

Basic Neuroscience

Full correlation matrix analysis (FCMA): An unbiased method for task-related functional connectivity

Yida Wang^{a,*}, Jonathan D. Cohen^{b,c}, Kai Li^a, Nicholas B. Turk-Browne^{b,c}^a Department of Computer Science, Princeton University, Princeton, NJ, 08544, United States^b Department of Psychology, Princeton University, Princeton, NJ, 08544, United States^c Princeton Neuroscience Institute, Princeton University, Princeton, NJ, 08544, United States

H I G H L I G H T S

- There are exponentially more correlations than voxels in brain imaging data.
- Such correlations are not commonly analyzed at full scale for computational reasons.
- We developed full correlation matrix analysis (FCMA) to overcome these bottlenecks.
- FCMA incorporates and refines parallel computing and machine learning algorithms.
- We evaluate the performance of FCMA and demonstrate its use with a sample dataset.

A R T I C L E I N F O

Article history:

Received 19 March 2015

Received in revised form 13 May 2015

Accepted 14 May 2015

Available online 22 May 2015

Keywords:

Functional magnetic resonance imaging

Machine learning

Multivariate pattern analysis

Parallel computing

A B S T R A C T

Background: The analysis of brain imaging data often requires simplifying assumptions because exhaustive analyses are computationally intractable. Standard univariate and multivariate analyses of brain activity ignore interactions between regions and analyses of interactions (functional connectivity) reduce the computational challenge by using seed regions of interest or brain parcellations.

New method: To meet this challenge, we developed full correlation matrix analysis (FCMA), which leverages and optimizes algorithms from parallel computing and machine learning to efficiently analyze the pairwise correlations of all voxels in the brain during different cognitive tasks, with the goal of identifying task-related interactions in an unbiased manner.

Results: When applied to a localizer dataset on a small compute cluster, FCMA accelerated a naive, serial approach by four orders of magnitude, reducing running time from two years to one hour. In addition to this performance gain, FCMA emphasized different brain areas than existing methods. In particular, beyond replicating known category selectivity in visual cortex, FCMA also revealed a region of medial prefrontal cortex whose selectivity derived from differential patterns of functional connectivity across categories.

Comparison with existing method(s): For benchmarking, we started with a naive approach and progressively built up to the complete FCMA procedure by adding optimized classifier algorithms, multi-threaded parallelism, and multi-node parallelism. To evaluate what can be learned with FCMA, we compared it against multivariate pattern analysis of activity and seed-based analysis of functional connectivity.

Conclusions: FCMA demonstrates how advances in computer science can alleviate computational bottlenecks in neuroscience. We have released a software toolbox to help others evaluate FCMA.

© 2015 Elsevier B.V. All rights reserved.

* Correspondence to: Princeton Neuroscience Institute, Princeton University, Princeton, NJ, 08544, USA. Tel.: +1 609 258 1759; fax: +1 609 258 1771.

E-mail addresses: yidawang@princeton.edu (Y. Wang), jdc@princeton.edu (J.D. Cohen), li@cs.princeton.edu (K. Li), ntb@princeton.edu (N.B. Turk-Browne).

1. Introduction

Functional magnetic resonance imaging (fMRI) studies often seek to associate cognitive processes with brain regions. A classic finding of this type is the category selectivity of occipital and temporal cortex during visual perception, with punctate subregions responsive to objects from particular categories (Kanwisher, 2010; Op de Beeck et al., 2008). However, this univariate approach has been challenged by multivariate pattern analysis (MVPA) methods

(Norman et al., 2006; Kriegeskorte et al., 2008). In the case of object perception, MVPA demonstrated that category selectivity is not restricted to specific areas, but is also reflected in patterns of activity throughout occipital and temporal cortex (Haxby et al., 2001). Such findings have been offered as evidence that representations in the brain can be distributed over multiple areas. Here we extend this view, suggesting that the neural mechanisms supporting cognitive processing may be distributed over the entire brain, outside of traditional domain-specific systems. It may have been difficult to fully recognize this with existing analysis methods because of two implicit biases in such approaches.

First, the basic currency of MVPA is typically the amplitude of blood oxygenation level-dependent (BOLD) activity across a set of voxels. Although such spatial patterns can reveal information about neural representations, they may fail to capture interactions between brain regions over time that support neural processes. Even if neurons show identical average firing rates across behavioral tasks, when and how they interact with each other can distinguish between tasks (Vaadia et al., 1995; Cohen and Newsome, 2008; Poulet and Petersen, 2008; Cohen and Kohn, 2011). In fMRI, the BOLD activity of a voxel may fluctuate synchronously with certain other voxels in one task and with different voxels in another task, without the time-locked average amplitude of activity in any of these voxels reliably differing between the tasks or from baseline. A brain region X, for example, which contributes to object recognition and contains voxels that behave in this way, may show either no evoked response or similar responses to all object categories, and therefore would not be identified as category selective by standard analyses that contrast or classify patterns of activity. If region X's interactions with other regions differ by category, however, it could show category-selective patterns of correlations.

Second, although the analysis of correlations in fMRI data, or functional connectivity (Fox and Raichle, 2007; Smith, 2012), has become prevalent, this approach is subject to certain limitations. Such analyses often involve first identifying a small set of “seed” regions of interest that show task-related activity and then examining correlations between these seeds and other voxels in the brain. When seed correlations are calculated during tasks, efforts are made to eliminate the spurious correlations induced by evoked responses, such as by modeling these responses and the task-related seed timecourse simultaneously (as in psychophysiological interaction analysis; Friston et al., 1997) or by removing these responses prior to estimating correlations (as in background connectivity analysis; Norman-Haignere et al., 2012). Regardless, by choosing seeds based on activity, these procedures are affected by the limitations of activity-based methods highlighted above and would fail to identify region X (since its activity is not category selective). Some studies have attempted to consider functional connectivity more broadly by using MVPA to classify patterns of correlations amongst multiple regions both at rest (e.g., Dosenbach et al., 2010) and during tasks (e.g., Pantazatos et al., 2012; Watanabe et al., 2012; Mokhtari and Hossein-Zadeh, 2013). However, in these studies, only a small number of regions were selected relative to the total number of voxels in the brain, and this selection was based on activity in the same or other studies, or on coarse anatomical parcellations of the brain. To avoid any selection bias and consider the full range of data when identifying task-related interactions, the correlation of every voxel with every other voxel may be needed at the outset.

Here we describe the technical details of a method, full correlation matrix analysis (FCMA), that surmounts these limitations by performing unbiased multivariate analyses of whole-brain functional connectivity. The basic currency of FCMA is the full correlation matrix: the temporal correlation in BOLD activity of every voxel in the brain with every other voxel. A separate matrix is

computed for each temporal epoch of interest in a task, such as trial or block, just as would be done for mean activity or activity patterns. These matrices are then labeled with the task condition for that epoch and submitted to MVPA. Critically, the input to the classifier is now correlation patterns, rather than activity patterns. This analysis determines, in an exhaustive manner, which correlations distinguish between conditions. Region X above can therefore be found because it has differential interactions as a function of category.

This analysis may sound straightforward, but it is currently computationally intractable using a naïve approach. Consider the small fMRI dataset that we use later for an example application, which contains 18 subjects, 12 epochs per subject over two experimental conditions, and 34,470 voxels in the brain. The first, and easiest, part of FCMA is calculating the pairwise voxel autocorrelation matrix of every epoch. On a machine with two 4-core 2.6 GHz Xeon CPUs, running Matlab's *corr* function in batch mode, the computation of all 216 matrices (each with more than 594 million unique entries) takes 2.5 h and requires 478 GB of disk space at the end (and much more memory at intermediate stages). Rewriting the Pearson correlation computation in C++ using matrix multiplication and optimized linear algebra packages shortens running time to 348 s on the same machine. Thus, computing correlation matrices from fMRI data *per se* is not the hard problem and, in fact, there are already efficient tools, such as *InstaCorr* in AFNI, that have been used for this purpose (Gotts et al., 2013).

