

Representational geometry, not topography, best characterizes human neural activity in transmodal brain areas

Bogdan Petre (bogdan.petre.gr@dartmouth.edu)

Department of Psychological and Brain Sciences, Dartmouth College
3 Maynard Street, Hanover, NH 03755 USA

Martin A Lindquist (mlindqui@jhsph.edu)

Department of Biostatistics, Johns Hopkins University
615 N Wolfe Street, Baltimore, MD 21205 USA

Tor D Wager (tor.d.wager@dartmouth.edu)

Department of Psychological and Brain Sciences, Dartmouth College
3 Maynard Street, Hanover, NH 03755 USA

Abstract

Conserved topographies (e.g. retinotopy, tonotopy, somatotopy) and idiosyncratic population codes provide complementary accounts of neural representation at different scales. We use fMRI to examine which best explains human cortical function at mesoscale (mm to cm) where these accounts intersect and remain unsettled. We compare interindividual functional correspondence up to a low dimensional topography preserving warp (diffeomorphisms) or a high dimensional projection that preserves population code feature spaces (hyperalignment). Unlike previous studies our comparison is matched on training data and spans the full brain. When alignments computed from spontaneous activity were applied to a battery of task evoked responses, diffeomorphic warps predicted responses better in unimodal cortical areas while hyperalignment predicted responses better in transmodal areas. This is consistent with classic topographic representations of peripheral sensation and control, but demonstrates conserved feature spaces rather than topographies provide a better account of higher order computations in biological neural networks.

Keywords: population codes; cortical maps; neural functional topology; interindividual alignment; representation; fMRI

Introduction

There is strong evidence brain function is spatially organized at fine scale (Penfield & Boldrey, 1937). This provides a compelling embodied account of neural representation (retinotopy, somatotopy and tonotopy), is predetermined by conserved genetic mechanisms (Grove & Fukuchi-Shimogori, 2003) and pre-sensory spontaneous activity (Torborg & Feller, 2005), and generalizes to higher order brain areas (Steel et al., 2024) and across architecturally distinct brain structures (Xue et al., 2021; Nambu, 2011). Thus, the relative location of neurons informs their function.

A complementary perspective holds that neurons respond to computationally relevant but abstract features (Hubel & Wiesel, 1962), and distinct attributes or action sequences are encoded by overlapping populations of neurons. Abstract feature spaces are defined by the most frequent patterns of activity, and information propagates downstream by joint population readouts (Churchland et al., 2012). Population codes are idiosyncratic, in part because they develop in an experience dependent manner (Wiesel & Hubel, 1963; Roe et al., 1990), but multiple codes can represent the same feature space and studies in both biological and artificial networks show they often do given similar learning conditions (Morcos et al., 2018; Saxe et al., 2019; Jiahui et al., 2023). These principles cannot be reconciled with ordered topographies and are instead thought to be nested within them, suggesting a qualitative transition in the organization of cortical representation at ≈ 1 mm scale (Hubel & Wiesel, 1977).

A body of work now claims feature spaces are more conserved across individuals than mesoscale (1mm-1cm) topographies (Haxby et al., 2011). Topographic consistency is

measured by spatial correlations of neural responses across brains after interindividual alignments that preserve spatial continuity ("diffeomorphic" alignment) (Robinson et al., 2018). Feature spaces can be compared on equal footing if population codes are reprojected between brains using high dimensional rotations ("hyperalignment"). Crucially, this disrupts spatial contiguity but not feature spaces, and reveals substantial interindividual functional correspondences that are otherwise hidden. Unfortunately, hyperalignment studies are often confounded by ensemble learning techniques (Guntupalli et al., 2016, 2018), never directly compare hyperalignment to diffeomorphisms with matched training data (Bazeille et al., 2021), a negative control, and neglect biological and computational diversity throughout the brain which might affect the extent of stereotyped & innate vs. idiosyncratic & experience dependent organization, oversights we now address.

We compare interindividual correspondence in fine scale topographies and feature spaces using multivariate diffeomorphic alignment and parcel-wise hyperalignment (respectively). We train alignments using matched spontaneous activity at rest and test on responses to a battery of sensory, motor and cognitive tasks. We specifically test for differences between unimodal areas with known topographies and higher order multimodal and transmodal areas (Mesulam, 1998) using the previously published principal gradient of cortical organization (Margulies et al., 2016).

