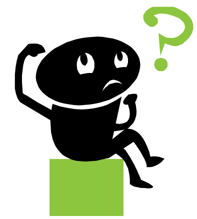
MULTIVARIATE ANALYSES WITH fMRI DATA

Sudhir Shankar Raman

Translational Neuromodeling Unit (TNU)
Institute for Biomedical Engineering
University of Zurich & ETH Zurich

Motivation

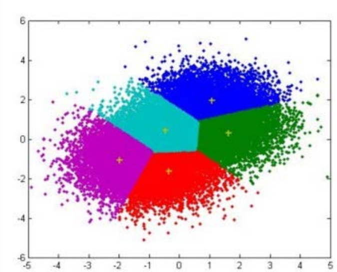


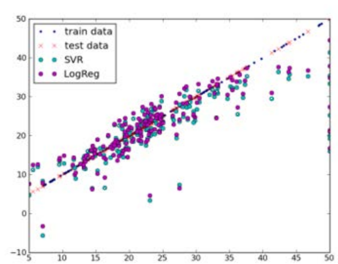
Modelling Concepts

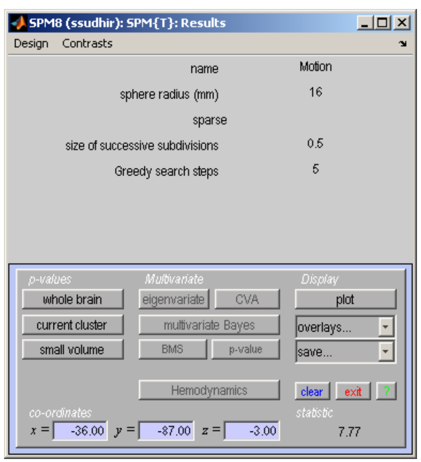