The more challenging problem arises when this massive amount of data needs to be analyzed. Typical seed-based functional connectivity maps are 3-D, reflecting the correlation of one seed voxel with all other voxels in the brain. Such data can be analyzed in a voxel-wise manner by examining which voxels have correlations with the seed that are reliably positive or negative, or that differ between conditions, using simple *t*-tests over subjects. The full correlation matrix, on the other hand, can be thought of as 6-D, reflecting one of these 3-D maps for every voxel in the 3-D brain. The sample dataset above produces 216 of these 6-D matrices, each with a correlation value for more than 594 million unique voxel pairs. That is, there are 4–5 orders of magnitude more variables to analyze, depending on the number of voxels. At that scale, algorithmic optimization and parallelization are necessary for the analysis to be tractable on current hardware, and machine learning techniques are needed to make sense of the data.

In particular, the set of full correlation matrices can be mined with MVPA to identify which pairs, and combinations of pairs, reliably discriminate between experimental conditions across subjects. To do so, the correlations are first preprocessed, including normalizing each coefficient with the Fisher transform and then *z*-scoring all coefficients across the matrix within subject. To classify two conditions, these correlation matrices are then divided into training and test sets, such as by leaving out one subject at a time, which enables random-effects cross-validation. Using this approach on the dataset above, the classifier would be trained on 204 matrices to find a boundary separating the conditions in a 594 million dimensional hyperspace and tested on the remaining 12 matrices to obtain a classification accuracy. With the data-driven feature selection approach described below, training and testing a basic linear classifier (e.g., linear support vector machine, SVM) on the correlation matrices in C++ would take 36 days on the machine described above. For all 18 subjects, this process must be repeated 17 more times for a total of 646 days.

We designed FCMA to address these challenges, by incorporating and refining algorithms for parallel computing and machine learning. Applying this method on a compute cluster, the time required for the analysis above was reduced from two years to one hour. Below we describe the technical details of the method, and then the results of an example application.

2. Methods

2.1. Description of FCMA

2.1.1. Overview

FCMA incorporates efficient algorithms for correlation computation, massive parallelization to manage the scale of the data, and on-demand analysis without the need for storage space or time. It runs on a compute cluster with any number of nodes, using a controller-worker architecture, and takes standard NIFTI-formatted images as input. A dataset consisting of fMRI activity over time in a set of voxels is divided in multiple temporal epochs and stored on a disk accessible to each node. These (relatively small) data are copied into the memory of each worker and the controller dynamically allocates a subset of voxels for computation when a node is available. A vector is computed for every epoch of each voxel using matrix multiplication, reflecting its correlation with every other voxel in that epoch. To handle the high dimensionality of these vectors without imposing biases, they are submitted to MVPA and the resulting cross-validation accuracy is assigned to the voxel. These accuracy scores can be used for nested feature selection, leading to a final round of MVPA over the correlation matrices of the automatically selected voxels.

2.1.2. Optimization of correlation computation

The Pearson correlation between several pairs of variables can be reduced to a matrix multiplication by normalizing the data (Worsley et al., 2005). Specifically, the timeseries for a given voxel and epoch is normalized by subtracting the mean and then dividing each value by the root sum of squares of the mean-centered data. The number of these operations scales linearly with the number of voxels, and thus computation poses little burden.

Matrix multiplication is significantly faster than other approaches for computing correlations. In addition, it can benefit from generally available technological advances in modern CPUs, such as the single instruction, multiple data (SIMD) set. FCMA implements advanced linear algebra algorithms from the GotoBlas library (Goto and van de Geijn, 2008) to exploit this hardware. At peak performance, with SSE instructions over 128-bit XMM registers, these algorithms allow for eight single-precision floating-point operations (four additions and four multiplications) in one CPU cycle.

2.1.3. Parallelization of correlation computation

FCMA was designed to run on a compute cluster with modern commercial machines and X86 architecture; each machine needs a reasonable amount of memory (e.g., 16 GB). Large-scale computing techniques are leveraged to accelerate computation and analysis. A controller/worker model is used (Fig. 1), in which a controller process coordinates numerous worker processes running on multiple machines with the Message Passing Interface (MPI). The controller allocates computation and analysis tasks to the workers; typically, one process is assigned to each node in order to fully utilize its resources. Each process consists of multiple threads to compute and analyze multiple voxels simultaneously within one node.

The first step of FCMA is to read in preprocessed fMRI data (e.g., corrected for head motion and other sources of noise) and text files specifying the experimental design. Based on this design, the data are partitioned into epochs. For every voxel and epoch, the data are normalized as described above. The controller then directs all available worker processes to load the full data from the storage device into memory and dynamically assigns each a subset of voxels to analyze. In other words, the full correlation matrix is automatically divided into groups of rows, and they are spread across worker processes. When the analysis in a worker finishes, the controller

collects the results, stores them in memory, and assigns another group of voxels.

By distributing the full correlation computation in this way over a 66-node cluster, and by using data normalization and optimized matrix multiplication algorithms in GotoBlas, total correlation computation time for our example dataset is reduced from 2.5 h to 0.73 s.

2.1.4. Parallelization of voxel-wise classifier analysis

To avoid the burden of storing full correlation matrices to disk (and associated write/read time), analysis is performed online within the nodes, immediately after correlation computation. After a worker process has computed a correlation matrix of the assigned subset of voxels with the rest of the voxels for each epoch, the same row of all matrices is extracted. Each of these rows comprises a vector of the correlations between a given voxel and all other voxels in the brain for one epoch. These correlation vectors are then labeled with the condition of the experimental design to which the epoch corresponds, and submitted to MVPA as training (or test) patterns. Each vector reflects a point in a high-dimensional space, and the goal of MVPA is to determine how accurately the points with different labels can be separated. MVPA is run with multiple threads; each thread processes the correlation vectors across different epochs for one voxel.

The default FCMA configuration uses a linear Support Vector Machine (SVM) classifier based on LibSVM (Chang and Lin, 2011). On a server with two 4-core 2.6 GHz Xeon CPUs, the standard version of LibSVM in serial mode takes about 90 s to perform cross-validation on a set of 204 correlation vectors from one voxel (each of length 34,470) in our example dataset. To speed up this process, we accelerated the original LibSVM algorithm by pre-computing the linear kernel matrices with GotoBlas. This reduced the running time to 2 s per voxel in a single thread.

2.1.5. Classification of correlation patterns

Applying the classifier to the correlation vectors produces a cross-validation accuracy for each voxel, reflecting how informative its correlations with the rest of the brain are about the task conditions (Fig. 2). FCMA can stop here, providing an unbiased, voxel-wise “map” of the extent to which every brain region shows task-related changes in whole-brain functional connectivity.

The results of the voxel-wise classifier analysis described above can also be used as feature selection for MVPA, by ranking voxels according to the accuracy with which their individual patterns of correlation differentiate conditions and then selecting a subset that exceeds some threshold (e.g., percentage of voxels, absolute accuracy level, statistical significance over cross-validation folds, etc.). FCMA implements this step as well: a chosen subset of voxels can be submitted again to obtain and analyze with MVPA the correlation matrix for only these voxels. This again requires that the original data be partitioned into training and test sets, with the feature selection step only applied to the training sets and the final classification accuracy derived from the held-out test sets. Using FCMA on our cluster, the total running time was reduced from 646 days to 72 min.

2.1.6. Code availability

We wrote a software toolbox that implements FCMA and it can be downloaded from GitHub: <http://princetonuniversity.github.io/fcma-toolbox>. The source code is written in C++ and requires OpenMP, MPI, and GotoBlas/Intel MKL packages. A compute cluster is recommended for running the toolbox, although it can also be run on a standalone workstation. The current stable version is 1.7 and it is being actively updated.

