Measuring prefrontal representational geometry: fMRI adaptation vs pattern analysis

Apoorva Bhandari (apoorva_bhandari@brown.edu)
CLPS, Brown University,
Providence, RI USA

Marcus Benna (mkb2162@cumc.columbia.edu)
Department of Neuroscience, Columbia University,
Centre for Theoretical Neuroscience, College of Physicians and Surgeons, Columbia University,
Mortimer B. Zuckerman Mind Brain Behavior Institute, Columbia University
New York, NY USA

Mattia Rigotti (mrigotti@us.ibm.com)
IBM Research AI,
Yorktown Heights, NY USA

Stefano Fusi (sf2237@columbia.edu)
Department of Neuroscience, Columbia University,
Centre for Theoretical Neuroscience, College of Physicians and Surgeons, Columbia University,
Mortimer B. Zuckerman Mind Brain Behavior Institute, Columbia University,
Kavli Institute for Brain Sciences, Columbia University,
New York, NY USA

David Badre (david_badre@brown.edu)
CLPS, Brown University, 190 Thayer Street
Robert J. and Nancy D. Carney Institute for Brain Science, 164 Angell Street
Providence, RI USA

Abstract:

  The prefrontal cortex (PFC) is necessary for the expression of flexible behavior. In over-trained monkeys, lateral PFC neurons represent a variety of task-relevant information in a high-dimensional code. In humans, the relatively low reliability of fMRI BOLD activity patterns and the difficulty of decoding their information content poses an obstacle to measuring PFC representational geometry. We systematically evaluated multi-voxel pattern analysis (MVPA) and the alternate method of fMRI adaptation for their reliability in estimating representational geometry and dimensionality in lateral PFC. Subjects solved a 3-dimension, audio-visual, parity task over 5 fMRI sessions. Leveraging the large amount of within-participant data, we estimated all pair-wise pattern distances and cross-condition adaptation effects in lateral PFC and visual cortex. We show that fMRI adaptation provides significantly more reliable estimates of the distances between task conditions in the lateral PFC's representational space compared to MVPA.

Keywords: prefrontal cortex; representations; dimensionality, repetition suppression, MVPA

Introduction

A central problem in cognitive neuroscience is understanding the nature and form of human prefrontal cortex (PFC) task representations. In several theories of PFC function, these task representations are posited to play a critical role in the expression of flexible, goal-directed behavior (Miller & Cohen, 2001). Previous fMRI studies of PFC have focused primarily on examining what information is encoded in PFC activity patterns (representational content), but far fewer studies have examined how this information is organized (representational format or geometry). Yet, this latter property is of paramount importance as the format determines the accessibility of encoded information to downstream circuits and so shapes network computation.

  A representation's geometry is defined as the arrangement of responses to task conditions in a multi-

dimensional space defined by neuronal firing rates (or, alternatively, voxel activity space), and can be recovered by estimating the distances between the responses in this space (Figure 2). An important property of this geometry is dimensionality. High-dimensional representation of task-relevant variables can support flexibility by enabling linear readouts of multiple, different conjunctions of task variables from the same representation (Fusi, Rigotti & Miller, 2016), providing a basis set for implementing rapid transitions between different task states. Indeed, in highly trained macaques, PFC representations of task variables approach maximum dimensionality, and this property predicts success on the task (Rigotti et al., 2013).

In contrast, macaque PFC neurons also simultaneously encode abstract task variables and categories (Bernardi et al, 2018). Similarly, in humans, a large body of research has linked PFC with abstract task representations that can be readily generalized and transferred to novel task settings. But, these abstract representations would be predicted to have low dimensionality. Therefore, to understand human PFC representations it is important to empirically estimate the dimensionality of these representations directly in humans, using non-invasive methods like fMRI.

Existing methods of estimating representational dimensionality with fMRI rely on the analysis of multi-voxel BOLD patterns (Diedrichsen et al., 2013; Ahlheilm et al., 2018). We have previously shown that multi-voxel BOLD patterns in the PFC have relatively low reliability (Bhandari et al., 2018), and it is difficult to decode their information content. This is an obstacle to the application of MVPA-based methods for estimating representational geometry of PFC. fMRI adaptation circumvents this problem by leveraging the phenomenon of repetition suppression to estimate the geometry and dimensionality (Rigotti et al., 2016). Here our aim is to implement both these approaches for estimating representational geometry in a single experiment and evaluate their reliability for estimating representational geometry.

Methods

Subjects

24 right-handed subjects (ages 18-30, 17 females) with no reported neurological or psychological disorders were recruited for an initial behavioral training session. Four of these subjects returned for an additional behavioral training session followed by 5 fMRI sessions, each on separate days. All subjects gave informed, written consent as approved by the Human Research Protections Office at Brown University, and were compensated for their participation.

