

Modeling the shared deep structure of information encoded in fine-scale cortical topographies

Jim Haxby

Center for Cognitive Neuroscience, Dartmouth College

Haxby Lab



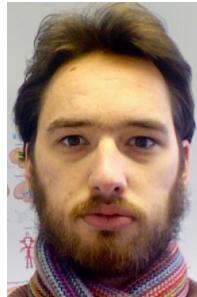
Hyper alignment
Swaroop Guntupalli
Vicarious



Individual differences
Feilong Ma
Graduate student



Attention
Sam Nastase
Graduate student



Action representation, computational methods
Nick Oosterhof
Emeritus



Forward encoding
Cara van Uden
Undergraduate student



Yaroslav Halchenko
Research professor, software



Analysis of similarity structure,
representation of biological classes
Andy Connolly
Asst prof, Geisel School of Medicine

With help from



Peter Ramadge
Electrical Engineering
Princeton University

and EE grad students, past and present



Mert Rory Sabuncu
now at MGH



Bryan Conroy
Philips Research



Alex Lorbert
Superfish, Israel



Hao Xu
Google



Cameron Chen
current

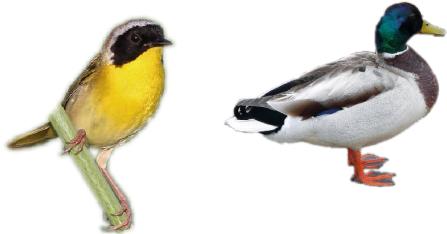
A common high-dimensional linear model of information spaces in human cortex

- Statement of the problem: capturing coarse- and fine-grained topographies in a common model
- Conceptual framework: high-dimensional representational spaces
- Deriving the common space and individual transformation matrices with hyperalignment
- Validation
- Connectivity hyperalignment
- Individual differences in fine-scale cortical functional architecture
- Conclusions

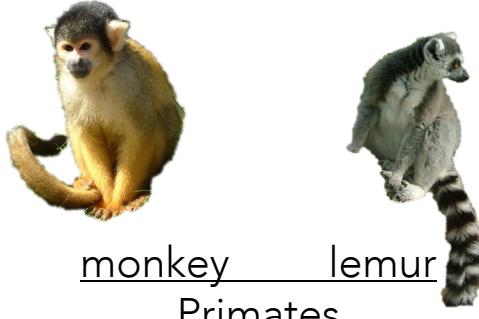
Multivariate Pattern Classification Example: Classifying responses during viewing of animal species (VT cortex, SVM)



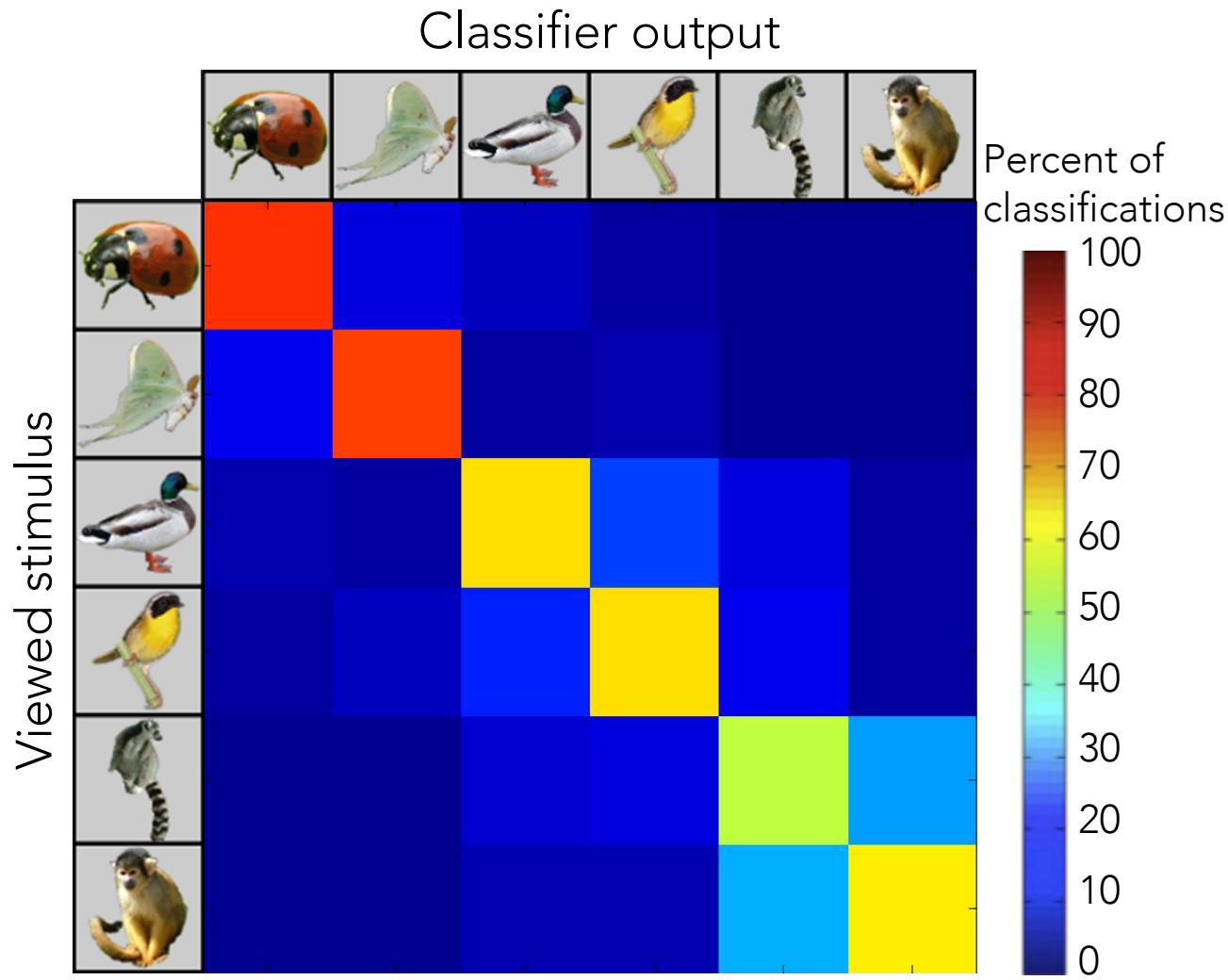
luna moth ladybug
Insects



warbler mallard
Birds

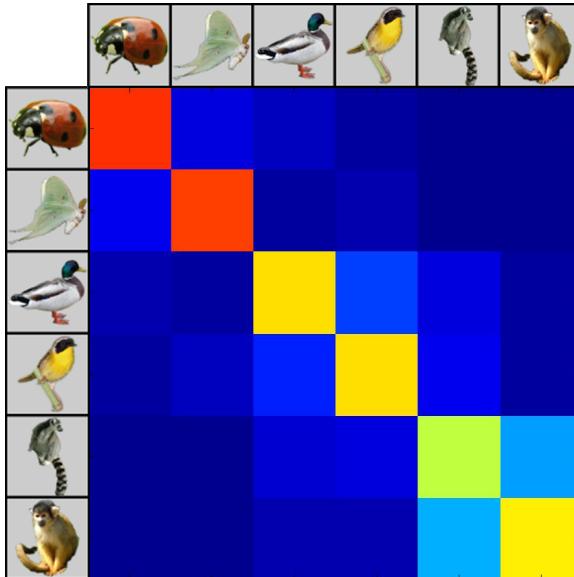


monkey lemur
Primates

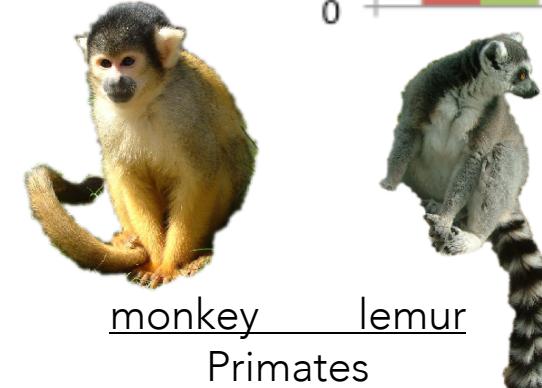
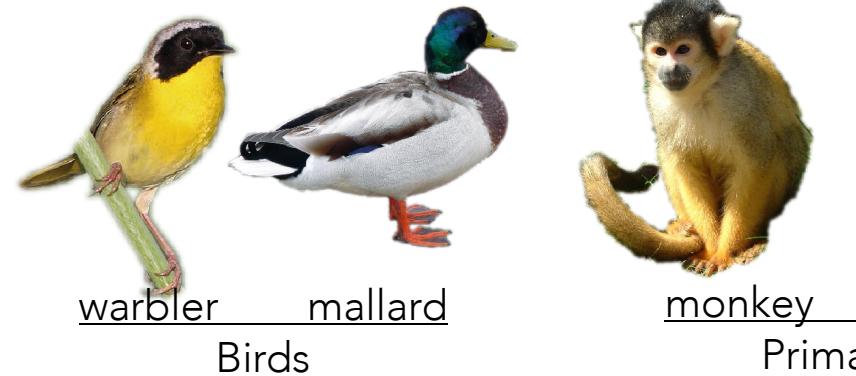
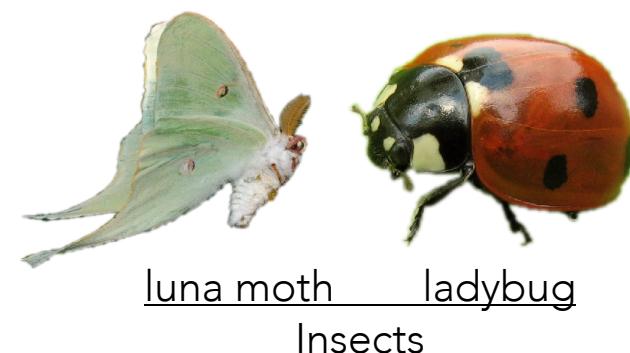
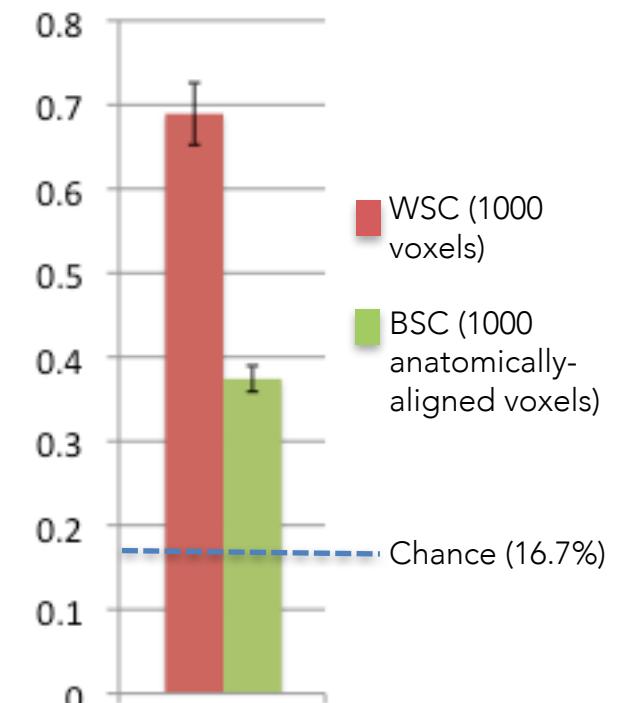
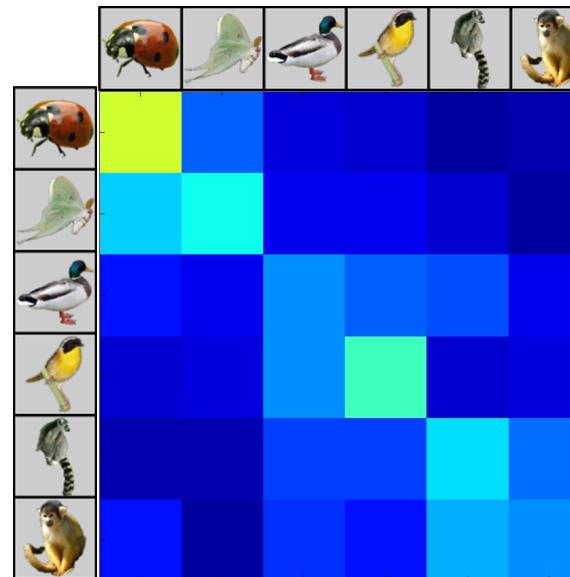


The problem: Loss of fine-grained distinctions among representations after anatomical alignment of brains

Within-subject classification
(new model for each subject)



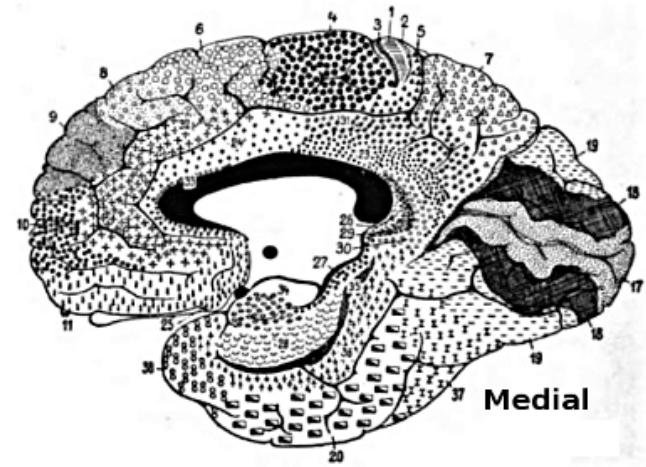
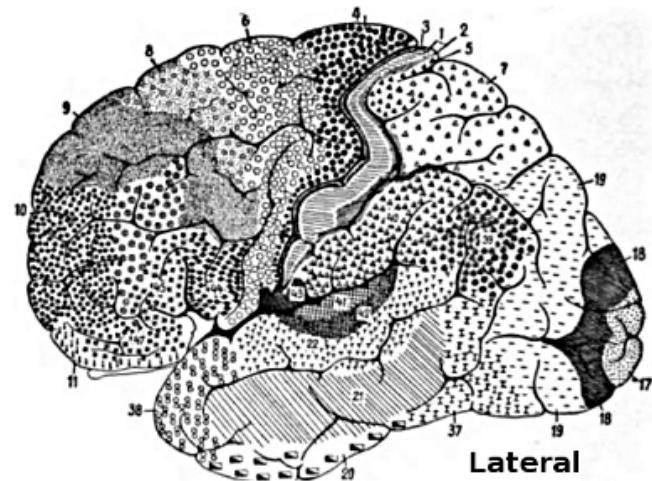
Between-subject classification
(common model based on anatomy)



Models of cortical architecture: parcellation of cortex into areas and systems

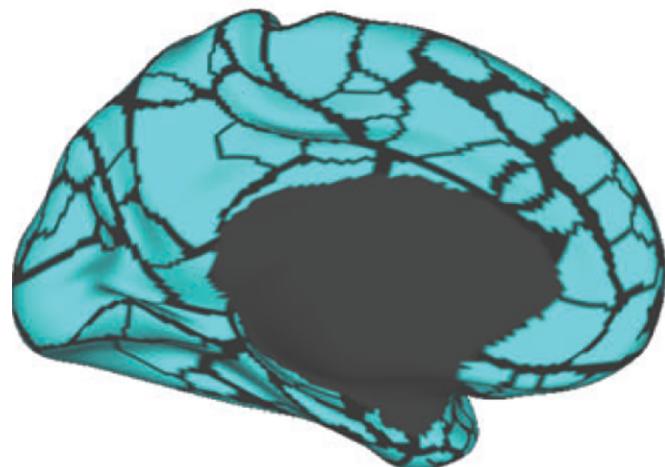
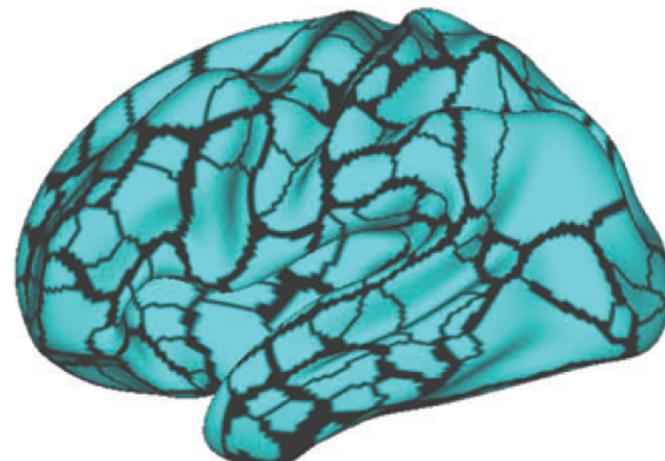
52 areas

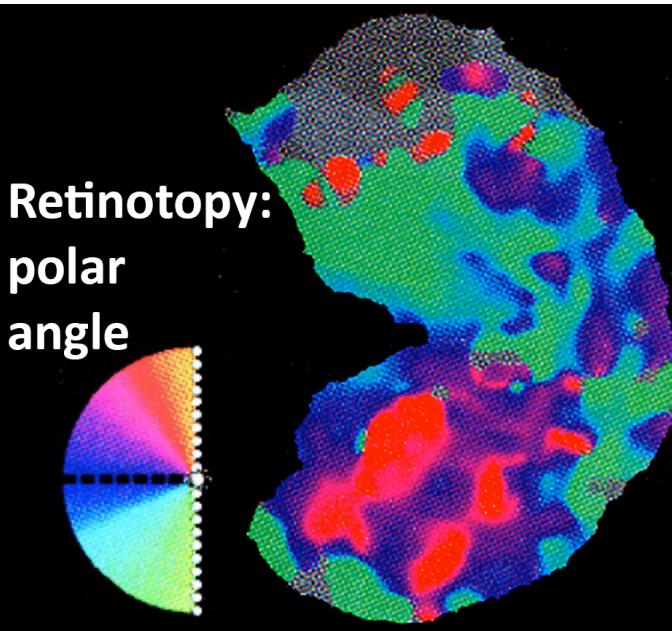
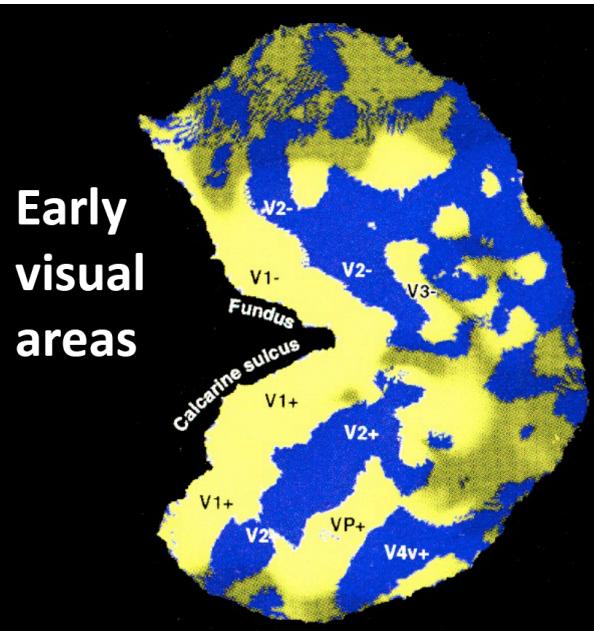
(Brodmann, 1909)



422 areas

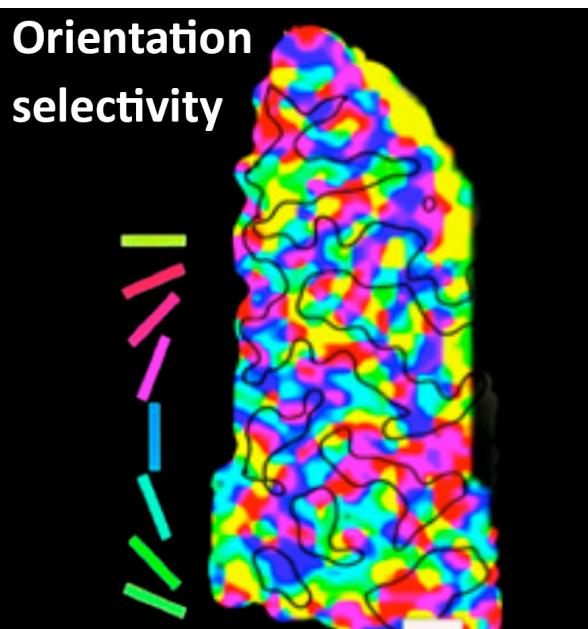
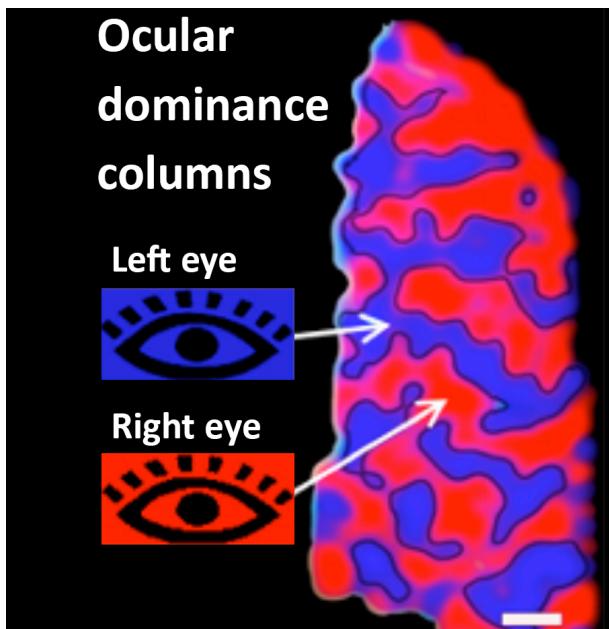
(Gordon et al. 2014)





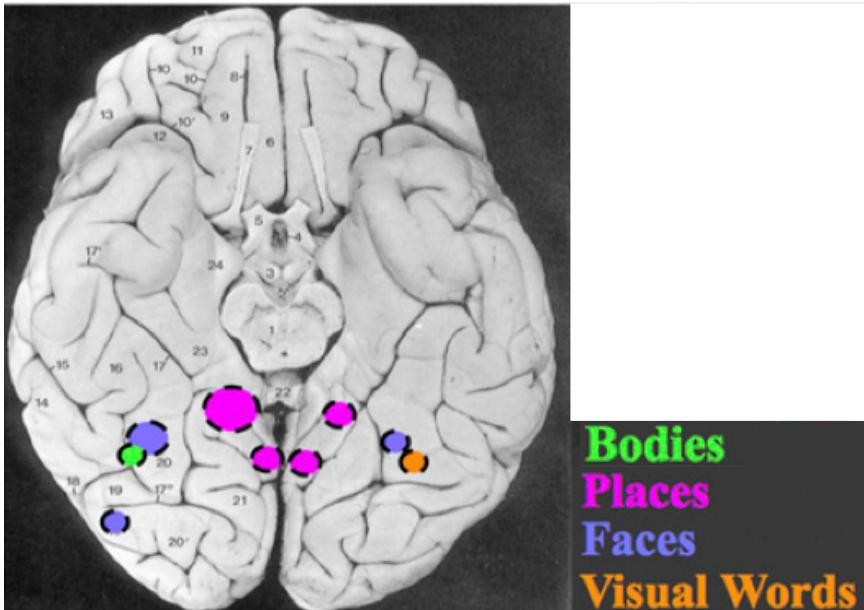
Known
topographies of
different scales in
early visual cortex

Sereno et al. 1995

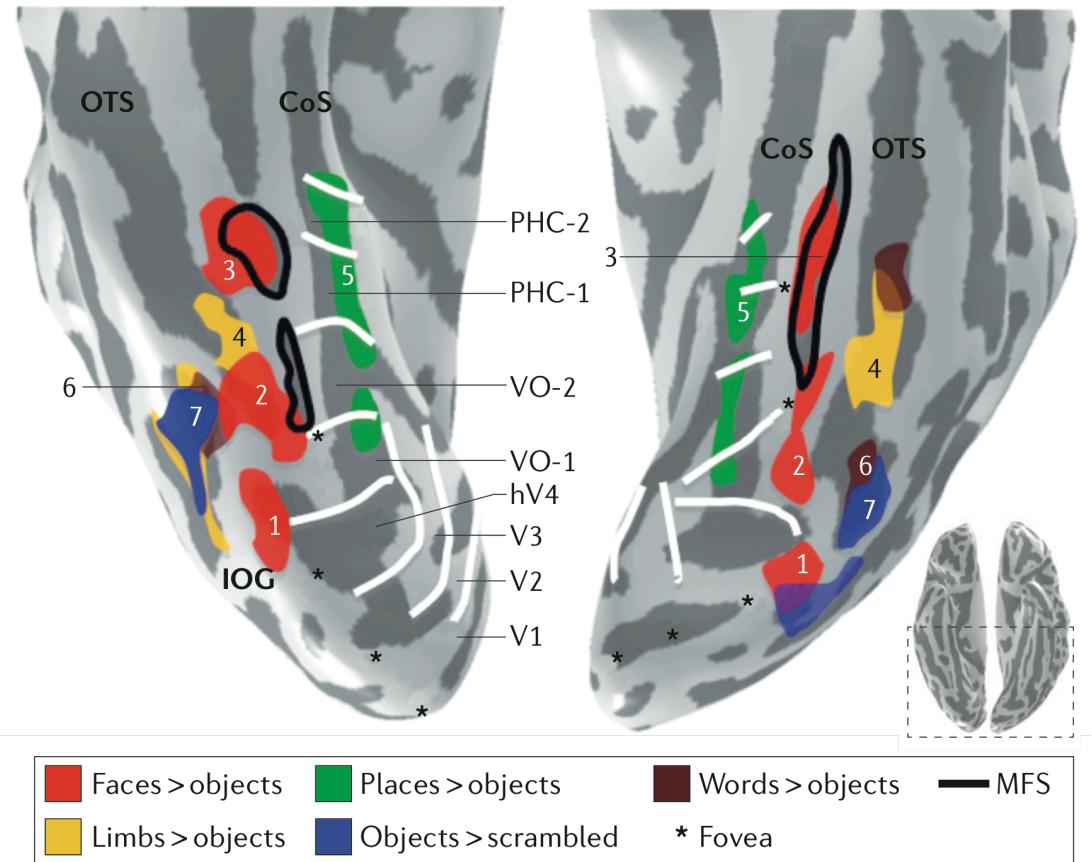


Yacoub et al. 2007, 2008

Category-selective regions in ventral temporal cortex also carry finer distinctions carried by fine-grained topographies

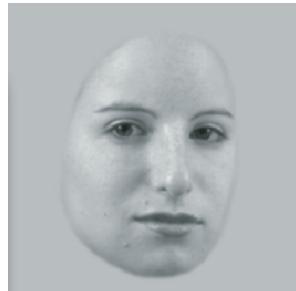


Kanwisher (2010)



Grill-Spector & Weiner (2014)

Category-selective regions in ventral temporal cortex
also carry finer distinctions carried by fine-grained topographies:
Distinctions that can be decoded from response patterns in the FFA



VS



VS



VS



VS



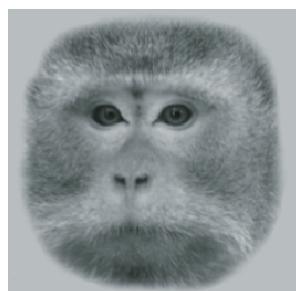
VS



VS



VS



Haxby et al. 2011; Connolly et al. 2012; Guntupalli et al. 2017

Question

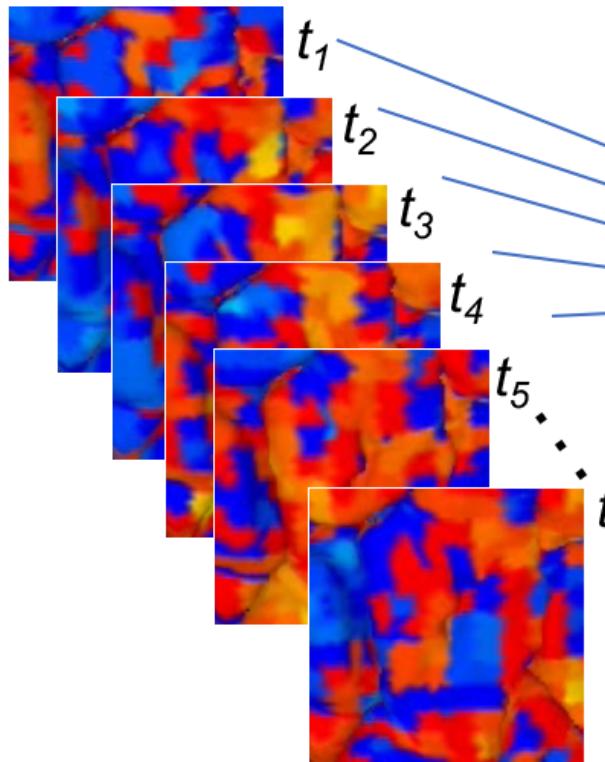
- Can the information that is encoded in fine-grained topographies be modeled in a computational framework that is common across brains?
 - i.e. Is there a common deep structure for these population responses, or does each brain develop an idiosyncratic code?