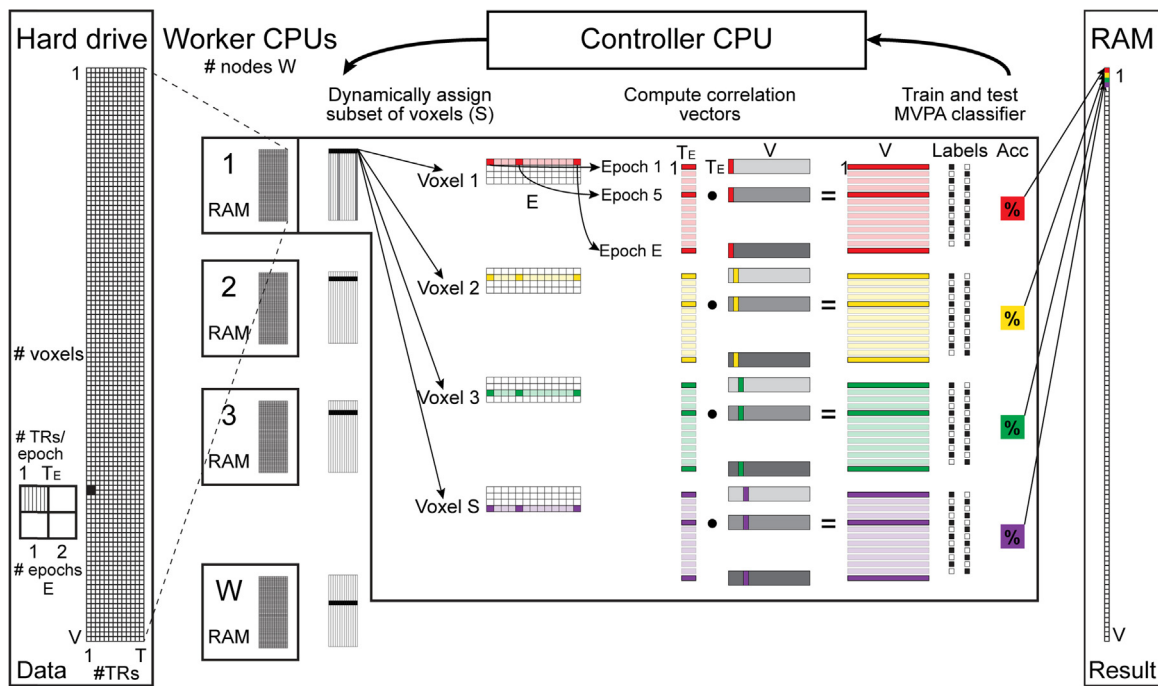


Fig. 1. Workflow overview. FCMA uses a controller/worker architecture, in which each worker first loads the full data into memory. The full data consist of a matrix with V voxels in rows and T timepoints in columns; the timepoints can be subdivided into E epochs, each with T_E timepoints (inset depicts two voxels and epochs). The controller process does the following: assigns a subset S of voxels to each of W workers; instructs the worker to compute the correlation between each of these voxels and the rest of the brain in each epoch; instructs the worker to analyze the correlation vectors for each voxel across epochs with MVPA and supplied condition labels; collects the analysis result (i.e., cross-validation accuracy) for each voxel and loads it into memory; and returns to the first step to assign another subset of voxels until there are none left. Finally, the controller writes the results to disk.

2.2. Example application

To show how FCMA works, we used it to analyze a face/scene localizer dataset. We focused on category-selective object perception because it is one of the most common and robust domains in cognitive neuroscience. In particular, GLM and MVPA approaches have revealed areas of occipital and temporal cortex that are activated by objects of particular categories (Kanwisher, 2010; Op de Beeck et al., 2008; Haxby et al., 2001). To examine whether FCMA can provide added value, we tested whether it emphasizes other, less-traditional brain areas as being category selective in terms of their patterns of correlation. Our hope is that releasing the software toolbox will prompt additional applications of FCMA, but this example illustrates the different steps involved and the variety of results that can be obtained.

2.2.1. Hardware

For performance benchmarking and the example application, we used a compute cluster with 66 nodes. Each node was equipped with two Intel Xeon E5430 processors, 16 GB memory, and 4 TB local disk, and could run eight threads simultaneously at peak.

2.2.2. Dataset

We used fMRI data from 18 subjects (localizer data from Turk-Browne et al., 2012), who viewed sequences of face and scene images in alternating blocks, while performing male/female and indoor/outdoor judgment tasks, respectively. Each subject completed one functional run that contained 12 blocks (six involving faces and six involving scenes). Each block contained 12 images from one category and lasted for 18 s, followed by blank fixation period of 12 s between blocks.

Data were acquired using a Siemens 3T Allegra scanner, with a T2*-weighted EPI sequence: volumes = 224, slices = 26, TR = 1500 ms, TE = 28 ms, matrix = 64, field of view = 224 mm, flip

angle = 64°, and thickness = 5 mm ($3.5 \times 3.5 \times 5 \text{ mm}^3$ voxels). Data were preprocessed prior to FCMA with corrections for head motion and slice-time acquisition, spatial smoothing (5 mm), and high-pass temporal filtering (128 s period). Brains were anatomically aligned to MNI space using standard linear registration methods (Jenkinson et al., 2002) and masked to remove non-brain voxels. After alignment, every brain contained 34,470 voxels.

2.2.3. Data preparation

Whole-brain activity and correlation patterns were calculated for each block. For activity patterns, the BOLD signal in each voxel was averaged over the 12 timepoints in that block, shifted forward by two TRs to account for the hemodynamic lag and capture peak response amplitude. We then normalized within subject by z-scoring these values across blocks within each voxel.

For correlation patterns, the BOLD signal in all pairs of voxels was correlated over the timecourse of each block. We used Pearson correlation and thus assume that the relationship between activity in one region and another is linear—a common assumption in fMRI studies of functional connectivity. It remains likely that meaningful interactions between regions are non-linear, and thus future development of FCMA will incorporate other more flexible relationship metrics (Reshef et al., 2011). In pilot testing, variance during the rise of the hemodynamic response meaningfully contributed to correlations and thus the timecourses were not shifted prior to being correlated. Although a sample size of 12 timepoints for computing correlations is small and leads to volatile estimates, we reasoned that the classifier would be able to extract meaningful correlation patterns across training examples. Any concerns about statistical power work against finding reliable classification performance. Additional preprocessing steps (mean-centering, division by root sum of squares) were applied to each timecourse to optimize correlation computation. The

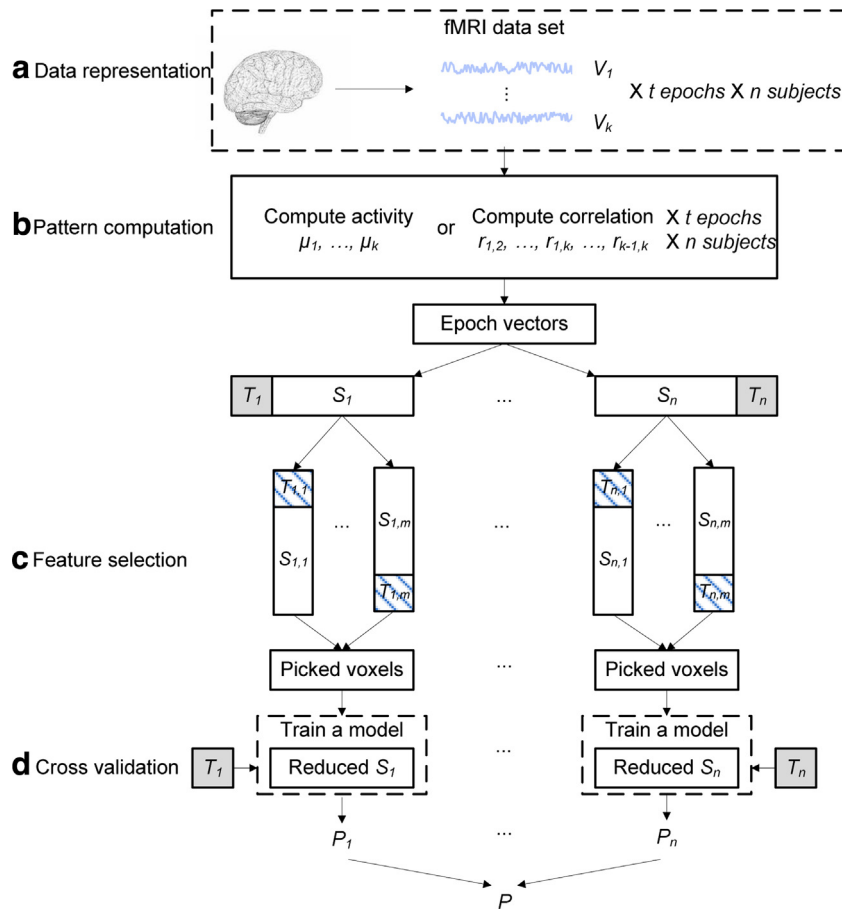


Fig. 2. Classification procedure. (a) The preprocessed fMRI data set contains n subjects, each represented with a k voxels by t epochs matrix. (b) For standard MVPA of activity patterns, vectors are defined for each subject and epoch as the average BOLD signal over time in every voxel (μ_i). For MVPA of correlation patterns, vectors are defined for each subject and epoch as the pairwise correlation of the BOLD signal over time between every voxel and every other voxel ($r_{i,j}$). (c) The same nested cross-validation pipeline can be applied to activity and correlation patterns. The inner loop serves to select features (voxels) for classification: A training set (S_i) is divided into m pieces to do an m -fold cross-validation that identifies the voxels with highest performance. (d) The outer loop is n -fold, with each fold leveraging the selected voxels to train a model on S_i and test it on the left-out test set (T_i). This results in a classification accuracy (P_i), which is then averaged across folds (P) to quantify overall performance.

resulting coefficients were Fisher-transformed and normalized within subject by z-scoring across blocks within each voxel pair.

