Undefined labels?
Try unsupervised approaches!

OHBM2020

ML4NI Educational Course

CENTRE for SLEEP & COGNITION

Valeria Kebets

NUS National University of Singapore

Outline

- Supervised vs. unsupervised approaches
- Partial Least Squares (PLS)
- Comparison with other unsupervised techniques
  - Principal component analysis, Canonical correlation analysis

Supervised vs. unsupervised learning

Data

Labels
defined

dogs

cats

Labels
not defined

Supervised

$$f : \boldsymbol{X} \rightarrow \boldsymbol{y}$$

Unsupervised

Supervised vs. unsupervised learning

Data

Labels
defined

dogs

cats

Supervised

$$f : \boldsymbol{X} \rightarrow \boldsymbol{y}$$

Labels
not defined

floppy ears

pointy ears

Unsupervised

Supervised vs. unsupervised learning

Data

brown

grey

Labels
defined

dogs

cats

Labels
not defined

Supervised

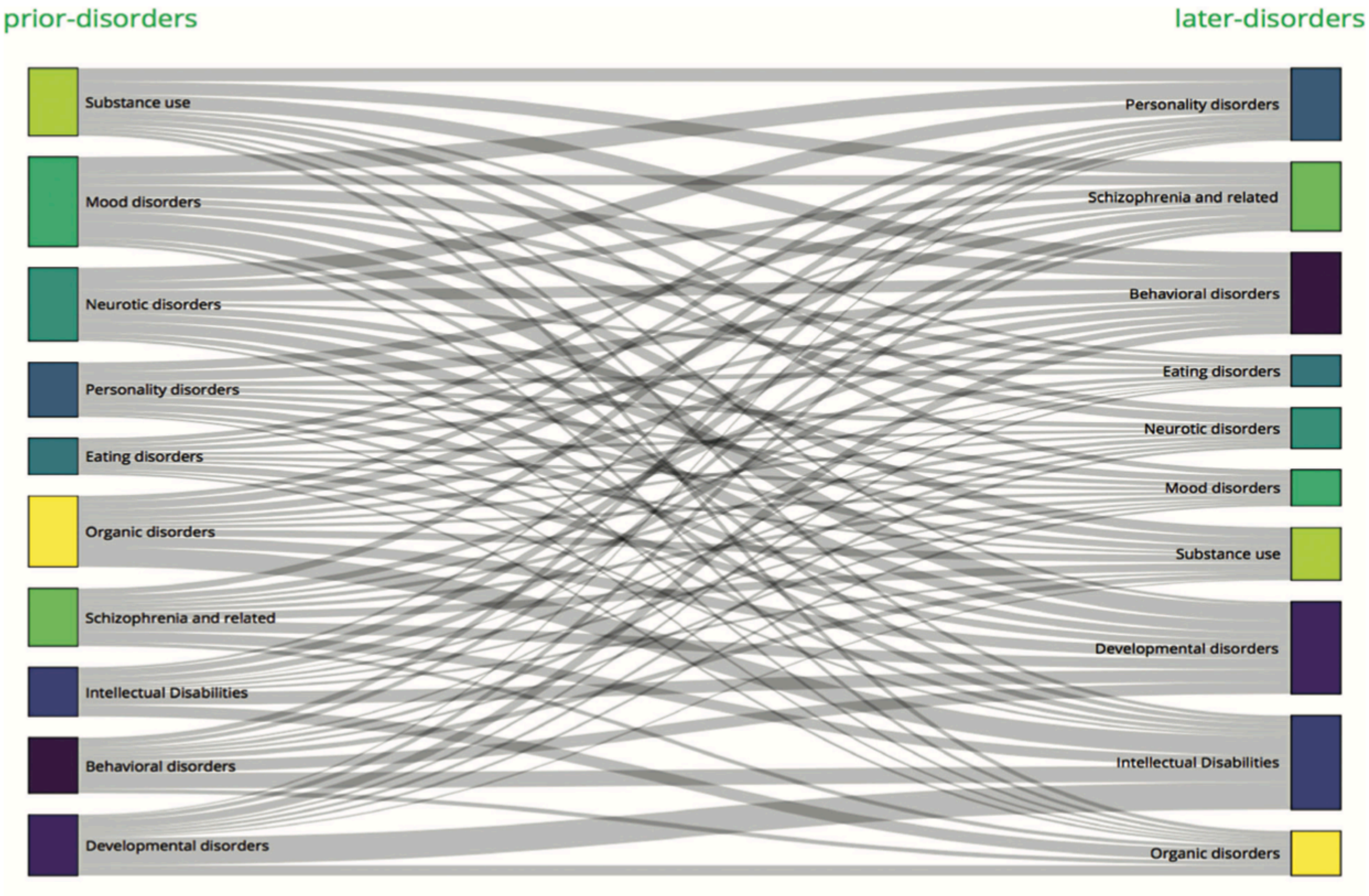
$$f : \boldsymbol{X} \rightarrow \boldsymbol{y}$$

Unsupervised

When is unsupervised learning useful?

- **When labels are not reliable**
  - In psychiatry, disorders are classified according to criteria from the DSM and ICD
  - However, clinical presentations are heterogeneous *within* a diagnostic category
  - There is also overlap in clinical symptoms, cognitive deficits, & genetic risk factors *across* diagnostic categories, and high comorbidity *among* disorders

When is unsupervised learning useful?



Plana-Ripoll et al. (2019)

When is unsupervised learning useful?

- **When labels are not reliable**
  - In psychiatry, disorders are classified according to criteria from the DSM and ICD
  - However, clinical presentations are heterogeneous *within* a diagnostic category
  - There is also overlap in clinical symptoms, cognitive deficits, & genetic risk factors *across* diagnostic categories, and high comorbidity *among* disorders

- Unsupervised learning can help uncover underlying neurobiological mechanisms that **transcend** diagnostic boundaries

Outline

- Partial Least Squares (PLS)

Partial least squares (PLS)

- PLS finds optimal associations between 2 matrices (X and Y)

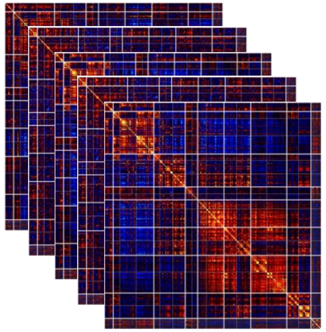
PLS variant	X	Y
Behavior PLS	Imaging measures	Behavior measures
PLS Discriminant Analysis	Imaging measures	Group labels
Multi-Block PLS	Imaging measures	Behavior + Conditions (+ ...)
Task / Spatiotemporal PLS	Brain activity x Timeseries	Contrasts / Task conditions
Seed PLS	Whole brain activity	Seed activity

Partial least squares (PLS)

Goal = Find the **shared** information between the 2 modalities

= Find imaging patterns **optimally** related to behavioral patterns

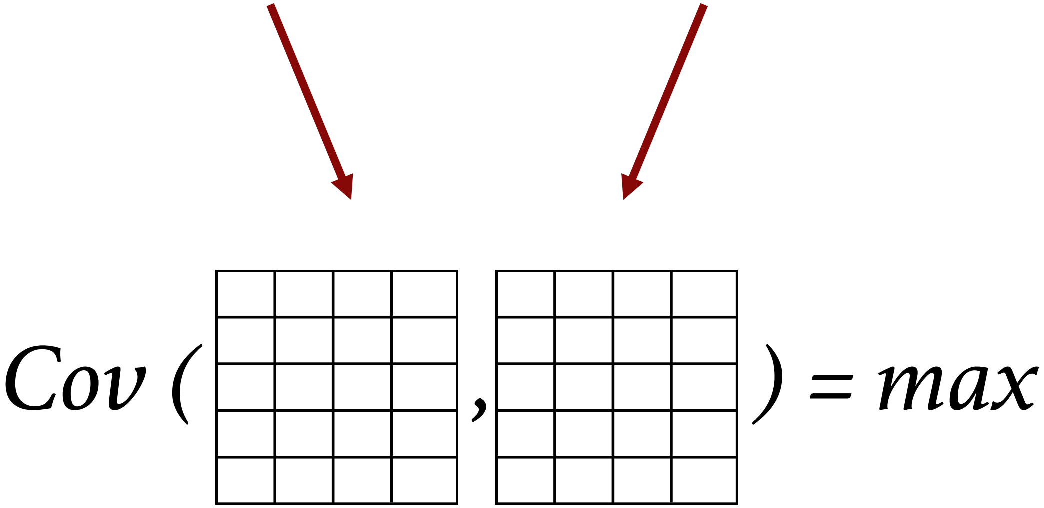
PLS finds low-dimensional **latent variables** showing maximal covariance