Learning From Data

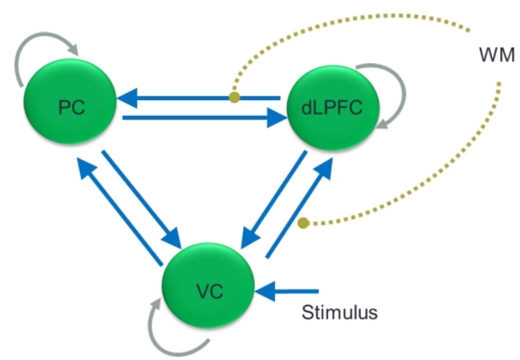
Multivariate Bayes in SPM

Generative Embedding



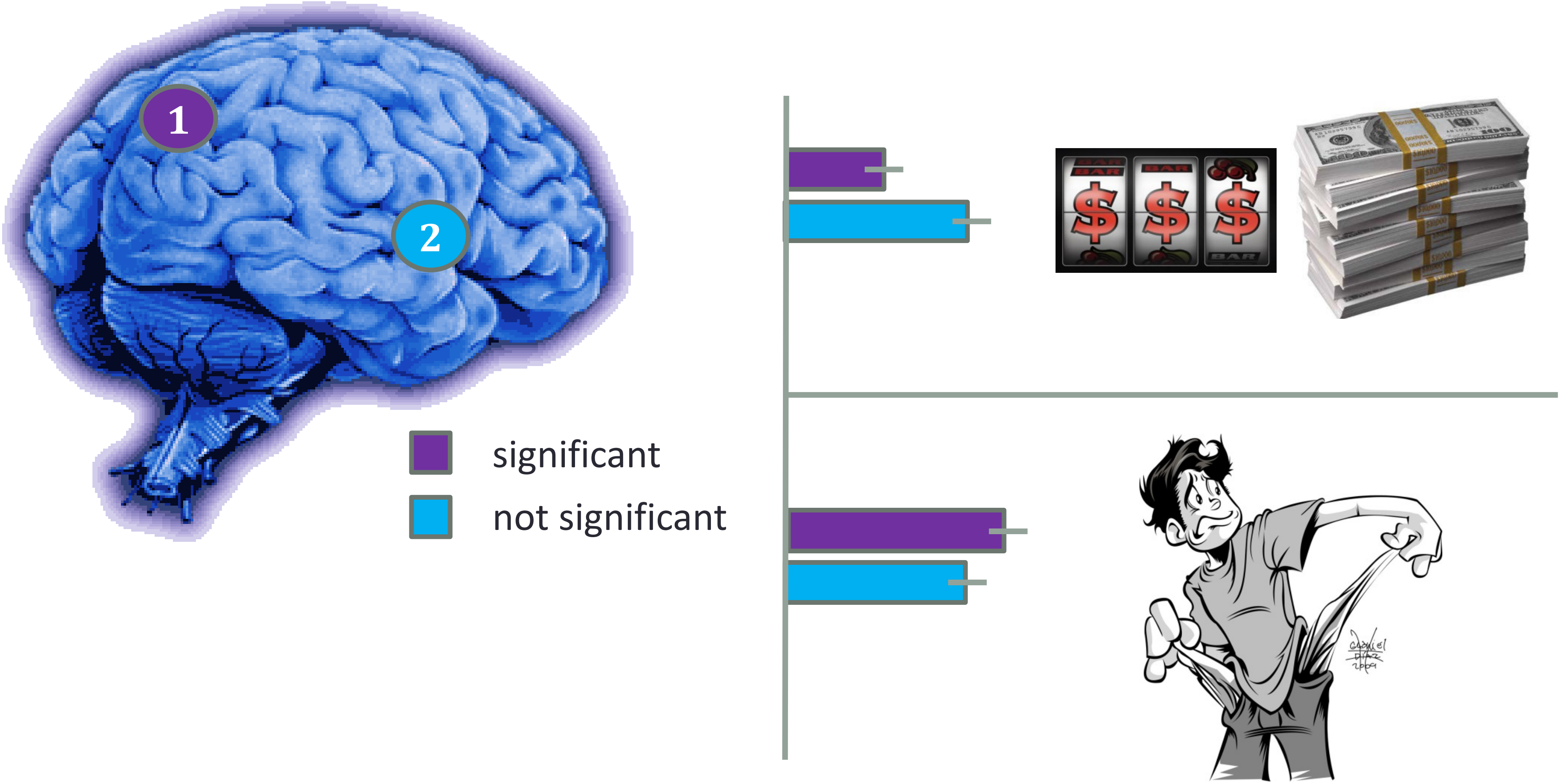




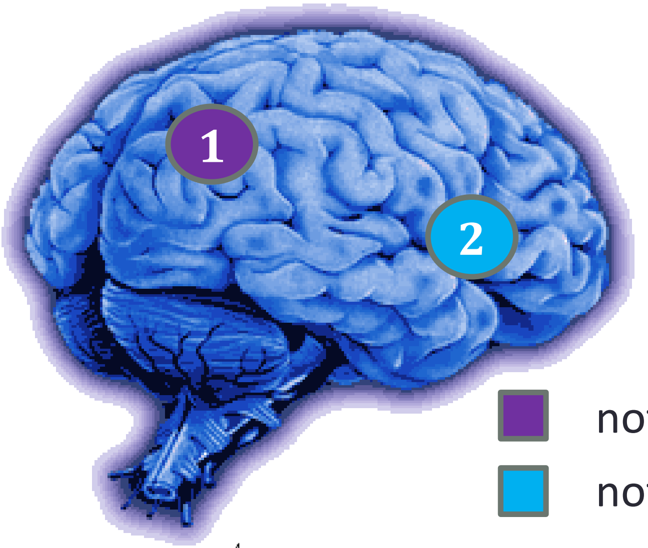


Motivation

- Local activations – Univariate approach

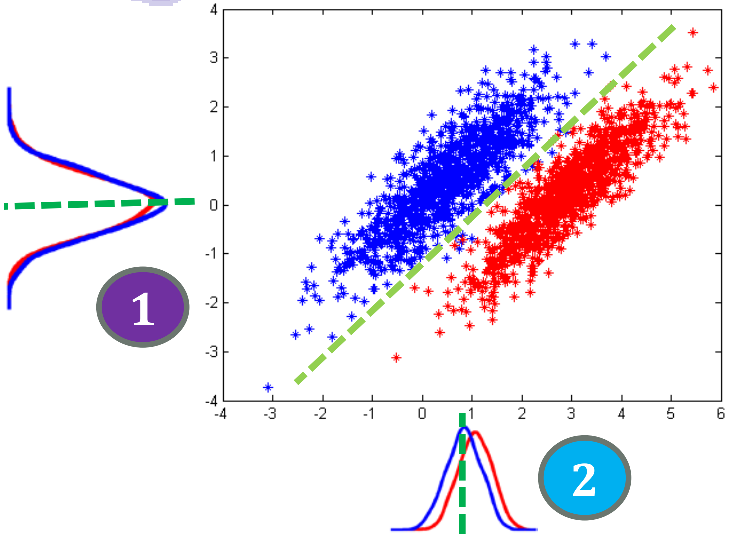


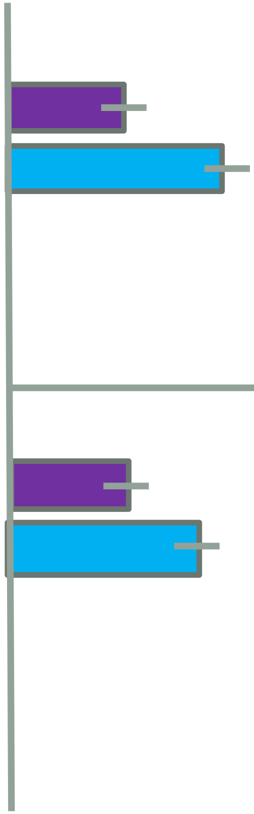