2.2.4. Classification

The SVM algorithm was applied to activity and correlation patterns with a leave-one-subject-out nested n -fold cross-validation approach (where $n = 18$ subjects). An inner loop was used for feature (voxel) selection on the training set, while an outer loop used the chosen voxels to train a final classifier on the training set for cross-validation against the test set. The goal of feature selection was to automatically identify voxels whose activity/correlation patterns distinguish between classes, as well as to reduce the dimensionality of final classifier training. For each training set from the outer loop, we did an inner leave-one-subject-out m -fold cross validation for this feature selection (where $m =$ the remaining 17 subjects). Note that leaving one subject out at a time is not a requirement of FCMA itself, which can be applied just as easily across trials or blocks within subject. This approach is convenient, however, as the resulting sample of cross-validation accuracies (one per subject) can be analyzed statistically treating subject as a random effect.

The feature selection for activity and correlation patterns was similar, but differed due to the nature of the data. For activity patterns, we used a searchlight approach (Kriegeskorte et al., 2006). For every block in the inner-loop training set ($m - 1$ or 16 subjects \times 12 blocks/subject = 192 blocks), we generated a spherical

pattern of activity surrounding each voxel, made up of the average activity over that block for each voxel in the sphere. We then trained a linear SVM classifier to distinguish these patterns as having been obtained from face vs. scene blocks. The cross-validation test performance on the m th (17th) subject was assigned to this voxel, and the process was repeated for all 34,470 voxels. For correlation patterns, the procedure was very similar except in terms of how the block patterns for each voxel were generated. Rather than the average pattern of activity of the surrounding voxels, we used the pairwise correlation of a given voxel with all other voxels, resulting in a vector of 34,470 coefficients for that voxel for every block. These vectors were then submitted to the same kind of classifier to distinguish face vs. scene blocks, and the test performance was assigned to that voxel. The process was repeated for all voxels.

This entire process was repeated m times for the activity searchlights and correlation vectors for each voxel, with final voxel-wise classifier performance computed as the average performance across folds. The k most informative voxels were chosen from these voxel-wise average classifier performance maps for each type of classification, to be used for the outer-loop training and test. Specifically, these k voxels were used to generate patterns of activity and correlation for the full set of m subjects in the training set of the current outer-loop fold. Activity patterns reflected the average activity for each block from all selected voxels (i.e., vector of length k). Correlation patterns consisted of a correlation matrix

for every block, with the selected voxels in the rows and columns (rearranged into a vector of length k^2). These activity and correlation patterns were used to train a linear SVM classifier, which was then tested on the n th (18th) subject. To evaluate the robustness of decoding with respect to the choice of k , the average outer-loop classification accuracy across folds was calculated for $k = 10, 20, 50, 100, 200, 500, 1000, 2000$ (used in main analysis), and 5000 voxels. Random-effects reliability against chance (50%) was assessed with a permutation test over the subject folds.

2.2.5. Information mapping

To determine which areas of the brain contributed to successful classification, we examined the frequency with which a given voxel was chosen during feature selection across the n outer-loop folds. These n sets of k top voxels were not independent from each other because any two training sets had exactly $n - 2$ subjects in common. Due to this dependence, binomial or parametric tests were not appropriate for assessing the statistical significance of the frequencies. Thus, we used a non-parametric method to simulate the null distribution. We independently randomized the spatial location of voxels in every subject's preprocessed data, and then repeated the same full suite of analyses, selecting the top 2000 voxels in each of the 18 folds. Since the voxels in all subjects were shuffled, this led to a null distribution of 34,470 counts of how many times a voxel would be chosen in feature selection by chance (Table S1). We ran this simulation separately for activity- and correlation-based analyses, and obtained nearly identical results: less than 5% of voxels had counts greater than or equal to 8/18 (i.e., $p < .05$).

This procedure produced a voxel-wise map of p values based on the frequencies for activity and correlation, corrected for the dependence of the training sets. To further correct for multiple comparisons across voxels, we used cluster-size thresholding (Forman et al., 1995). The initial cluster-forming threshold for individual voxels was set at a frequency of 8/18 ($p < .05$ uncorrected). For each of 10,000 iterations, the true distribution of frequencies was randomized over the brain, the cluster-forming threshold was applied, and the number of clusters of each possible extent was counted. This procedure resulted in a null distribution of cluster sizes (Table S2), with clusters of size eight or more voxels corresponding to a whole-brain corrected threshold of $p < .05$. We attempted to correct for spatial autocorrelation by estimating the smoothness of the true frequency map and applying a matched Gaussian kernel after the randomization step of each iteration. However, due to the strength of the results and the relative sparseness of the frequencies, no amount of smoothing resulted in a more conservative cluster-size threshold. Thus, we retained the cluster-size threshold of eight voxels.

2.2.6. Vector change analysis

To examine how the pattern of correlations changed across categories (based on the results of feature selection), we isolated significant clusters in medial prefrontal cortex (mPFC; 81 voxels) and occipital/temporal cortex (1337 “visual” voxels). For each block and subject, we computed the correlation between every mPFC voxel and every visual voxel (81×1337 pairs). This matrix can be re-represented as 81 mPFC correlation vectors of length 1337 (one for each mPFC voxel) and 1337 visual vectors of length 81 for each block (one for each visual voxel). These two sets of vectors were then averaged across blocks from the same category, resulting in 81 face mPFC vectors, 81 scene mPFC vectors, 1337 face visual vectors, and 1337 scene visual vectors.

We calculated two metrics by comparing the face and scene vectors for a given mPFC or visual voxel: cosine distance (an estimate of the angle change) and L^2 -norm difference (an estimate of the length change). Since it is unknown how these metrics are distributed, we first normalized them with respect to their own null distribution.

Specifically, we randomized the block labels 1000 times within each subject, calculating the cosine distance and L^2 -norm difference for every voxel in each iteration. This generated null distributions against which the true values could be converted to z scores. For each subject, the 81 z values for mPFC voxels and the 1337 z values for visual voxels for each of cosine distance and L^2 -norm difference were averaged within region to produce four values: average mPFC angle change, average visual angle change, average mPFC length change, and average visual length change. These values were compared across subjects with permutation tests.

2.2.7. FFA/PPA analysis

We tested whether the mPFC pattern of correlations across face and scene categories could be explained by selective coupling with the FFA and PPA, respectively. To define these areas, we applied a standard univariate GLM to the data with FSL, modeling face and scene blocks with a canonical hemodynamic response function. The resulting parameter estimates were contrasted within subject, and the reliability of these contrasts was assessed at the group level with a higher-level FLAME mixed-effects model. We identified the FFA and PPA bilaterally by finding the peak face and scene selective voxel in the group analysis within the posterior lateral fusiform gyrus and the collateral sulcus/parahippocampal gyrus, respectively. We intersected a 10-mm radius sphere around these peaks with the visual voxels obtained from FCMA, obtaining 13 FFA voxels and 89 PPA voxels.

We then examined the average correlation between these voxels and mPFC, and tested reliability across subjects with permutation tests. Although these correlations were already known to differ between categories, the purpose of this analysis was to characterize the specific pattern of the correlations. Indeed, there are many patterns of correlations that could support the binary classification performed by FCMA. What we obtained was contrary to our initial hypothesis—and to what was suggested from prior findings (Al-Aidroos et al., 2012; Chadick and Gazzaley, 2011)—demonstrating the potential for FCMA to reveal unexpected findings.

