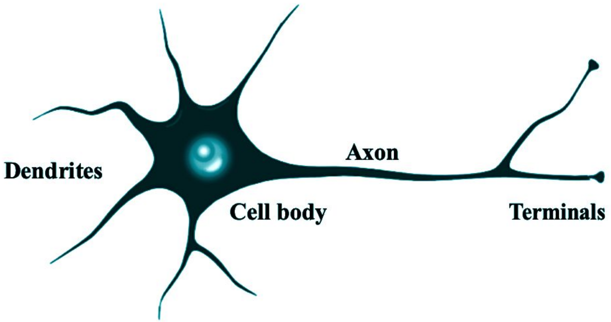
Aligning representations across individual models

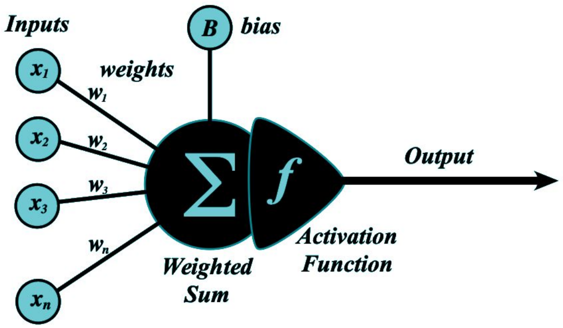
Elizabeth DuPre

Wu Tsai interdisciplinary postdoctoral fellow
Stanford University

OHBM2023 : NeuroAI Educational
22 July, 2023

In NeuroAI, we are rarely comparing apples to apples.

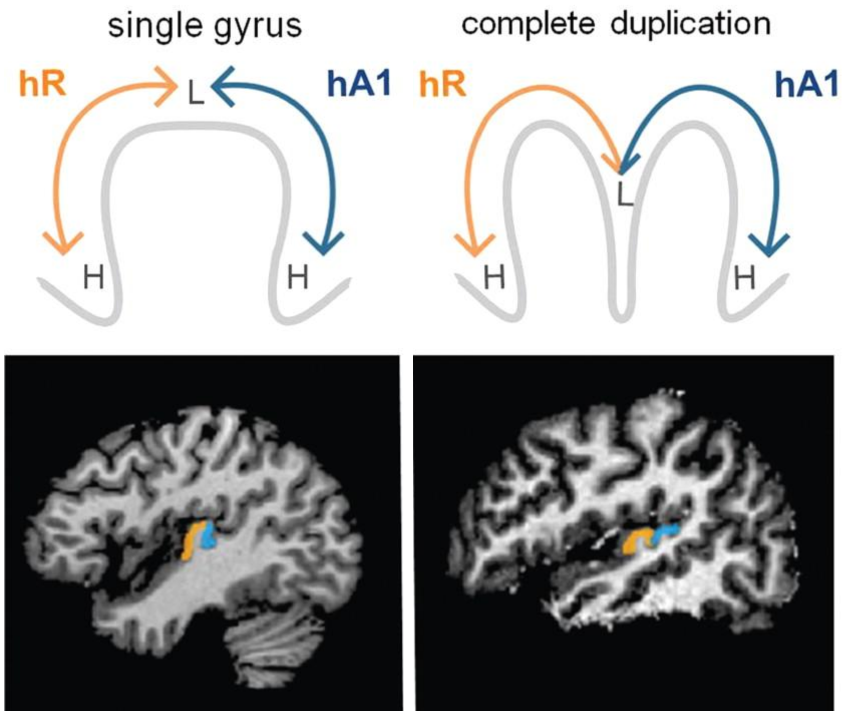




[Courtesy of Allen D. Elster, MRIquestions.com]()