Experiment Design

Experimental task Subjects performed a discrimination task based on a 3-dimension parity problem as described in Figure 1. Solving a parity problem requires a high-dimensional, conjunctive representation of the stimulus variables in at least one brain region.
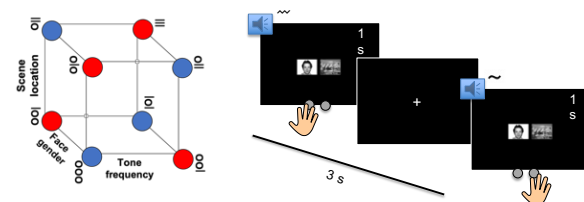
Figure 1 Experimental task. Stimuli consisted of 3-dimensions – a face (male/female), a scene (indoor/outdoor), and an auditory tone (low/high pitch). Graphic in left panel summarizes the task contingencies which implemented a parity problem (XOR in 3-dimensions). The 2^3 task conditions were mapped onto two categories (shown by red and blue circles). Right panel shows two example trials and timings employed.

Training Subjects were explicitly instructed on the categorization rule and then trained to perform the behavioral task (~700 trials with feedback). Subjects that achieved > 90% terminal accuracy in session 1 went through a second round of training on a separate day (~700 trials with feedback). These participants then returned for five additional fMRI sessions. Before being scanned, the subjects received refresher training (270 trials with feedback, 90 trials without feedback).

fMRI Whole-brain imaging was performed using a Siemens 3T Prisma MRI system. In each scanning session, we acquired 1) a high-resolution T1 weighted 3D MPRAGE image 2) 5-10 runs (185 volumes/run) of functional data while subject performed the task using a gradient-echo, echo-planar (EPI) pulse sequence (TR=2s, TE=30ms, 32 axial slices, 3x3x3 mm) 3) 1 run (124 volumes) of resting functional data was acquired (TR=3 s, TE=30 ms, FA=85°, 3 x 3 x 3-mm voxels, FOV=216, 47 axial slices, no skip, no dummy slices). In addition, 6 runs (150 volumes/run) of function data were acquired while subjects observed face and place stimuli in a block design for localizing face-sensitive voxels.
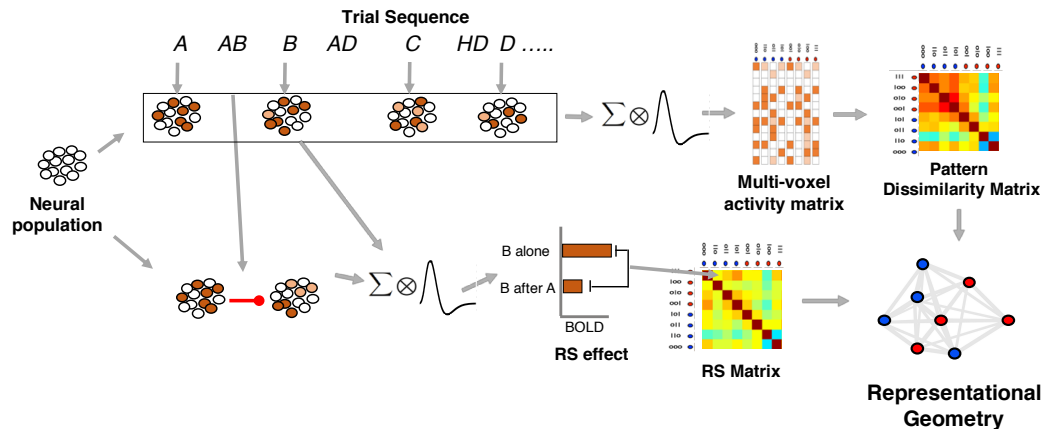
Figure 2: Overview of pattern analysis & adaptation approaches to recovering representational geometry. Pattern analysis relies on biased sampling of underlying neural population by different voxels, thus preserving the underlying similarity structure in multi-voxel activity patterns. Adaptation relies on repetition suppression operating at the level of individual neurons, with effects reflecting representational overlap summed at the level of the voxel. Both methods can be employed to obtain representational distances which are summarized in a pattern dissimilarity or RS-based dissimilarity matrix. Dimensionality is equivalent to the true rank of the dissimilarity matrix.

Trial sequence design For the functional task runs, 50 Type-1-Index-1 stimulus sequences with 2nd order counterbalancing were generated using methods described by Aguirre (2007). These sequences ensure each task condition is preceded by every other task condition (and a null), and each possible pair of task conditions is preceded by every task condition allowing the separate estimation of the BOLD response to each condition, as well as the 8x8 matrix of repetition suppression effects. In total, we collected more than 45 runs of data, each of which provided 1 sample of the 64 pair-wise repetition suppression effects, and 80 samples of the task condition patterns.

Regions-of-interest Two ROIs were selected prior to analyses: i) Fusiform face area (ffa) ii) the dorsolateral PFC. The ffa was defined as the face-sensitive voxels identified by a face vs place contrast obtained within an anatomically defined fusiform gyrus. The dlPFC ROI was defined as the task activity in each subject identified by an All task trials v/s implicit baseline contrast obtained within the multiple-demand dlPFC regions (Fedorenko, Duncan & Kanwisher, 2013).

fMRI Analysis

Preprocessing and general linear modelling: After standard preprocessing of the functional data in SPM (slice timing correction, registration, normalization and smoothing for the RS model) we fit general linear models to estimate a) the multi-voxel patterns

associated with the 8 conditions in every separate run and b) an 8 x 8 matrix of repetition-suppression effects.