2.2.8. Background connectivity analysis

The analyses above were applied to data that had only been preprocessed. To eliminate the potential confounding effects of stimulus-evoked responses on task-related functional connectivity, we repeated all analyses on residualized data from which these evoked responses had been removed. This background connectivity approach, which we have used previously (Al-Aidroos et al., 2012), contained four steps: (1) preprocessing the data; (2) fitting a “nuisance” GLM to the preprocessed data with regressors of no interest for 15 nuisance variables (six degrees of freedom for head motion, four seeds in bilateral anterior and posterior ventricles, four seeds in bilateral anterior and posterior white matter, global mean signal), as is done in analyses of resting connectivity; (3) fitting a finite impulse response (FIR) model to the residuals of the nuisance GLM with separate candlestick functions for each condition and timepoint in the block, to capture stimulus-evoked responses in a data-driven and voxel-specific manner; and (4) applying FCMA in the same way as for the preprocessed data, but to the residuals of the FIR model.

2.2.9. Seed-based analysis

To examine whether the FCMA results would be evident with a standard seed-based approach, we constructed a GLM with two regressors based on FFA and PPA activity: (1) the “selective” regressor consisted of the concatenated timeseries of activity from FFA in face blocks and PPA in scene blocks, and (2) the “non-selective” regressor consisted of the concatenated timeseries of activity from PPA in face blocks and from FFA in scene blocks. The

Table 1
Benchmark times for different optimizations in minutes. Each time refers to the training and testing of an SVM classifier model, in which one subject is held out for the final test, and cross-validation is performed among the remaining training subjects to select voxels whose correlations are used for the final test classifier. The baseline code was written in C++ using GotoBLAS library and LibSVM in single-thread mode. Improved linear SVM code pre-computed the linear kernel matrices used in LibSVM. Multi-core parallelism code used OpenMP to launch eight shared memory threads in one cluster node to run the program in parallel. Multi-node parallelism code runs MPI to coordinate 66 cluster nodes to work together in controller-worker mode. Long elapsed times were estimated via extrapolation.

Optimization	Running time in minutes			
	One voxel, one test subject (34,470 correlations)	All voxels, one test subject (34,470 voxels)	All voxels, all test subjects (18 subjects)	Performance speedup (x Baseline)
Baseline	1.5	51,705	930,528	–
Improved linear SVM	0.03	1174	21,138	44
Multi-core parallelism (eight cores)	0.03 ^a	201	3618	257
Multi-node parallelism (66 nodes)	0.03 ^a	4	72.3	12,855

^a Note that the benefit of parallelism is only realized when multiple voxels' correlation matrices are analyzed conjointly and can take advantage of multi-threading; a single thread is recruited for analyses restricted to one voxel.

resulting parameter estimates were contrasted (selective > non-selective), to identify voxels that correlated more strongly with category-selective regions during blocks of their preferred stimuli. Statistical significance was calculated across subjects using a randomization test and corrected for multiple comparisons with threshold-free cluster enhancement. This analysis was applied to both the preprocessed data and residualized data from the background connectivity analysis.

3. Results

3.1. Analysis benchmarks

FCMA accelerates the unbiased, exhaustive analysis of functional connectivity by optimizing correlation and machine learning algorithms and by leveraging multi-thread parallelism within nodes and multi-node parallelism across nodes. For each step of optimization, the performance of FCMA when applied to the example dataset with the sample hardware is provided in Table 1.

3.2. Classification performance

Eighteen subjects viewed sequences of faces or scenes in separate, recurring blocks of trials. For each subject and block, we identified two types of brain patterns: (1) the pattern of average activity over time for every voxel (basis of standard activity-based MVPA), and (2) the pattern of correlations over time of every voxel with every other voxel (basis of FCMA). We classified these patterns of activity and correlation from face vs. scene blocks with a linear SVM. A feature-selection procedure identified the 2000 most informative voxels. These voxels were used to generate final activity and correlation patterns for each block, which were in turn used to estimate overall classification performance. Having conducted feature selection and classifier training on $n - 1$ subjects, we tested the ability of the classifier to predict block category in the remaining subject. For activity patterns, average cross-validation accuracy was 98% ($p = 2.0 \times 10^{-6}$). This is consistent with previous findings (Haxby et al., 2001). For correlation patterns, average cross-validation accuracy was 95% ($p = 7.0 \times 10^{-6}$). Although less accurate in relative terms, the absolute level was nevertheless quite high and the correlation analysis started with lower statistical power in two ways. First, the unit of measurement was a single correlation value per block, compared to the multiple repeated samples of activity per block that went into the average activity. Second, the correlation-based classifier involved 4,000,000 inputs (2000 selected voxels squared), compared to 2000 inputs for the activity-based classifier, which increases the risk of overfitting. Finally, for both activity and correlation classifiers, performance was robust over a range of features (Fig. S1).

3.3. Information mapping

The similarity in performance for the two classifiers raises the question: Are they identifying the same neural mechanisms? To address this issue, we examined the frequency with which each voxel was chosen during both types of feature selection. That is, we summed the number of times that each voxel was chosen across outer-loop classification folds and assessed the reliability of these counts using a non-parametric procedure that accounted for the inherent overlap in the training sets. This produced voxel-wise maps of significance values for these counts, which were corrected for multiple comparisons.

Consistent with many previous findings, classification of activity patterns for faces and scenes was supported by regions of ventral temporal cortex, as well as known category-selective areas of dorsal occipital and retrosplenial cortex (Fig. 3a). Classification of correlation patterns was partly supported by overlapping regions. Additionally, however, they were joined by an area of mPFC, as well as by part of early visual cortex and the precuneus (Fig. 3b).

3.4. Vector change analysis

The involvement of mPFC in the classification of correlation but not activity patterns fits the description of the hypothetical region X in the Introduction. This area may participate in category-selective object perception not in terms of its pattern of activity, but rather in terms of its pattern of correlation. There are two explanations for how this could arise (Fig. 4): by mPFC interacting preferentially with different areas for different categories ("switching" hypothesis), or by mPFC interacting with the same areas to a greater or lesser degree based on category ("strength" hypothesis).

We tested these two possibilities, which are not mutually exclusive, by examining the correlations between mPFC and the 1337 voxels in occipital and temporal cortex that showed significant category selectivity in the correlation-based analysis. For any given block, the pattern of these mPFC correlations can be considered a point in a 1337-dimensional space, or a vector whose angle and length can be defined relative to the origin. According to the switching hypothesis, mPFC should be correlated with a subset of occipitotemporal voxels for one category and a different subset for the other category. Therefore, certain dimensions will have higher values for one category and other dimensions will have higher values for the other, which can be quantified as the difference in the angle of the vectors for each category. In contrast, the strength hypothesis predicts that the mPFC should always be correlated with the same subset of occipitotemporal voxels, just to different degrees. Therefore, the same dimensions will have higher values for one category and lower values for the other, which can

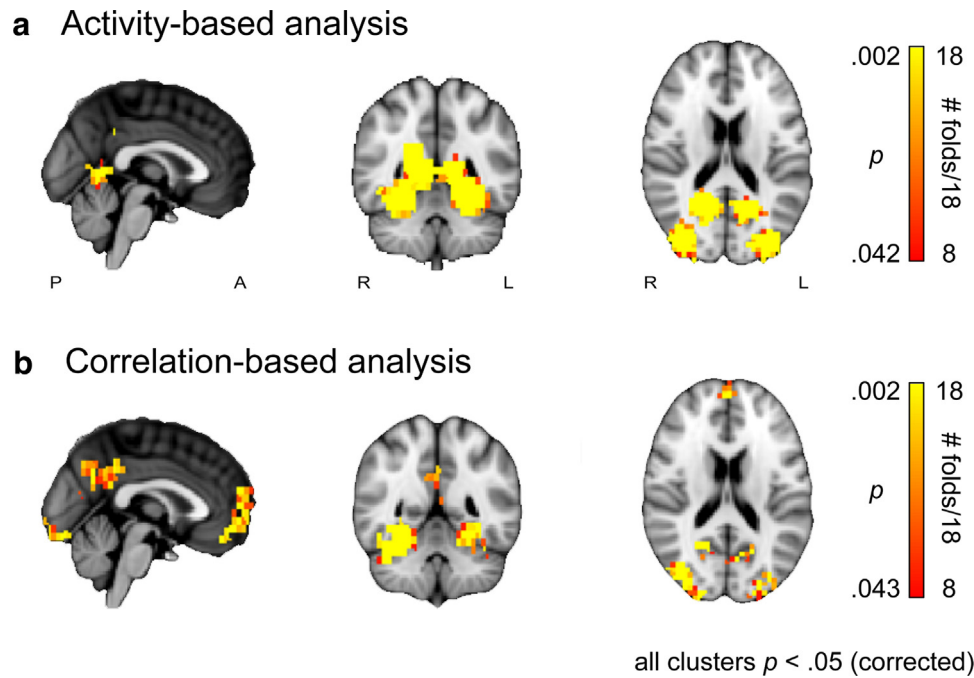


Fig. 3. (a) Activity-based analysis. Sagittal, coronal, and axial sections depicting voxels in which surrounding activity led to reliable classification of object category. These voxels were found in areas of ventral temporal, dorsal occipital, and retrosplenial cortex. (b) Correlation-based analysis. The same sections depicting voxels whose correlations with all other voxels led to reliable classification of object category. These voxels were found in areas overlapping with the activity-based analysis, but also in mPFC, early visual cortex, and precuneus. The color of each voxel reflects the frequency with which it was selected across cross-validation folds. P = posterior, A = anterior, R = right, L = left.

be quantified as the difference in the length of the vectors for each category.