Challenges to inter-individual comparisons in biology

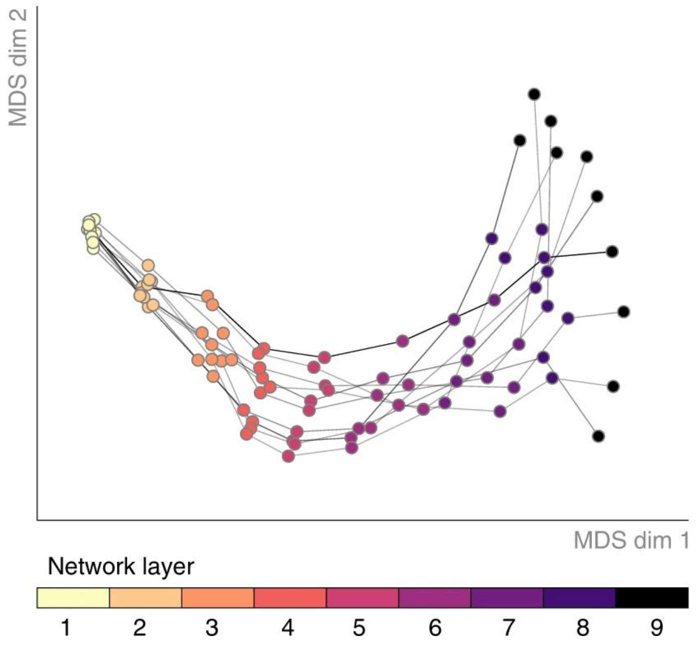
- Normalizing to a standardized anatomical space does not fully address inter-subject variability (Rademacher et al., 1993; Thirion et al., 2006)



Da Costa et al. (2011). *J Neuro.*

Challenges to comparing trained artificial network instances

- Convolutional (Mehrer et al., 2020) and recurrent (Maheswaranathan et al., 2019) networks both show individual differences in learned representations
  - These differences persist with the same architecture and training set



Mehrer et al. (2020). *Nat Comms.*

Outline

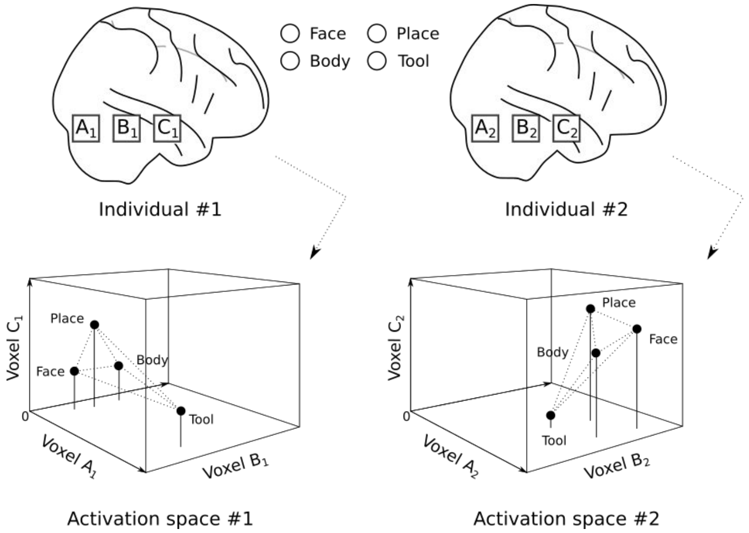
- Measuring dissimilarity with metrics
- Using alignment to leverage similarity

Outline

- Measuring dissimilarity with metrics
- Using alignment to leverage similarity

Measuring dissimilarity with metrics

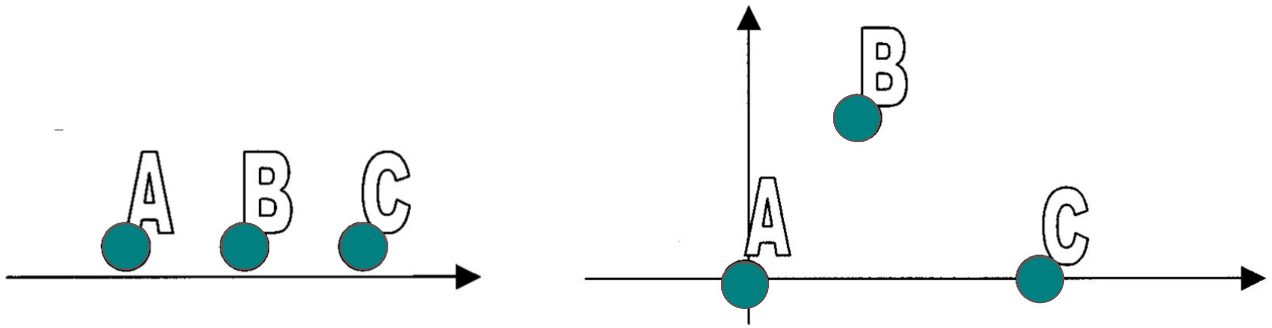
- To quantify differences between systems, individuals, or conditions, we can use metrics that measure their dissimilarity