Methods

Brain Parcellation

We defined an HCP 91k grayordinate space, full brain, composite atlas (CANLab2024; includes (Glasser et al., 2016) and others) of 518 parcels designed to delineated the finest structural and functional distinctions available in the human brain so that parcel-wise hyperalignment would not map signals across distinct neural architectures.

Data and Participants

We used 3T BOLD fMRI data from unrelated participants in the Human Connectome Project (N=278), including 4x 15 minute resting state (Smith et al., 2013) and 7 task scans. Task evoked response estimates are mean evoked BOLD contrast above baseline and were precomputed by the HCP (Barch et al., 2013). Functionally aligned "MSMAll" data was not used.

Resting state connectomes Resting state data was detrended, corrected for motion, global signal and csf, bandpass filtered (0.1-0.008Hz), had volumes 1-4 removed, and time-series z-scored and concatenated across scans. Functional connectivity was computed as Pearson correlation between each voxel/vertex and the mean timeseries of each of brain parcel (518 x 91k connectome).

Alignment

Pairs of resting state connectomes were aligned (N=139 pairs). Two diffeomorphic alignments were enchainned, and

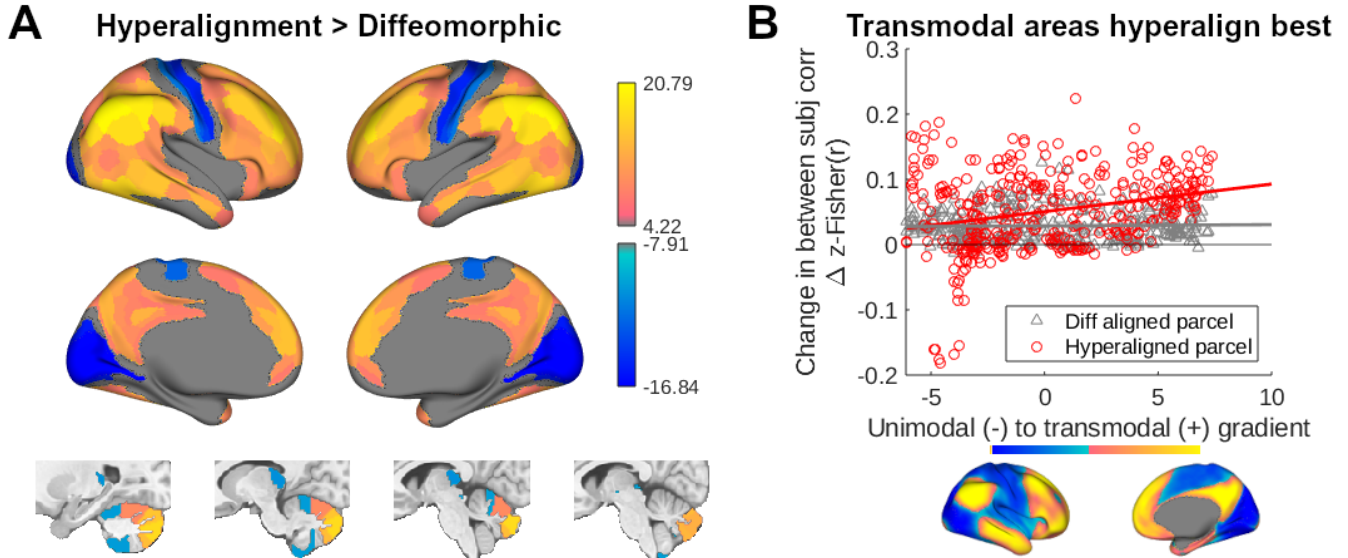


Figure 1: A. Change in between subject correlations (Δ BSC) of task evoked responses after parcel-wise hyperalignment on resting state geometry relative to multivariate diffeomorphic alignment ($t > 4.02$, $p < 1e-4$, Holm-Sidak corrected; t-stat shown). B. Hyperalignment is more effective in transmodal than unimodal areas (gradient loading regressed on Δ BSC, random subject effects, $t_{139} = -15.3$, $p < 1e-6$). Group mean Δ BSC shown. Improvement with diffeomorphic alignment over structural alignment shown as negative control (gray). Below: Principal gradient of macroscale cortical organization (Margulies et al., 2016).

in between connectomes were recomputed. Hyperalignment came last. Obtained transformations were applied to task evoked response data in the same order.