We defined these correlation vectors for every subject, block, and mPFC voxel. We then averaged them, within subject, over all mPFC voxels and over all blocks of the same category. This resulted in two correlation vectors for each subject, one for face blocks and the other for scene blocks. From these vectors, we extracted two metrics: angle change (cosine distance) and length change (difference in L^2 norms). We normalized these values within subject with respect to their null distributions (obtained by shuffling the block labels) and then tested reliability across subjects.

In support of the switching hypothesis, there was a significant change in angle ($p = 1.6 \times 10^{-4}$) of the mPFC vectors for face and scene blocks. Although length also changed reliably ($p = .02$), the angle change in mPFC was significantly stronger than the

length change ($p = .002$). Finally, the angle change in mPFC was greater than when the analysis was inverted, to examine the change in the angle of occipital and temporal correlation vectors with mPFC ($p = .02$). This finding suggests that some of mPFC's category selectivity derives from interactions with different voxels for each category. It also provides an example of how FCMA not only identifies task-related interactions between brain regions, but also permits more in-depth analysis of the nature of these interactions.

3.5. Relationship of correlations to activity-based category selectivity

To further characterize the interactions between mPFC and posterior areas, we tested a natural hypothesis: that mPFC correlates with face-selective regions (e.g., fusiform face area, FFA) during

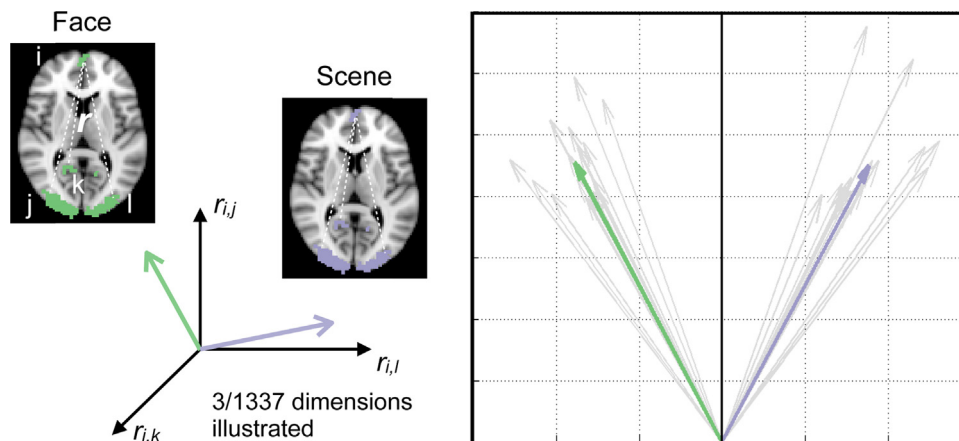


Fig. 4. Analysis of mPFC correlation vectors. For each subject and category, a vector of the average correlation between mPFC and each of the 1337 voxels in occipital and temporal cortex was computed. A 2-D projection of these vectors for each subject is depicted in gray, with the average across subjects in color. Each vector pair belonging to the same subject was placed symmetrically around the vertical meridian, and the angle between them was the real angle in the high-dimensional space. The lengths were rendered proportional to the group mean.

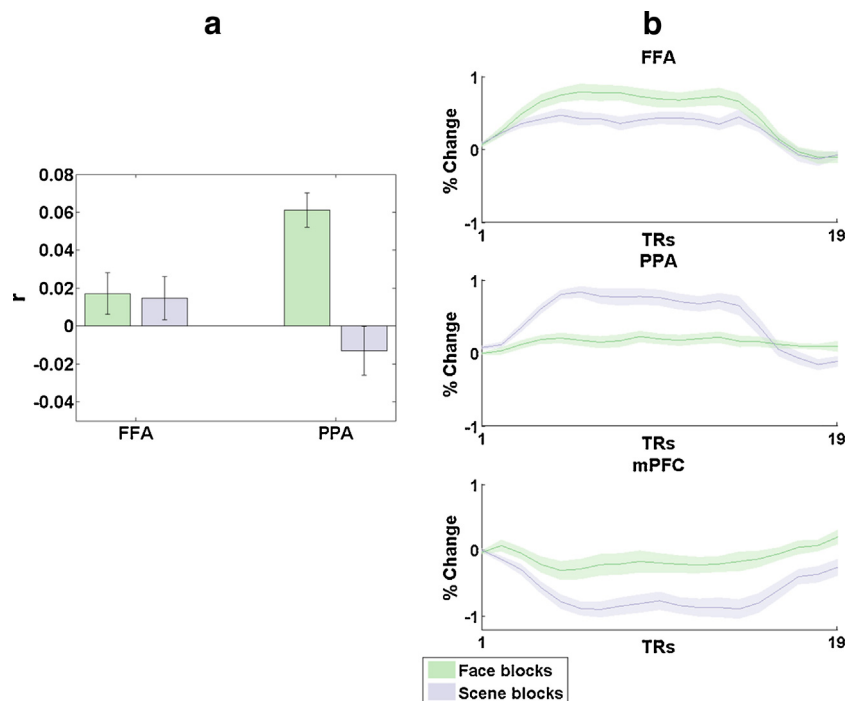


Fig. 5. (a) The average correlations of FFA and PPA voxels with mPFC plotted as a function of block type. (b) The average evoked BOLD activity of FFA, PPA, and mPFC voxels plotted as a function of block type. The FFA and PPA voxels were selected as face- and scene-selective, respectively, using a GLM. They were thus guaranteed to show this pattern of activity, which we display here only for visualization purposes.

face blocks, and with scene-selective regions (e.g., parahippocampal place area, PPA) during scene blocks. We did this by defining FFA and PPA voxels with a traditional univariate GLM of average activity (Kanwisher, 2010), and then examining the correlations between these voxels and mPFC for each subject (Fig. 5a).

The pattern of results across subjects was different than would be expected based on the category selectivity of FFA and PPA. There was no correlation on average between mPFC and FFA during face blocks ($p = .15$) or between mPFC and PPA during scene blocks ($p = .33$). Indeed, the only reliable correlation was between mPFC and PPA during face blocks ($p = 2.8 \times 10^{-5}$), which was also stronger than the mPFC-FFA correlation in these blocks ($p = 8.3 \times 10^{-4}$). The mPFC-FFA correlation during scene blocks was not reliable ($p = .24$). More generally, mPFC was not obtained in a standard seed-based whole-brain analysis of functional connectivity with FFA and PPA (Fig. S3a), highlighting the potential added value of FCMA.

Considering these correlations in the context of the average activity of FFA, PPA, and mPFC (Fig. 5b), helps rule out alternative explanations for the category-selective correlations in mPFC based on evoked activity. For example, if mPFC responded to both faces and scenes, and FFA and PPA responded selectively to faces and scenes, respectively, then greater mPFC-FFA correlations for faces and mPFC-PPA correlations for scenes would be spurious and unrelated to their interactions. The actual evoked activity and correlations in these regions are inconsistent with this kind of explanation in several ways: (1) FFA showed a stronger response than PPA to faces but a weaker correlation with mPFC; (2) FFA strongly responded to faces above baseline but did not correlate with mPFC; (3) mPFC did not respond to faces above baseline but strongly correlated with PPA; and (4) PPA strongly responded to scenes above baseline and mPFC strongly responded to scenes below baseline but there was no (negative) correlation between PPA and mPFC. Moreover, when we calculated correlations over the average timecourses in Fig. 5b (rather than the concatenated raw timeseries, as used to generate Fig. 5a), FFA was again less

correlated with mPFC than PPA in face blocks ($p = 3.1 \times 10^{-4}$), despite its greater activity.