A common high-dimensional linear model of representational spaces in human cortex

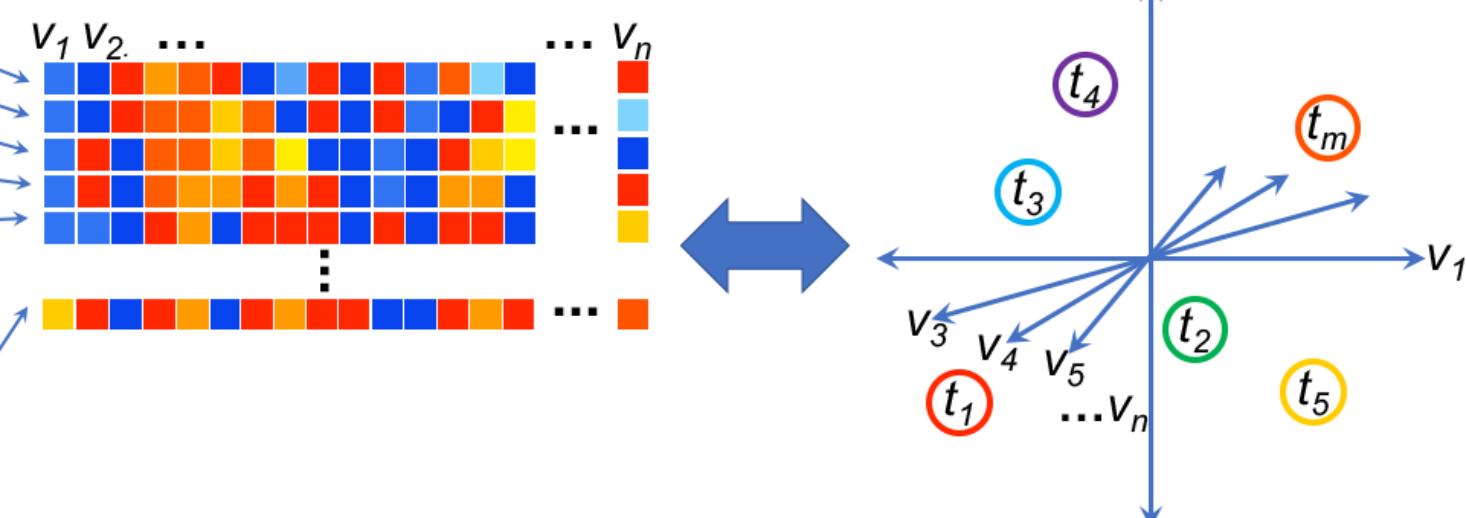
- Statement of the problem: capturing fine-grained distinctions in a common model
- Conceptual framework: high-dimensional representational spaces
 - A model space based on features (dimensions) with common tuning profiles
- Deriving the common space and individual transformation matrices with hyperalignment
- Validation
- Connectivity hyperalignment
- Individual differences in fine-scale cortical functional architecture
- Conclusions

Conceptual framework: High-dimensional information spaces

A. Cortical patterns

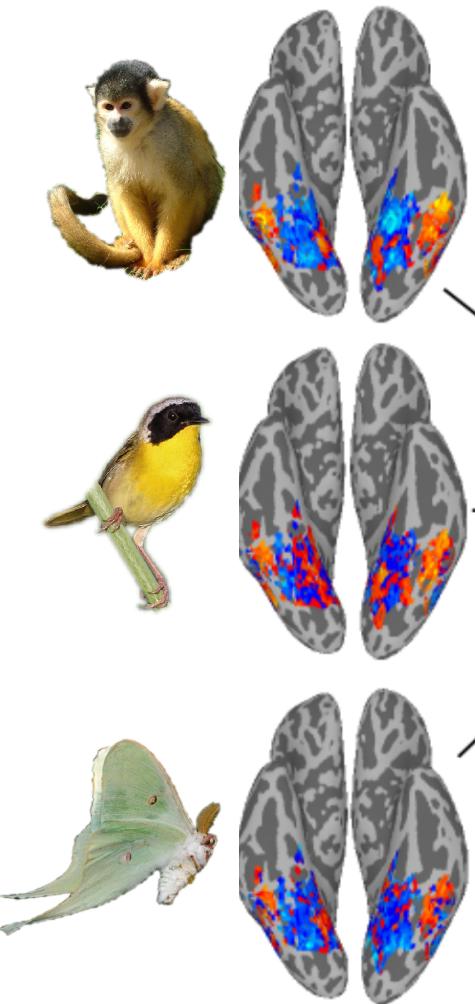


B. Individual data matrices and high-dimensional spaces

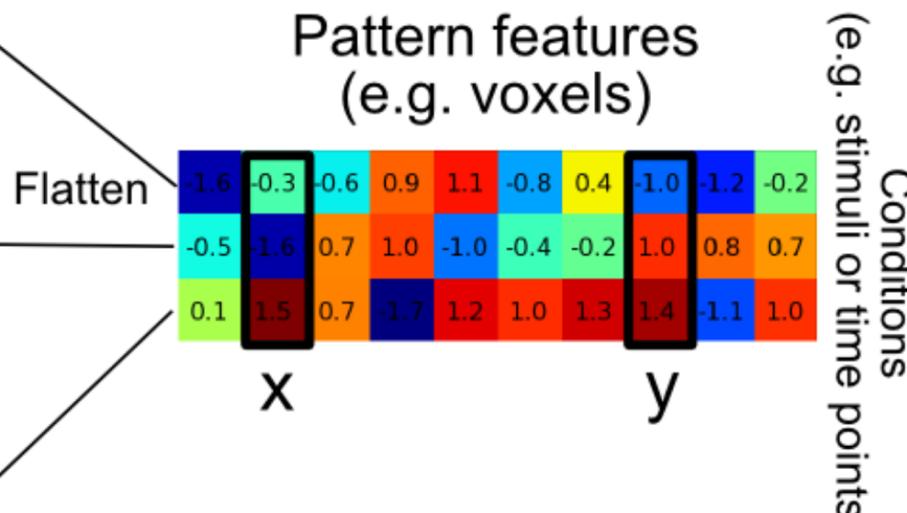


Conceptual framework: High-dimensional information spaces

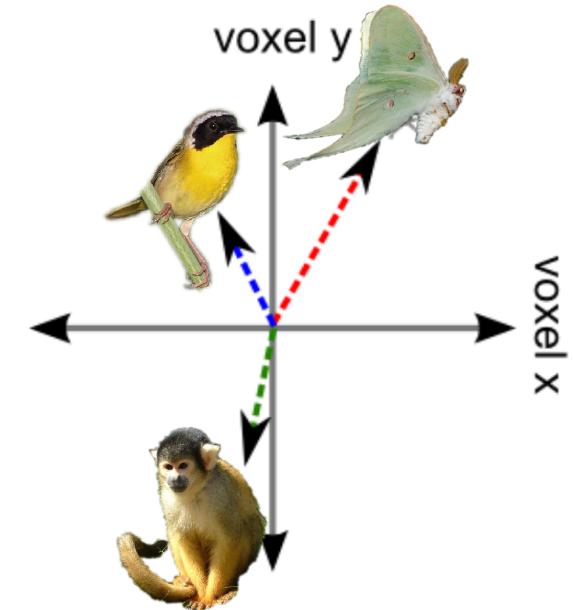
Brain activation patterns



Data matrix



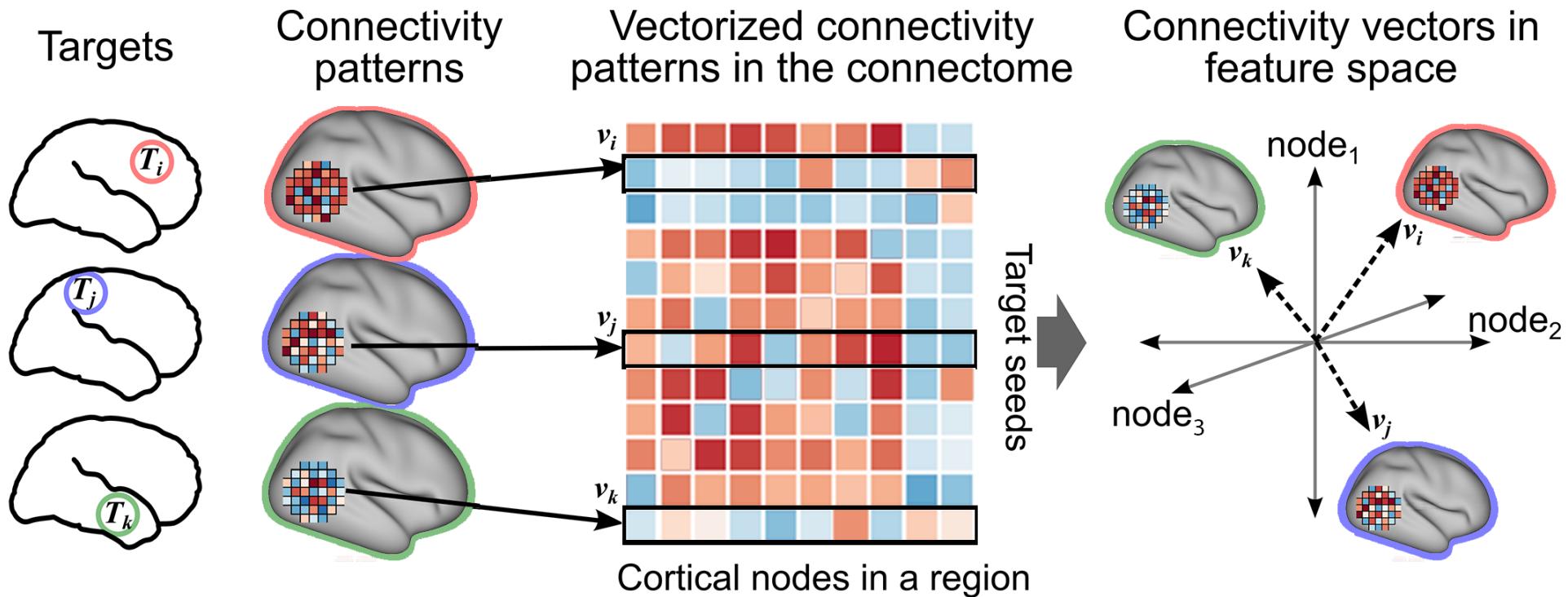
Information space



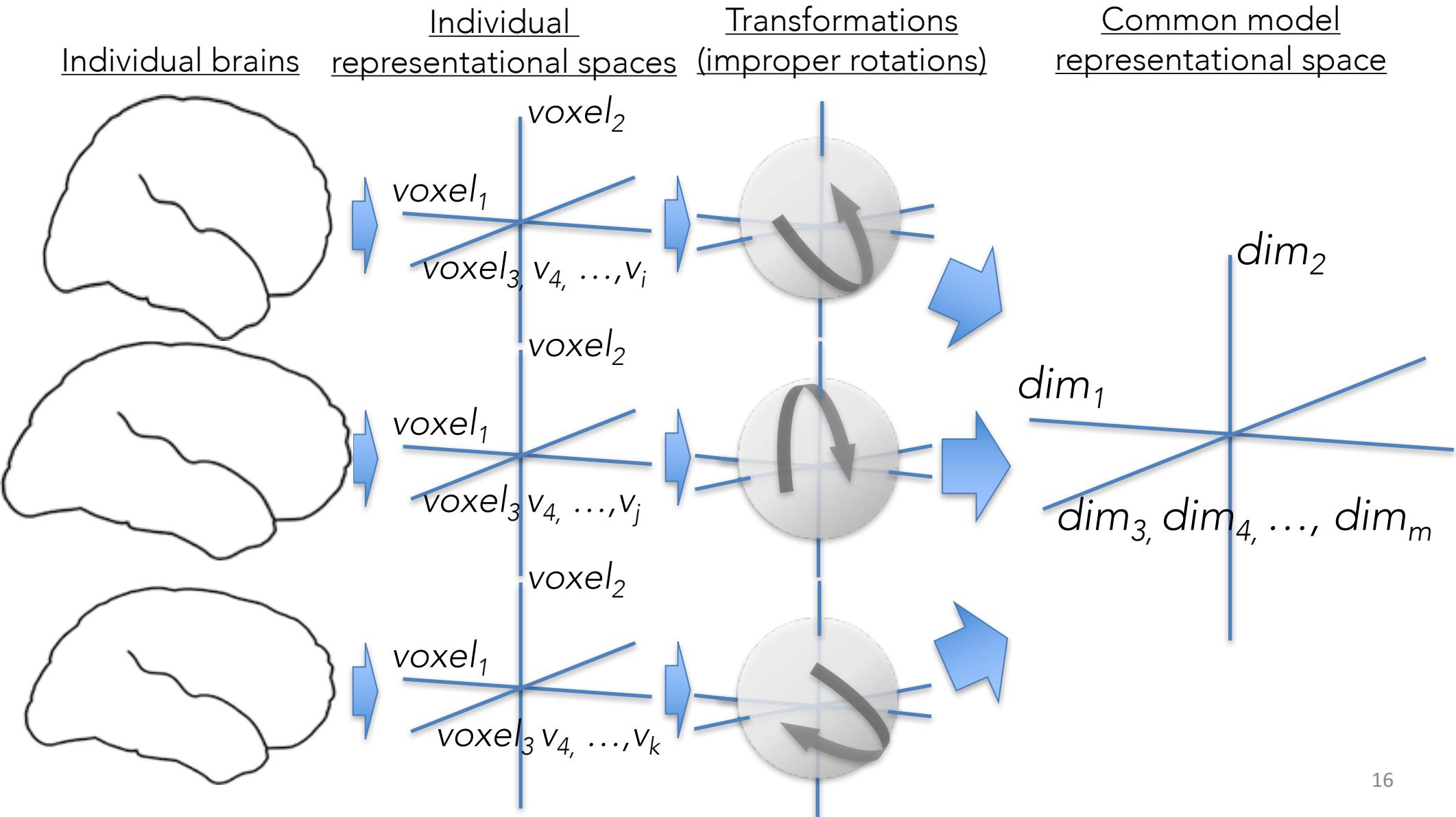
Conceptual framework: High-dimensional information spaces also can accommodate connectivity and network data

Connectivity vectors

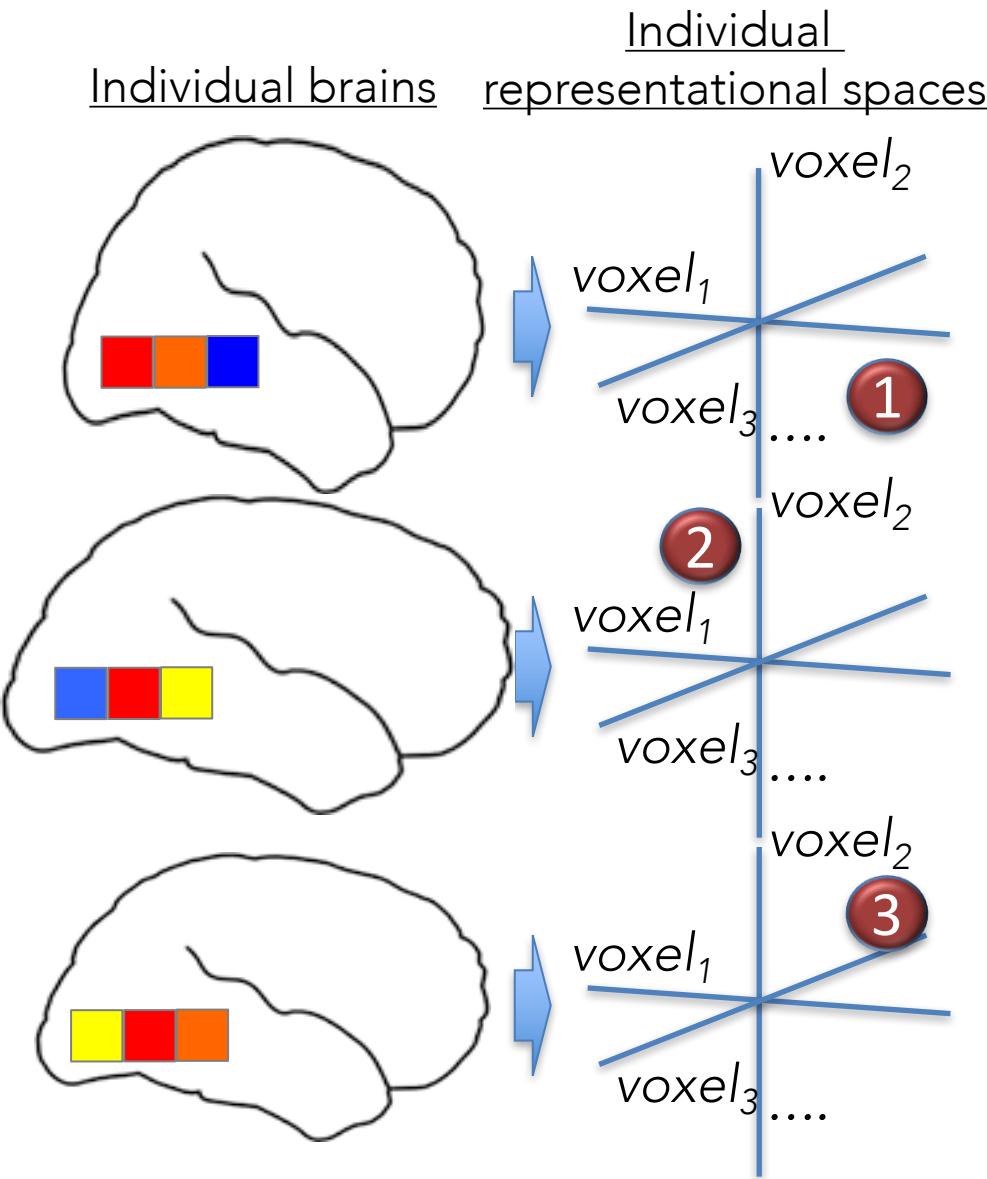
A



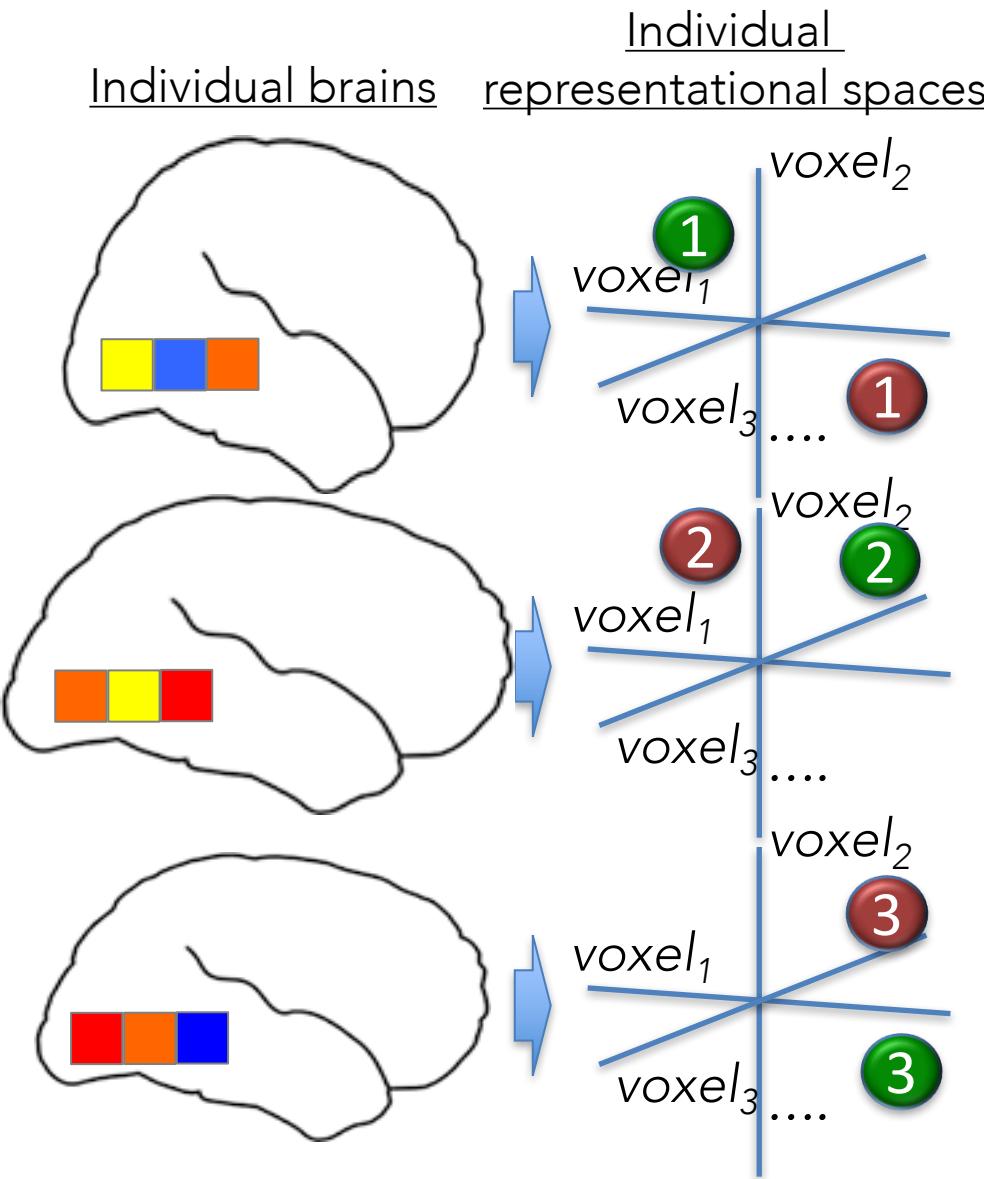
Modeling functional architecture of the human cortex: Individual representational spaces \Leftrightarrow common representational space



Modeling functional architecture of the human cortex: Individual representational spaces \Leftrightarrow common representational space