Adapted from [Churchland (1998). *J Philo.*]

Representational Similarity Analysis

- Representational Similarity Analysis (RSA; Kriegeskorte et al., 2008) is a popular approach to compare observed representational geometries



Laasko and Cottrell (2000)

Representational Similarity Analysis

- Representational Similarity Analysis (RSA; Kriegeskorte et al., 2008) is a popular approach to compare observed representational geometries
- While RSA was originally proposed using correlation distance, this is not a true metric (Williams et al., 2021), complicating downstream analyses (Thirion et al., 2015)

The measure of a metric

1. Equivalence if $A \longrightarrow B = \emptyset$; $A = B$

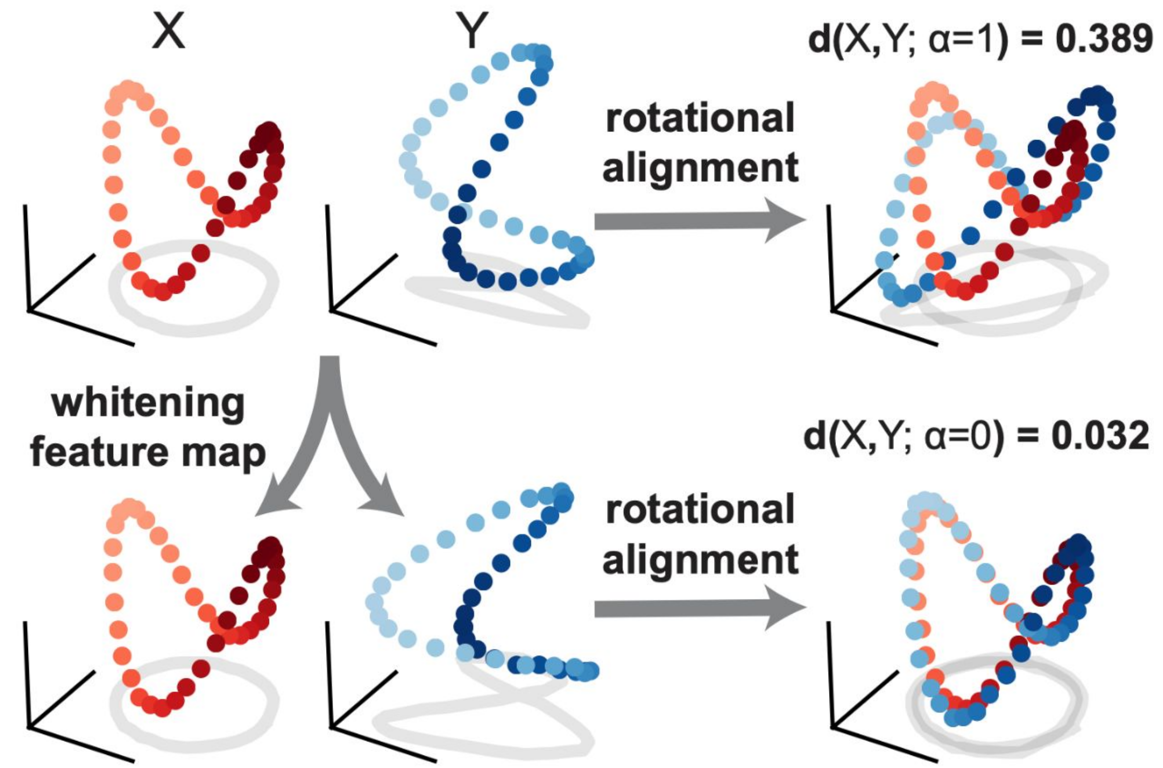
2. Symmetry $A \longrightarrow B = A \longleftarrow B$

3. Triangle inequality

$$\begin{tikzcd} & A \arrow[dl, leftrightarrow] \arrow[dr, leftrightarrow] & \\ B & ? \arrow[l, dashed] \arrow[r, dashed] & C \end{tikzcd}$$