Diffeomorphic alignment Multivariate diffeomorphic alignment was implemented using MSM (Robinson et al., 2018) or ANTs using hyperparameters adapted from (Glasser et al., 2016) or sMRIprep for surfaces and volumes (respectively).

Hyperalignment We introduce a novel subspace hyperalignment algorithm that is mean-invariant. Traditional hyperalignment (Haxby et al., 2011) preserves angular distance and vector norm, but not the mean, e.g. a global sign flip is permitted, but this is a significant biological parameter independent from our question. We define a mean operator as a vector $x_i = x_j \forall i, j$, and compute the forward and reverse orthogonal transformations, \mathbf{F} and \mathbf{F}^{-1} , that rotate this vector to and from the component vector x_1 . We apply \mathbf{F} to moving and fixed images and then compute an orthogonal transformation, \mathbf{R} , to align the residual vertices $x_{j \neq 1}$. \mathbf{FRF}^{-1} yields mean invariant hyperalignment. To compute \mathbf{R} we use FastSRM (Richard & Thirion, 2023) which outperforms other orthogonal hyperalignment algorithms (Bazeille et al., 2021).

Results and Discussion

Hyperalignment achieved better intersubject correspondence in transmodal areas than diffeomorphic alignment, but diffeomorphic alignment was superior in unimodal areas like visual cortex and somatosensory strip (Figure 1A). When compared with an *a priori* unimodal to transmodal gradient map

(Margulies et al., 2016), hyperalignment success was significantly correlated with the degree of representational abstraction. While diffeomorphic alignment also produced consistent and widespread improvements in intersubject correspondence relative to structural alignment, no similar association was apparent (Figure 1B). This indicates that improved intersubject correspondence in these regions is not simply a consequence of functional alignment with task-free resting state data, but rather is specific to functional hyperalignment. Notably, a similar pattern emerged in subcortical areas. Diffeomorphic alignment outperformed hyperalignment in anterior cerebellum and vice versa in posterior cerebellum, consistent with known somatotopic maps in the anterior lobules (Xue et al., 2021).

While topographic organization offers a useful account throughout the brain, the relative superiority of hyperalignment in transmodal brain areas shows fine scale topographies in these regions may be less informative and recommends measures of shared representational features instead. Conversely, the disruption of functional organization of sensory motor regions by hyperalignment suggests topographic organization in unimodal areas may be uniquely precise and important.

Acknowledgments

We thank Feilong Ma for providing the python implementation of Procrustes analysis which formed the basis of our algorithm, and Emma Robinson and Renato Besenczi for help implementing efficient multivariate surface alignment.