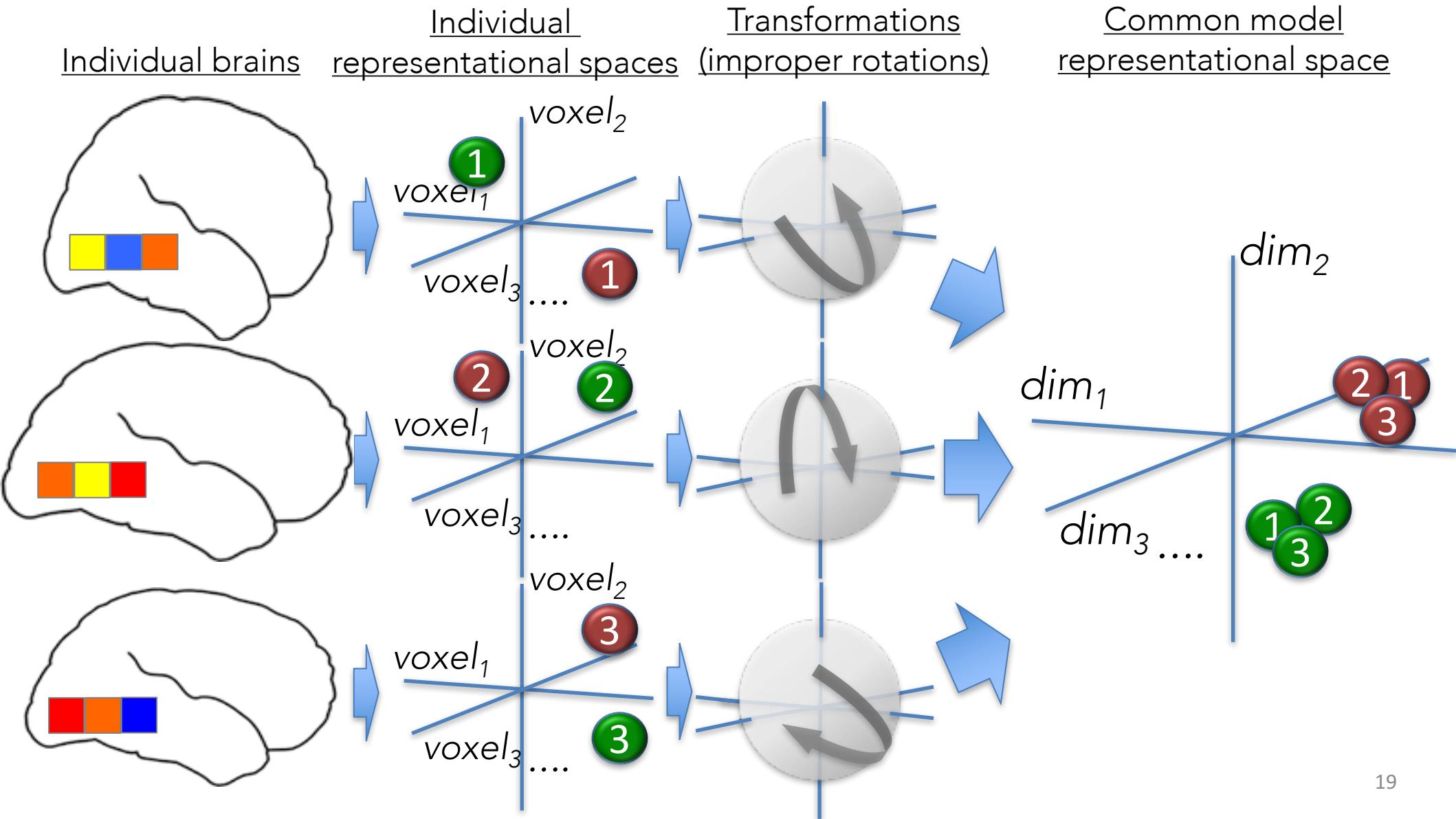
Modeling functional architecture of the human cortex: Individual representational spaces \Leftrightarrow common representational space



Another stimulus

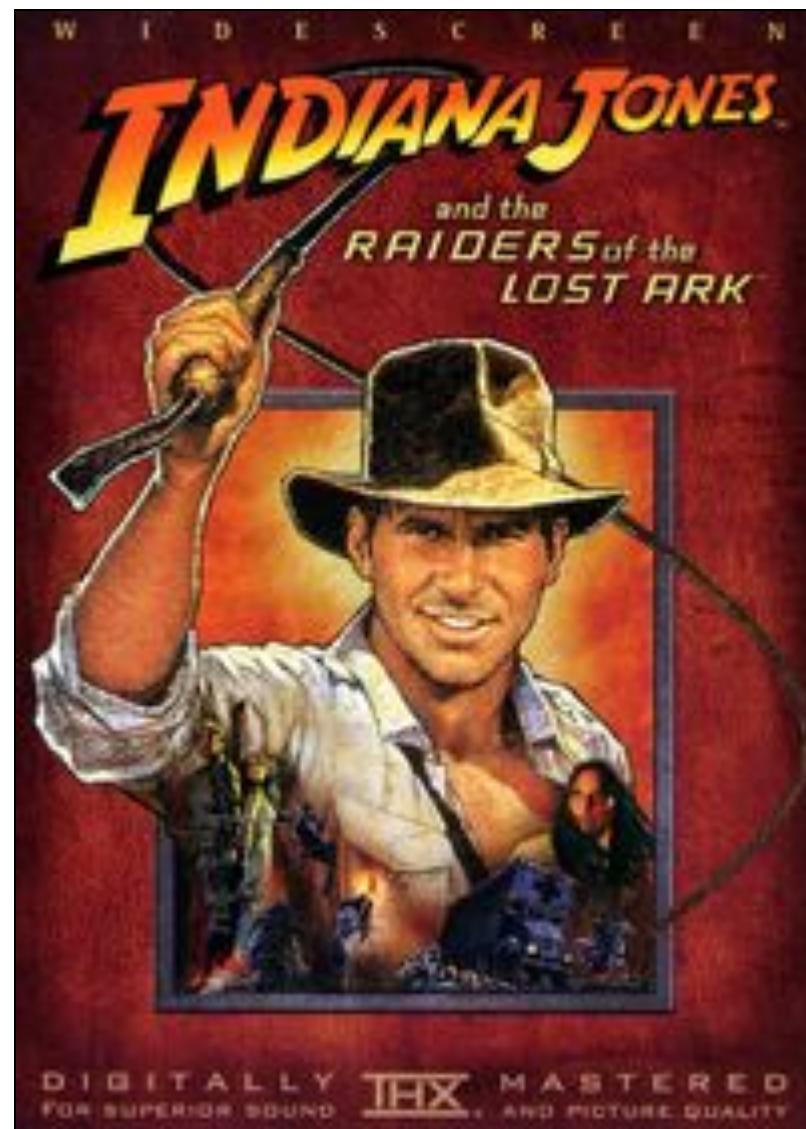


Modeling functional architecture of the human cortex: Individual representational spaces \Leftrightarrow common representational space



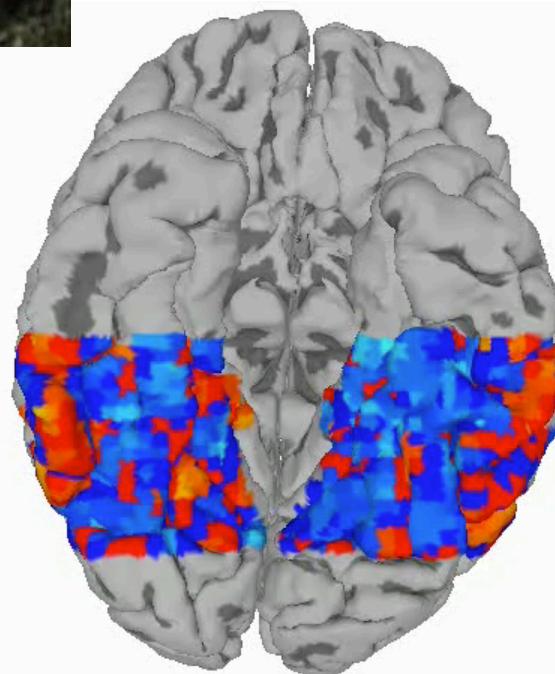
A common high-dimensional linear model of representational spaces in human cortex

- Statement of the problem: capturing fine-grained distinctions in a common model
- Conceptual framework: high-dimensional representational spaces
- Deriving the common space and individual transformation matrices with hyperalignment
 - Hyperalignment algorithm based on Procrustes transformations
 - A rich sampling of response vectors using natural stimulus
- Validation
- Connectivity hyperalignment
- Individual differences in fine-scale cortical functional architecture
- Conclusions

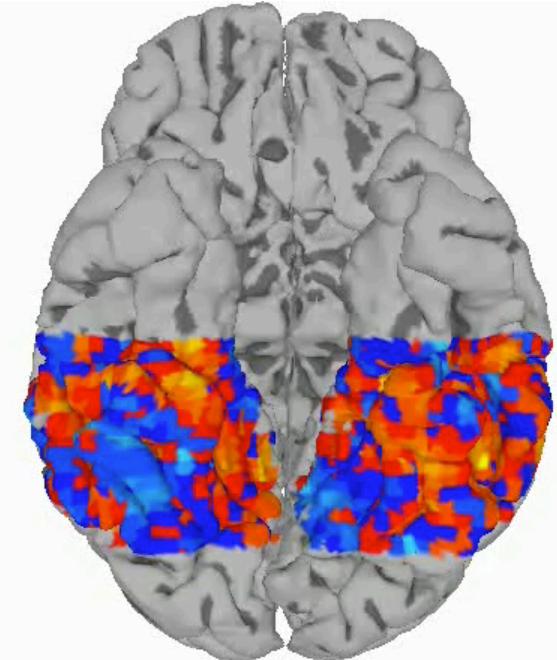




Subject 1



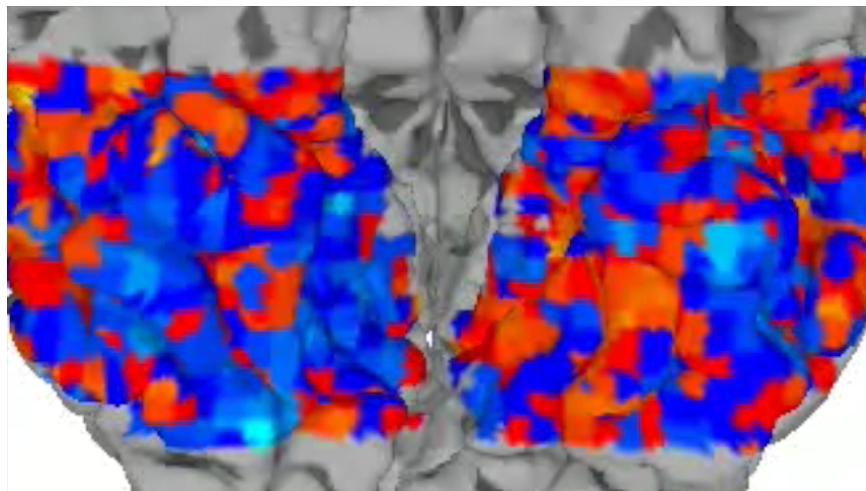
Subject 2



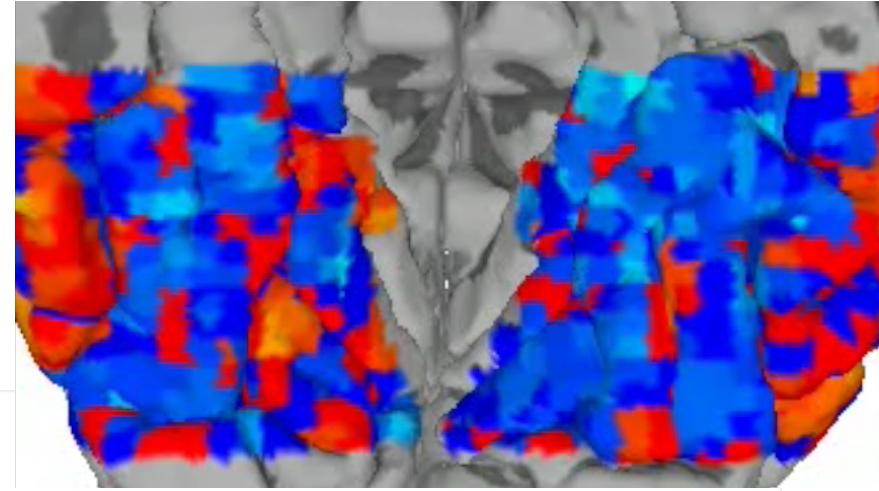
Broad sampling of a neural representational space with a movie

Response patterns in cortex

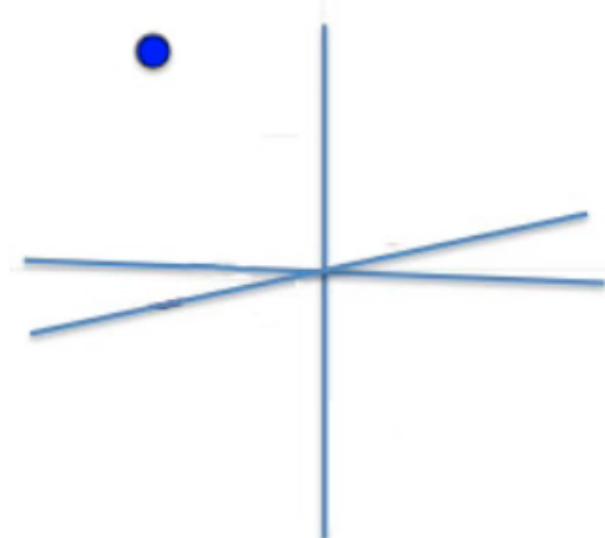
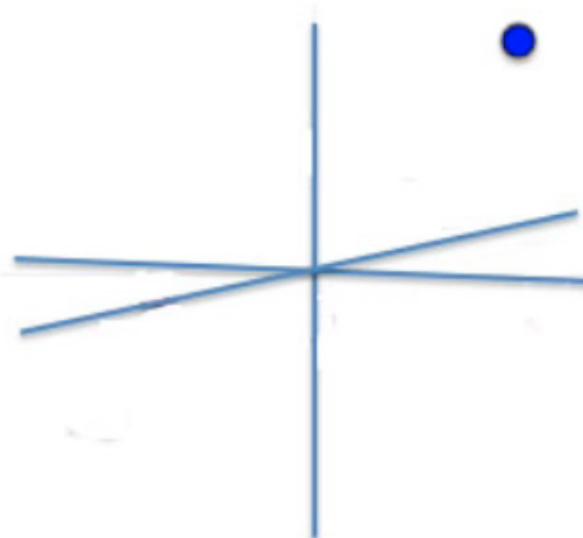
Subject 1



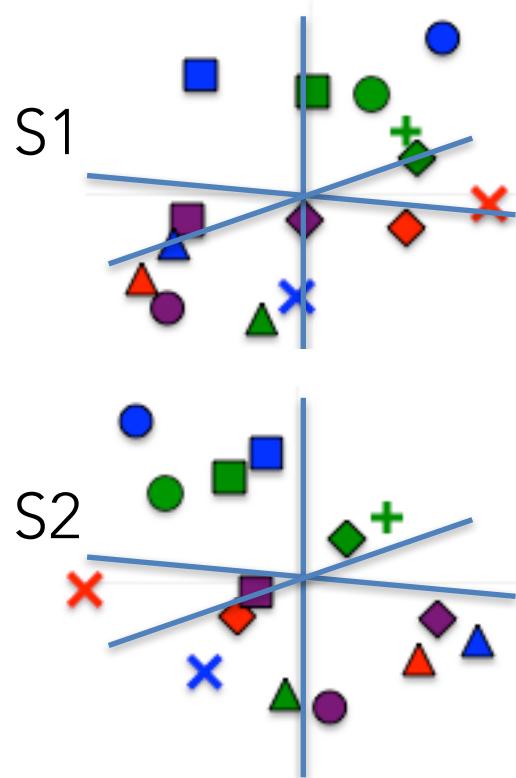
Subject 2



15 response pattern vectors in individual 3D representational spaces
(full exp't has >2600 vectors in >50,000D space)

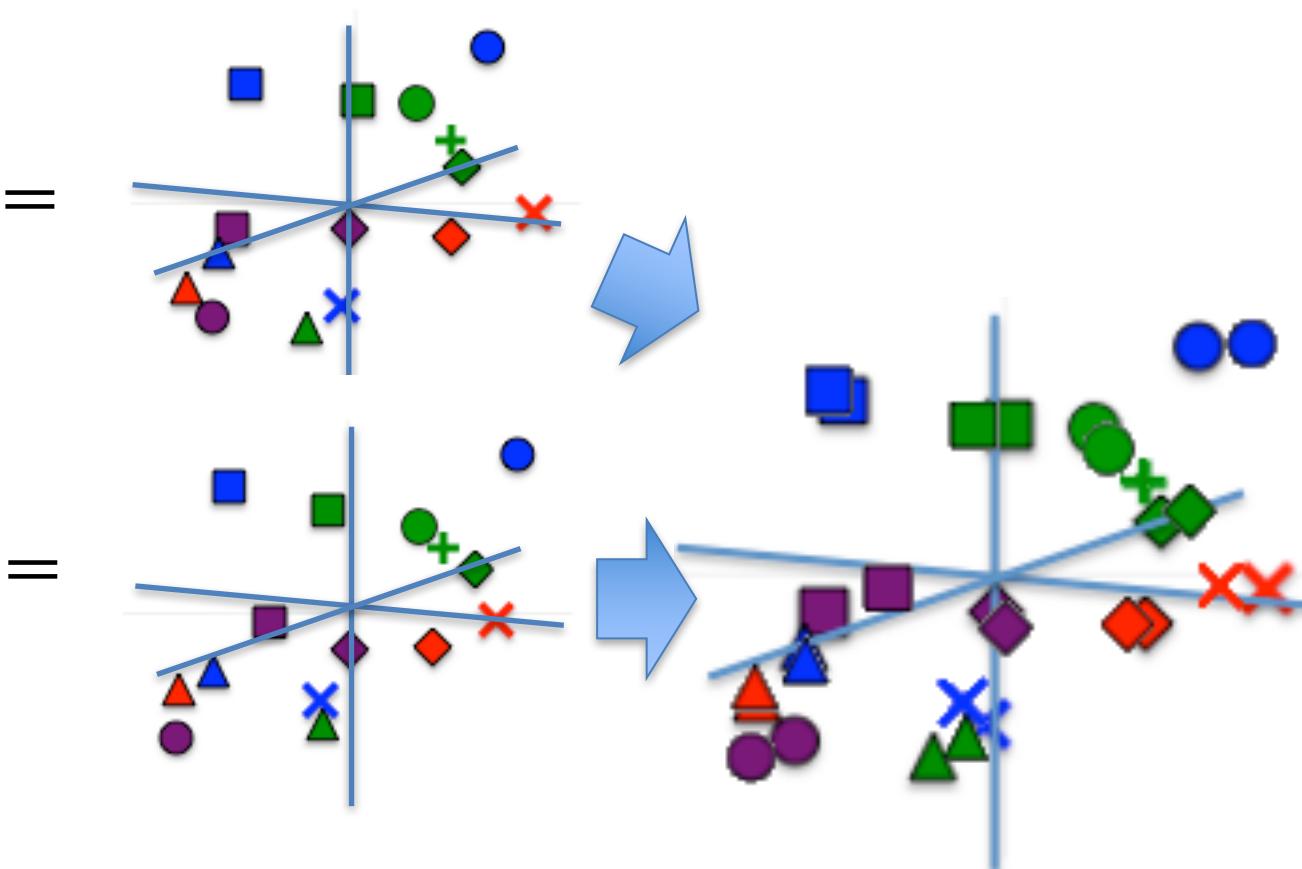


Individual
representational spaces



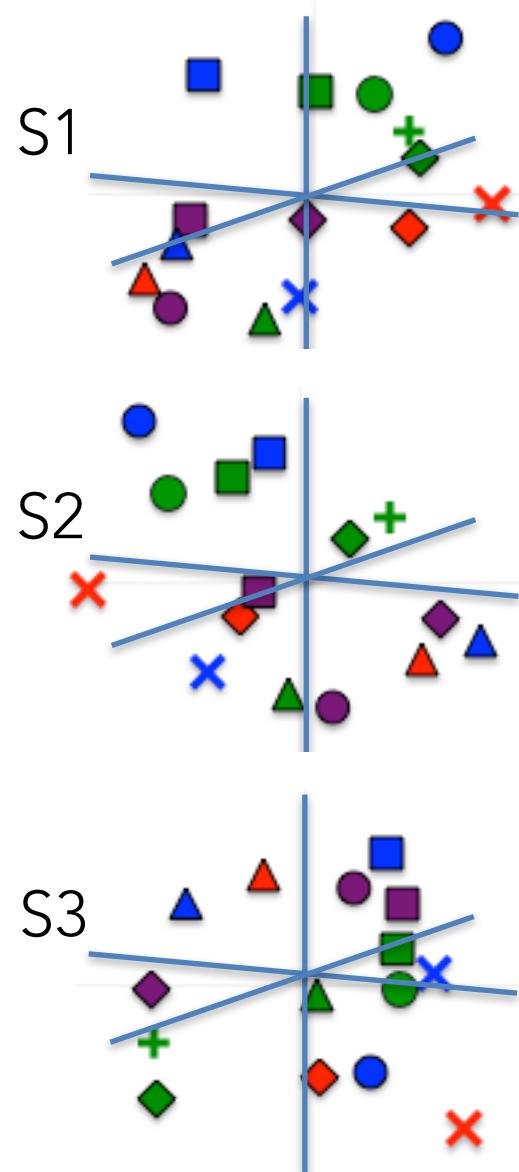
Procrustes transformations
(improper rotations)

$$X [] =$$

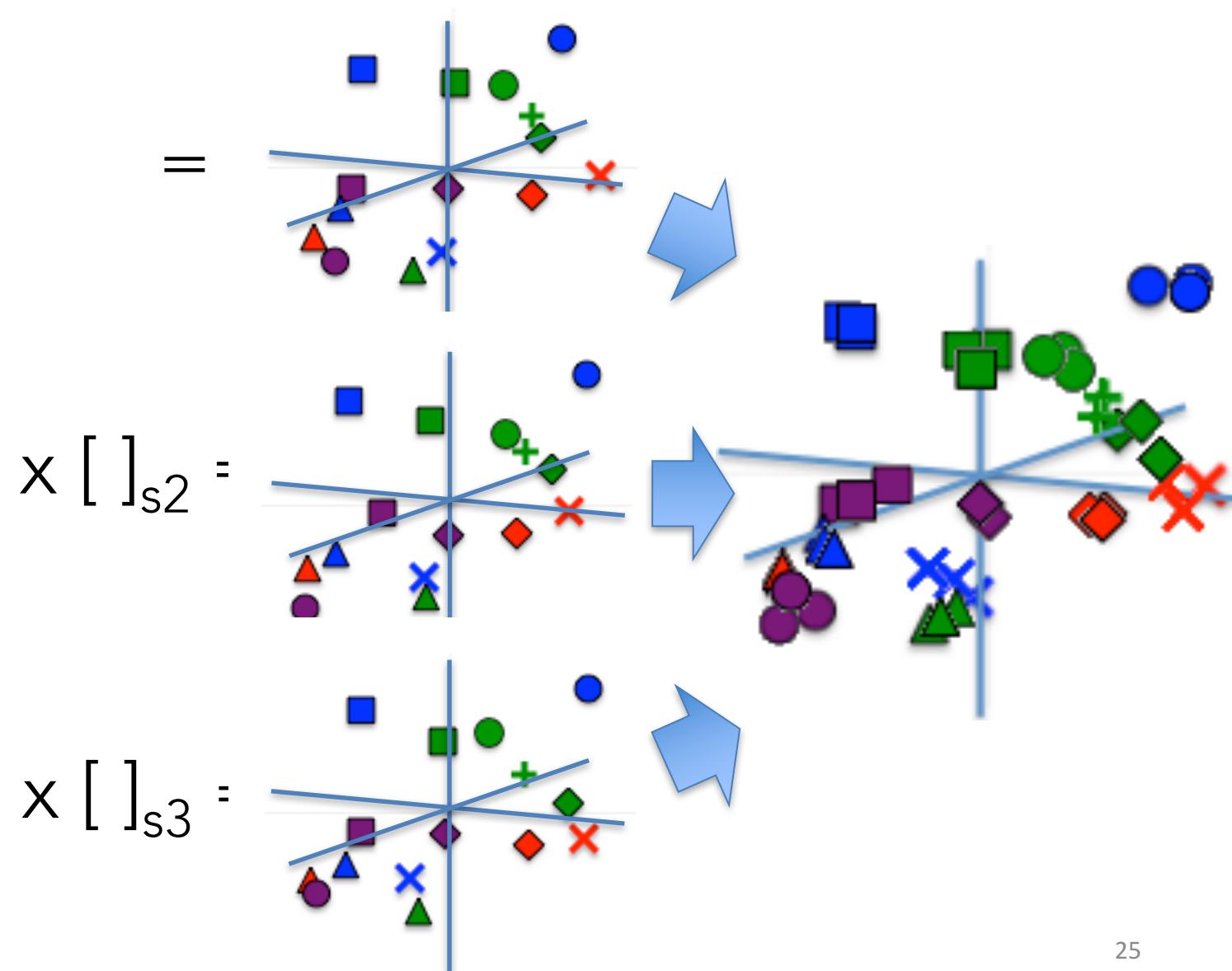


Common model
representational space

Individual
representational spaces



Procrustes transformations
(improper rotations)



