, *43*(12), 939–950. https://doi.org/10.1016/j.tins.2020.09.006

Feilong, M., Guntupalli, J. S., & Haxby, J. V. (2021). The neural basis of intelligence in fine-grained cortical topographies. *ELife*, *10*, e64058. https://doi.org/10.7554/eLife.64058

Fernandez, C., Madore, K. P., & Wagner, A. D. (2022). *Encoding and the medial temporal lobe. to appear in: Oxford handbook of human memory (m. j. kahana & a. d. wagner, eds). oxford university press.* PsyArXiv. https://doi.org/10.31234/osf.io/za6gw

Finn, E. S., Shen, X., Scheinost, D., Rosenberg, M. D., Huang, J., Chun, M. M., Papademetris, X., & Constable, R. T. (2015). Functional connectome fingerprinting: Identifying individuals using patterns of brain connectivity. *Nature Neuroscience*, *18*(11), 1664–1671. https://doi.org/10.1038/nn.4135

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507854; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Fornito, A., Harrison, B. J., Zalesky, A., & Simons, J. S. (2012). Competitive and cooperative dynamics of large-scale brain functional networks supporting recollection. *Proceedings of the National Academy of Sciences*, *109*(31), 12788–12793. https://doi.org/10.1073/pnas.1204185109

Fritch, H. A., Spets, D. S., & Slotnick, S. D. (2021). Functional connectivity with the anterior and posterior hippocampus during spatial memory. *Hippocampus*, *31*(7), 658–668. https://doi.org/10.1002/hipo.23283

Fruchterman, T. M. J., & Reingold, E. M. (1991). Graph drawing by force-directed placement. *Software: Practice and Experience*, *21*(11), 1129–1164. https://doi.org/10.1002/spe.4380211102

Gazzaley, A., & Nobre, A. C. (2012). Top-down modulation: Bridging selective attention and working memory. *Trends in Cognitive Sciences*, *16*(2), 129–135. https://doi.org/10.1016/j.tics.2011.11.014

Geib, B. R., Stanley, M. L., Dennis, N. A., Woldorff, M. G., & Cabeza, R. (2017). From hippocampus to whole-brain: The role of integrative processing in episodic memory retrieval. *Human Brain Mapping*, *38*(4), 2242–2259. https://doi.org/10.1002/hbm.23518

Gilmore, A. W., Nelson, S. M., & McDermott, K. B. (2016). The contextual association network activates more for remembered than for imagined events. *Cerebral Cortex*, *26*(2), 611–617. https://doi.org/10.1093/cercor/bhu223

Golland, Y., Golland, P., Bentin, S., & Malach, R. (2008). Data-driven clustering reveals a fundamental subdivision of the human cortex into two global systems. *Neuropsychologia*, *46*(2), 540–553. https://doi.org/10.1016/j.neuropsychologia.2007.10.003

Gonzalez-Castillo, J., Hoy, C. W., Handwerker, D. A., Robinson, M. E., Buchanan, L. C., Saad, Z. S., & Bandettini, P. A. (2015). Tracking ongoing cognition in individuals using brief, whole-brain functional connectivity patterns. *Proceedings of the National Academy of Sciences*, *112*(28), 8762–8767. https://doi.org/10.1073/pnas.1501242112

Gorgolewski, K., Burns, C. D., Madison, C., Clark, D., Halchenko, Y. O., Waskom, M. L., & Ghosh, S. S. (2011). Nipype: A flexible, lightweight and extensible neuroimaging data processing framework in python. *Frontiers in Neuroinformatics*, *5*. https://doi.org/10.3389/fninf.2011.00013

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507854; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Guderian, S., Schott, B. H., Richardson-Klavehn, A., & Düzel, E. (2009). Medial temporal theta state before an event predicts episodic encoding success in humans. *Proceedings of the National Academy of Sciences*, *106*(13), 5365–5370. https://doi.org/10.1073/pnas.0900289106

Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, *143*(1), 29–36. https://doi.org/10.1148/radiology.143.1.7063747

Hasselmo, M. E., Wyble, B. P., & Wallenstein, G. V. (1996). Encoding and retrieval of episodic memories: Role of cholinergic and GABAergic modulation in the hippocampus. *Hippocampus*, *6*(6), 693–708. https://doi.org/10.1002/(SICI)1098-1063(1996)6:6<693::AID-HIPO12>3.0.CO;2-W