Imaging data

	Depression	Mania	IQ	Verbal memory
Subject 1	12	3	82	12
Subject 2	4	6	114	25
Subject 3	6	2	108	19
Subject 4	3	7	120	21
Subject 5	10	2	95	16

Behavior data

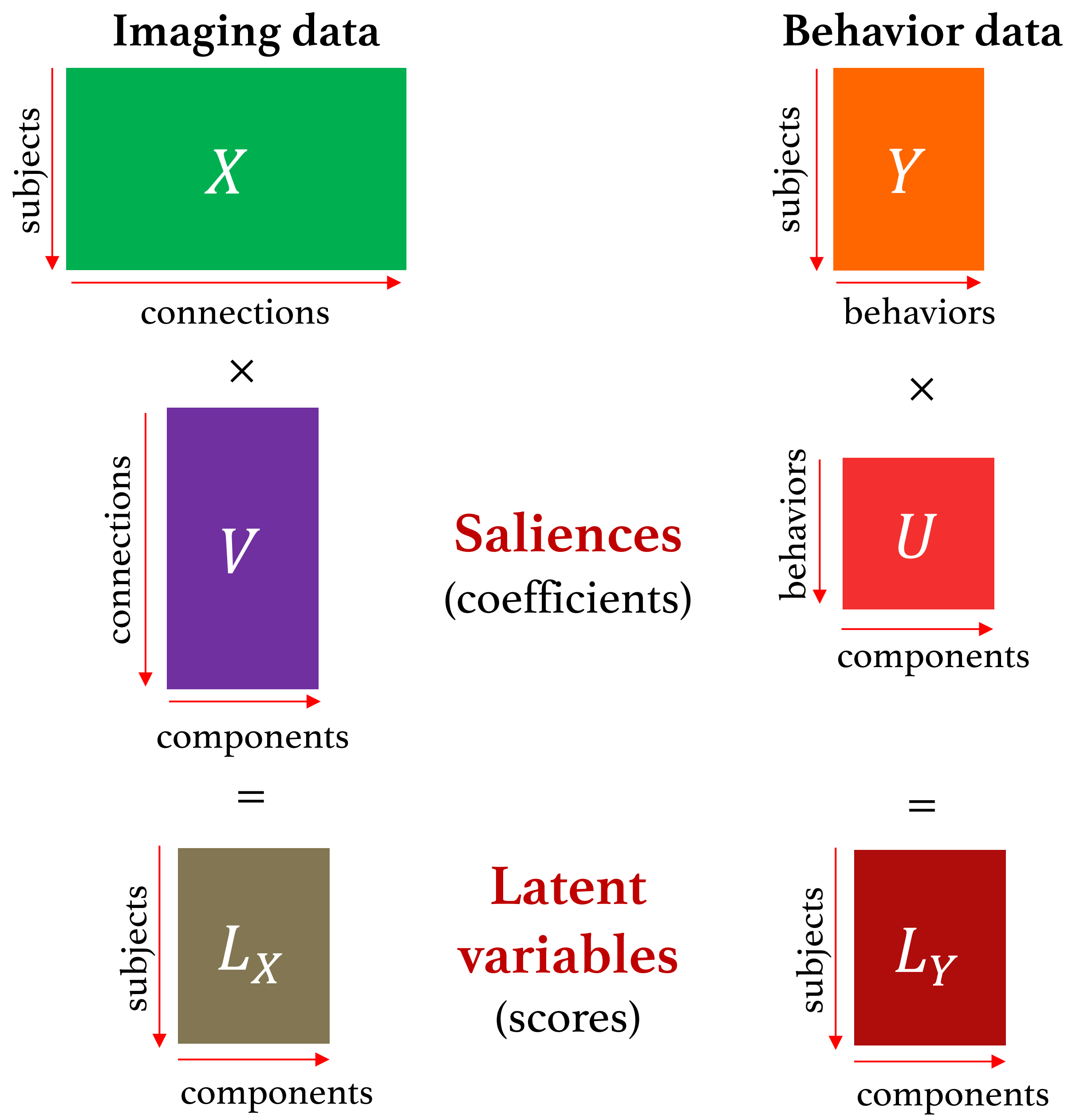


PLS

PLS finds

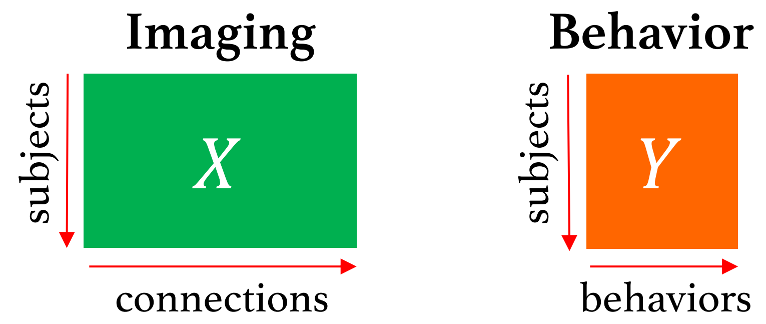
weighted pairs of vectors

whose *projection*
on original data yields
maximal **covariance**



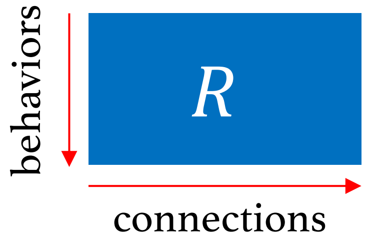
PLS

Original data



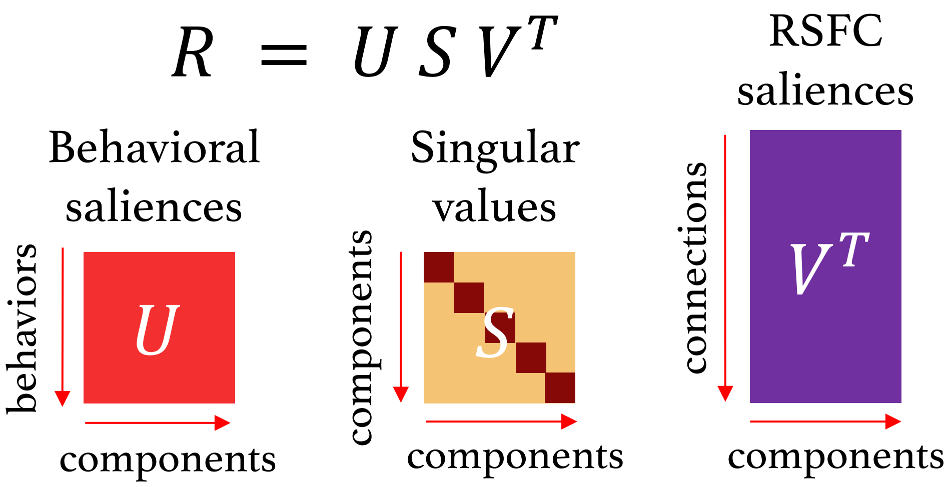
Cross-covariance matrix

$$Y^T X = R$$



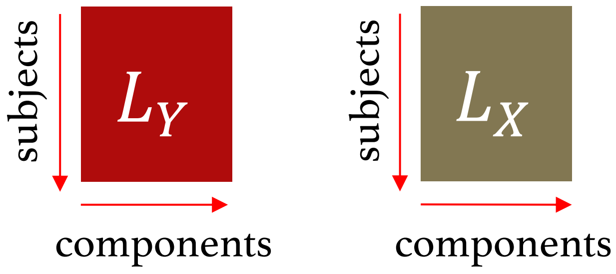
Singular value decomposition

$$R = U\,S\,V^T$$

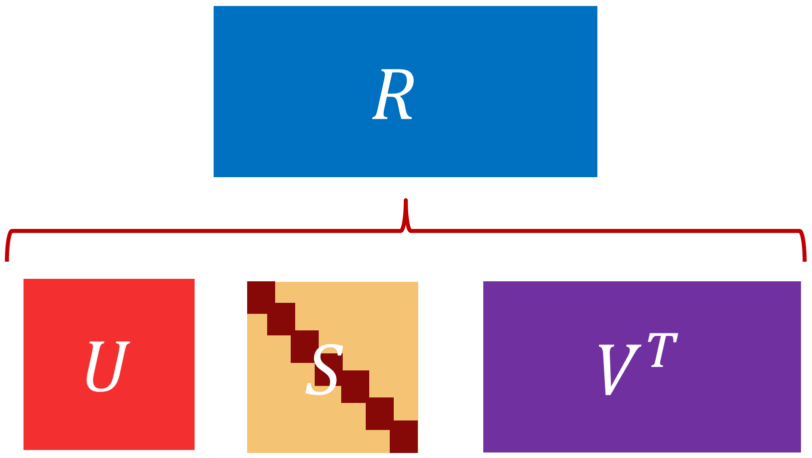


Latent variables

Behavioral subjects' scores	RSFC subjects' scores
$Y\,U$	$X\,V$