Univariate to Multivariate



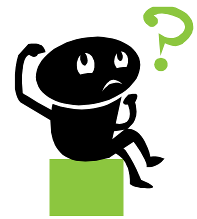
not significant

not significant





Motivation

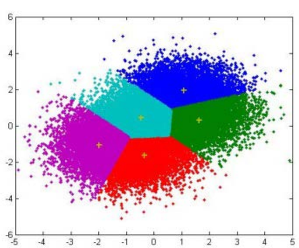


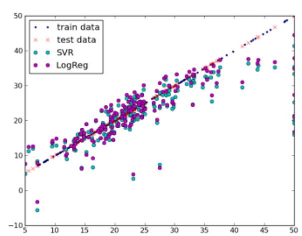
Modelling Terminology

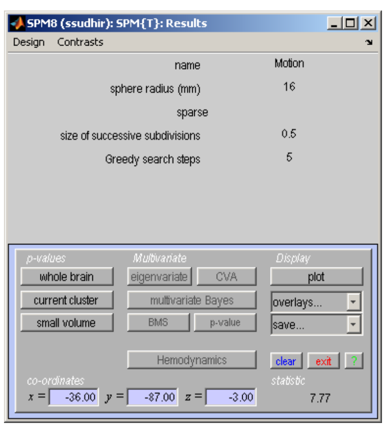
Learning from data

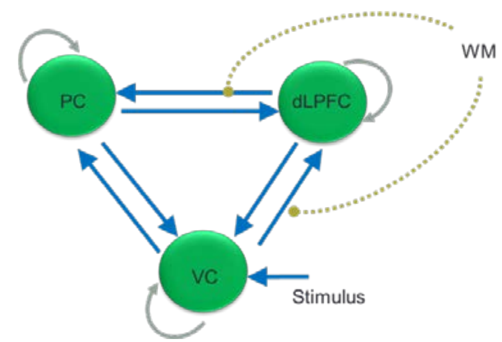
Multivariate Bayes in SPM