Haxby, James. V. (2012). Multivariate pattern analysis of fMRI: The early beginnings. *Neuroimage*, *62*(2), 852–855. https://doi.org/10.1016/j.neuroimage.2012.03.016

Honey, C. J., Newman, E. L., & Schapiro, A. C. (2017). Switching between internal and external modes: A multiscale learning principle. *Network Neuroscience*, *1*(4), 339–356. https://doi.org/10.1162/NETN_a_00024

Hutchinson, J. B., Uncapher, M. R., Weiner, K. S., Bressler, D. W., Silver, M. A., Preston, A. R., & Wagner, A. D. (2014). Functional Heterogeneity in Posterior Parietal Cortex Across Attention and Episodic Memory Retrieval. *Cerebral Cortex*, *24*(1), 49–66. https://doi.org/10.1093/cercor/bhs278

Ji, J. L., Spronk, M., Kulkarni, K., Repovš, G., Anticevic, A., & Cole, M. W. (2019). Mapping the human brain's cortical-subcortical functional network organization. *NeuroImage*, *185*, 35–57. https://doi.org/10.1016/j.neuroimage.2018.10.006

Kim, D.-Y., Yoo, S.-S., Tegethoff, M., Meinlschmidt, G., & Lee, J.-H. (2015). The inclusion of functional connectivity information into fMRI-based neurofeedback improves its efficacy in the reduction of cigarette cravings. *Journal of Cognitive Neuroscience*, *27*(8), 1552–1572. https://doi.org/10.1162/jocn_a_00802

Kim, H. (2013). Differential neural activity in the recognition of old versus new events: An Activation Likelihood Estimation Meta-Analysis. *Human Brain Mapping*, *34*(4), 814–836. https://doi.org/10.1002/hbm.21474

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507854; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

King, D. R., Chastelaine, M. de, Elward, R. L., Wang, T. H., & Rugg, M. D. (2015). Recollection-related increases in functional connectivity predict individual differences in memory accuracy. *Journal of Neuroscience*, *35*(4), 1763–1772. https://doi.org/10.1523/JNEUROSCI.3219-14.2015

Konkle, T., Brady, T. F., Alvarez, G. A., & Oliva, A. (2010). Scene memory is more detailed than you think: The role of categories in visual long-term memory. *Psychological Science*, *21*(11), 1551–1556. https://doi.org/10.1177/0956797610385359

Kosslyn, S. M., Thompson, W. L., Klm, I. J., & Alpert, N. M. (1995). Topographical representations of mental images in primary visual cortex. *Nature*, *378*(6556), 496–498. https://doi.org/10.1038/378496a0

Krienen, F. M., Yeo, B. T. T., & Buckner, R. L. (2014). Reconfigurable task-dependent functional coupling modes cluster around a core functional architecture. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *369*(1653), 20130526. https://doi.org/10.1098/rstb.2013.0526

Kumar, M., Anderson, M. J., Antony, J. W., Baldassano, C., Brooks, P. P., Cai, M. B., Chen, P.-H. C., Ellis, C. T., Henselman-Petrusek, G., Huberdeau, D., Hutchinson, J. B., Li, Y. P., Lu, Q., Manning, J. R., Mennen, A. C., Nastase, S. A., Richard, H., Schapiro, A. C., Schuck, N. W., … Norman, K. A. (2022). Brainiak: The brain imaging analysis kit. *Aperture Neuro*, *2021*(4), 42. https://doi.org/10.52294/31bb5b68-2184-411b-8c00-a1dacb61e1da

Li, Y. P., Cooper, S. R., & Braver, T. S. (2021). The role of neural load effects in predicting individual differences in working memory function. *NeuroImage*, *245*, 118656. https://doi.org/10.1016/j.neuroimage.2021.118656

Manning, J. R., Zhu, X., Willke, T. L., Ranganath, R., Stachenfeld, K., Hasson, U., Blei, D. M., & Norman, K. A. (2018). A probabilistic approach to discovering dynamic full-brain functional connectivity patterns. *NeuroImage*, *180*, 243–252. https://doi.org/10.1016/j.neuroimage.2018.01.071

Marek, S., & Dosenbach, N. U. F. (2018). The frontoparietal network: Function, electrophysiology, and importance of individual precision mapping. *Dialogues in Clinical Neuroscience*, *20*(2), 133–140.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507854; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

connectionist models of learning and memory. *Psychological Review*, *102*(3), 419–457. https://doi.org/10.1037/0033-295X.102.3.419