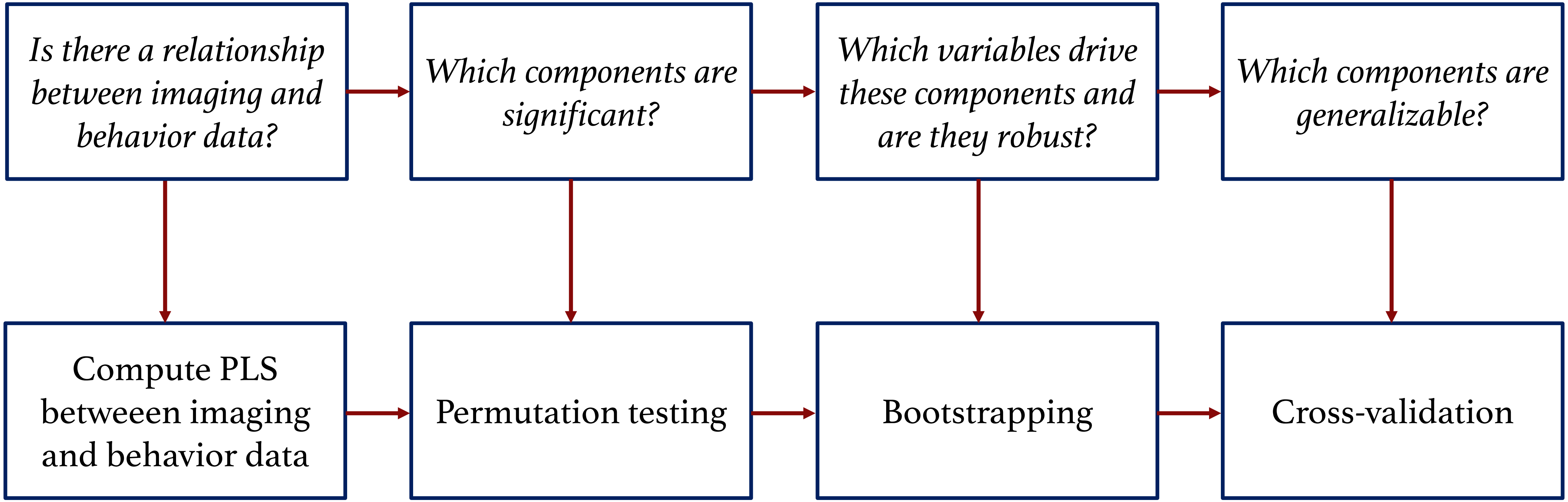
PLS | Singular value decomposition



s_i : *singular value of component i*

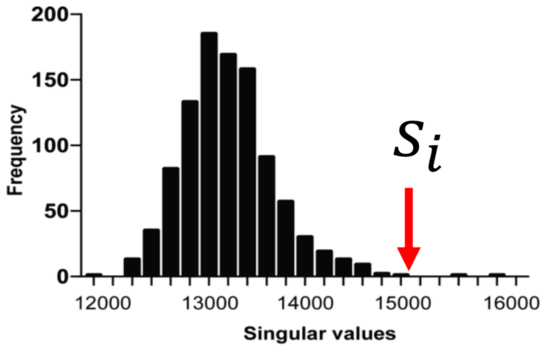
- Rank of cross-covariance matrix determines the number of components
- Components are ordered by effect size
  - Amount of covariance explained by each component $= \dfrac{s_i^{\,2}}{\sum s^2}$
- Components are orthogonal
  - Each component explains a different part of the covariance between imaging and behavior data

PLS | Analysis flowchart



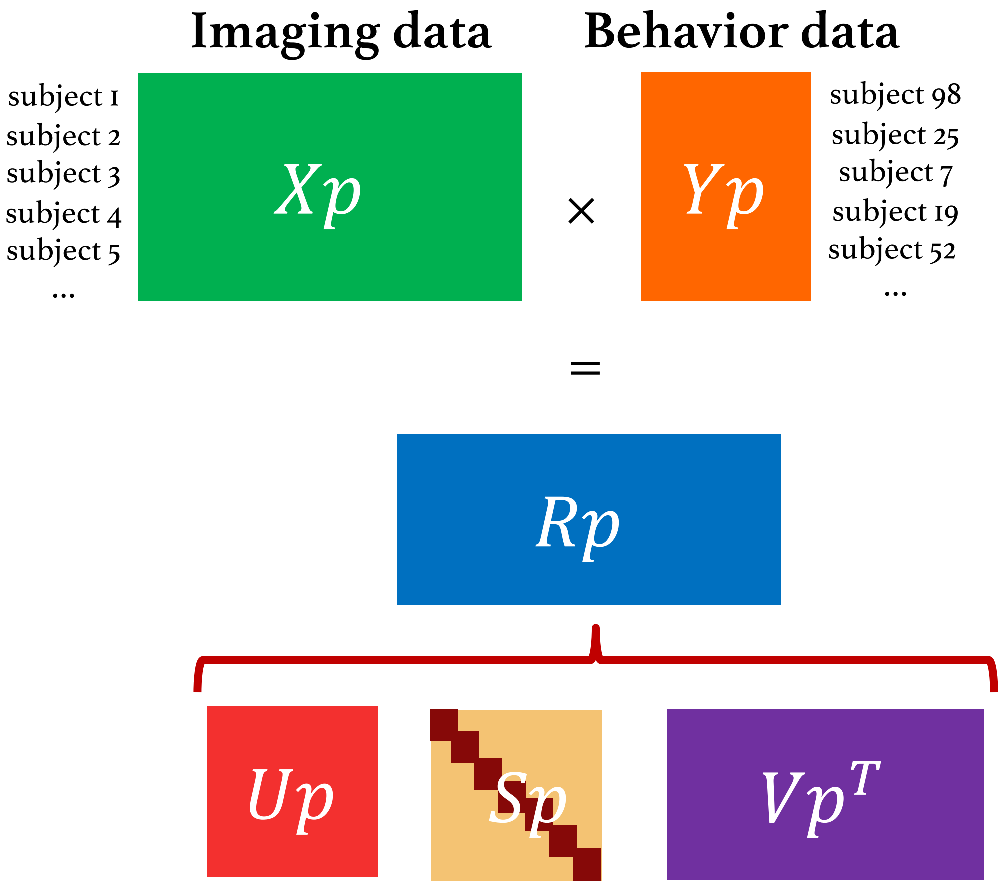
PLS | Permutation testing

- **Which components are significant?**
  - Permute rows (subjects) in Y
  - Distribution of singular values under the null hypothesis