Generative Embedding

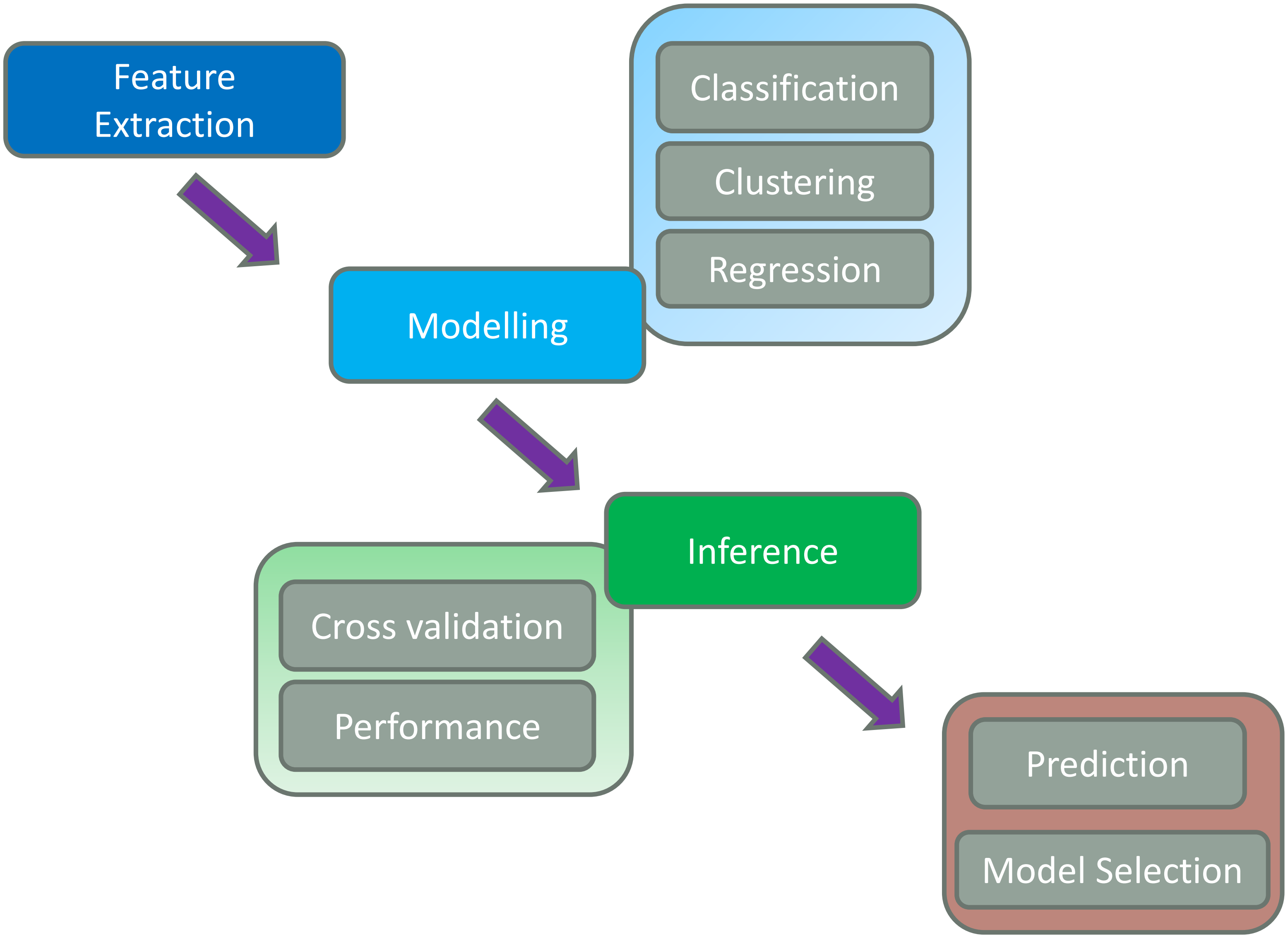








Steps for analysis



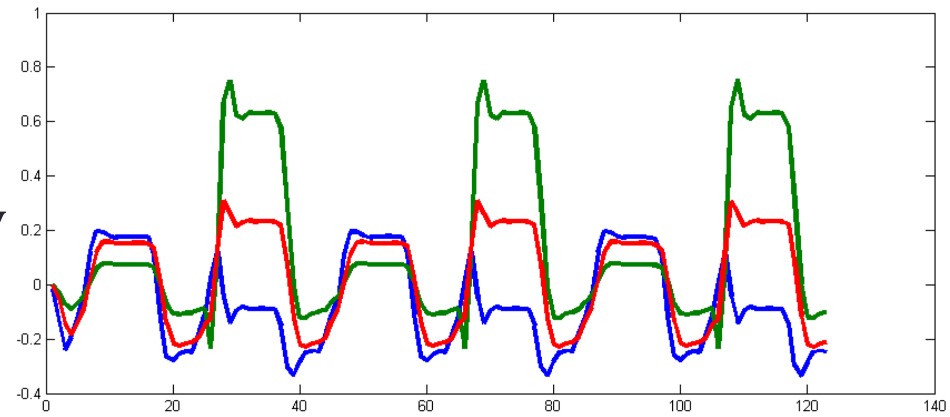
Feature Spaces

	F_1	F_2	. . .	F_P
S_1				
S_2	Data Point or Feature Vector			
.				
.				
S_N				

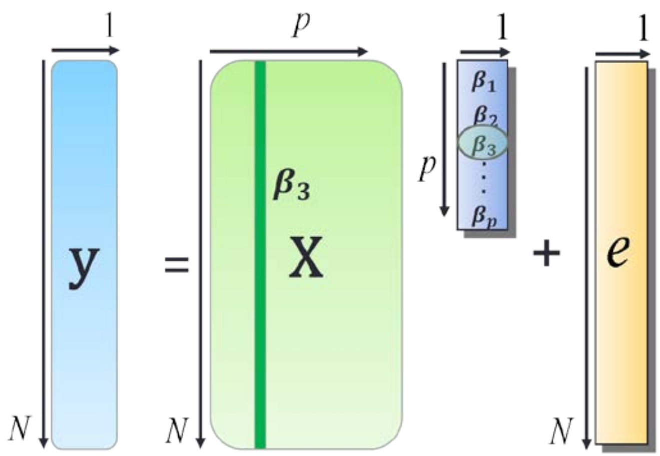
- High dimensionality
- Class/Cluster distributions
- Interpretation