Common model
representational space

Modeling representational spaces in human cortex with hyperalignment

Basic equation:

$$M_j = (1/N) \sum_{i=1}^N (B_{ij} R_{ij})$$

minimizing:

$$R_{ij} = \operatorname{argmin}_R \sum_{i=1}^N \|B_{ij} R_{ij} - M_j\|_F$$

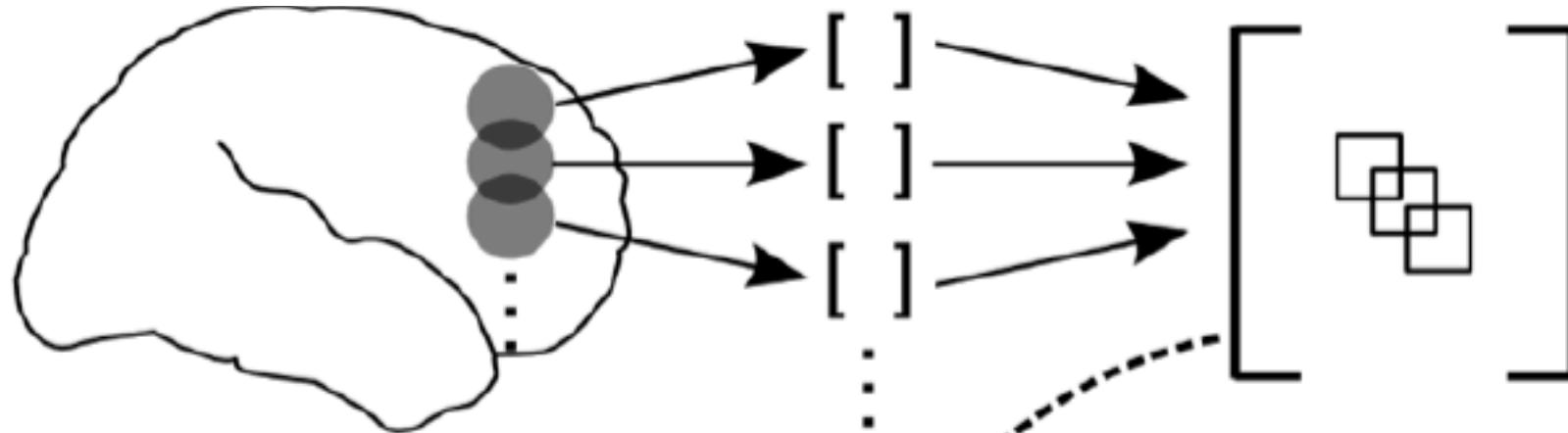
iteratively using:

- Procrustes transformations (Haxby et al. 2011, Guntupalli et al. 2016)
- Regularized CCA (Xu, Lorbert et al. 2012)
- Probabalistic hyperalignment (Chen et al. 2015)

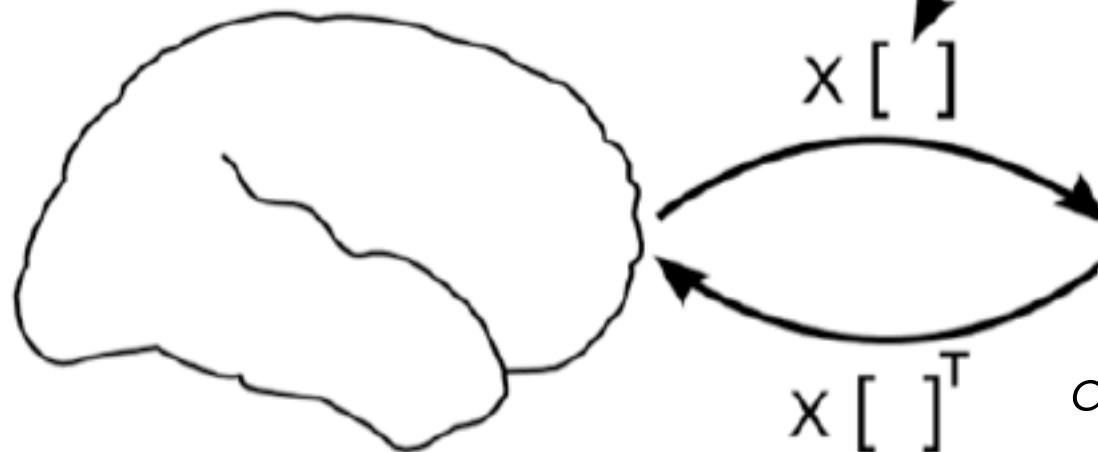
Modeling representational spaces in all human cortex with searchlight hyperalignment

Voxels in overlapping searchlights are hyperaligned across subjects

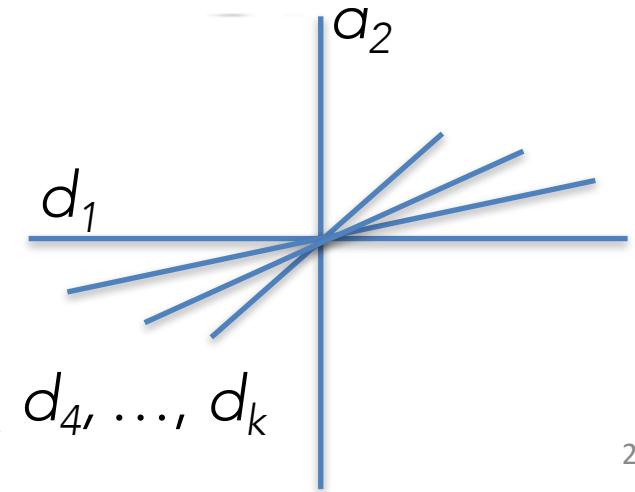
Overlapping searchlight transformation matrices are aggregated into a whole cortex matrix



Data in individual brain anatomy



Data in common model space



A common high-dimensional linear model of representational spaces in human cortex

- Statement of the problem: capturing fine-grained distinctions in a common model
- Conceptual framework: high-dimensional representational spaces
- Deriving the common space and individual transformation matrices with hyperalignment
- Validation
 - Between-subject classification of movie time segments
 - Between-subject correlations of local similarity structures
 - Estimation spatial granularity of common model features
- Connectivity hyperalignment
- Individual differences in fine-scale cortical functional architecture
- Conclusions

Data in subject brain voxel x time-points data matrix

$$\begin{array}{c}
 \text{Time-points} \\
 | \\
 \begin{array}{c}
 \text{Voxels} \\
 | \\
 \mathbf{B} \quad v_1 \quad v_2 \quad v_3 \quad \dots \quad v_l \\
 | \\
 t_1 \quad x_{1,1} \quad x_{2,1} \quad x_{3,1} \quad \dots \quad x_{l,1} \\
 t_2 \quad x_{1,2} \quad x_{2,2} \quad x_{3,2} \quad \dots \quad x_{l,2} \\
 t_3 \quad x_{1,3} \quad x_{2,3} \quad x_{3,3} \quad \dots \quad x_{l,3} \\
 t_4 \quad x_{1,4} \quad x_{2,4} \quad x_{3,4} \quad \dots \quad x_{l,4} \\
 \dots \quad \dots \quad \dots \quad \dots \quad \dots \\
 t_j \quad x_{1,j} \quad x_{2,j} \quad x_{3,j} \quad \dots \quad x_{l,j}
 \end{array}
 \end{array}
 \times
 \begin{array}{c}
 \text{Voxels} \\
 | \\
 \mathbf{R} \quad d_1 \quad d_2 \quad d_3 \quad \dots \quad d_m \\
 | \\
 v_1 \quad w_{1,1} \quad w_{2,1} \quad w_{3,1} \quad \dots \quad w_{m,1} \\
 v_2 \quad w_{1,2} \quad w_{2,2} \quad w_{3,2} \quad \dots \quad w_{m,2} \\
 v_3 \quad w_{1,3} \quad w_{2,3} \quad w_{3,3} \quad \dots \quad w_{m,3} \\
 \dots \quad \dots \quad \dots \quad \dots \quad \dots \\
 v_l \quad w_{1,l} \quad w_{2,l} \quad w_{3,l} \quad \dots \quad w_{m,l}
 \end{array}
 =
 \begin{array}{c}
 \text{Time-points} \\
 | \\
 \mathbf{M} \quad d_1 \quad d_2 \quad d_3 \quad \dots \quad d_m \\
 | \\
 t_1 \quad y_{1,1} \quad y_{2,1} \quad y_{3,1} \quad \dots \quad y_{m,1} \\
 t_2 \quad y_{1,2} \quad y_{2,2} \quad y_{3,2} \quad \dots \quad y_{m,2} \\
 t_3 \quad y_{1,3} \quad y_{2,3} \quad y_{3,3} \quad \dots \quad y_{m,3} \\
 t_4 \quad y_{1,4} \quad y_{2,4} \quad y_{3,4} \quad \dots \quad y_{m,4} \\
 \dots \quad \dots \quad \dots \quad \dots \quad \dots \\
 T_j \quad y_{1,j} \quad y_{2,j} \quad y_{3,j} \quad \dots \quad y_{m,j}
 \end{array}$$

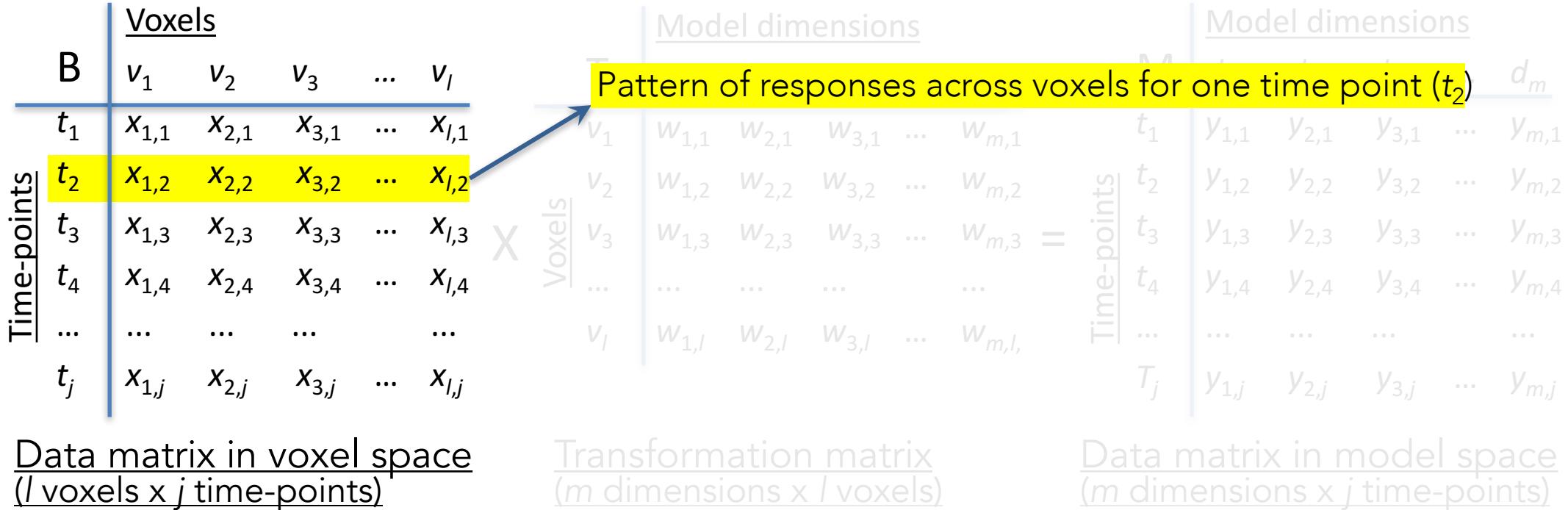
X

Data matrix in voxel space
(l voxels x j time-points)

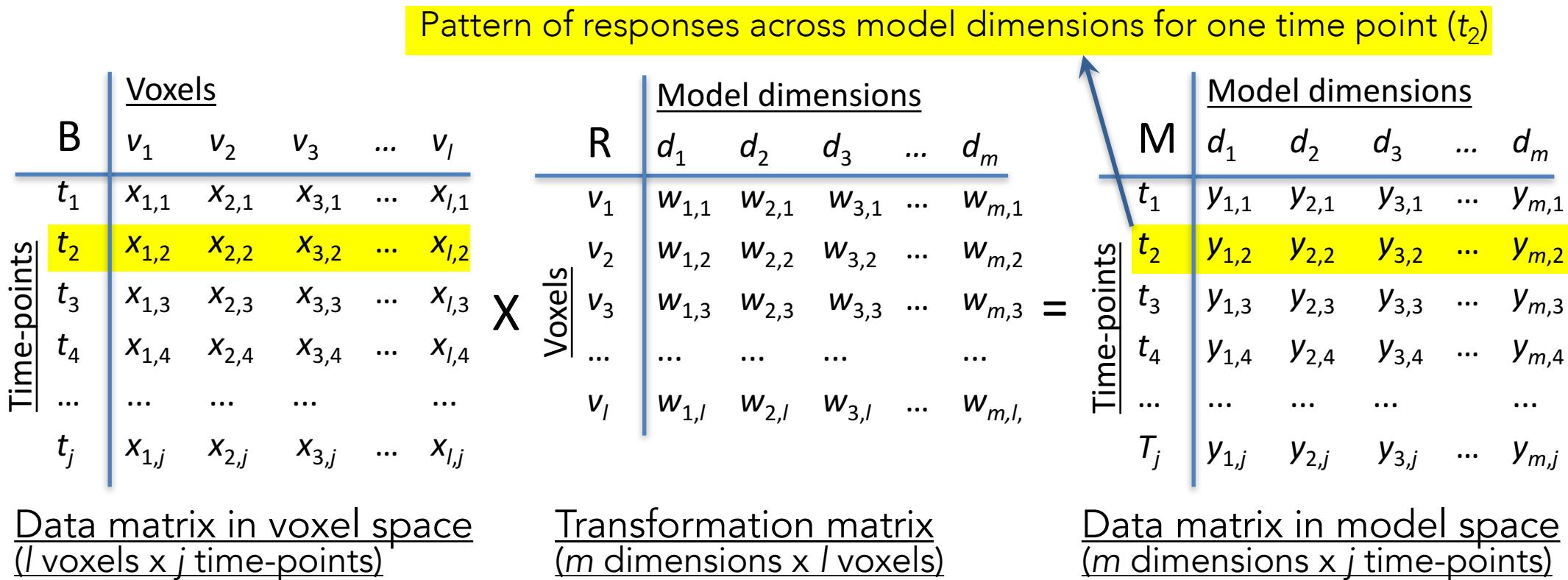
Transformation matrix
(m dimensions x l voxels)

Data matrix in model space
(m dimensions x j time-points)

Data in subject brain voxel x time-points data matrix

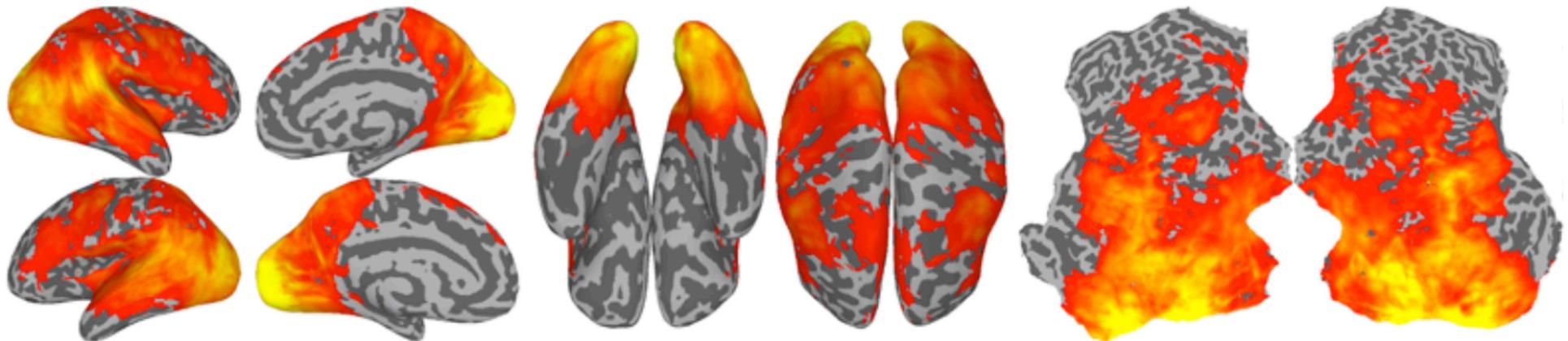


Transforming patterns and time-series into common model space

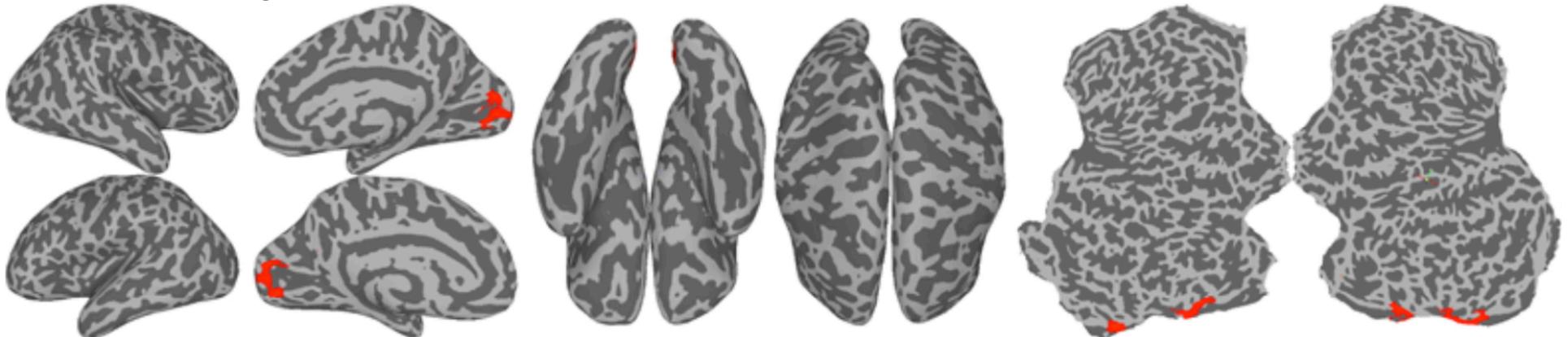


Whole-brain hyperalignment increases between-subject classification of 15 s movie time segments in occipital, temporal, parietal, and frontal cortices

Hyperalignment



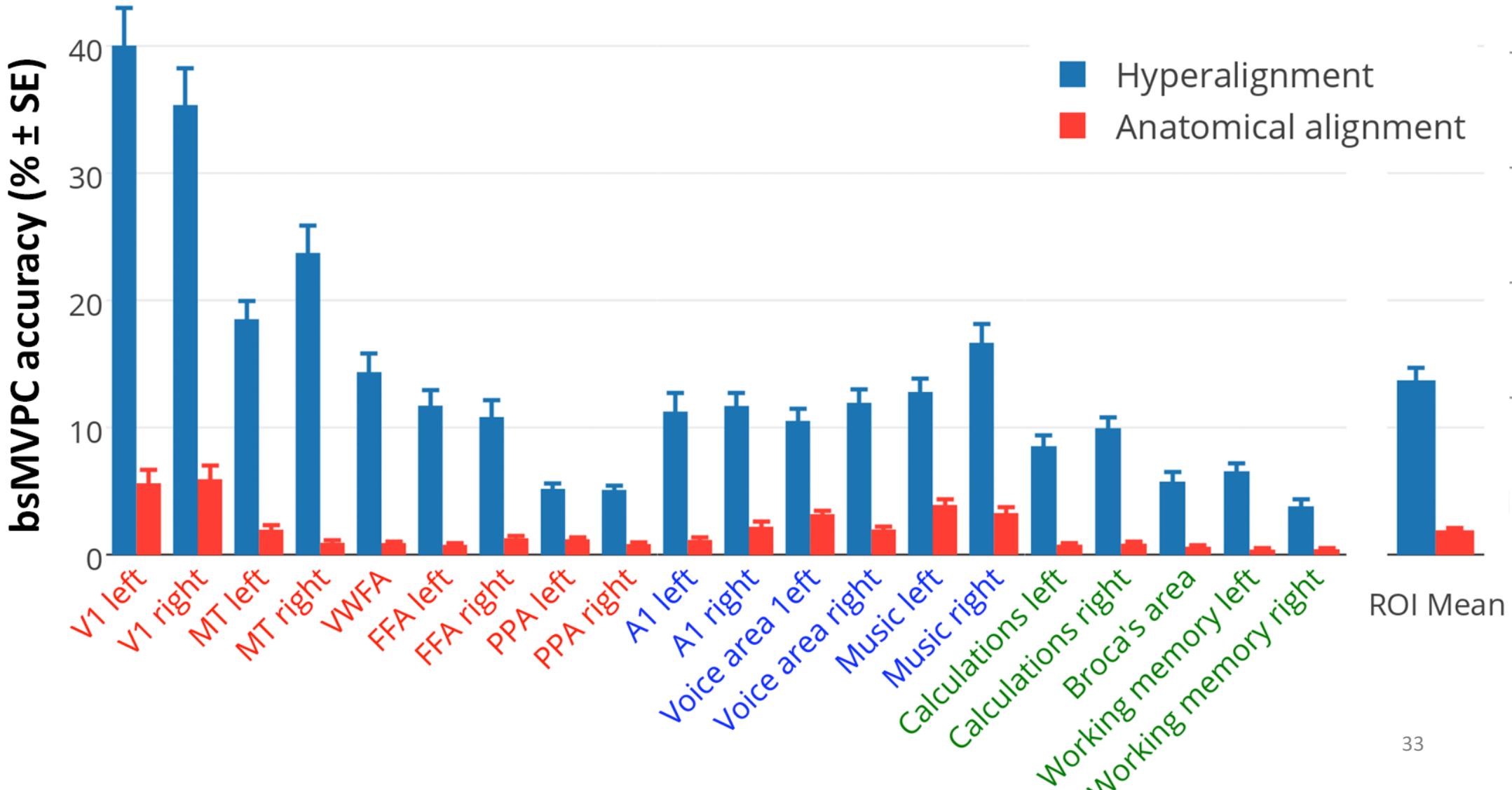
Anatomical alignment



Classification accuracy (%)



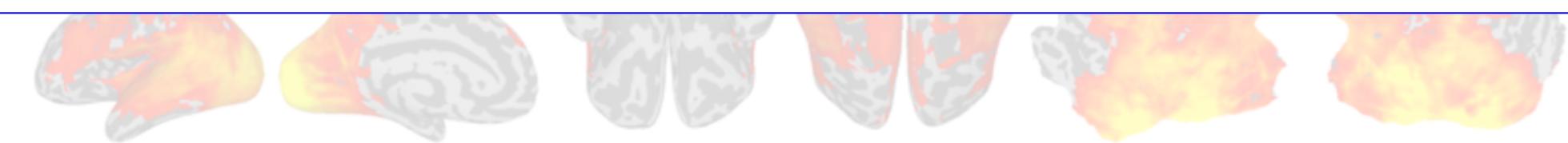
Increased bsMVPc of movie time-segments in **visual**, **auditory**, and **cognitive** regions of interest (ROIs) (coordinates from NeuroSynth)



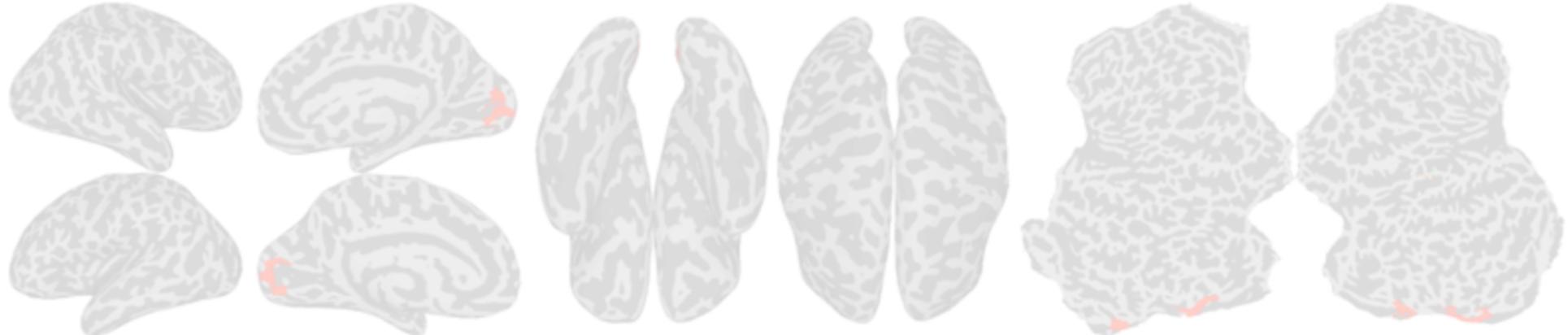
Whole-brain hyperalignment increases between-subject classification of 15 s movie time segments in occipital, temporal, parietal, and frontal cortices

Hyperalignment

Shared response patterns in common model dimensions
afford >7X higher bsMVPC accuracies



Anatomical alignment



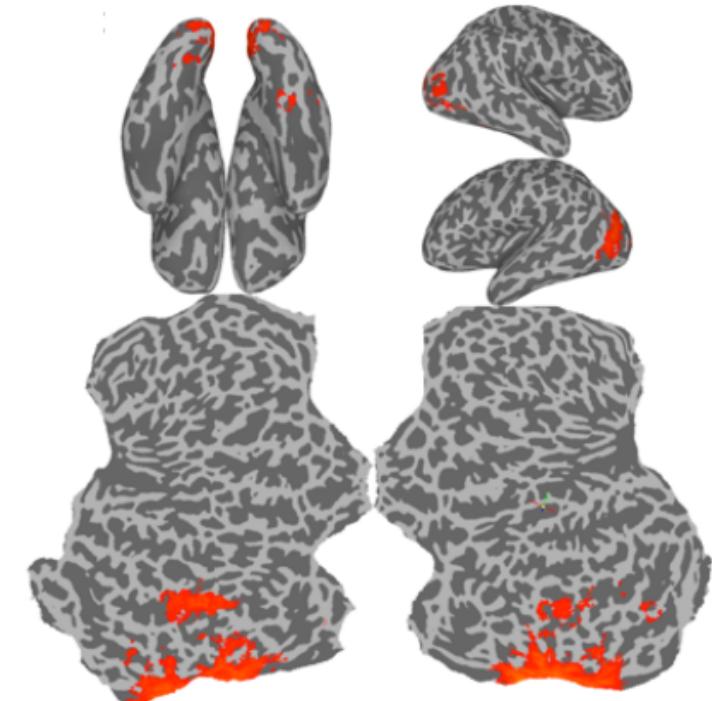
Classification accuracy (%)