$$p = \frac{1 + \sum_{c=1}^{C} sp_{ic} \geq s_i}{1 + C}$$

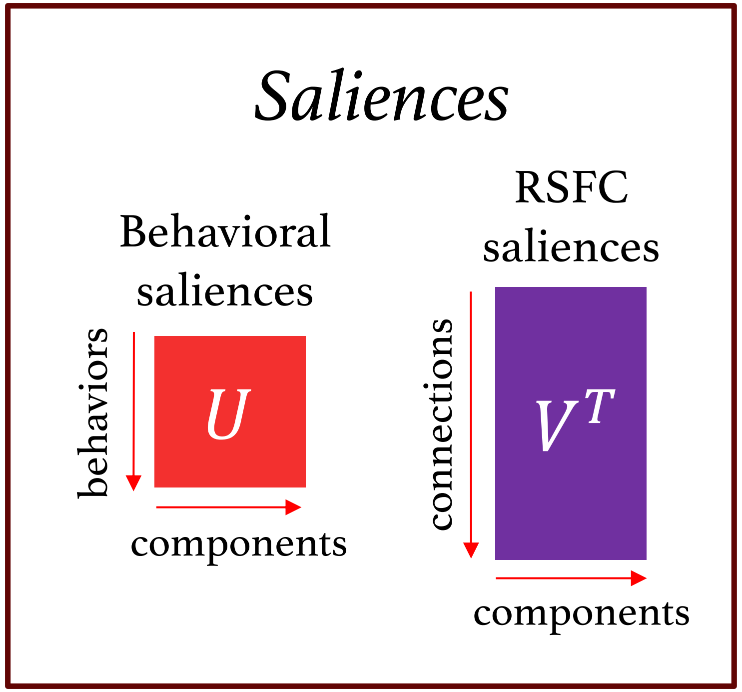
C = number of permutations

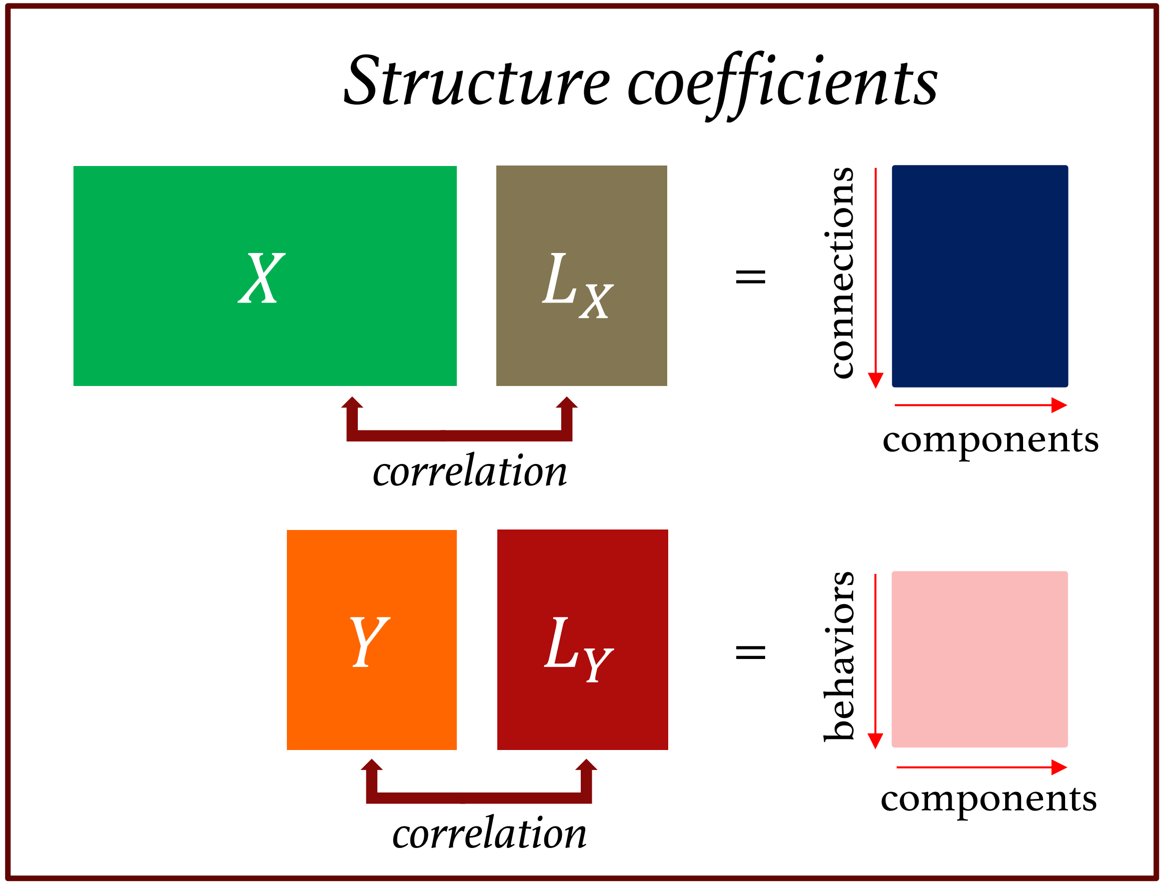


sp_{ic} : *permuted singular values of component i*

PLS | Loadings

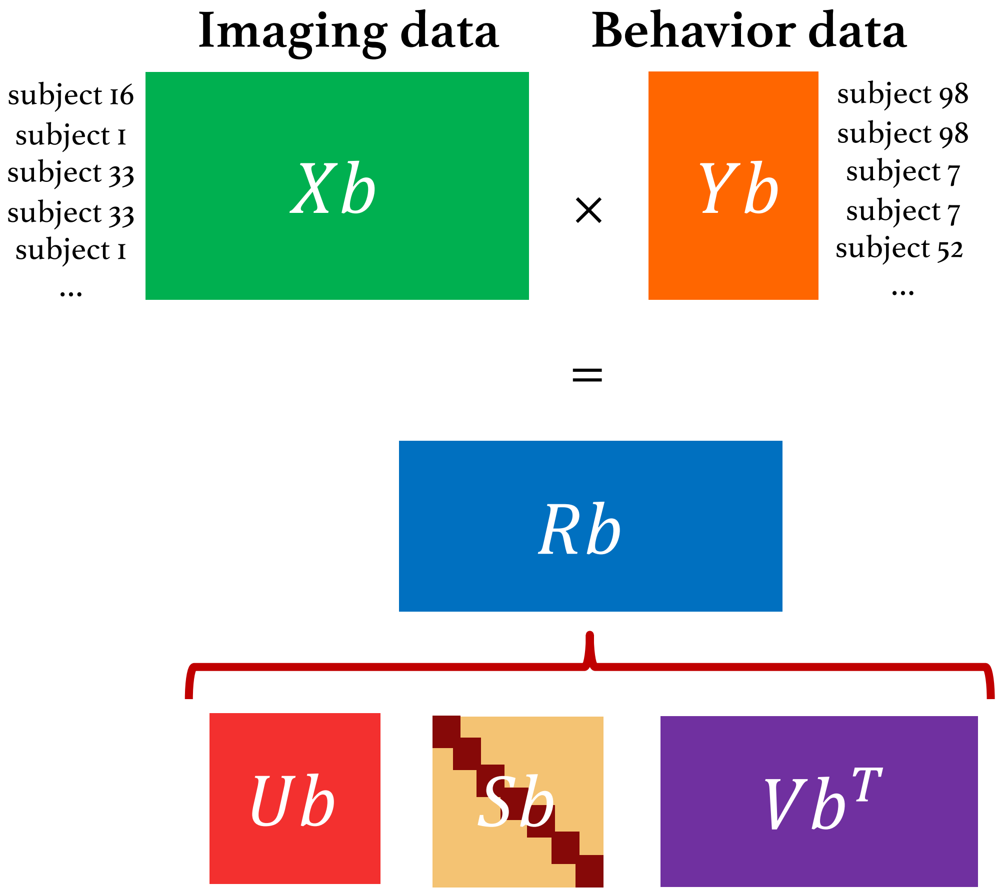
- **Which variables characterize these components?**
  - Loadings can either be :





PLS | Bootstrapping

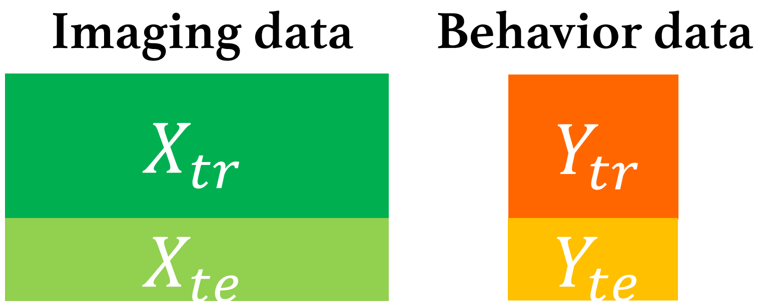
- **Which loadings are stable (irrespective of the sample) ?**
  - Sample rows (subjects) in X and Y with replacement
  - Bootstrap ratio: $\dfrac{u_i}{\hat{\sigma}(ub_i)}$ and $\dfrac{v_i}{\hat{\sigma}(vb_i)}$
  - High loadings with low standard error are considered stable