This complex pattern of correlations between mPFC, FFA, and PPA can be visualized over all selected voxels in visual cortex using a circular graph (Fig. 6). One possible explanation for the overall trends in this graph is that mPFC is more focused on face information and thus sends excitatory input to visual cortex during face blocks (leading to positive correlations) and inhibitory input to visual cortex during scene blocks (leading to negative correlations). This occurred over a broad swath of visual cortex, which would be plausible if face- and scene-selective neurons were widely distributed (Haxby et al., 2001). Regardless, what determined the nature of mPFC correlations was the perceptual process being engaged rather than the activity-based identity of the target region (e.g., mPFC was strongly correlated with PPA on face blocks).

This result was unexpected and will require further study. However, it at least provides a demonstration that findings from FCMA need not align with those of seed-based analyses, which in this case would have assumed that category-selective correlations would align with category-selective activity.

3.6. Background connectivity analysis

Although the mPFC correlation results are difficult to explain by way of the pattern of evoked responses in mPFC, FFA, and PPA, a more general concern about the contribution of evoked responses to correlation-based classification remains. In particular, a voxel that responds during both face and scene blocks would be more correlated with other face-responsive voxels during face blocks and with other scene-responsive voxels during scene blocks. This voxel's category-selective correlations would thus be spurious, reflecting shared responses to a common influence rather than intrinsic interactions. If the successful correlation-based classification we observed reflects only spurious correlations, then removing such evoked responses from the data should result in chance-level

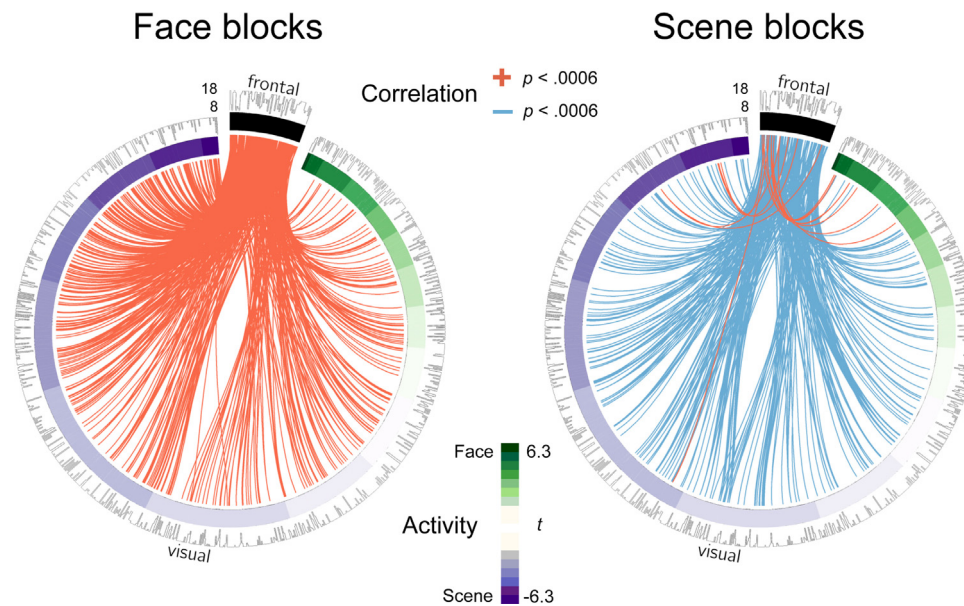


Fig. 6. mPFC correlation patterns for face and scene blocks. Each point on the circle represents a voxel, with mPFC voxels shown in black and visual voxels arranged clockwise according to their activity-based selectivity, from face (green) to scene (purple). Links are drawn for correlations between pairs of voxels that were reliably positive (red) or negative (blue) in a one-sample t -test across subjects. The surrounding histogram reflects the frequency (8–18) with which each voxel was selected across cross-validation folds (as in Fig. 3). Created with the Circos graphical tool (Krzywinski et al., 2009).

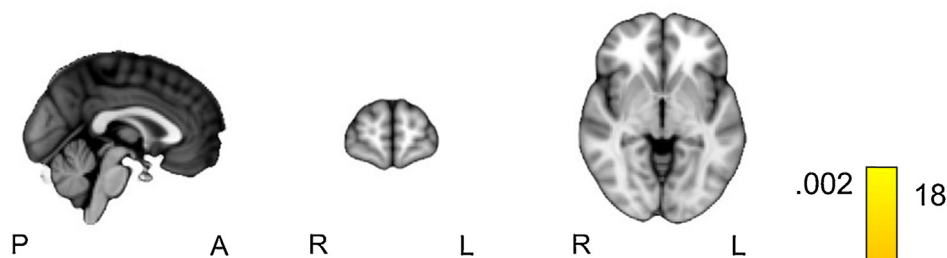
decoding. We tested this possibility using background connectivity analysis (Al-Aidroos et al., 2012), in which we scrubbed the data of evoked responses prior to applying FCMA.

As a check that this procedure worked, we re-ran activity-based classification, which previously produced 98% accuracy, and obtained chance performance (Fig. S2), with no voxels reliably selected (Fig. 7a). The critical question was whether FCMA would still be able to classify object category from correlations after

evoked responses were removed. Indeed, cross-validation accuracy was still reliably above chance ($p = .002$). Moreover, overlapping mPFC and precuneus regions were obtained (Fig. 7b). These regions were again not obtained in a seed-based analysis of the residualized data (Fig. S3b).

Unlike mPFC and precuneus, early visual cortex was no longer observed. Thus, we cannot rule out the possibility that the category-selective correlations in this region were a byproduct of evoked

a Activity-based analysis of residuals



b Correlation-based analysis of residuals

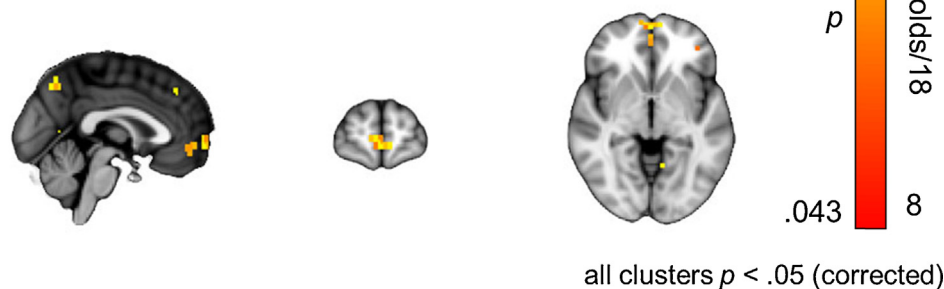


Fig. 7. Sagittal, coronal, and axial sections depicting voxels whose (a) surrounding activity patterns in a searchlight, and (b) correlations with all other voxels led to reliable classification of object category in residuals scrubbed of evoked responses. The color of each voxel reflects the frequency with which it was selected over cross-validation folds. P = posterior, A = anterior, R = right, L = left.

responses. This point extends to later visual areas in ventral temporal cortex, dorsal occipital cortex, and retrosplenial cortex as well, which were robust in the original analysis but much weaker in the background connectivity analysis. These findings indicate that parts of frontal and parietal cortices may fit the profile of regions whose correlations carry unique information, beyond what is reflected in activity.

4. Discussion

FCMA takes advantage of high-performance computing to efficiently analyze the full correlation matrix of fMRI data. It deals gracefully with enormous amounts data, optimally splitting and scheduling problems based on the latest techniques in large-scale computing and minimizing the need for slow data transfer by managing memory intelligently during online analysis. Beyond flexible parallelization, FCMA accelerates analysis by improving standard algorithms for correlation and classification.