Whole-brain hyperalignment based on movie affords between-subject classification of responses in a visual category experiment (6 animal species) at levels of accuracy that exceed within-subject classification

Between-subject classification

Anatomical alignment



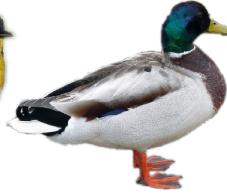
luna moth



ladybug

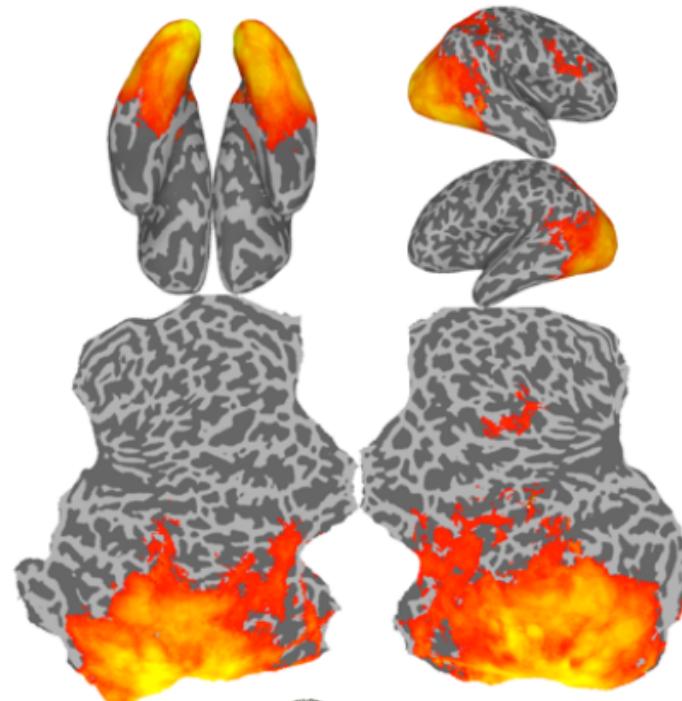


warbler



mallard

Hyperalignment

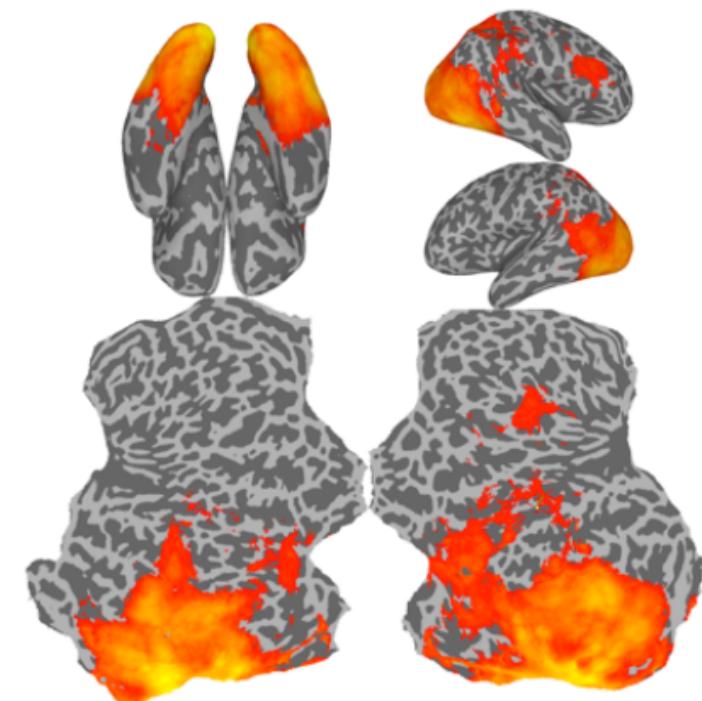


monkey



lemur
Primates

Within-subject classification



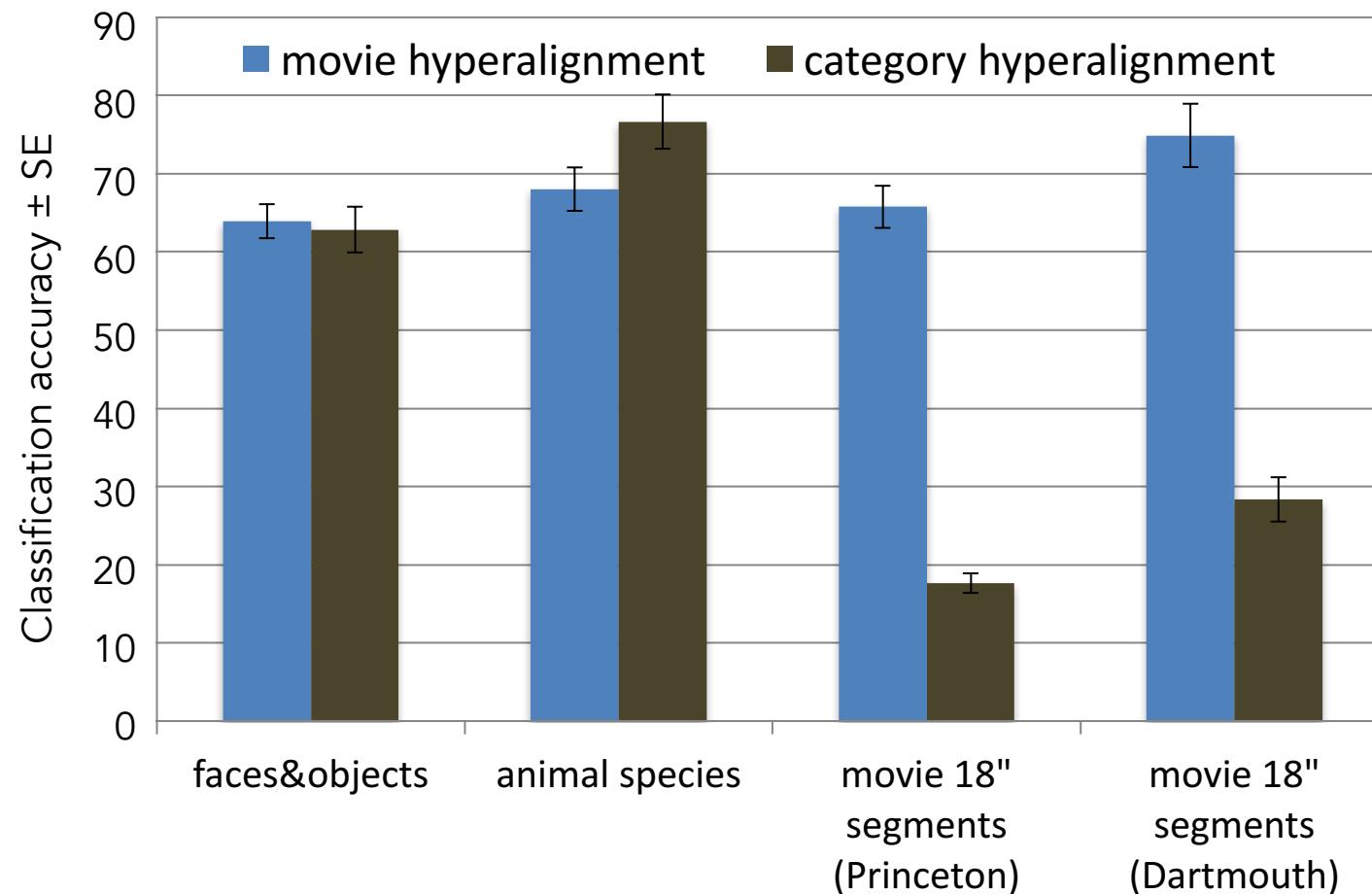
Classification accuracy (%)

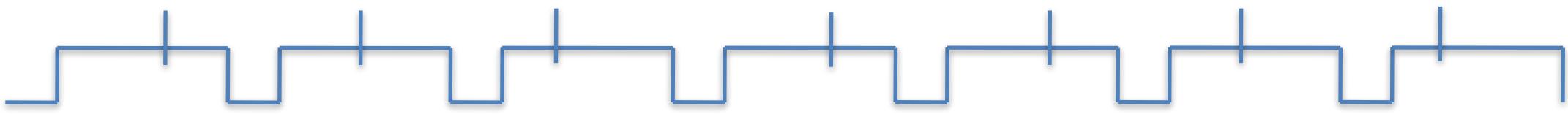
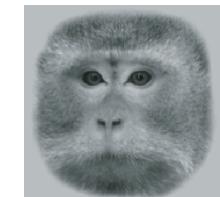
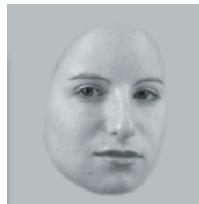
30% 60%³⁵

Why the movie?

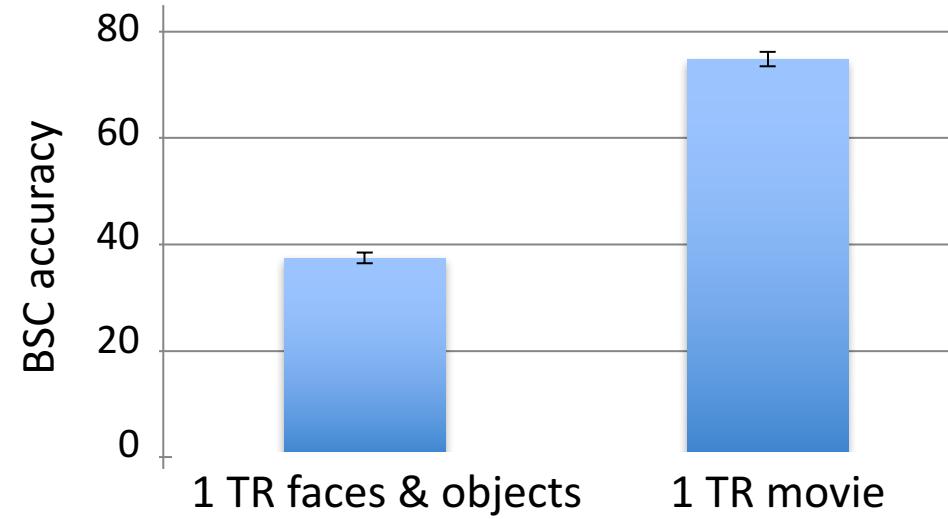
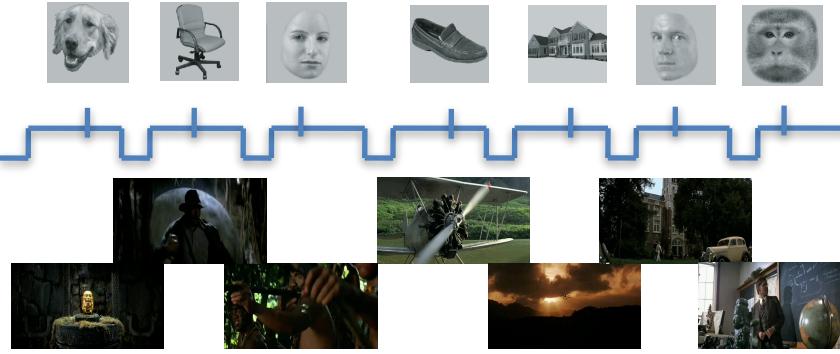
Does it add anything? **YES**

Hyperalignment based on the responses to face & object images or on the responses to animal species images works but has less general validity





Between-subject classification of single time-points from the movie
vs single TRs from the faces and objects experiment



Between-subject classification of single time-points from the movie
vs single TRs from the faces and objects experiment:

Movie stimuli evoke patterns of response that are more distinct

Representational geometry captures information in patterns across time for the whole data set

$$\begin{array}{c}
 \text{Time-points} \\
 | \\
 \text{B} \quad \text{Voxels} \\
 | \\
 t_1 \quad x_{1,1} \quad x_{2,1} \quad x_{3,1} \quad \dots \quad x_{l,1} \\
 t_2 \quad x_{1,2} \quad x_{2,2} \quad x_{3,2} \quad \dots \quad x_{l,2} \\
 t_3 \quad x_{1,3} \quad x_{2,3} \quad x_{3,3} \quad \dots \quad x_{l,3} \\
 t_4 \quad x_{1,4} \quad x_{2,4} \quad x_{3,4} \quad \dots \quad x_{l,4} \\
 \dots \quad \dots \quad \dots \quad \dots \quad \dots \\
 t_j \quad x_{1,j} \quad x_{2,j} \quad x_{3,j} \quad \dots \quad x_{l,j}
 \end{array}
 \times
 \begin{array}{c}
 \text{Model dimensions} \\
 | \\
 \text{R} \quad d_1 \quad d_2 \quad d_3 \quad \dots \quad d_m \\
 | \\
 v_1 \quad w_{1,1} \quad w_{2,1} \quad w_{3,1} \quad \dots \quad w_{m,1} \\
 v_2 \quad w_{1,2} \quad w_{2,2} \quad w_{3,2} \quad \dots \quad w_{m,2} \\
 v_3 \quad w_{1,3} \quad w_{2,3} \quad w_{3,3} \quad \dots \quad w_{m,3} \\
 \dots \quad \dots \quad \dots \quad \dots \quad \dots \\
 v_l \quad w_{1,l} \quad w_{2,l} \quad w_{3,l} \quad \dots \quad w_{m,l}
 \end{array}
 =
 \begin{array}{c}
 \text{Model dimensions} \\
 | \\
 \text{M} \quad d_1 \quad d_2 \quad d_3 \quad \dots \quad d_m \\
 | \\
 t_1 \quad y_{1,1} \quad y_{2,1} \quad y_{3,1} \quad \dots \quad y_{m,1} \\
 t_2 \quad y_{1,2} \quad y_{2,2} \quad y_{3,2} \quad \dots \quad y_{m,2} \\
 t_3 \quad y_{1,3} \quad y_{2,3} \quad y_{3,3} \quad \dots \quad y_{m,3} \\
 t_4 \quad y_{1,4} \quad y_{2,4} \quad y_{3,4} \quad \dots \quad y_{m,4} \\
 \dots \quad \dots \quad \dots \quad \dots \quad \dots \\
 t_j \quad y_{1,j} \quad y_{2,j} \quad y_{3,j} \quad \dots \quad y_{m,j}
 \end{array}$$

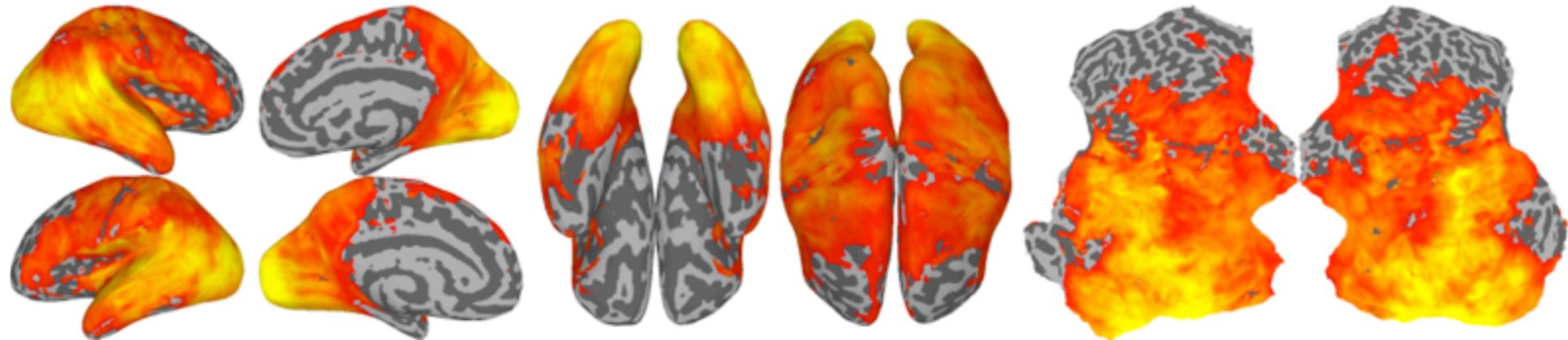
Data matrix in voxel space
(l voxels x j time-points)

Transformation matrix
(m dimensions x l voxels)

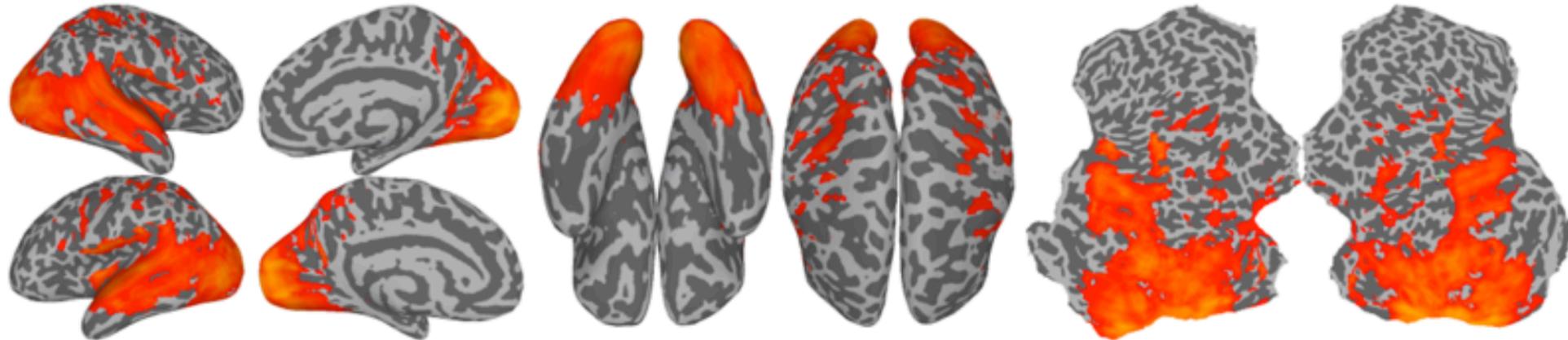
Data matrix in model space
(m dimensions x j time-points)

Whole-brain hyperalignment increases intersubject correlation of
high-dimensional representational geometries
(correlations between movie time-points)

Hyperalignment



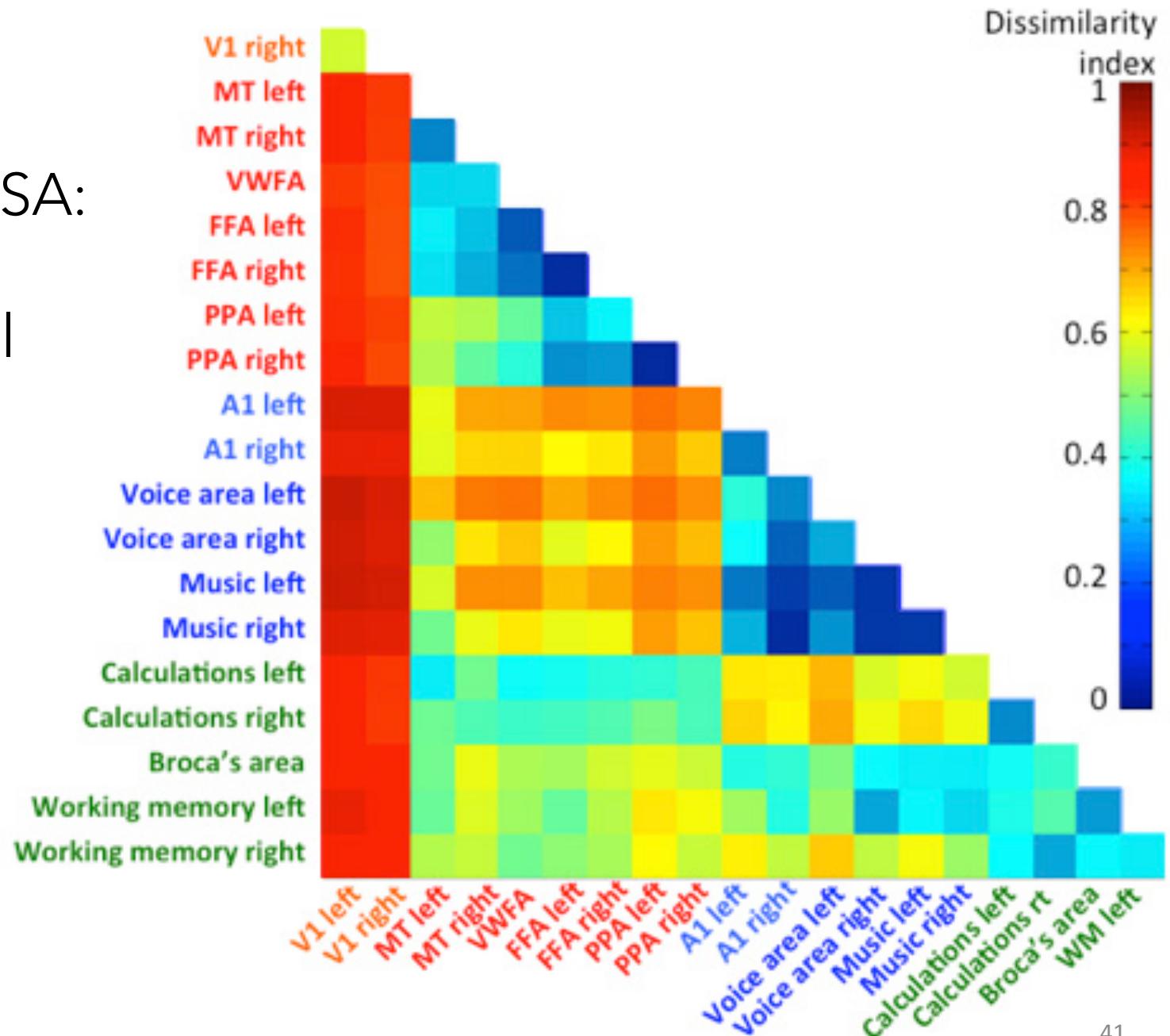
Anatomical alignment



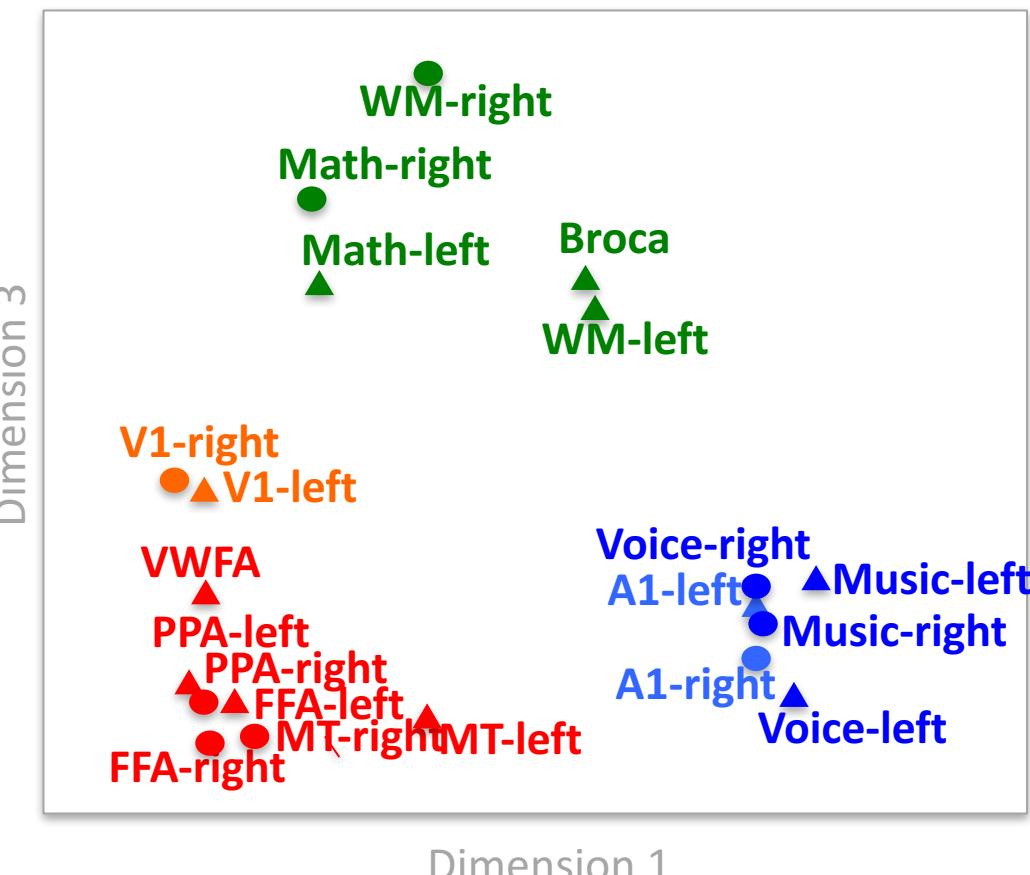
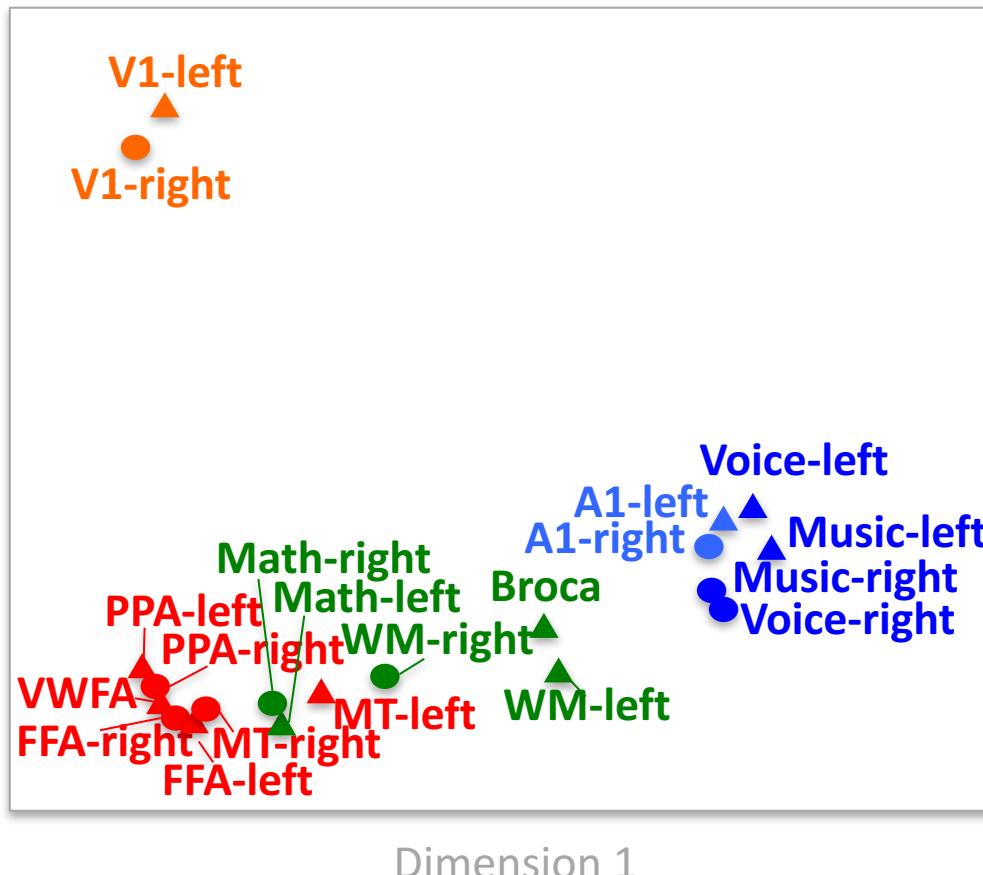
Intersubject correlation