ub_i : *bootstrapped behavior saliences of component i*

vb_i : *bootstrapped imaging saliences of component i*

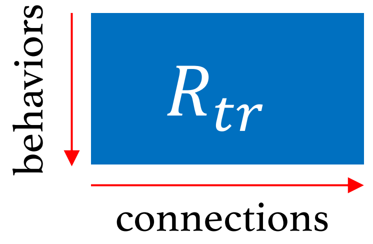
PLS | Cross-validation



- **Which components are generalizable to unseen data?**

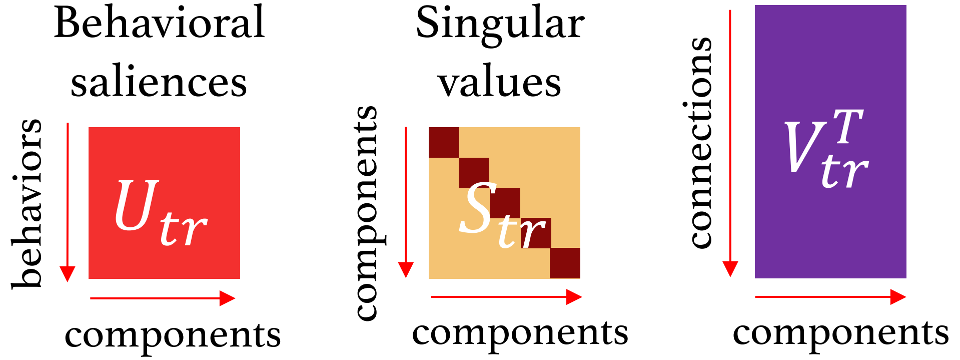
Cross-covariance matrix

$$Y_{tr}^T X_{tr} = R_{tr}$$



Singular value decomposition

$$R_{tr} = U_{tr} S_{tr} V_{tr}^T$$



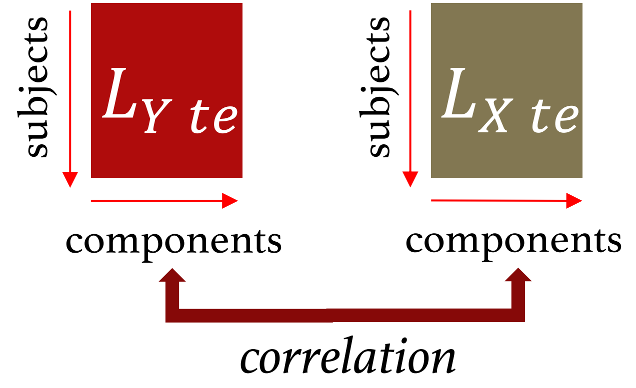
Latent variables

Behavioral subjects' scores

$$Y_{te} U_{tr}$$

RSFC subjects' scores

$$X_{te} V_{tr}$$

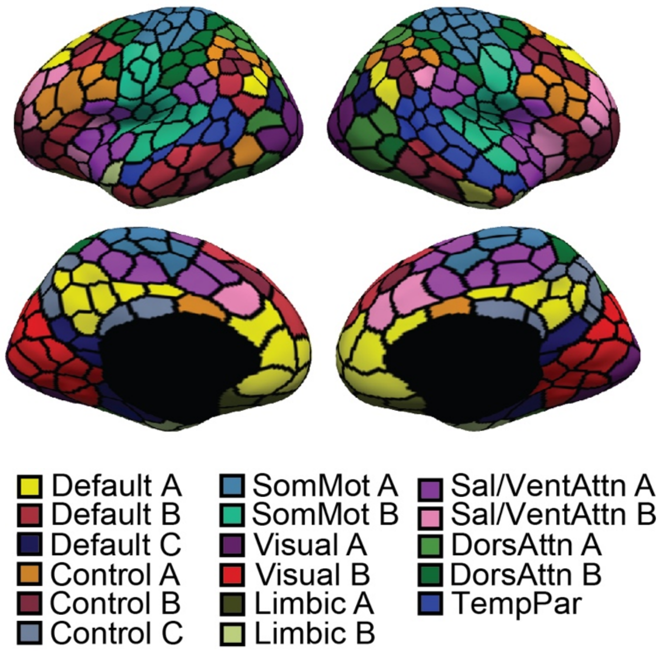


PLS | Illustration

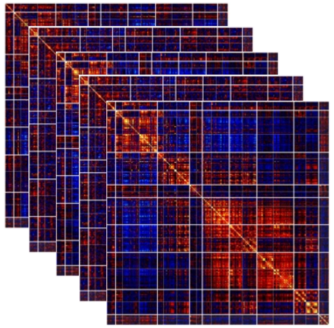
Kebets et al.
Biol Psych (2019)

UCLA Consortium for Neuropsychiatric Phenomics dataset
224 subjects
(psychiatric patients & controls)

400 cortical ROIs
(Schaefer et al., 2018)



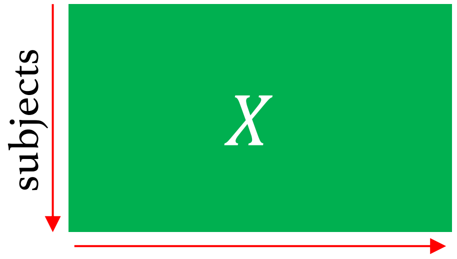
+ 19 subcortical ROIs
(Fischl et al., 2002)



	Hallucinations	Vocabulary	Impulsivity
Subject 1	12	33	15
Subject 2	4	26	11
Subject 3	6	29	21

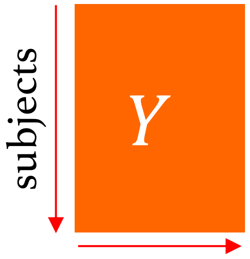
- Clinical symptoms (e.g., hallucinations)
- Cognitive measures (e.g., vocabulary)
- Personality measures (e.g., impulsivity)

Imaging data



87'571 connections

Behavior data

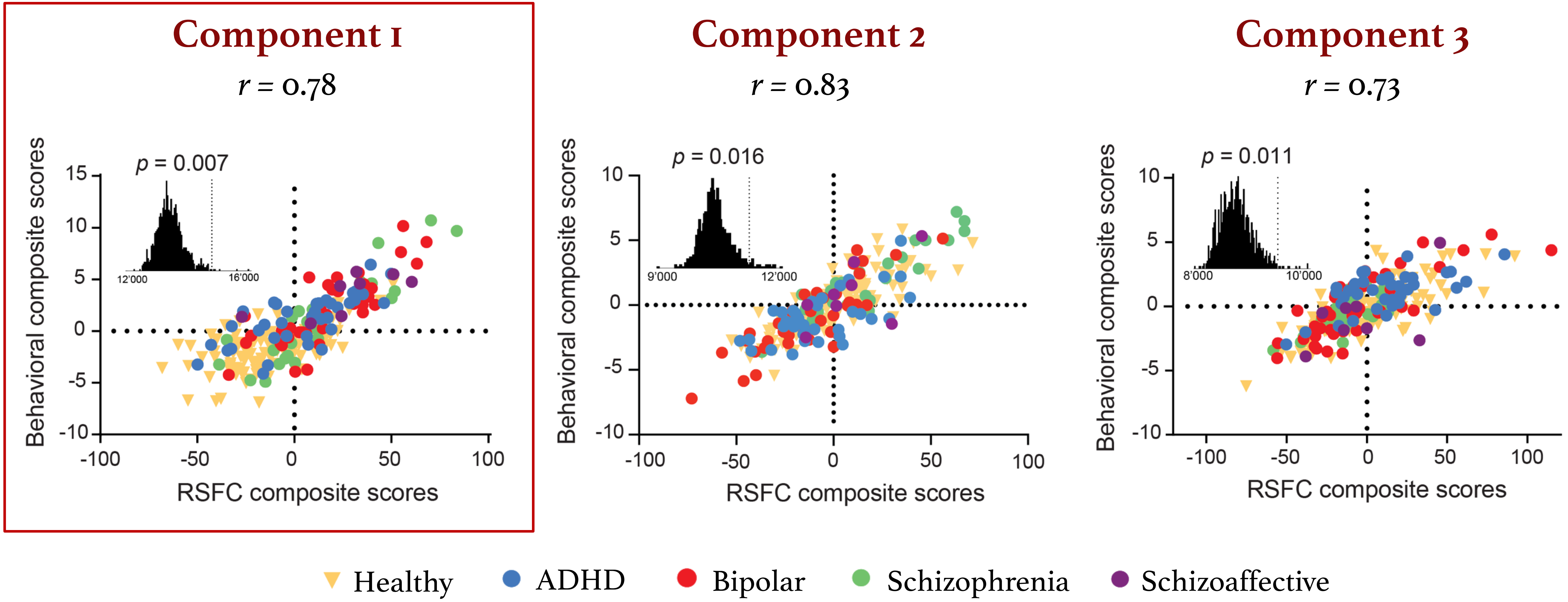


54 behaviors

PLS | Components' significance

Kebets et al.
Biol Psych (2019)