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*(1), 167–202. https://doi.org/10.1146/annurev.neuro.24.1.167

Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: Multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences*, *10*(9), 424–430. https://doi.org/10.1016/j.tics.2006.07.005

Norman-Haignere, S. V., McCarthy, G., Chun, M. M., & Turk-Browne, N. B. (2012). Category-selective background connectivity in ventral visual cortex. *Cerebral Cortex*, *22*(2), 391–402. https://doi.org/10.1093/cercor/bhr118

Oppenheim, A. V., & Lim, J. S. (1981). The importance of phase in signals. *Proceedings of the IEEE*, *69*(5), 529–541. https://doi.org/10.1109/PROC.1981.12022

O'Reilly, R. C., & McClelland, J. L. (1994). Hippocampal conjunctive encoding, storage, and recall: Avoiding a trade-off. *Hippocampus*, *4*(6), 661–682. https://doi.org/10.1002/hipo.450040605

Pantazatos, S. P., Talati, A., Pavlidis, P., & Hirsch, J. (2012). Decoding unattended fearful faces with whole-brain correlations: An approach to identify condition-dependent large-scale functional connectivity. *PLOS Computational Biology*, *8*(3), e1002441. https://doi.org/10.1371/journal.pcbi.1002441

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, *12*(85), 2825–2830.

Polyn, S. M., Natu, V. S., Cohen, J. D., & Norman, K. A. (2005). Category-specific cortical activity precedes retrieval during memory search. *Science (New York, N.Y.)*, *310*(5756), 1963–1966. https://doi.org/10.1126/science.1117645

Ranganath, C., & Ritchey, M. (2012). Two cortical systems for memory-guided behaviour. *Nature Reviews Neuroscience*, *13*(10), 713–726. https://doi.org/10.1038/nrn3338

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507854; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Rissman, J., & Wagner, A. D. (2012). Distributed representations in memory: Insights from functional brain imaging. *Annual Review of Psychology*, *63*(1), 101–128. https://doi.org/10.1146/annurev-psych-120710-100344

Ritchey, M., & Cooper, R. A. (2020). Deconstructing the posterior medial episodic network. *Trends in Cognitive Sciences*, *24*(6), 451–465. https://doi.org/10.1016/j.tics.2020.03.006

Rosen, M. L., Stern, C. E., Devaney, K. J., & Somers, D. C. (2018). Cortical and Subcortical Contributions to Long-Term Memory-Guided Visuospatial Attention. *Cerebral Cortex*, *28*(8), 2935–2947. https://doi.org/10.1093/cercor/bhx172

Rosen, M. L., Stern, C. E., Michalka, S. W., Devaney, K. J., & Somers, D. C. (2016). Cognitive Control Network Contributions to Memory-Guided Visual Attention. *Cerebral Cortex*, *26*(5), 2059–2073. https://doi.org/10.1093/cercor/bhv028

Rosenberg, M. D., Finn, E. S., Scheinost, D., Papademetris, X., Shen, X., Constable, R. T., & Chun, M. M. (2016). A neuromarker of sustained attention from whole-brain functional connectivity. *Nature Neuroscience*, *19*(1), 165–171. https://doi.org/10.1038/nn.4179

Rosenberg, M. D., Scheinost, D., Greene, A. S., Avery, E. W., Kwon, Y. H., Finn, E. S., Ramani, R., Qiu, M., Constable, R. T., & Chun, M. M. (2020). Functional connectivity predicts changes in attention observed across minutes, days, and months. *Proceedings of the National Academy of Sciences*, *117*(7), 3797–3807. https://doi.org/10.1073/pnas.1912226117

Rubinov, M., & Sporns, O. (2011). Weight-conserving characterization of complex functional brain networks. *NeuroImage*, *56*(4), 2068–2079. https://doi.org/10.1016/j.neuroimage.2011.03.069

Rugg, M. D., & Vilberg, K. L. (2013). Brain networks underlying episodic memory retrieval. *Current Opinion in Neurobiology*, *23*(2), 255–260. https://doi.org/10.1016/j.conb.2012.11.005

Schaefer, A., Kong, R., Gordon, E. M., Laumann, T. O., Zuo, X.-N., Holmes, A. J., Eickhoff, S. B., & Yeo, B. T. T. (2018). Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity mri. *Cerebral Cortex*, *28*(9), 3095–3114. https://doi.org/10.1093/cercor/bhx179