Second-order RSA: Between-ROI representational geometry dissimilarities



Multidimensional scaling (MDS) of similarity structures in **visual**, **auditory**, and **cognitive** regions of interest* (ROIs)



* ROI coordinates from Neurosynth



Multidimensional scaling (MDS) of regions in the face processing system

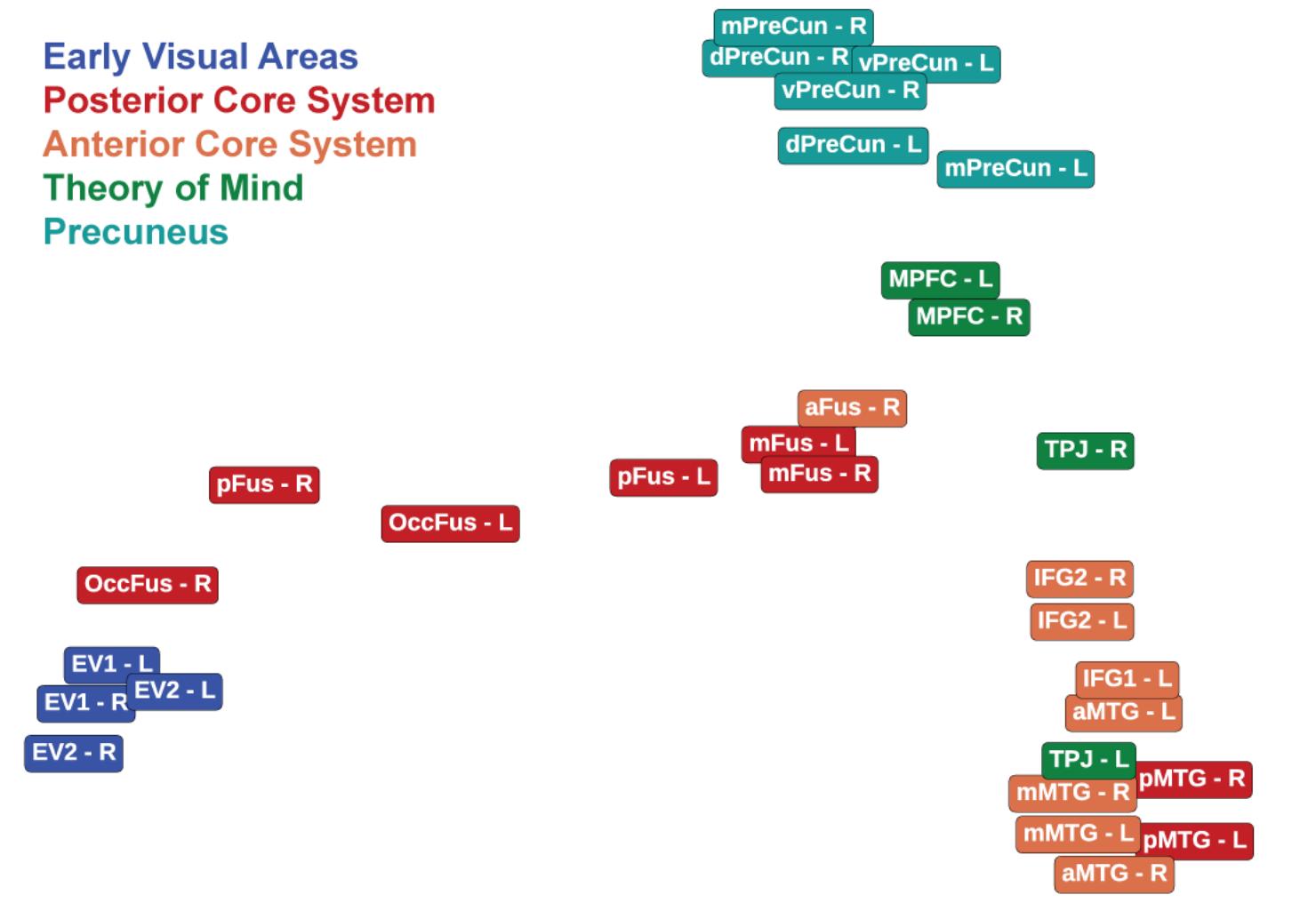
(Visconti di Oleggio Castello, Halchenko, ..., Gobbini, 2017)



Dimension 2

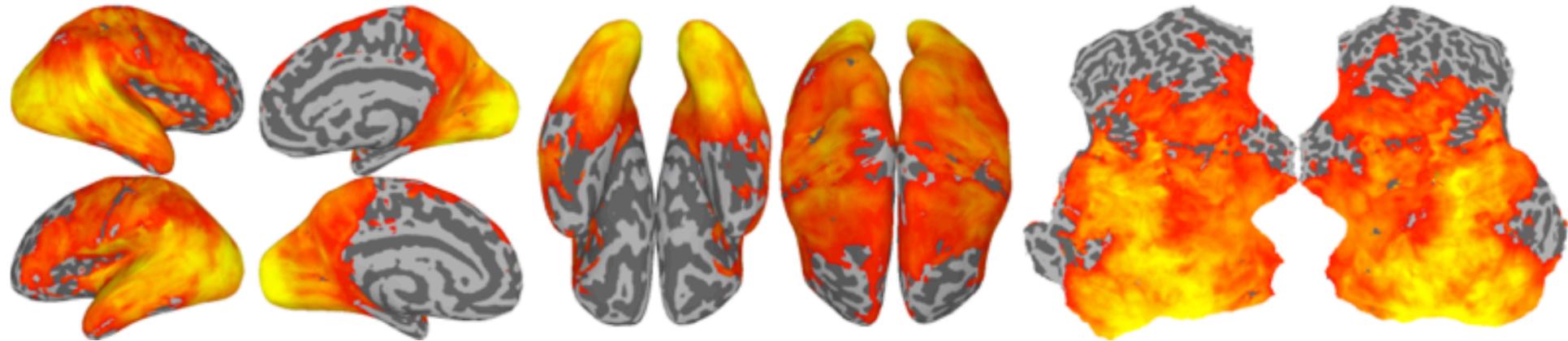
Early Visual Areas
Posterior Core System
Anterior Core System
Theory of Mind
Precuneus

Dimension 1

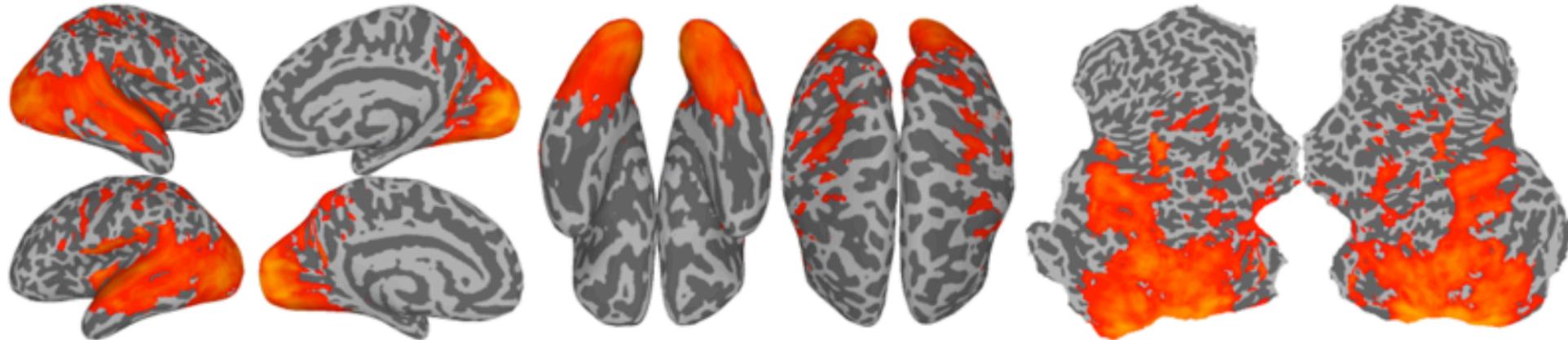


Whole-brain hyperalignment increases intersubject correlation of high-dimensional representational geometries that reflect widely divergent domains of information

Hyperalignment



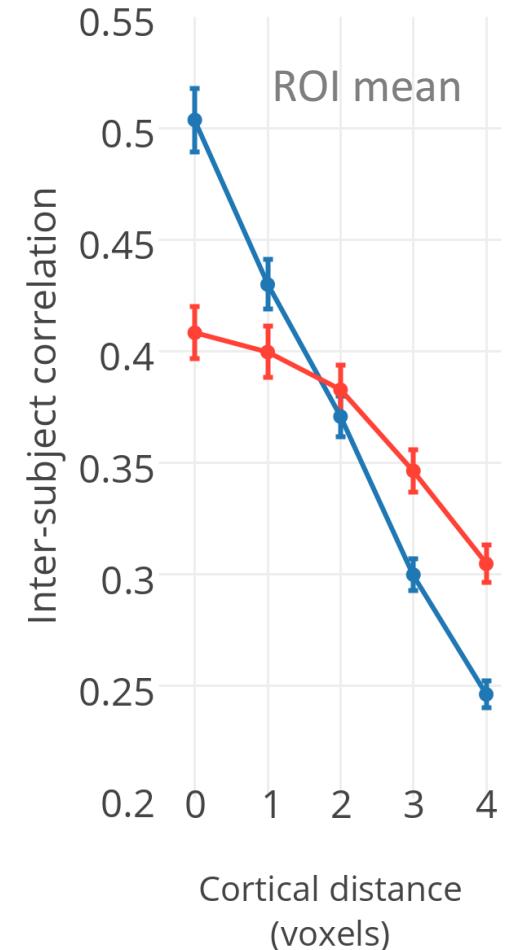
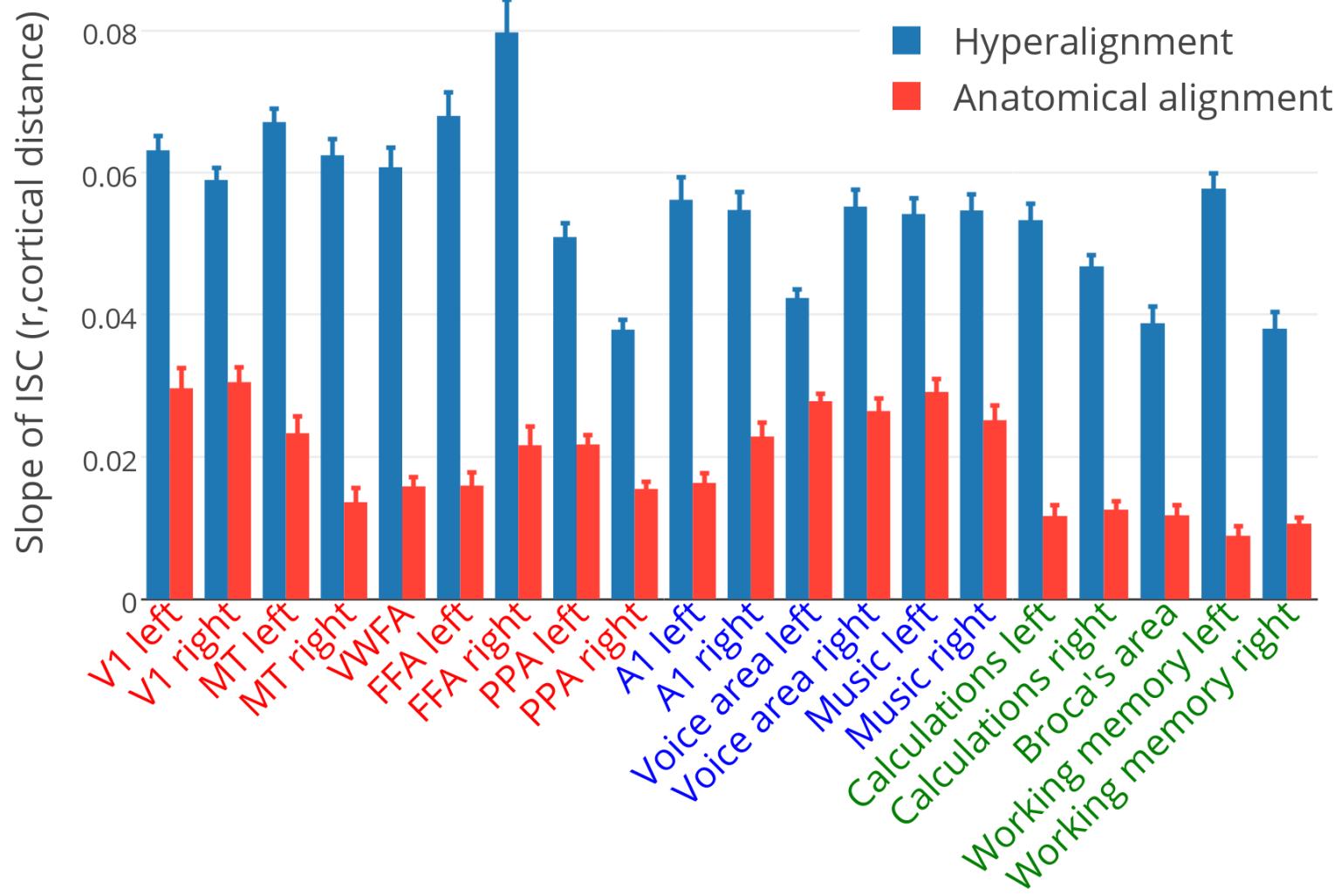
Anatomical alignment



Intersubject correlation



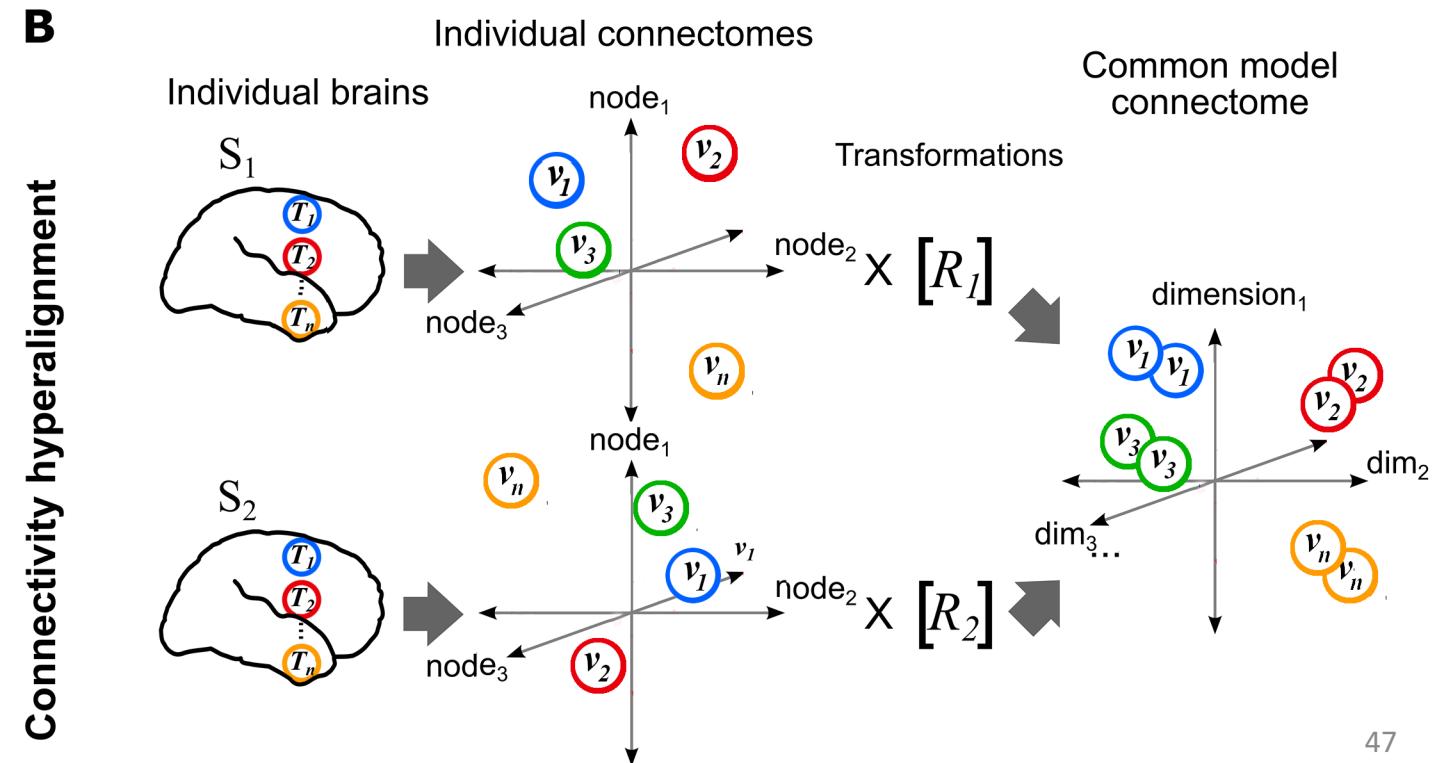
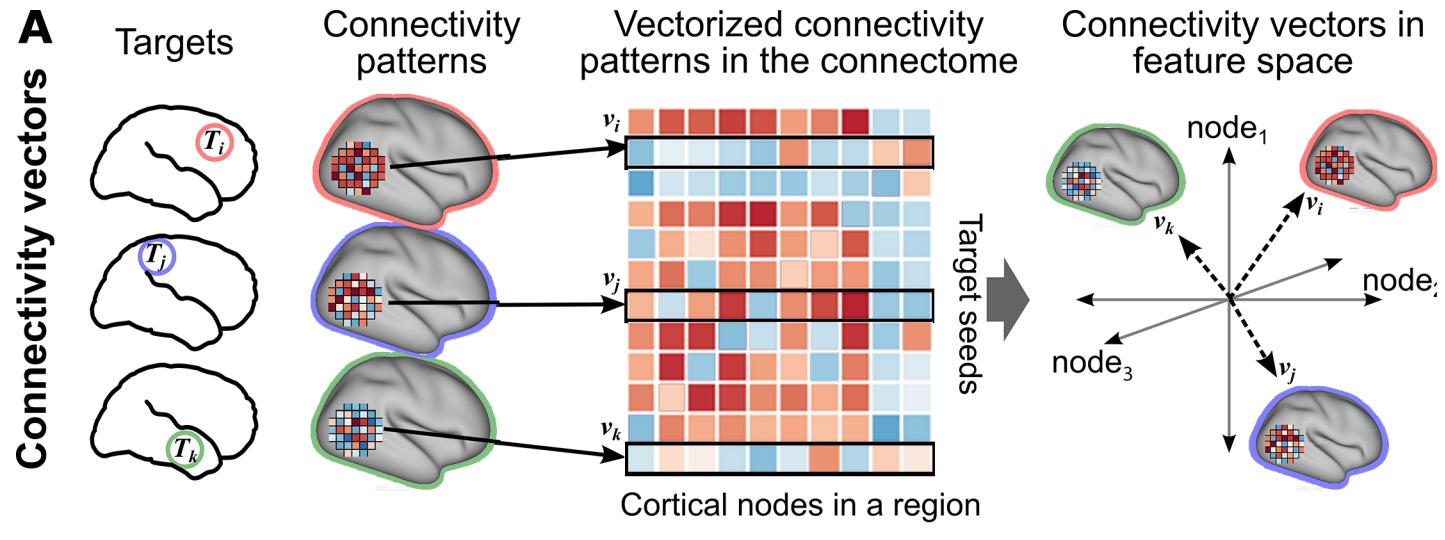
Point spread function (intersubject correlations of movie time series):
 Fine spatial scale of alignment of local variation in response tunings



A common high-dimensional linear model of representational spaces in human cortex

- Statement of the problem: capturing coarse- and fine-grained topographies in a common model
- Conceptual framework: high-dimensional representational spaces
- Deriving the common space and individual transformation matrices with hyperalignment
- Validation
- **Connectivity hyperalignment**
- Individual differences in fine-scale cortical functional architecture
- Conclusions

Connectivity hyperalignment



Connectivity data in subject brain voxel data (B_s)

<u>Voxels</u>		<u>Model dimensions</u>		<u>Model dimensions</u>	
Connectivity targets	B	$v_1 \quad v_2 \quad v_3 \quad \dots \quad v_l$	$w_{1,1} \quad w_{2,1} \quad w_{3,1} \quad \dots \quad w_{m,1}$	$t_1 \quad y_{1,1} \quad y_{2,1} \quad y_{3,1} \quad \dots \quad y_{m,1}$	d_m
	t_1	$x_{1,1} \quad x_{2,1} \quad x_{3,1} \quad \dots \quad x_{l,1}$	$v_1 \quad w_{1,2} \quad w_{2,2} \quad w_{3,2} \quad \dots \quad w_{m,2}$	$t_2 \quad y_{1,2} \quad y_{2,2} \quad y_{3,2} \quad \dots \quad y_{m,2}$	$y_{m,2}$
	t_2	$x_{1,2} \quad x_{2,2} \quad x_{3,2} \quad \dots \quad x_{l,2}$	$v_2 \quad w_{1,3} \quad w_{2,3} \quad w_{3,3} \quad \dots \quad w_{m,3}$	$t_3 \quad y_{1,3} \quad y_{2,3} \quad y_{3,3} \quad \dots \quad y_{m,3}$	$y_{m,3}$
	t_3	$x_{1,3} \quad x_{2,3} \quad x_{3,3} \quad \dots \quad x_{l,3}$	$v_3 \quad \dots \quad \dots \quad \dots \quad \dots$	$t_4 \quad y_{1,4} \quad y_{2,4} \quad y_{3,4} \quad \dots \quad y_{m,4}$	$y_{m,4}$
	t_4	$x_{1,4} \quad x_{2,4} \quad x_{3,4} \quad \dots \quad x_{l,4}$	$v_l \quad w_{1,l} \quad w_{2,l} \quad w_{3,l} \quad \dots \quad w_{m,l}$	\dots	\dots
	\dots	$\dots \quad \dots \quad \dots \quad \dots \quad \dots$			
	t_j	$x_{1,j} \quad x_{2,j} \quad x_{3,j} \quad \dots \quad x_{l,j}$		$T_j \quad y_{1,j} \quad y_{2,j} \quad y_{3,j} \quad \dots \quad y_{m,j}$	

Data matrix in voxel space
(l voxels $\times j$ time-points)