Representational distance estimation
Correlation distance (1-r) was employed as the metric for estimating pair-wise representational distances between task condition using multi-voxel patterns (Kriegeskorte, Mur, Bandettini, 2008). With adaptation, the repeat vs non-repeat trial contrast for a pair of task conditions was employed as a distance metric. We assume both repetition suppression and enhancement effects reflect representational overlap. Distance matrices were constructed for each run. For adaptation, error trials are treated as missing values.

Reliability estimation
For both pattern-based and adaptation-based distances, split-half reliabilities of the representational geometry recovered from n trials was estimated by sampling 2n single-trial estimates of task condition patterns or repetition suppression effects and computing the rank correlation between 1000 random half-splits of these estimates. Thus, the reliabilities reflect the reproducibility of the rank order of distances.

Results

In Figure 3, split-half reliabilities of distance estimates for adaptation-based and pattern-based distances in both dorsolateral PFC and fusiform visual cortex are plotted as a function of the number of trials employed to obtain distance estimates. We observed a significant

region-by-method interaction with reliabilities being significantly higher in dlPFC, but not in visual cortex.

Note that for n conditions, fMRI adaptation requires n times more samples than a pattern-based approach for estimating a full representational dissimilarity matrix. Therefore, for a fair comparison, we also plot the pattern-based estimate from dlPFC obtained with 23x8 = 184 trials. This value is comparable to the adaptation
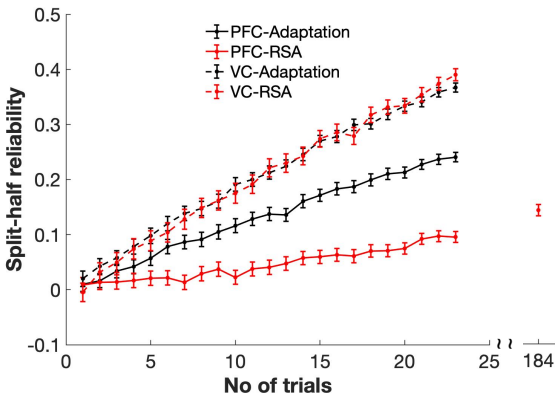
Figure 3 Split-half reliabilities (Spearman rank correlation) for adaptation-based (black) and pattern-based (red) distances in dorsolateral PFC (solid lines) and fusiform visual cortex (dashed lines) as a function of the amount of data used to estimate distances.

-based estimates for 23 trials (the maximum possible given our method). In other words, after controlling for the different amounts of scanning time required by the two methods, the adaptation-based methods still outperformed the pattern-based method in dlPFC.

In summary, our result demonstrates that fMRI adaptation may be a more reliable method for recovering the geometry of representations in lateral PFC. We note that this same property may hold for other regions of the brain not tested here. In the future, we will employ this approach to empirically estimate the dimensionality of human PFC representations.

Acknowledgments

References

Ahlheim, C., & Love, B. C. (2018). Estimating the functional dimensionality of neural representations. *NeuroImage*, *179*, 51-62.

Aguirre, G. K. (2007). Continuous carry-over designs for fMRI. *Neuroimage*, *35*(4), 1480-1494.

Bernardi, S., Benna, M. K., Rigotti, M., Munuera, J., Fusi, S., & Salzman, D. (2018). The geometry of abstraction in hippocampus and prefrontal cortex. *bioRxiv*, 408633.

Bhandari, A., Gagne, C., & Badre, D. (2018). Just above chance: is it harder to decode information from prefrontal cortex hemodynamic activity patterns? *Journal of cognitive neuroscience*, *30* (10), 1473-1498.

Diedrichsen, J., Wiestler, T., & Ejaz, N. (2013). A multivariate method to determine the dimensionality of neural representation from population activity. *Neuroimage*, *76*, 225-235.

Fedorenko, E., Duncan, J., & Kanwisher, N. (2013). Broad domain generality in focal regions of frontal and parietal cortex. *Proceedings of the National Academy of Sciences*, *110*(41), 16616-16621.

Fusi, S., Miller, E. K., & Rigotti, M. (2016). Why neurons mix: high dimensionality for higher cognition. *Current opinion in neurobiology*, *37*, 66-74.

Kriegeskorte, N., Mur, M., & Bandettini, P. A. (2008). Representational similarity analysis-connecting the branches of systems neuroscience. *Frontiers in systems neuroscience*, *2*, 4.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual review of neuroscience*, *24*(1), 167-202.

Rigotti, M., Barak, O., Warden, M. R., Wang, X. J., Daw, N. D., Miller, E. K., & Fusi, S. (2013). The importance of mixed selectivity in complex cognitive tasks. *Nature*, *497*(7451), 585.

Rigotti, M., & Fusi, S. (2016). Estimating the dimensionality of neural responses with fMRI Repetition Suppression. *arXiv preprint arXiv:1605.03952*.