Raw data

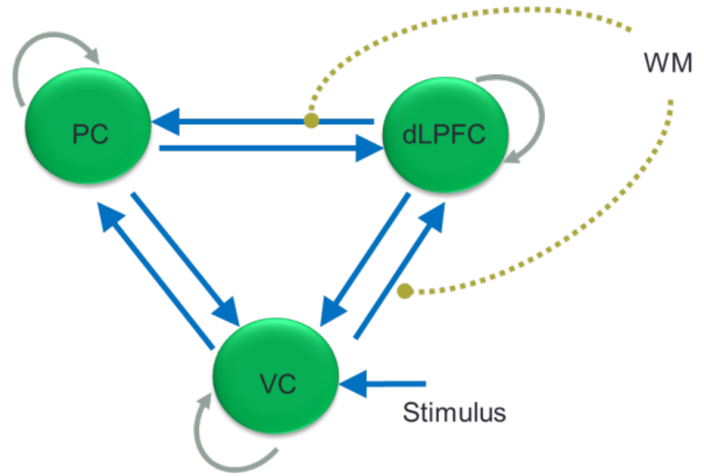
Voxel activity



Model parameters

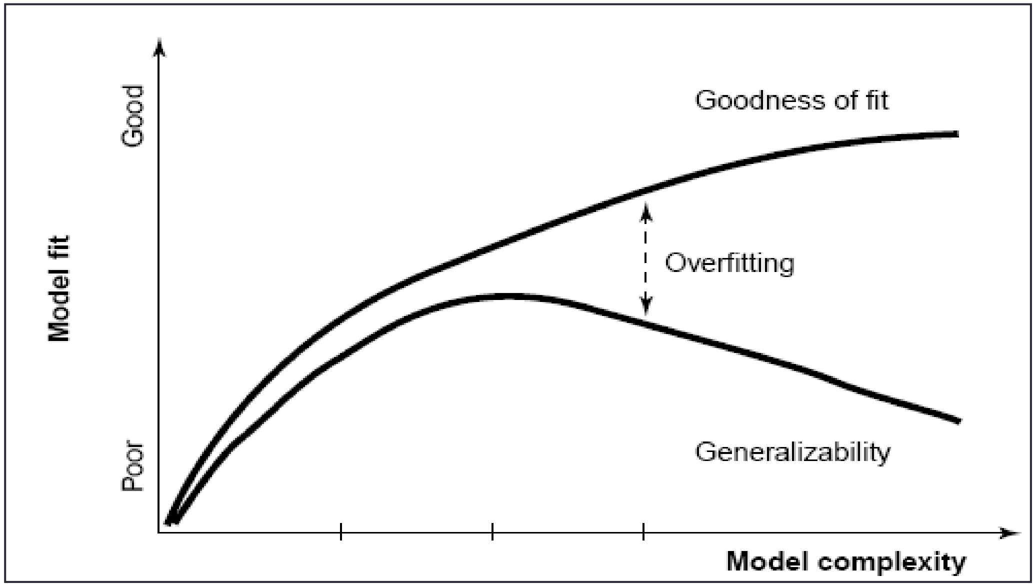


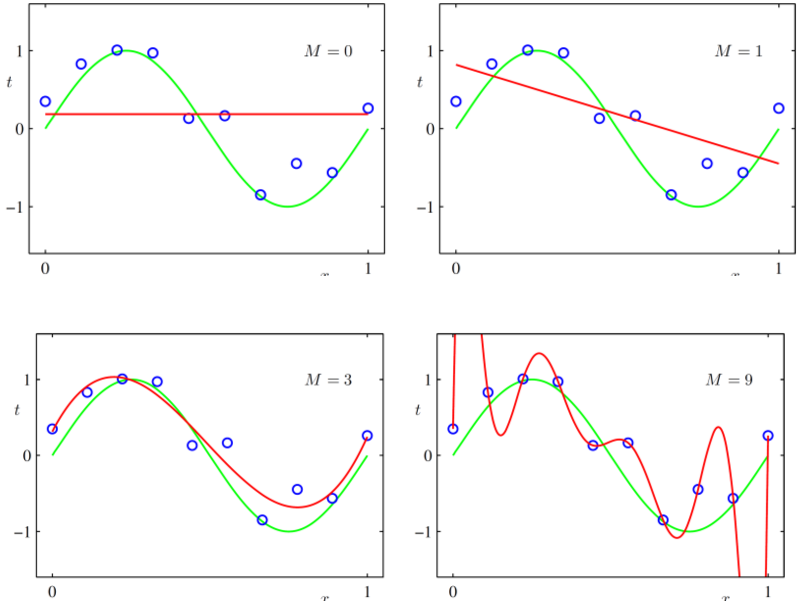
GLM – regression coefficients



DCM – connection parameters

Model Selection - Generalizability