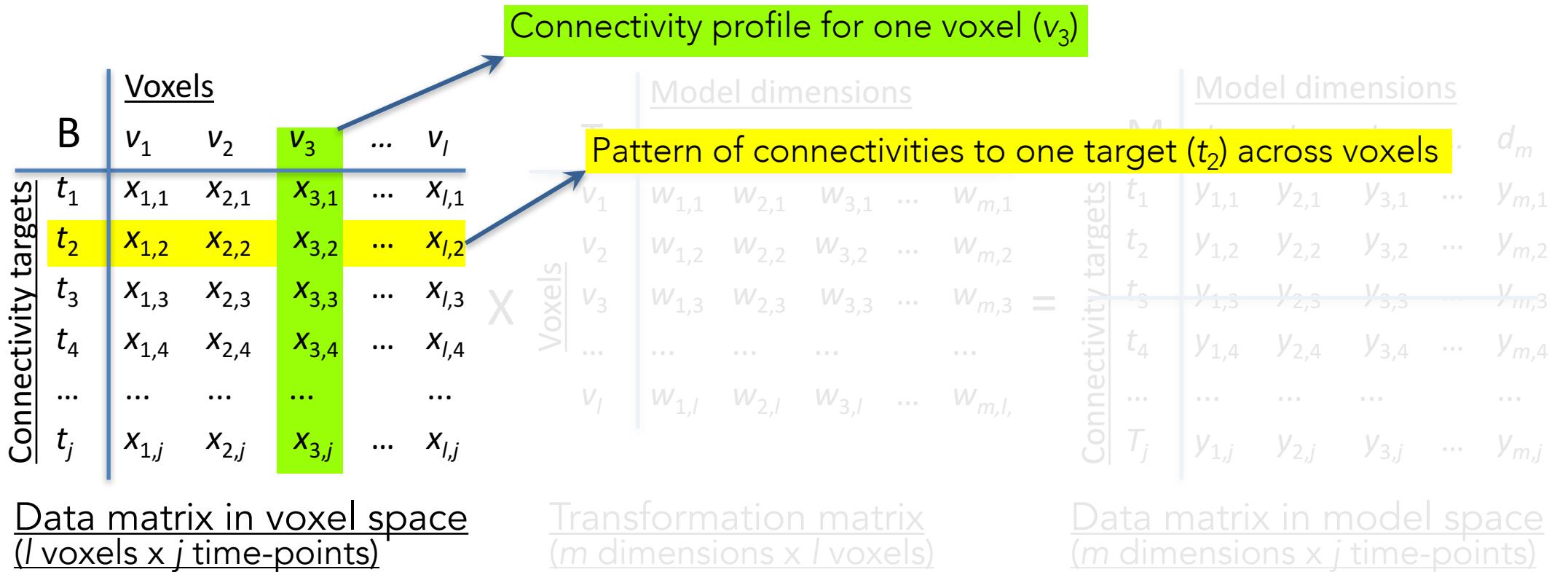
Transformation matrix
(m dimensions $\times l$ voxels)

Data matrix in model space
(m dimensions $\times j$ time-points)

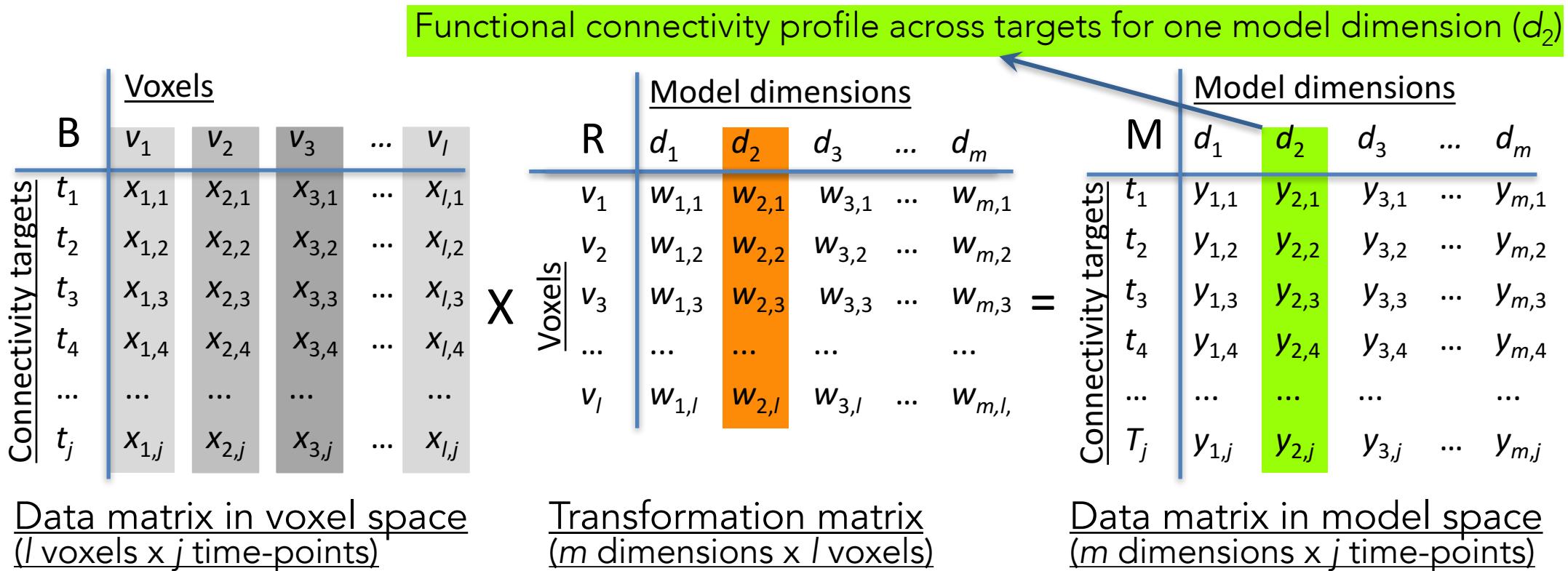
Pattern of connectivities to one target (t_2) across voxels



Connectivity data in subject brain voxel data (B_s)

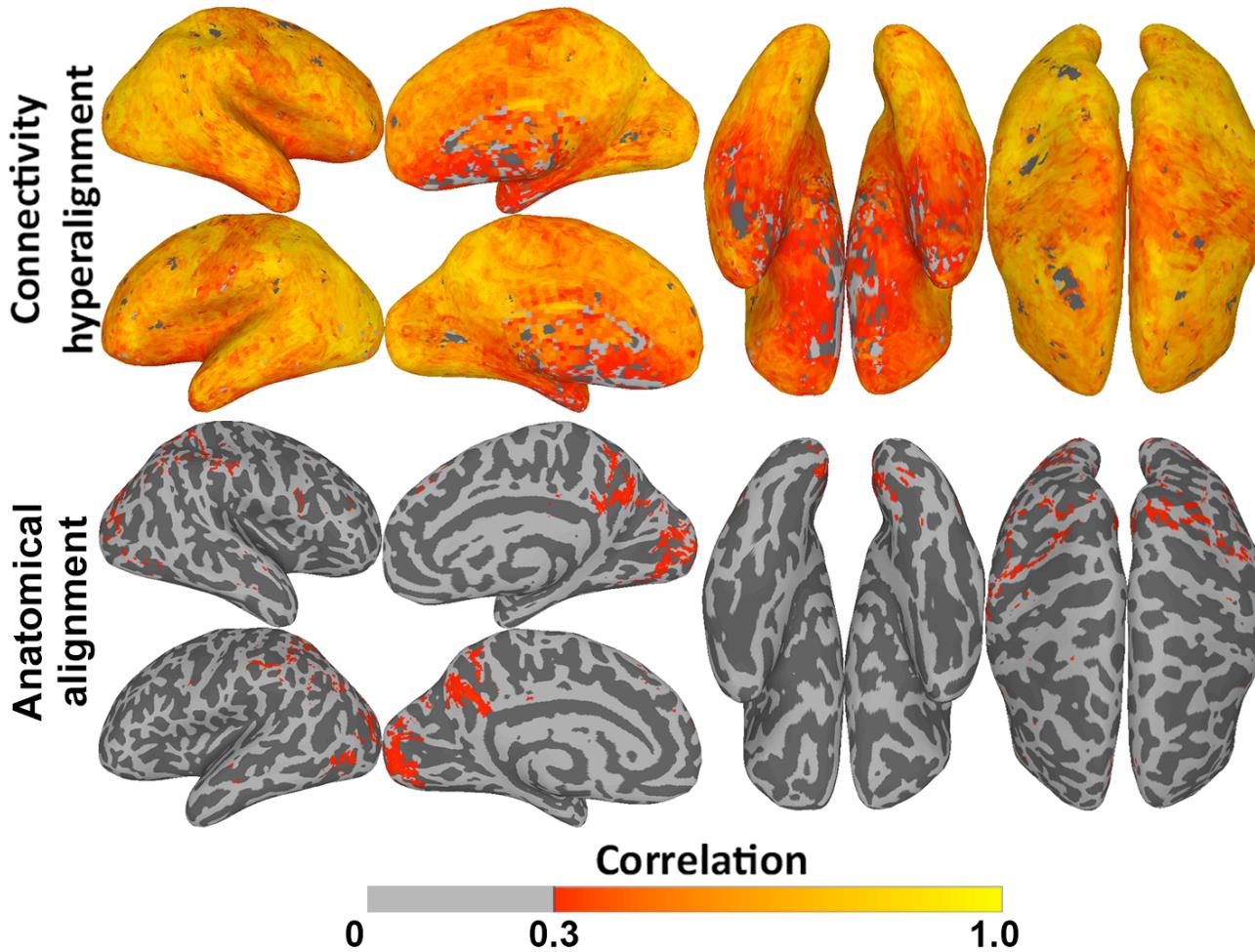


Deriving response tuning profiles for model space (M_s) dimensions (d_i)

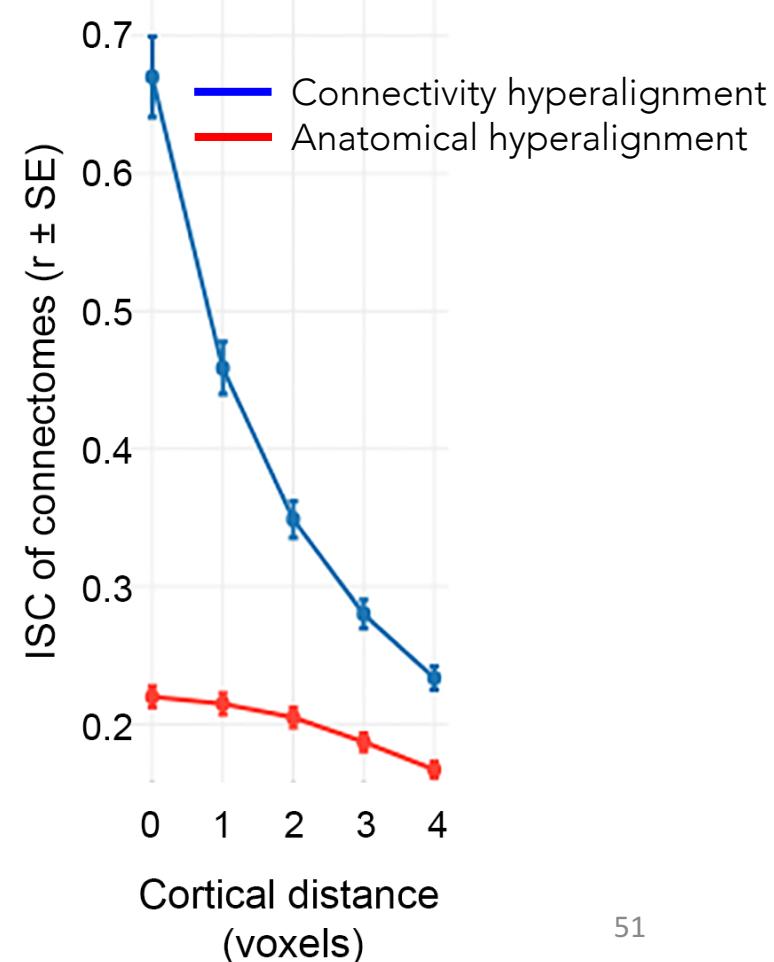


Connectivity hyperalignment of movie fMRI data increases ISC of connectivity vectors and reveals a fine-scale architecture

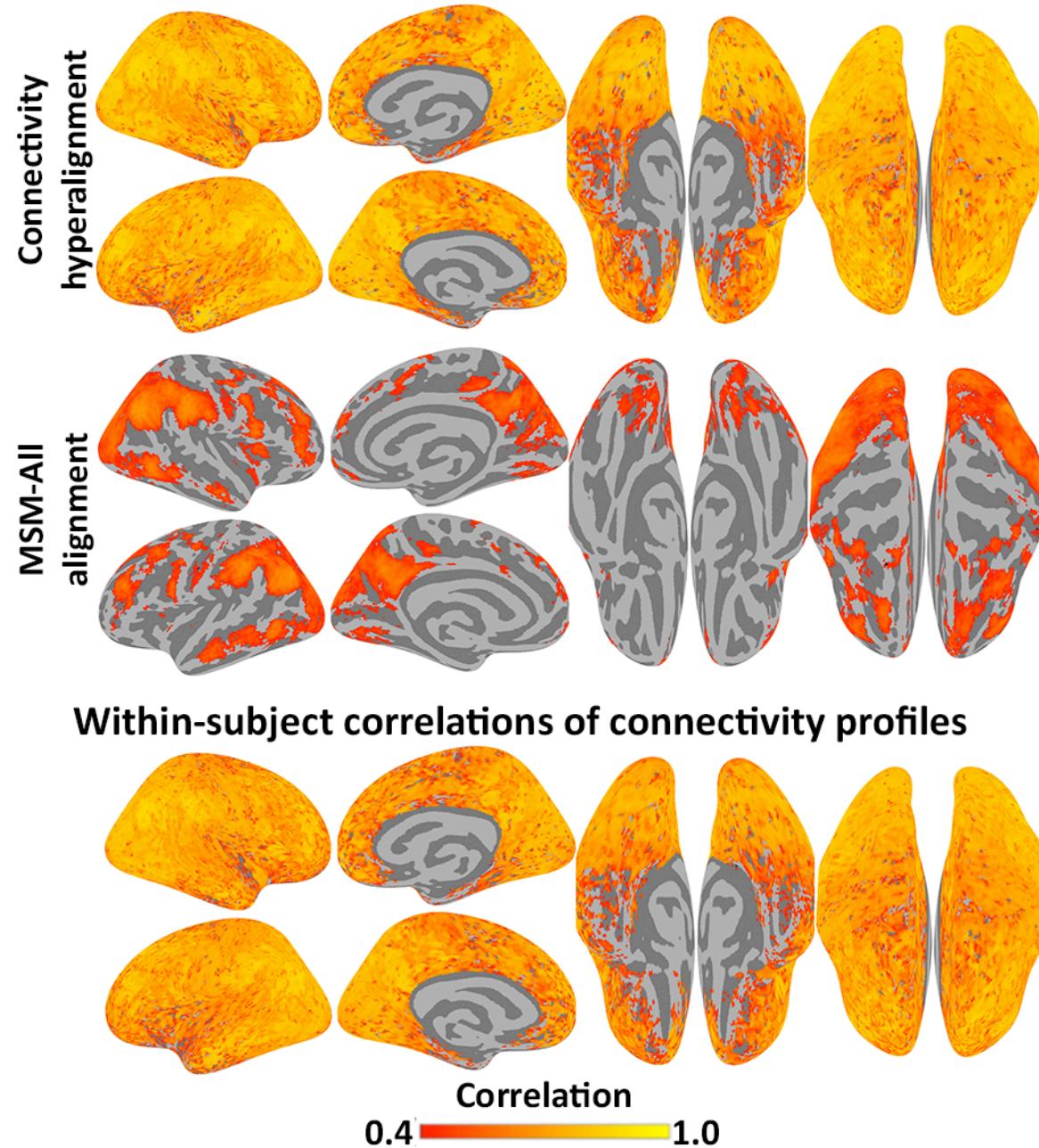
Intersubject correlations of connectivity profiles



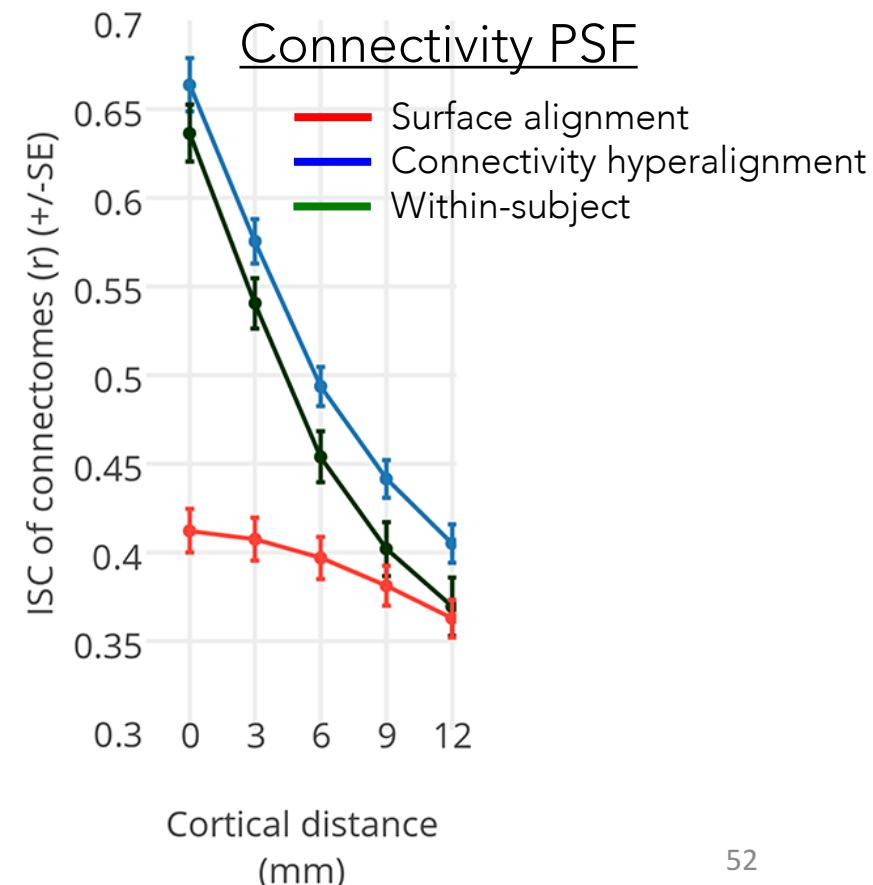
Connectivity PSF



Intersubject correlations of connectivity profiles



Connectivity hyperalignment of Human Connectome Project (HCP) resting state fMRI data ($N=20$) increases ISC of connectivity vectors and reveals a fine-scale architecture

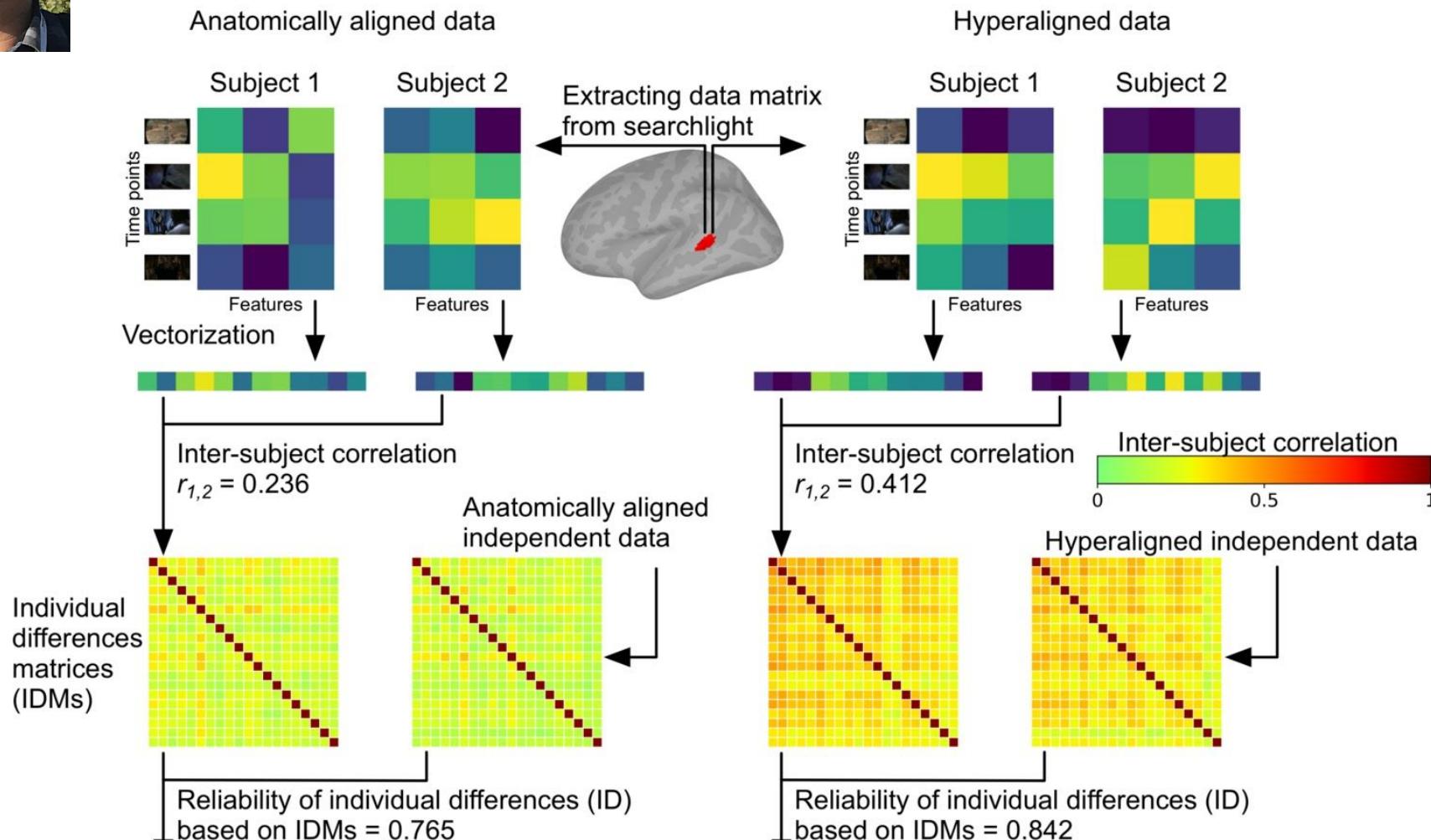


A common high-dimensional linear model of representational spaces in human cortex

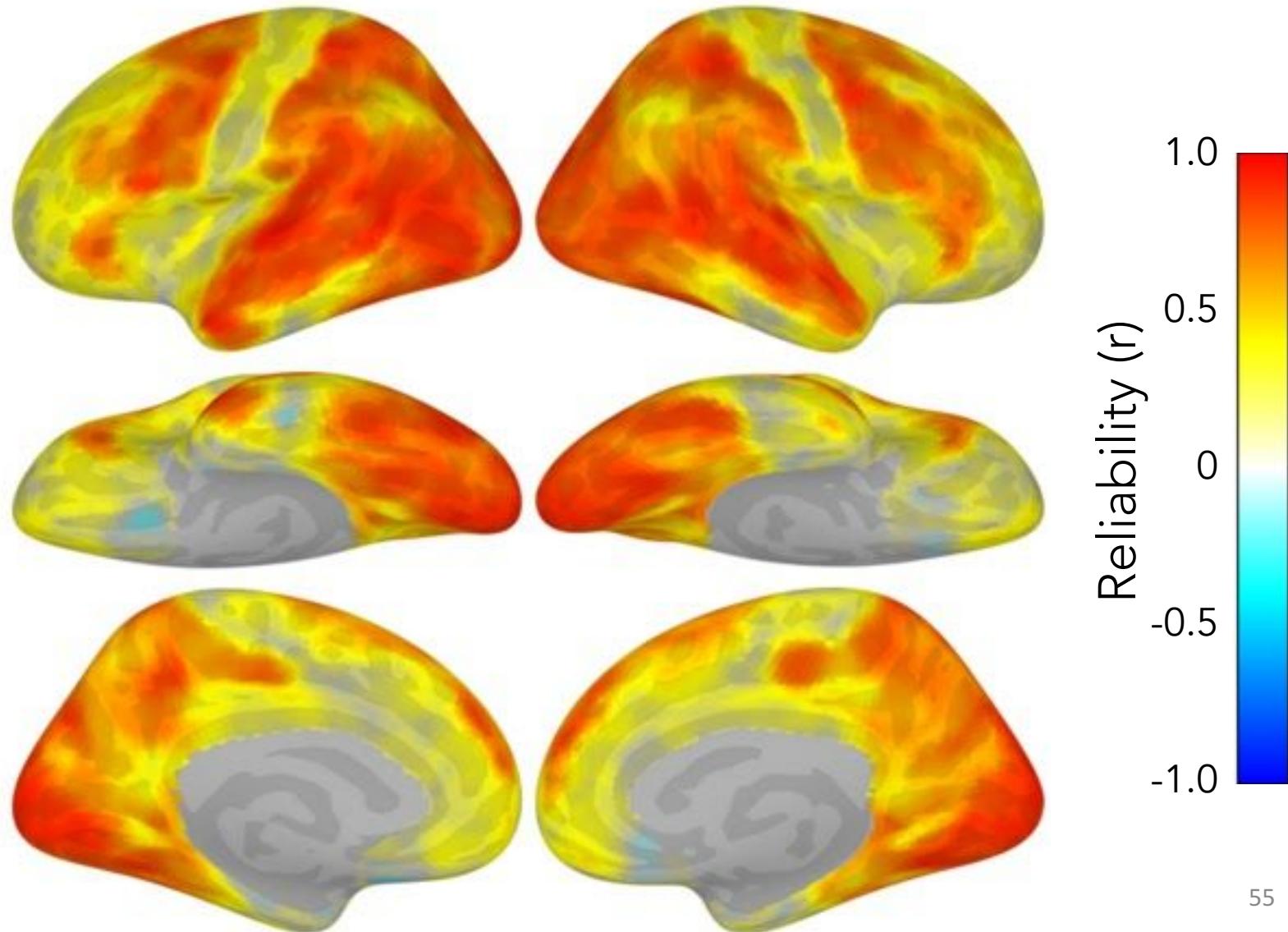
- Statement of the problem: capturing coarse- and fine-grained topographies in a common model
- Conceptual framework: high-dimensional representational spaces
- Deriving the common space and individual transformation matrices with hyperalignment
- Validation
- Connectivity hyperalignment
- Individual differences in fine-scale cortical functional architecture
- Conclusions