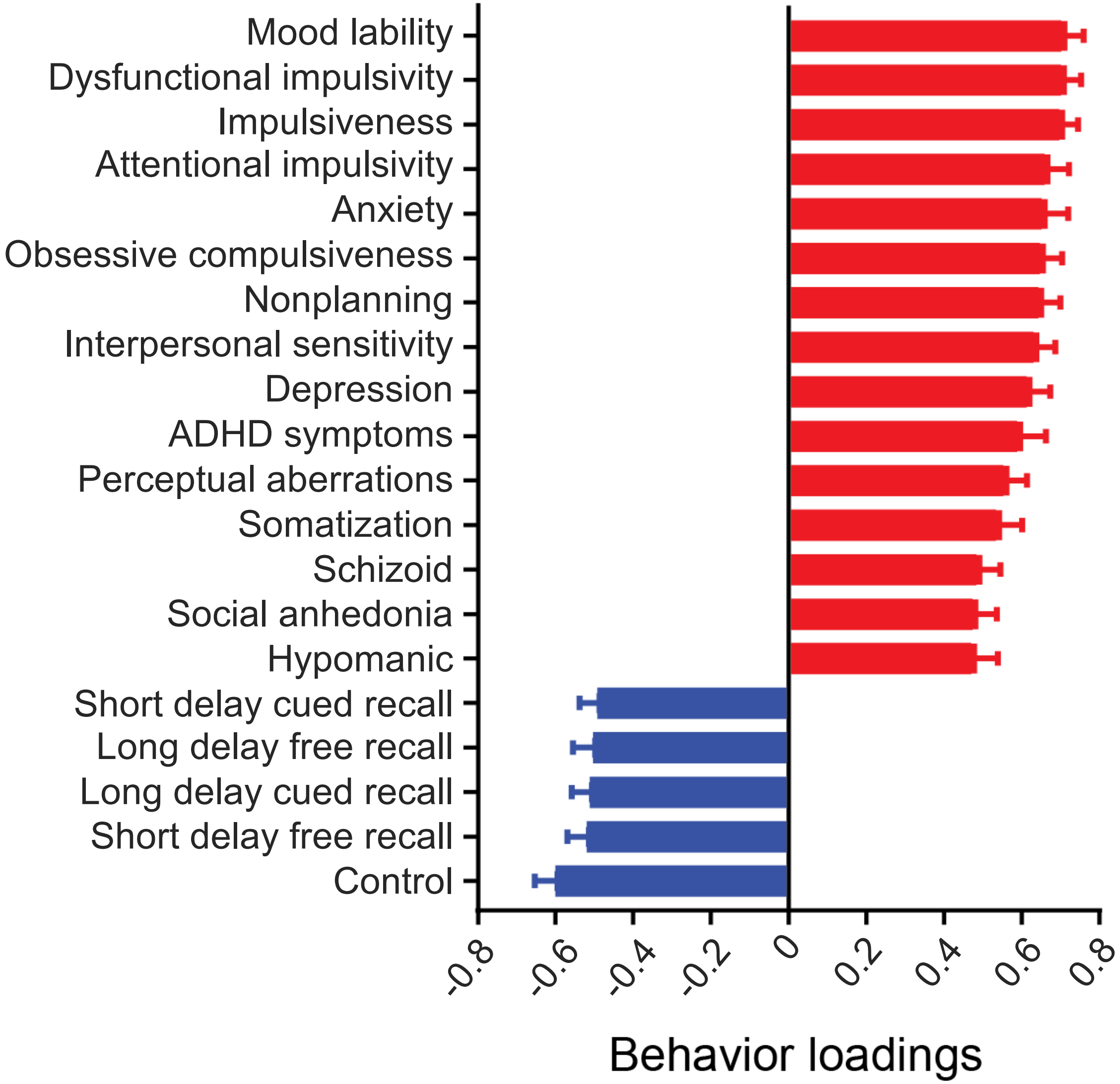
- 3 significant components found using permutation testing (1'000 permutations)



Kebets et al.
Biol Psych (2019)

PLS | Loadings

- Behavior loadings (structure coefficients)
- Top 20 behavior loadings characterized by higher clinical symptoms
  → **General psychopathology**

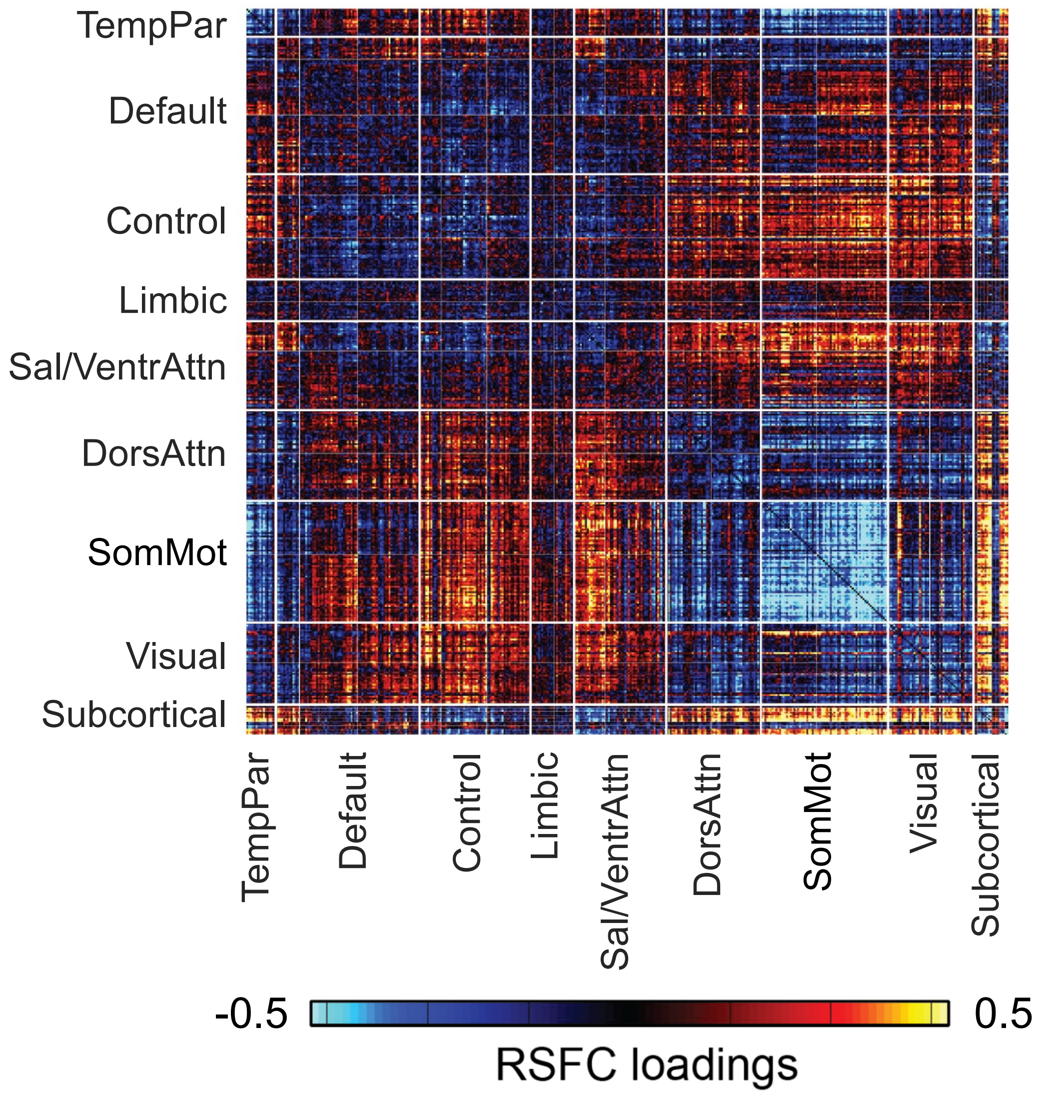


Kebets et al.
Biol Psych (2019)

PLS | Loadings

- RSFC loadings (structure coefficients)