Shen, X., Finn, E. S., Scheinost, D., Rosenberg, M. D., Chun, M. M., Papademetris, X., & Constable, R. T. (2017). Using connectome-based predictive modeling to predict individual behavior from brain connectivity. *Nature Protocols*, *12*(3), 506–518. https://doi.org/10.1038/nprot.2016.178

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507854; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Shirer, W. R., Ryali, S., Rykhlevskaia, E., Menon, V., & Greicius, M. D. (2012). Decoding subject-driven cognitive states with whole-brain connectivity patterns. *Cerebral Cortex*, *22*(1), 158–165. https://doi.org/10.1093/cercor/bhr099

Smith, S. M. (2012). The future of FMRI connectivity. *NeuroImage*, *62*(2), 1257–1266. https://doi.org/10.1016/j.neuroimage.2012.01.022

Song, H., Finn, E. S., & Rosenberg, M. D. (2021). Neural signatures of attentional engagement during narratives and its consequences for event memory. *Proceedings of the National Academy of Sciences*, *118*(33). https://doi.org/10.1073/pnas.2021905118

Song, H., & Rosenberg, M. D. (2021). Predicting attention across time and contexts with functional brain connectivity. *Current Opinion in Behavioral Sciences*, *40*, 33–44. https://doi.org/10.1016/j.cobeha.2020.12.007

Stojanoski, B., & Cusack, R. (2014). Time to wave good-bye to phase scrambling: Creating controlled scrambled images using diffeomorphic transformations. *Journal of Vision*, *14*(12), 6. https://doi.org/10.1167/14.12.6

Summerfield, C., Greene, M., Wager, T., Egner, T., Hirsch, J., & Mangels, J. (2006). Neocortical connectivity during episodic memory formation. *PLOS Biology*, *4*(5), e128. https://doi.org/10.1371/journal.pbio.0040128

Thomaz, C. E., & Giraldi, G. A. (2010). A new ranking method for principal components analysis and its application to face image analysis. *Image and Vision Computing*, *28*(6), 902–913. https://doi.org/10.1016/j.imavis.2009.11.005

Tompary, A., Al-Aidroos, N., & Turk-Browne, N. B. (2018). Attending to what and where: Background connectivity integrates categorical and spatial attention. *Journal of Cognitive Neuroscience*, *30*(9), 1281–1297. https://doi.org/10.1162/jocn_a_01284

Turk-Browne, N. B. (2013). Functional interactions as big data in the human brain. *Science*, *342*(6158), 580–584. https://doi.org/10.1126/science.1238409

Wang, Y., Cohen, J. D., Li, K., & Turk-Browne, N. B. (2015). Full correlation matrix analysis (FCMA): An unbiased method for task-related functional connectivity. *Journal of Neuroscience Methods*, *251*, 108–119. https://doi.org/10.1016/j.jneumeth.2015.05.012

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.14.507854; this version posted September 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Watanabe, T., Kimura, H. M., Hirose, S., Wada, H., Imai, Y., Machida, T., Shirouzu, I., Miyashita, Y., & Konishi, S. (2012). Functional dissociation between anterior and posterior temporal cortical regions during retrieval of remote memory. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *32*(28), 9659–9670. https://doi.org/10.1523/JNEUROSCI.5553-11.2012

Wheeler, M. E., Petersen, S. E., & Buckner, R. L. (2000). Memory's echo: Vivid remembering reactivates sensory-specific cortex. *Proceedings of the National Academy of Sciences*, *97*(20), 11125–11129. https://doi.org/10.1073/pnas.97.20.11125

Woolrich, M. W., Ripley, B. D., Brady, M., & Smith, S. M. (2001). Temporal autocorrelation in univariate linear modeling of fmri data. *NeuroImage*, *14*(6), 1370–1386. https://doi.org/10.1006/nimg.2001.0931

Yeo, B. T. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., Roffman, J. L., Smoller, J. W., Zöllei, L., Polimeni, J. R., Fischl, B., Liu, H., & Buckner, R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, *106*(3), 1125–1165. https://doi.org/10.1152/jn.00338.2011

Yeshurun, Y., Nguyen, M., & Hasson, U. (2021). The default mode network: Where the idiosyncratic self meets the shared social world. *Nature Reviews Neuroscience*, 1–12. https://doi.org/10.1038/s41583-020-00420-w