This research was funded by NIBIB R01EB026549

References

- Barch, D. M., Burgess, G. C., Harms, M. P., Petersen, S. E., Schlaggar, B. L., Corbetta, M., ... Essen, D. C. V. (2013, 10). Function in the human connectome: Task-fMRI and individual differences in behavior. *NeuroImage*, *80*, 169-189. doi: 10.1016/j.neuroimage.2013.05.033
- Bazeille, T., DuPre, E., Richard, H., Poline, J.-B., & Thirion, B. (2021, December). An empirical evaluation of functional alignment using inter-subject decoding. *NeuroImage*, *245*, 118683.
- Churchland, M. M., Cunningham, J. P., Kaufman, M. T., Foster, J. D., Nuyujukian, P., Ryu, S. I., & Shenoy, K. V. (2012, July). Neural population dynamics during reaching. *Nature*, *487*(7405), 51–56.
- Glasser, M. F., Coalson, T. S., Robinson, E. C., Hacker, C. D., Harwell, J., Yacoub, E., ... Van Essen, D. C. (2016, August). A multi-modal parcellation of human cerebral cortex. *Nature*, *536*(7615), 171–178.
- Grove, E. A., & Fukuchi-Shimogori, T. (2003). *Generating the cerebral cortical area map* (Vol. 26). doi: 10.1146/annurev.neuro.26.041002.131137
- Guntupalli, J. S., Feilong, M., & Haxby, J. V. (2018). A computational model of shared fine-scale structure in the human connectome. *PLoS Comput. Biol.*, *14*(4), e1006120.
- Guntupalli, J. S., Hanke, M., Halchenko, Y. O., Connolly, A. C., Ramadge, P. J., & Haxby, J. V. (2016). A model of representational spaces in human cortex. *Cereb. Cortex*, *26*(6), 2919–2934.
- Haxby, J. V., Guntupalli, J. S., Connolly, A. C., Halchenko, Y. O., Conroy, B. R., Gobbini, M. I., ... Ramadge, P. J. (2011, October). A common, high-dimensional model of the representational space in human ventral temporal cortex. *Neuron*, *72*(2), 404–416.
- Hubel, D. H., & Wiesel, T. N. (1962, January). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.*, *160*, 106–154.
- Hubel, D. H., & Wiesel, T. N. (1977, July). Ferrier lecture. functional architecture of macaque monkey visual cortex. *Proc. R. Soc. Lond. B Biol. Sci.*, *198*(1130), 1–59.
- Jiahui, G., Feilong, M., di Oleggio Castello, M. V., Nastase, S. A., Haxby, J. V., & Gobbini, M. I. (2023, 10). Modeling naturalistic face processing in humans with deep convolutional neural networks. *Proceedings of the National Academy of Sciences*, *120*. doi: 10.1073/pnas.2304085120
- Margulies, D. S., Ghosh, S. S., Goulas, A., Falkiewicz, M., Huntenburg, J. M., Langs, G., ... Smallwood, J. (2016, 11). Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proceedings of the National Academy of Sciences of the United States of America*, *113*, 12574-12579. doi: 10.1073/pnas.1608282113
- Mesulam, M.-M. (1998). *From sensation to cognition* (Vol. 121).
- Morcos, A., Raghu, M., & Bengio, S. (2018). Insights on representational similarity in neural networks with canonical correlation. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, & R. Garnett (Eds.), (Vol. 31). Curran Associates, Inc.
- Nambu, A. (2011). Somatotopic organization of the primate basal ganglia. *Frontiers in Neuroanatomy*, *5*, 1-9. doi: 10.3389/fnana.2011.00026
- Penfield, W., & Boldrey, E. (1937, 12). Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation. *Brain*, *60*(4), 389-443. doi: 10.1093/brain/60.4.389
- Richard, H., & Thirion, B. (2023, 10). FastSRM: A fast, memory efficient and identifiable implementation of the shared response model. *Aperture Neuro*, *3*. doi: 10.52294/001c.87954
- Robinson, E. C., Garcia, K., Glasser, M. F., Chen, Z., Coalson, T. S., Makropoulos, A., ... Rueckert, D. (2018). Multimodal surface matching with higher-order smoothness constraints. *NeuroImage*, *167*, 453-465. doi: 10.1016/j.neuroimage.2017.10.037
- Roe, A. W., Pallas, S. L., Hahm, J. O., & Sur, M. (1990). A map of visual space induced in primary auditory cortex. *Science*, *250*, 818–820.
- Saxe, A. M., McClelland, J. L., & Ganguli, S. (2019, June). A mathematical theory of semantic development in deep neural networks. *Proc. Natl. Acad. Sci. U. S. A.*, *116*(23), 11537–11546.
- Smith, S. M., Beckmann, C. F., Andersson, J., Auerbach, E. J., Bijsterbosch, J., Douaud, G., ... Glasser, M. F. (2013, 10). Resting-state fMRI in the human connectome project. *NeuroImage*, *80*, 144-168. doi: 10.1016/j.neuroimage.2013.05.039
- Steel, A., Silson, E. H., Garcia, B. D., & Robertson, C. E. (2024, 1). A retinotopic code structures the interaction between perception and memory systems. *Nature Neuroscience*. doi: 10.1038/s41593-023-01512-3
- Torborg, C. L., & Feller, M. B. (2005, 7). *Spontaneous patterned retinal activity and the refinement of retinal projections* (Vol. 76). doi: 10.1016/j.pneurobio.2005.09.002
- Wiesel, T. N., & Hubel, D. H. (1963, November). SINGLE-CELL RESPONSES IN STRIATE CORTEX OF KITTENS DEPRIVED OF VISION IN ONE EYE. *J. Neurophysiol.*, *26*, 1003–1017.
- Xue, A., Kong, R., Yang, Q., Eldaief, M. C., Angeli, P. A., DiNicola, L. M., ... Yeo, B. T. T. (2021, 2). The detailed organization of the human cerebellum estimated by intrinsic functional connectivity within the individual. *Journal of Neurophysiology*, *125*, 358-384. doi: 10.1152/jn.00561.2020