Measuring dissimilarity with metrics

- To quantify differences between individuals or conditions, we can use metrics that measure their dissimilarity
- Each metric provides different insight into the underlying structure or dynamics (Ostrow et al., 2023)



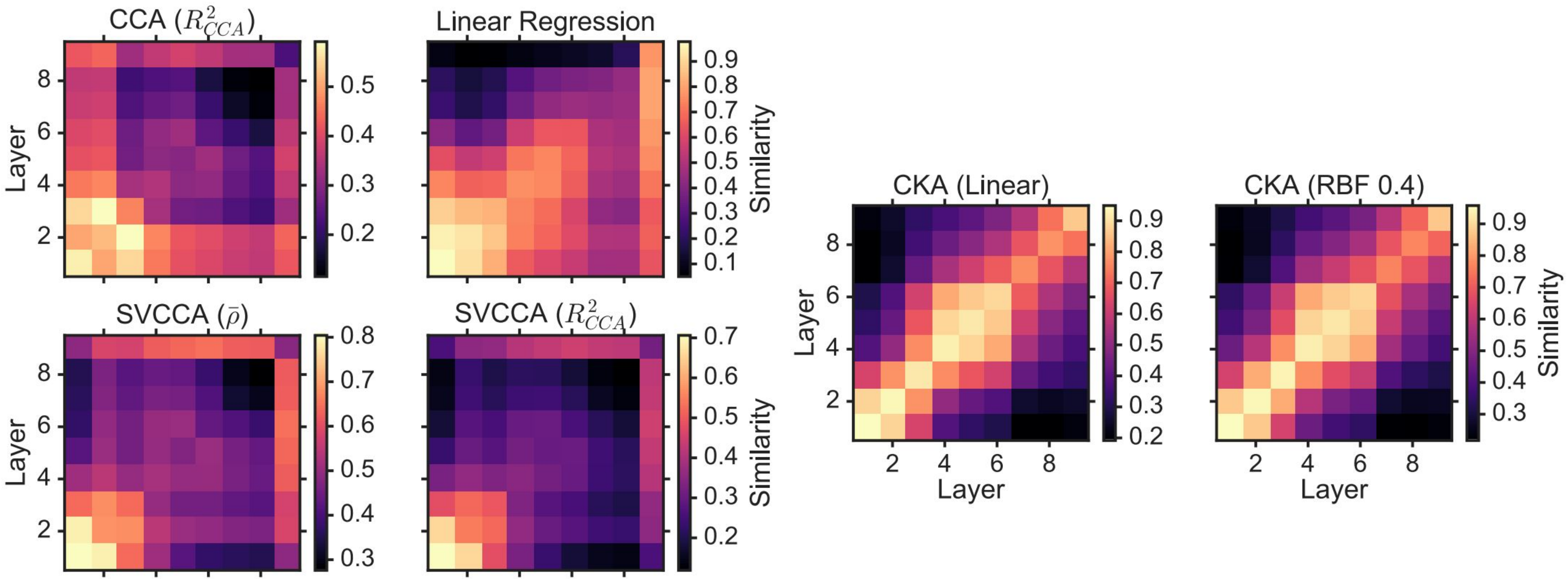
Orthogonal Procrustes distance

(invariant to rotations)

CCA-based distance

(invariant to linear transformations)

Williams et al., *NeurIPS*, 2021



Kornblith et al., *ICML*, 2019

More than metrics

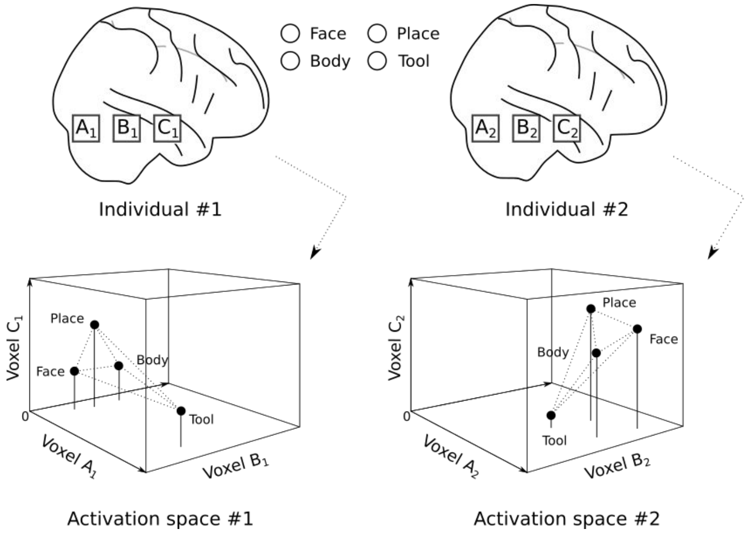
- While metric-based results have driven significant research—and work to define rigorous metrics is ongoing; e.g., Duong et al., 2022—metrics are intended to provide insight at a given level of analysis
- In some contexts, we can also leverage the similarity of the neural systems (e.g., individual brain activations) to improve downstream inferences