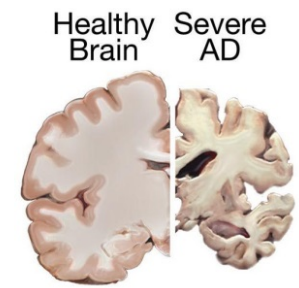




Bishop (2006), Pitt & Miyung (2002), *TICS*

Modelling goals

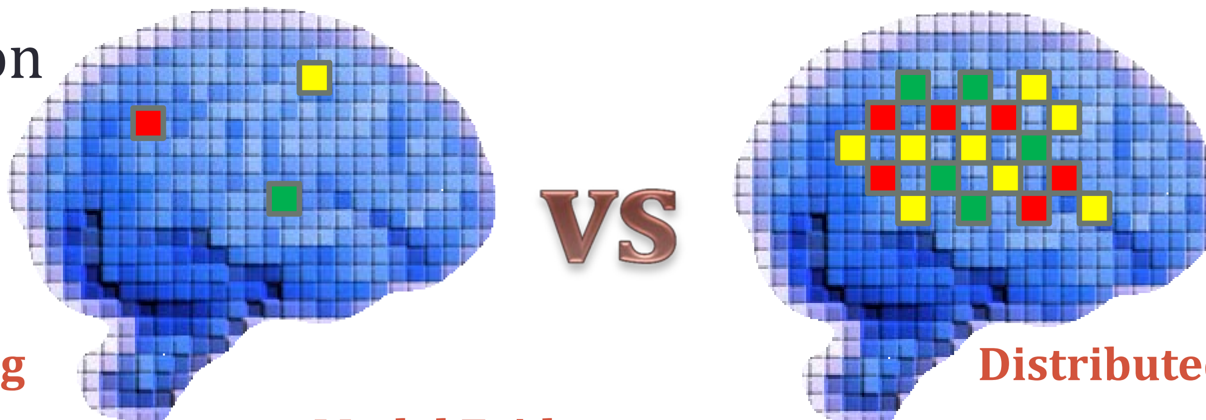
- Prediction



Predictive Density

$$p(X_{\text{new}}|Y_{\text{new}}, X, Y) = \int p(X_{\text{new}}|\theta, Y_{\text{new}})q(\theta)d\theta$$