As an example of how FCMA might produce findings that are less apparent with existing methods, applying it to a face/scene dataset revealed the involvement of mPFC and precuneus in category-selective object perception. After controlling for evoked responses, FCMA emphasized these regions more than the occipital and temporal regions typically linked to object perception. The mPFC has previously been implicated in the perception of social and emotional information (Martin and Weisberg, 2003; Sabatinelli et al., 2011), including self-concept, attitudes, and mentalizing (Mitchell, 2009). As an important social stimulus, faces have also been shown to elicit activity in mPFC (Allison et al., 1999; Kuhl et al., 2012). Although we did not obtain mPFC in our whole-brain activity-based classification, and the category selectivity of mPFC correlations persisted despite regressing out average activity, the mPFC may play a role in modulating face processing in ways that can manifest in both local activity and long-range correlations. This is consistent with our finding that mPFC showed robust positive correlations with visual cortex (both face- and scene-preferring voxels) during face but not scene blocks.

Though applied to voxel pairs in fMRI, FCMA could help uncover relationships between variables across a range of applications (Reshef et al., 2011). For instance, our example application involved the autocorrelation of one dataset, but two or more datasets (e.g., different temporal offsets, regions, or brains) could be submitted to FCMA, for unbiased analysis of cross-correlations. Moreover, although developed to work efficiently at the scale individual voxels, FCMA also works with any larger spatial unit of measurement that reduces the problem size, such as anatomical (Wang et al., 2009) or functional (Power et al., 2011) parcellations. Finally, different classifier algorithms could be specified, and any division of the data into training and test sets is possible (e.g., within rather than between subjects). By releasing our code, we hope that researchers will explore these and other avenues, to help further evaluate the utility of this approach. Ultimately, FCMA complements other approaches for the analysis of functional connectivity, helping identify task-related correlations in an unbiased manner, and providing a high-performance platform for applying machine learning techniques to correlational data.

Acknowledgments

The authors thank Moses Charikar for helpful discussions, Benjamin Singer and Alexa Tomparry for technical support, and Per Sederberg and Mason Simon for assistance with data collection. This work was supported by the J. Insley Blair Pyne fund at Princeton University, the John Templeton Foundation, the National Science Foundation (MRI BCS1229597), and the National Institutes

of Health (R01 EY021755). Intel and Fusion-io provided significant in-kind donations of computer hardware. The opinions expressed in this publication do not necessarily reflect the views of these agencies.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.jneumeth.2015.05.012>

References

- Al-Aidroos N, Said CP, Turk-Browne NB. Top-down attention switches coupling between low-level and high-level areas of human visual cortex. *Proc Natl Acad Sci U S A* 2012;109:14675–80.
- Allison T, Puce A, Spencer DD, McCarthy G. Electrophysiological studies of human face perception. I: potentials generated in occipitotemporal cortex by face and non-face stimuli. *Cereb Cortex* 1999;9:415–30.
- Chadick JZ, Gazzaley A. Differential coupling of visual cortex with default or frontal-parietal network based on goals. *Nat Neurosci* 2011;14:830–2.
- Chang CC, Lin CJ. LIBSVM: a library for support vector machines. *ACM Trans Intell Syst Technol* 2011;2:1–27, <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- Cohen MR, Kohn AK. Measuring and interpreting neuronal correlations. *Nat Neurosci* 2011;14:811–9.
- Cohen MR, Newsome WT. Context-dependent changes in functional circuitry in visual area MT. *Neuron* 2008;60:162–73.
- Dosenbach NU, Nardos B, Cohen AL, Fair DA, Power JD, Church JA, et al. Prediction of individual brain maturity using fMRI. *Science* 2010;329:1358–61.
- Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, Noll DC. Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. *Magn Resonance Med* 1995;33:636–47.
- Fox MD, Raichle ME. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat Rev Neurosci* 2007;8:700–11.
- Friston KJ, Buechel C, Fink GR, Morris J, Rolls E, Dolan RJ. Psychophysiological and modulatory interactions in neuroimaging. *NeuroImage* 1997;6:218–29.
- Goto K, van de Geijn RA. Anatomy of high-performance matrix multiplication. *ACM Trans Math Softw* 2008;34:1–25, <http://www.tacc.utexas.edu/tacc-software/gotoblas2>.
- Gotts SJ, Jo HJ, Wallace GL, Saad ZS, Cox RW, Martin A. Two distinct forms of functional lateralization in the human brain. *Proc Natl Acad Sci U S A* 2013;110:E3435–44.
- Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 2001;293:2425–30.
- Jenkinson M, Bannister P, Brady M, Smith S. Improved methods for the registration and motion correction of brain images. *NeuroImage* 2002;17:825–41.
- Kanwisher NG. Functional specificity in the human brain: a window into the functional architecture of the mind. *Proc Natl Acad Sci U S A* 2010;107:11163–70.
- Kuhl BA, Rissman J, Wagner AD. Multi-voxel patterns of visual category representation during episodic encoding are predictive of subsequent memory. *Neuropsychologia* 2012;50:458–69.
- Kriegeskorte N, Goebel R, Bandettini P. Information-based functional brain mapping. *Proc Natl Acad Sci U S A* 2006;103:3863–8.
- Kriegeskorte N, Mur M, Bandettini P. Representational similarity analysis: connecting the branches of systems neuroscience. *Front Syst Neurosci* 2008;2:1–28.
- Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, et al. Circo: an information aesthetic for comparative genomics. *Genome Res* 2009;19:1639–45.
- Martin A, Weisberg J. Neural foundations for understanding social and mechanical concepts. *Cogn Neuropsychol* 2003;20:575–87.
- Mitchell JP. Social psychology as a natural kind. *Trends Cogn Sci* 2009;13:246–51.
- Mokhtari F, Hossein-Zadeh GA. Decoding brain states using backward edge elimination and graph kernels in fMRI connectivity networks. *J Neurosci Methods* 2013;212:259–68.
- Norman KA, Polyn SM, Detre G, Haxby JV. Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends Cogn Sci* 2006;10:424–30.
- Norman-Haignere SV, McCarthy G, Chun MM, Turk-Browne NB. Category-selective background connectivity in ventral visual cortex. *Cereb Cortex* 2012;22:391–402.
- Op de Beeck HP, Haushofer J, Kanwisher NG. Interpreting fMRI data: maps, modules and dimensions. *Nat Rev Neurosci* 2008;9:123–35.
- Pantazatos SP, Talati A, Pavlidis P, Hirsch J. Decoding unattended fearful faces with whole-brain correlations: an approach to identify condition-dependent large-scale functional connectivity. *PLOS Comput Biol* 2012;8:e1002441.
- Poulet JF, Petersen CC. Internal brain state regulates membrane potential synchrony in barrel cortex of behaving mice. *Nature* 2008;454:881–5.
- Power JD, Cohen AL, Nelson SM, Wig GS, Barnes KA, Church JA, et al. Functional network organization of the human brain. *Neuron* 2011;72:665–78.
- Reshef DN, Reshef YA, Finucane HK, Grossman SR, McVean G, Turnbaugh PJ, et al. Detecting novel associations in large data sets. *Science* 2011;334:1518–24.

- Sabatinelli D, Fortune EE, Li Q, Siddiqui A, Krafft C, Oliver WT, et al. [Emotional perception: meta analyses of face and natural scene processing.](#) *NeuroImage* 2011;54:2524–33.
- Smith SM. [The future of fMRI connectivity.](#) *NeuroImage* 2012;62:1257–66.
- Turk-Browne NB, Simon MG, Sederberg PB. [Scene representations in parahippocampal cortex depend on temporal context.](#) *J Neurosci* 2012;32:7202–7.
- Vaadia E, Haalman I, Abeles M, Bergman H, Prut Y, Slovin H, et al. [Dynamics of neuronal interactions in monkey cortex in relation to behavioural events.](#) *Nature* 1995;373:515–8.
- Wang J, Wang L, Zang Y, Yang H, Tang H, Gong Q, et al. [Parcellation-dependent small-world brain functional networks: A resting-state fMRI study.](#) *Hum Brain Mapp* 2009;30:1511–23.
- Watanabe T, Kimura HM, Hirose S, Wada H, Imai Y, Machida T, et al. [Functional dissociation between anterior and posterior temporal cortical regions during retrieval of remote memory.](#) *J Neurosci* 2012;32:9659–70.
- Worsley KJ, Chen J-I, Lerch J, Evans AC. [Comparing functional connectivity via thresholding correlations and singular value decomposition.](#) *Philos Trans R Soc Lond Ser B Biol Sci* 2005;360:913–20.