Outline

- Measuring dissimilarity with metrics
- Using alignment to leverage similarity

Using alignment to leverage similarity

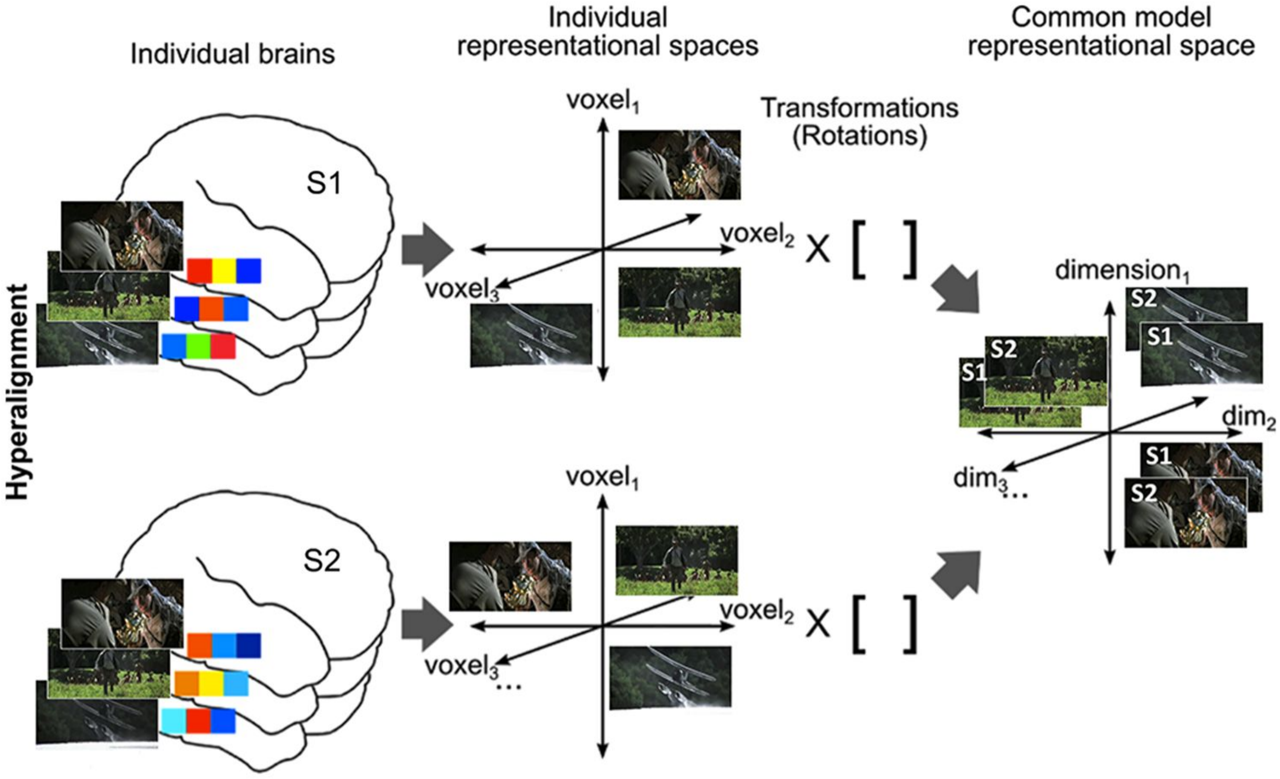
- For a given measure of similarity, we can also calculate transformations that maximize that similarity metric