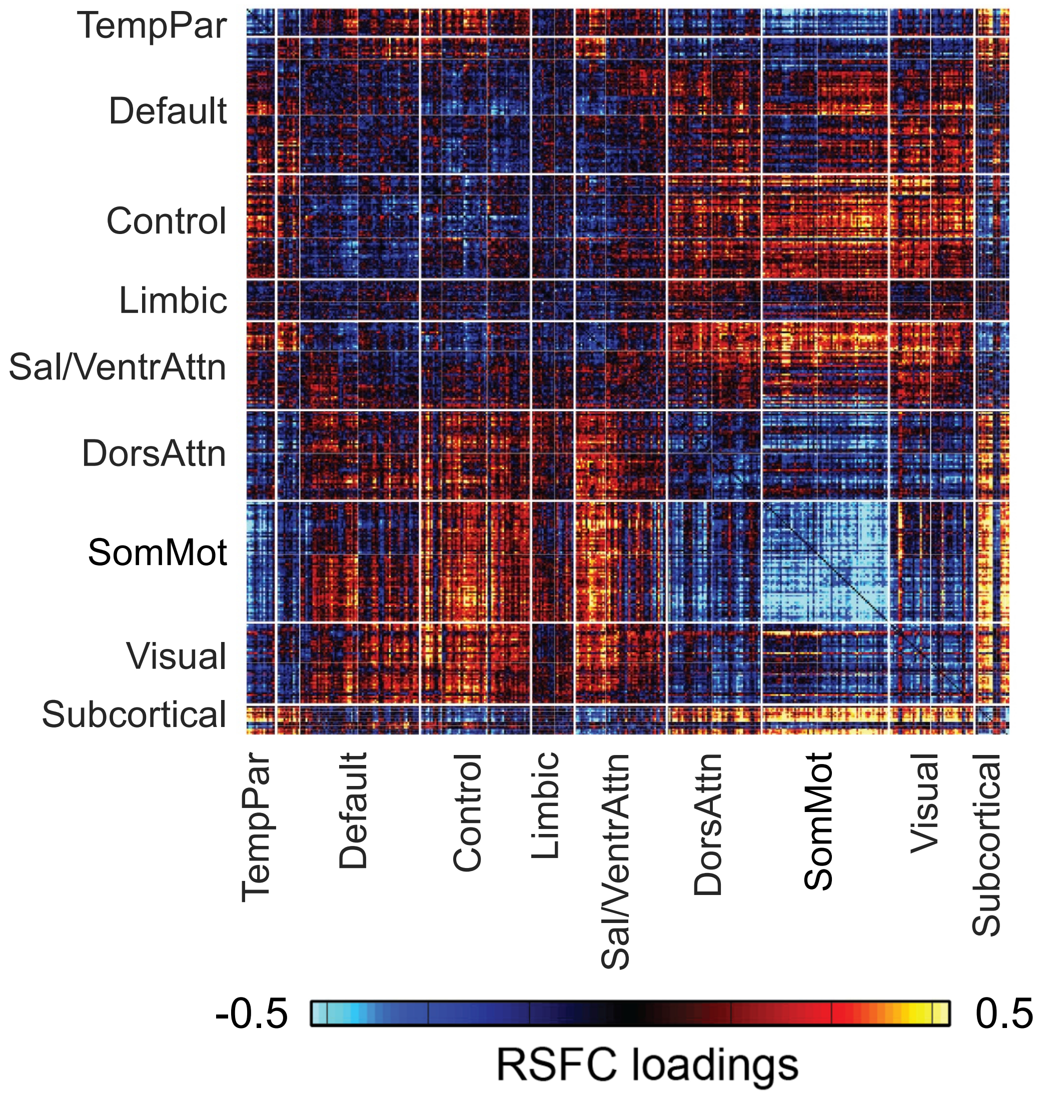
- With greater psychopathology
  - FC within the Somatomotor network is decreased
  - The Dorsal attention, Somatomotor and Visual networks show increased FC with the Default, Control, Salience and Subcortical networks



PLS | Loadings' robustness

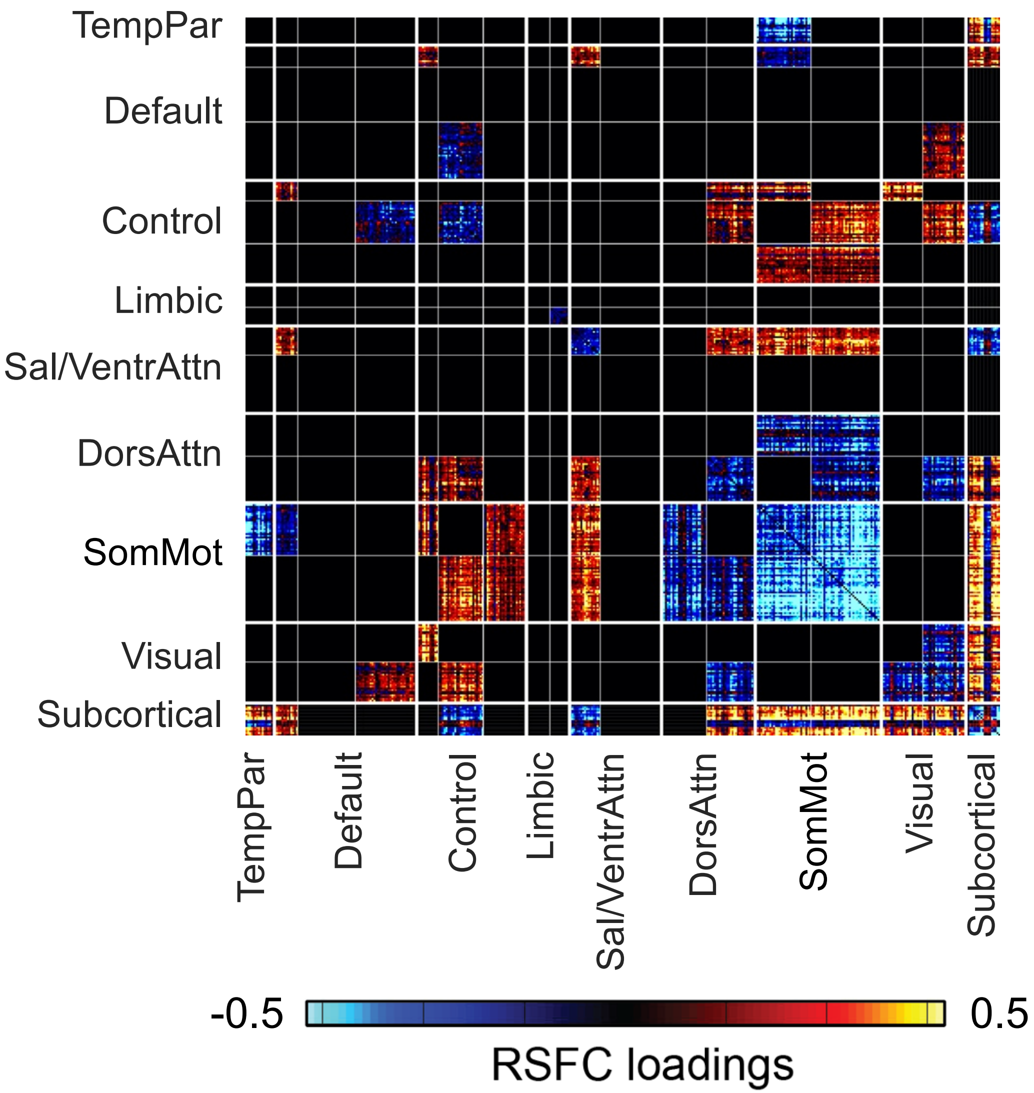
Kebets et al.
Biol Psych (2019)

- Bootstrap resampling (500 samples)
  - Bootstrap estimation of loadings' standard errors
  - Bootstrap ratios $= \dfrac{loading_{\,i}}{\hat{\sigma}(loading_{\,i})}$
  - Bootstrap ratios averaged within/between networks
  - Bootstrap ratios (≈ z-scores) converted to p-values & FDR-corrected



PLS | Loadings' robustness

- Bootstrap resampling (500 samples)
  - Bootstrap estimation of loadings' standard errors
  - Bootstrap ratios $= \dfrac{loading_i}{\hat{\sigma}(loading_i)}$
  - Bootstrap ratios averaged within/between networks
  - Bootstrap ratios (≈ z-scores) converted to p-values & FDR-corrected



PLS | Cross-validation

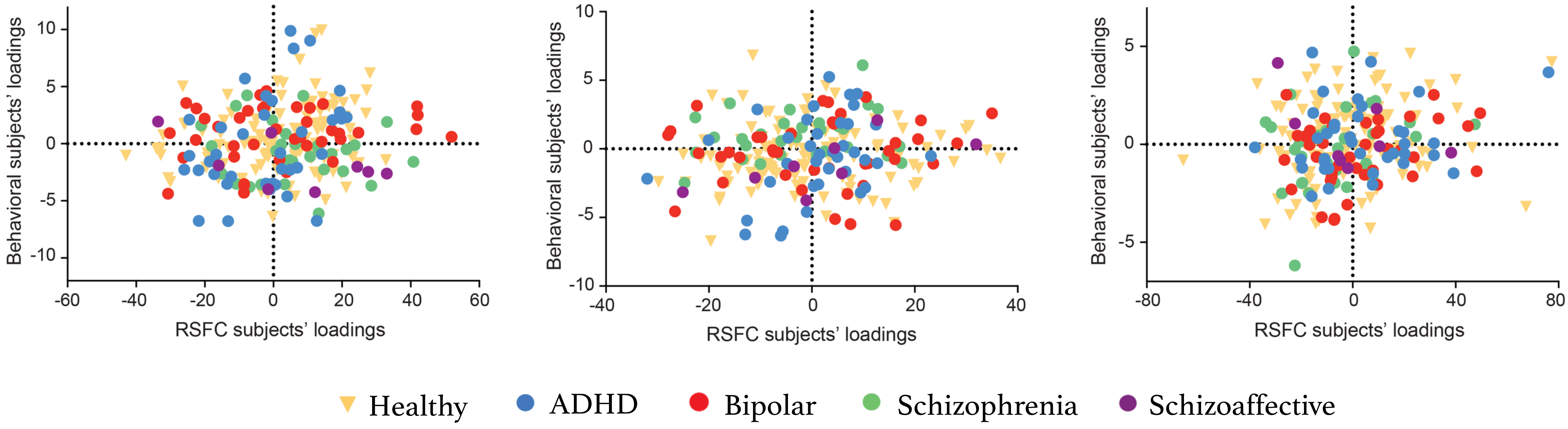
Kebets et al.
Biol Psych (2019)