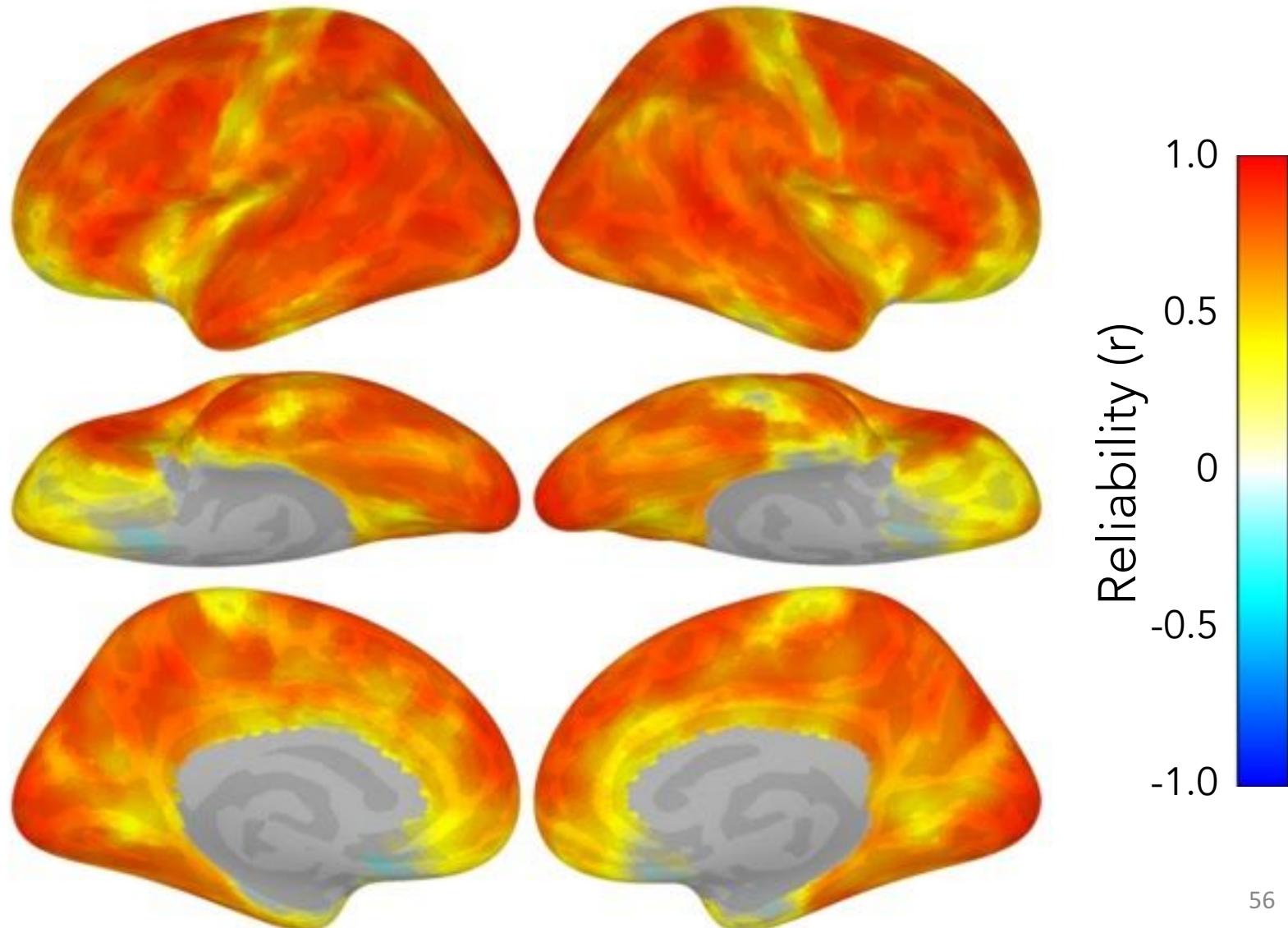
Individual differences dissimilarity matrix (IDM) (Feilong M, 2018)



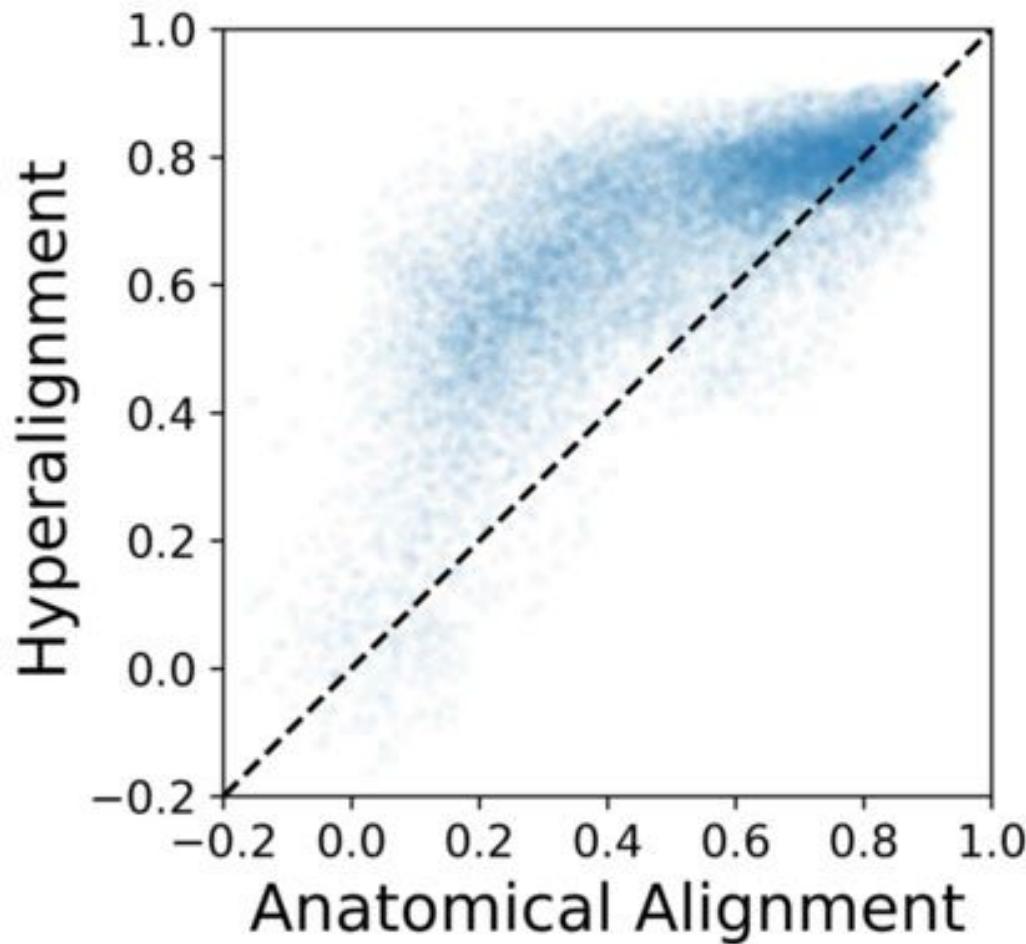
Reliability of IDMs (anatomical alignment)



Reliability of IDMs (hyperaligned common model space)

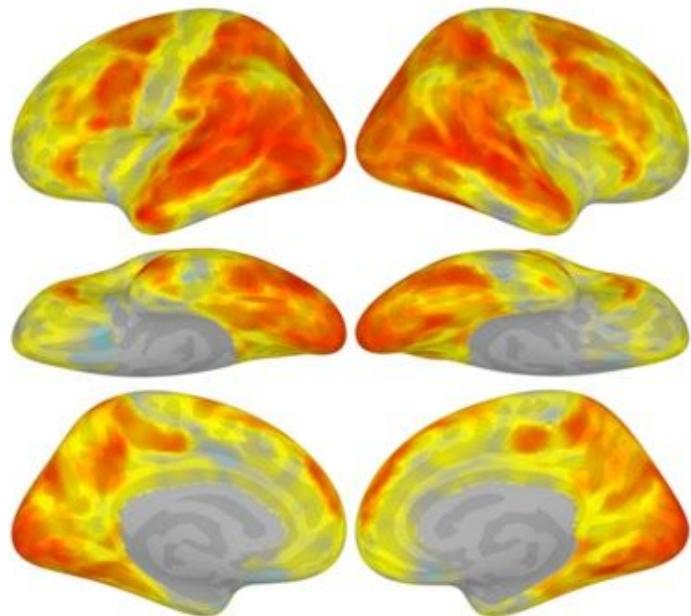


Reliability of IDMs

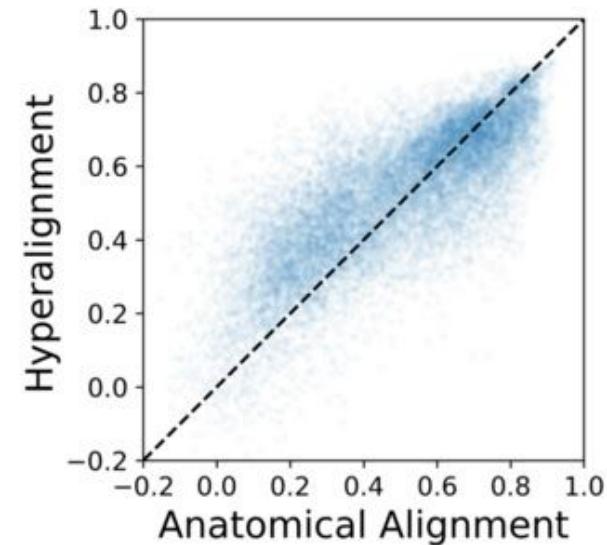
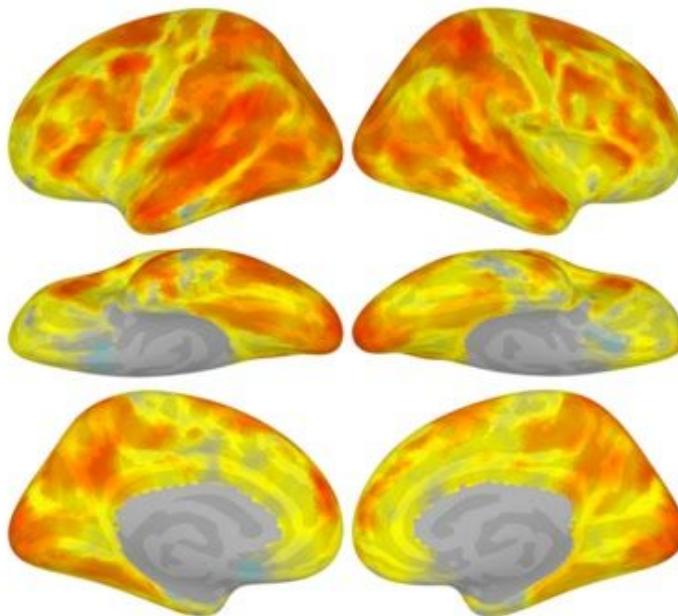


Reliability of coarse-scale IDMs (based on searchlight mean time-series)

Anatomical Alignment

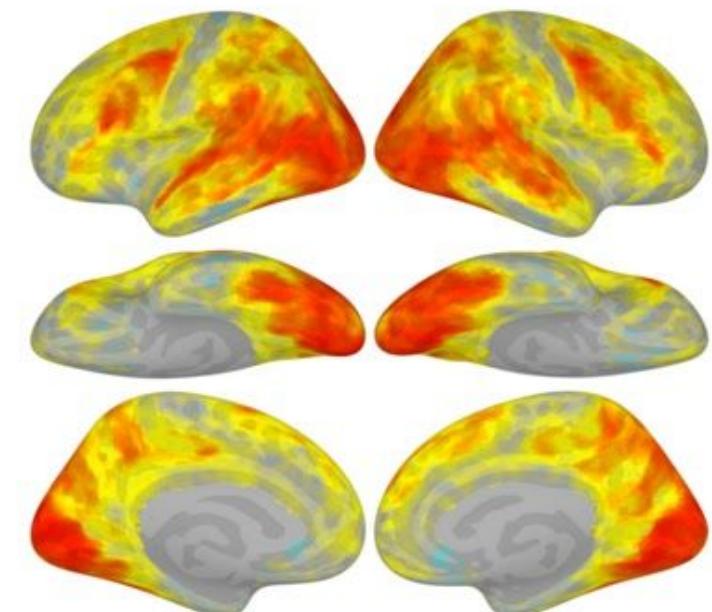


Hyperalignment

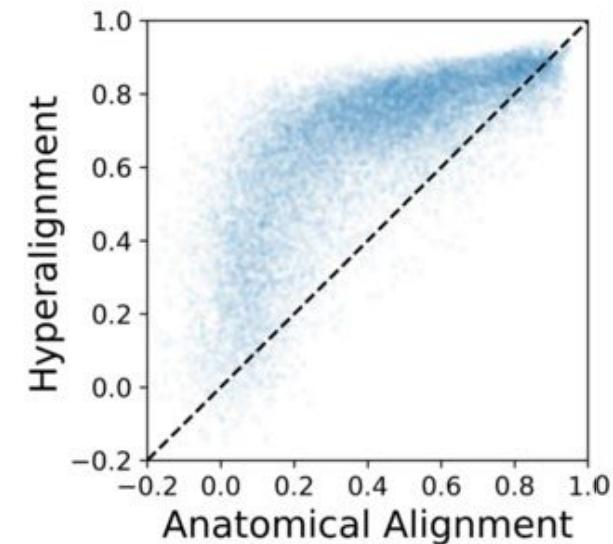
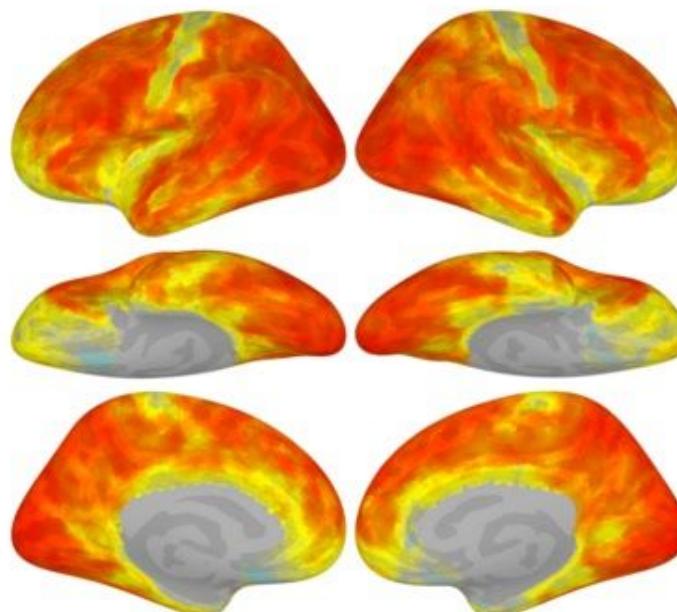


Reliability of fine-scale IDMs (based on residual time-series after removing searchlight mean)

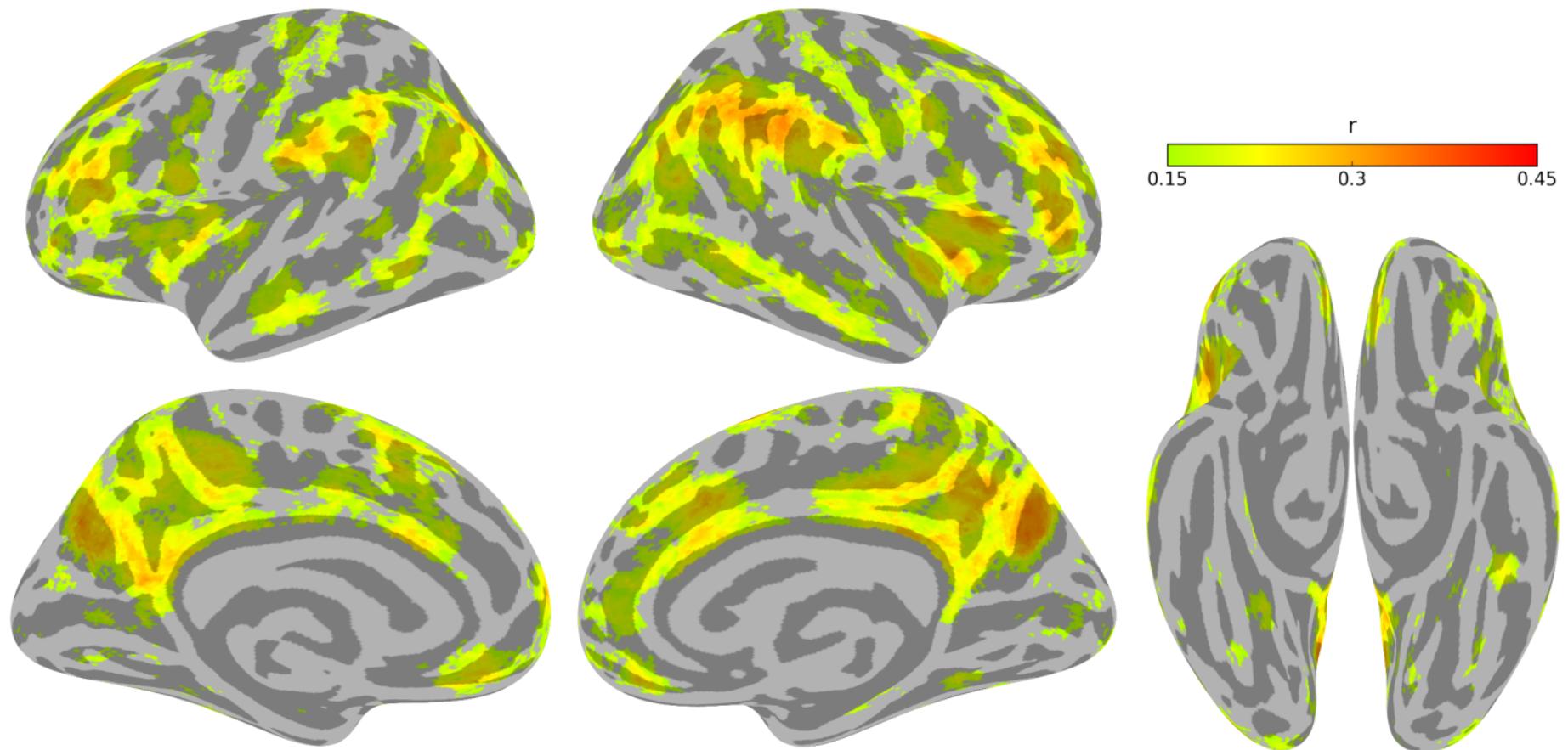
Anatomical Alignment



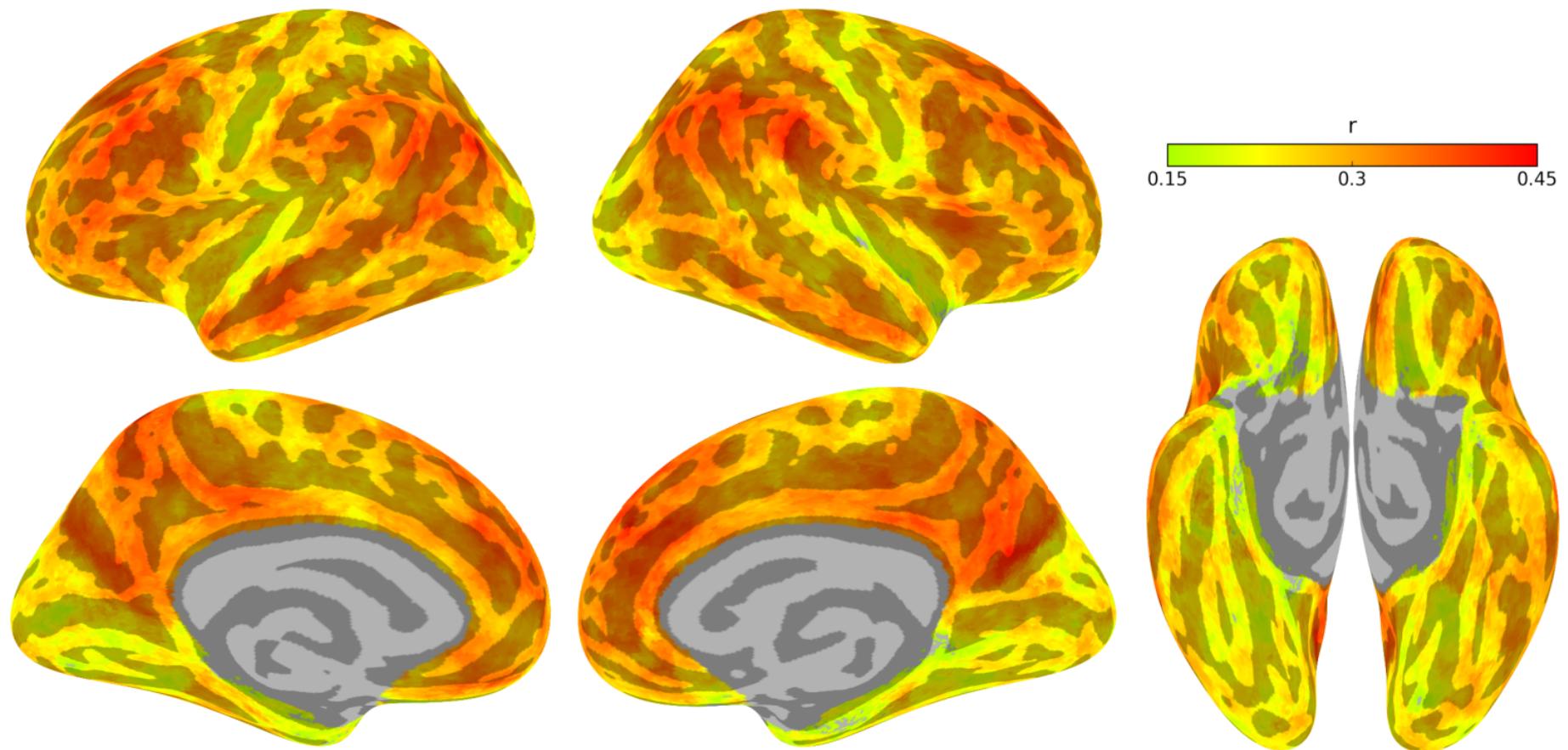
Hyperalignment



Correlation of functional connectivity IDs with verbal ability (picture vocabulary) (anatomical alignment)



Correlation of functional connectivity IDs with verbal ability (picture vocabulary) (hyperaligned data)



A common high-dimensional linear model of representational spaces in human cortex

- Statement of the problem: capturing coarse- and fine-grained topographies in a common model
- Conceptual framework: high-dimensional representational spaces
- Deriving the common space and individual transformation matrices with hyperalignment
- Validation
- Connectivity hyperalignment
- Conclusions

Common model: Basic components

- Common, high-dimensional representational/connectivity space
- Individual transformation matrices
- Algorithm for calculating transformation matrices and deriving the common model space (hyperalignment)

Common model dimensions

- Response tuning basis functions
 - Common across brains
 - Model population responses that carry fine distinctions
 - Capture variations with a fine spatial granularity
 - Valid for diverse domains of information
- Functional connectivity vectors
 - Common across brains
 - Capture variations with a fine spatial granularity
- Topographic basis functions
 - Individual specific
 - Model individual variations in functional topography with high fidelity

Fine-scale structure in cortical functional architecture: How big a factor is it?

The common model

- affords 7X higher bsMVPC accuracies
- greatly increases ISCs of
 - response tuning profiles
 - local representational geometries
 - connectivity profiles

Fine-scale structure in cortical functional architecture: How big a factor is it?

The common model

- affords **7X higher bsMVPC accuracies**
- greatly increases ISCs of
 - response tuning profiles
 - local representational geometries
 - connectivity profiles

*Fine-scale structure is the “dark matter”
of cortical functional architecture*

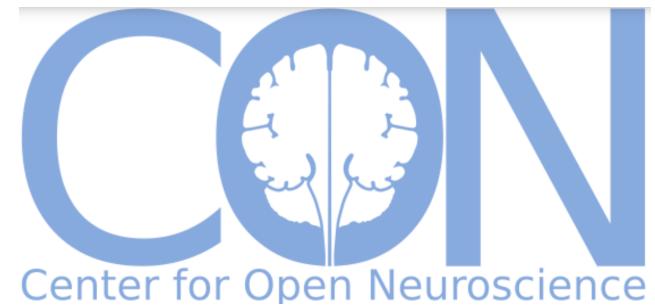
The Common Model captures the shared information that is encoded in idiosyncratic fine-scale functional topographies

- Captures common bases for the fine-grained structure of local variations in response tuning and functional connectivity profiles
- There is no canonical brain
 - Each individual brain has idiosyncratic fine-scale topographies
 - The common model finds the shared deep structure of information that is encoded in idiosyncratic topographies

Individual differences in the Common Model capture how the representation of fine-scale information varies across brains

- Hyperalignment increases reliability of individual differences by capturing variation in the deep structure of information that is not accessible in anatomically aligned data.
- Individual differences in hyperaligned data better predict cognitive differences.
- The granularity of cortical architecture that underlies cognitive differences is matched better by the scale of fine-scale topographies, than by the coarse scale of systems of functional areas.

Software for hyperalignment and data are on PyMVPA (www.pymvpa.org)



Guntupalli et al. (2016) A model of representational spaces in human cortex. *Cerebral Cortex*.

Guntupalli & Haxby (2018) Computational model of shared fine-scale structure in the human connectome. *PLoS Computational Biology*.

Feilong, Nastase, Haxby (2018) Reliable individual differences in fine-grained cortical functional architecture. *bioRxiv*.