- Model Selection



Sparse Coding

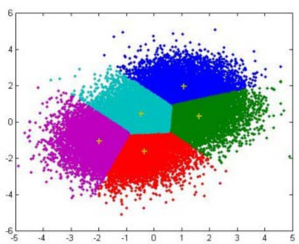
Distributed Coding

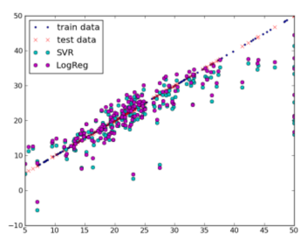
Model Evidence

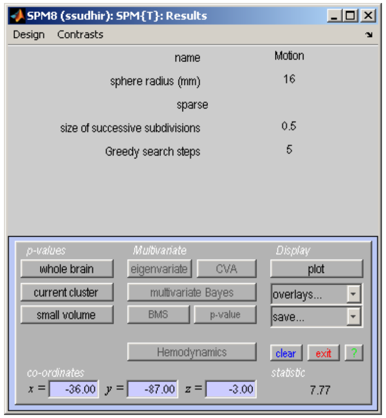
$$p(Y|X) = \int p(Y, \theta|X)d\theta$$

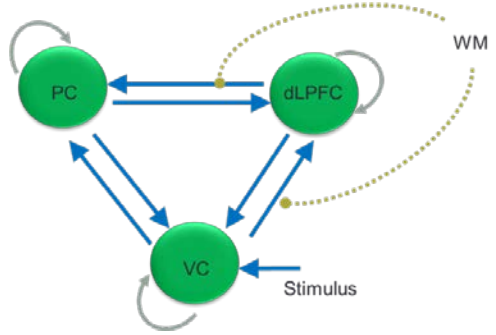
Motivation

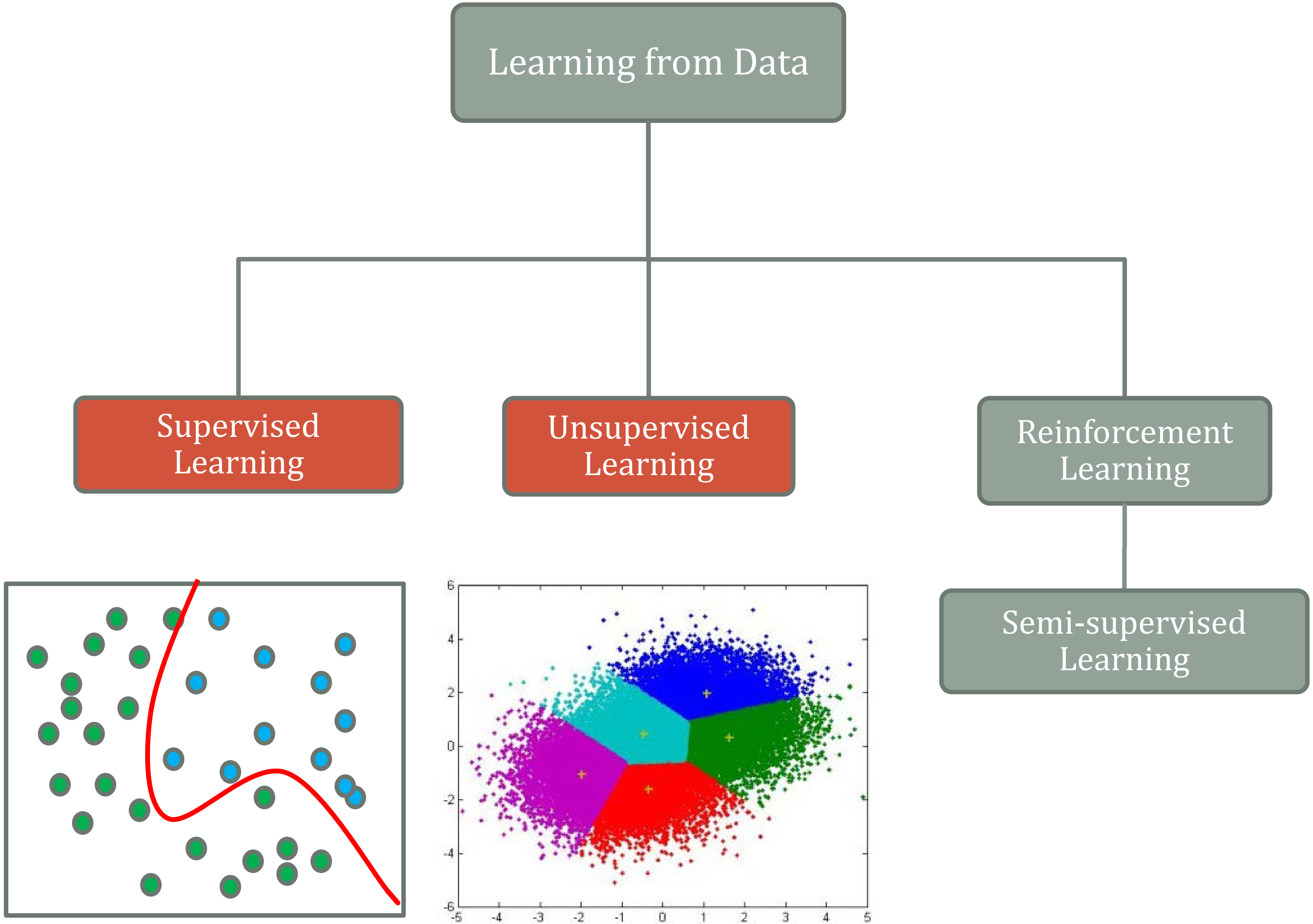
- Modelling Concepts
- Learning From Data
- Multivariate Bayes in SPM
- Generative Embedding