- The 3 components estimated from 80% successfully generalized to the remaining 20%

Component 1
r = 0.15 (mean)
p = 0.001-0.002 (range)

Component 2
r = 0.12 (mean)
p = 0.002-0.004 (range)

Component 3
r = 0.18 (mean)
p = 0.002-0.002 (range)



Outline

- Comparison with other unsupervised techniques
  - Principal component analysis, Canonical correlation analysis

Principal component analysis (PCA)

- Principal component analysis (PCA) creates a set of new variables, called **principal components**, which are linear combinations of the original variables in X

- Principal components are mutually **uncorrelated** and capture unique, non-overlapping portions of variance

- Principal components are ordered by the magnitude of their squared singular values, which are proportional to the portion of variance accounted for by the component

PLS vs. PCA

PLS

SVD of covariance matrix

$$Y'X = U\ S\ V'$$

saliences

Latent variables

$$L_X = X\ V$$

$$L_Y = Y\ U$$

PCA

SVD of covariance matrix

$$X'X = U\ S\ V'$$

coefficients

Component scores

$$L_X = X\ V$$

Canonical correlation analysis (CCA)

- Conceptually, canonical correlation analysis is very similar to PLS
  - CCA aims to create pairs of new variables, called **canonical variates**, which are linear combinations of the original variables (X and Y), that are maximally **correlated**

- Mathematically, there's an extra step in CCA
  - X and Y are first adjusted for within-set correlations before computing the cross-correlation matrix
  - However, often $p > n$, so the matrix inverse $X'X^{-1}$ doesn't exist because $X'X$ is rank deficient
  - Therefore, **dimensionality reduction is usually applied before computing CCA**

PLS vs. CCA

SVD of covariance matrix

$$Y'X = U\ S\ V'$$

saliences

Latent variables

$$L_X = X\ V$$

$$L_Y = Y\ U$$

SVD of correlation matrix

$$(Y'Y)^{-\frac{1}{2}}\ Y'X\ (X'X)^{-\frac{1}{2}} = U\ S\ V'$$

canonical variates

Canonical scores

$$L_X = X\ V$$

$$L_Y = Y\ U$$

PLS | Implementation

- myPLS - MATLAB - Daniela Zöller & Valeria Kebets
  - https://github.com/danizoeller/myPLS
  - Behavior PLS with 1D, 2D, 3D imaging data

→ → **POSTER 1111**

- PLS - MATLAB - Rotman Baycrest
  - https://www.rotman-baycrest.on.ca/index.php?section=84
  - Behavior PLS, Seed PLS, Task PLS with 3D, 4D imaging data
- PYLS - Python - Ross Markello
  - https://github.com/rmarkello/pyls
  - Behavior PLS, Mean-centered PLS, PLS regression

References

PLS

- Krishnan A, Williams LJ, McIntosh AR, Abdi H (2011). Partial Least Squares (PLS) methods for neuroimaging: a tutorial and review. *Neuroimage, 56(2), 455-75.*

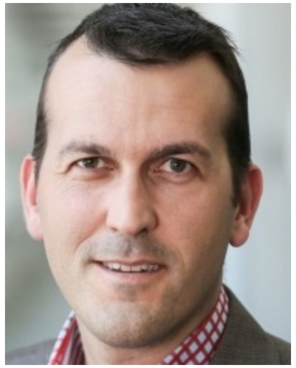
- McIntosh AR, Lobaugh NJ (2004). Partial least squares analysis of neuroimaging data: applications and advances. *Neuroimage, 23(Suppl 1), S250-63.*

Related methods

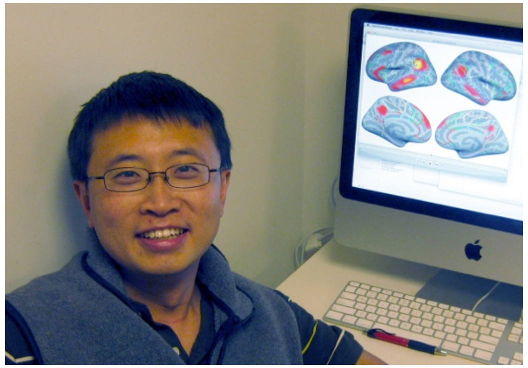
- McIntosh AR, Misic B (2013). Multivariate statistical analyses for neuroimaging data. *Annu Rev Psychol, 64, 499-525*

- Sui J, Adali T, Yu Q, Chen J, Calhoun VD (2012). A review of multivariate methods for multimodal fusion of brain imaging data. *J Neurosci Methods, 204(1),68-81*

Stay safe &
thank you for your attention !



Dimitri Van De Ville



B.T. Thomas Yeo

NUS National University of Singapore | N.1 INSTITUTE FOR HEALTH

CENTRE for SLEEP & COGNITION

MINISTRY OF EDUCATION SINGAPORE

Ministry of Education SINGAPORE

NMRC National Medical Research Council Singapore

Memory Network Program

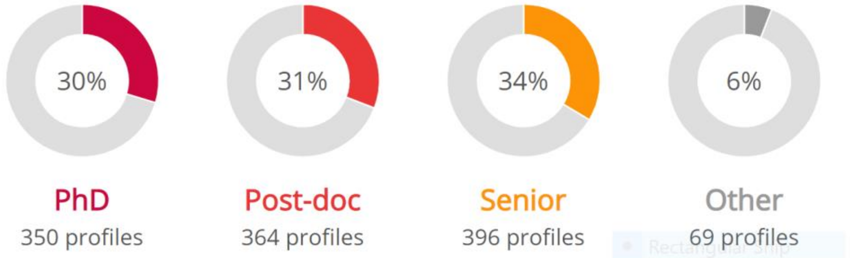
NATIONAL RESEARCH FOUNDATION
PRIME MINISTER'S OFFICE
SINGAPORE

CLINICAL IMAGING RESEARCH CENTRE
SINGAPORE
A joint venture between the Agency for Science, Technology And Research (A*STAR) and the National University of Singapore (NUS)

Women in Neuroscience

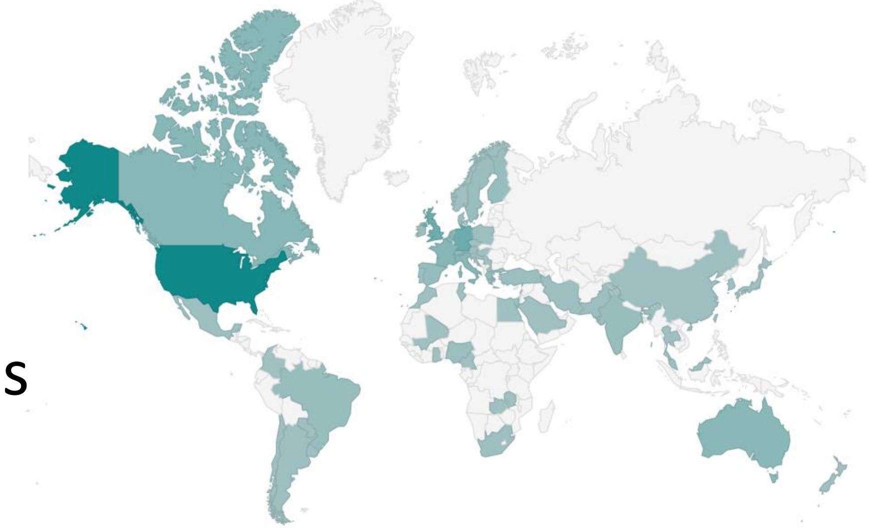
Repository for Women in Neuroscience

- www.winrepo.org
- over 1,100 profiles
- easy search
- recommendations



Support the project:

- ➢sign up
- ➢spread the word
- ➢submit recommendations



WINRePo1
www.facebook.com/WiNRepository/