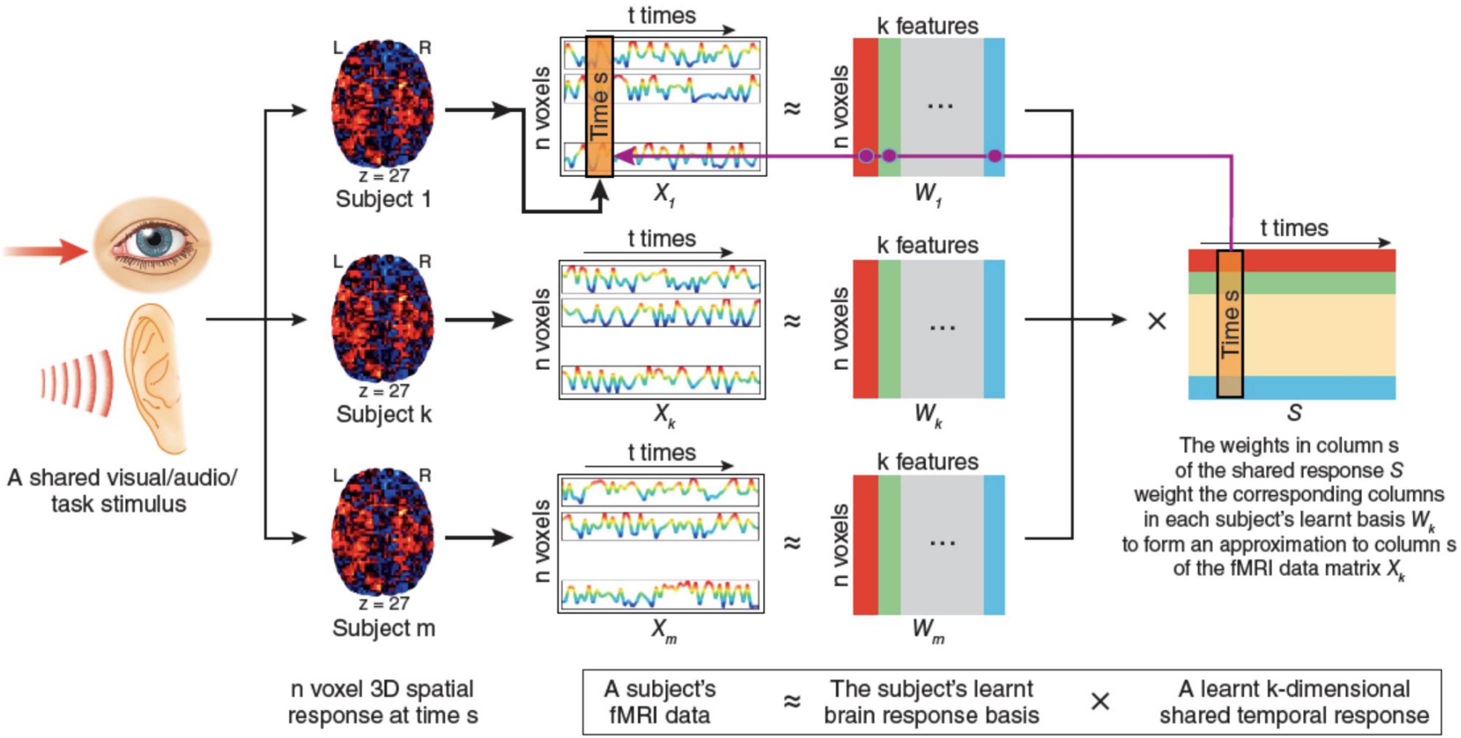


Adapted from [Churchland (1998). *J Philo.*]

Alignment as a rich alternative

- Using alignment, we can directly bring data from different individuals or different experimental conditions into the same functional space
- This can be done:
    - In high- or low-dimensional space
    - Using labelled or or unlabeled experimental data; i.e., with or without knowing correspondence between time points
- These alignments can be re-used in new data


Individual brains
Individual
representational spaces
Common model
representational space
Transformations
(Rotations)
Hyperalignment
S1
S2
voxel$_1$
voxel$_2$
voxel$_3$
X [ ]
dimension$_1$
dim$_2$
dim$_3$...
S1
S2



Cohen et al., *Nat Neuro*, 2017

Choosing an alignment method

- Much as for metrics, *which* alignment method to choose is data dependent and remains largely guided by field norms
- In the cases considered thus far, we assume alignment to a real, known target
  - Calculating such a template is an active research area, with initial methods proposing Generalized Procrustes to an inferred average (Haxby et al., 2020) or Wasserstein barycenters (Thual et al., 2022)

Thank you !

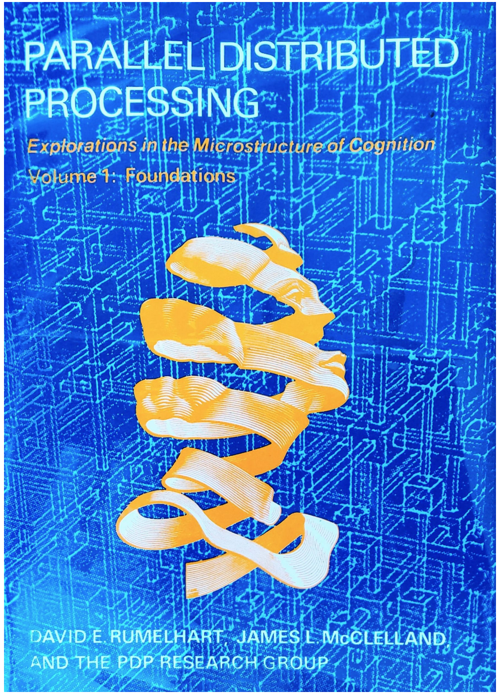
Poldrack Lab

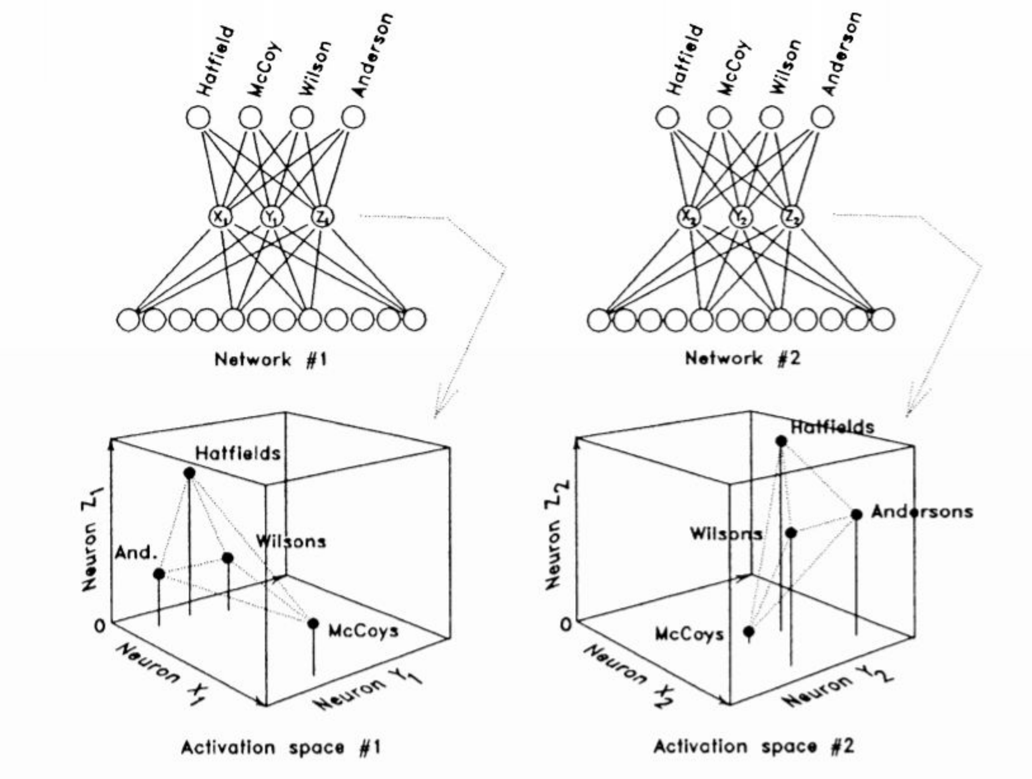
Linderman Lab

Wu Tsai Neuroscience Institute

Let's run some alignments !

Supplementary slides





Churchland, *J Philo*, 1998