Learning from Data
Supervised Learning
Unsupervised Learning
Reinforcement Learning
Semi-supervised Learning
6
4
2
0
-2
-4
-6
-5
-4
-3
-2
-1
0
1
2
3
4
5

Supervised Learning

Independent variables

X

Function - f

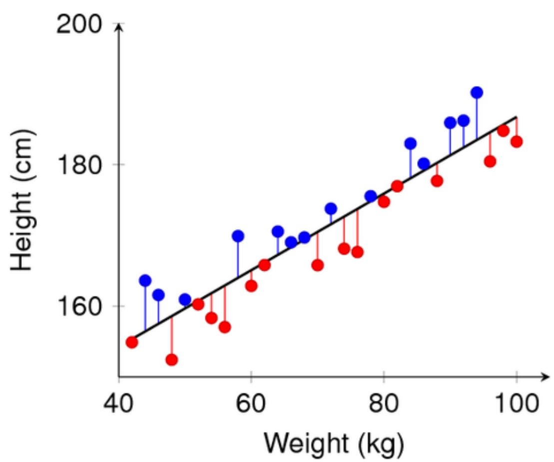
Regression

dependent variable

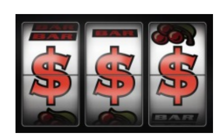
Y

Classification

Continuous

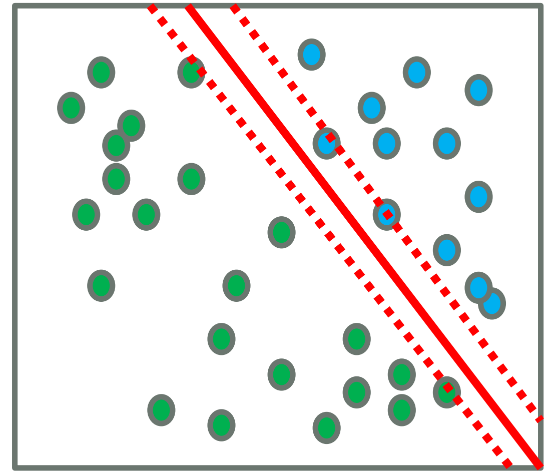


Categorical

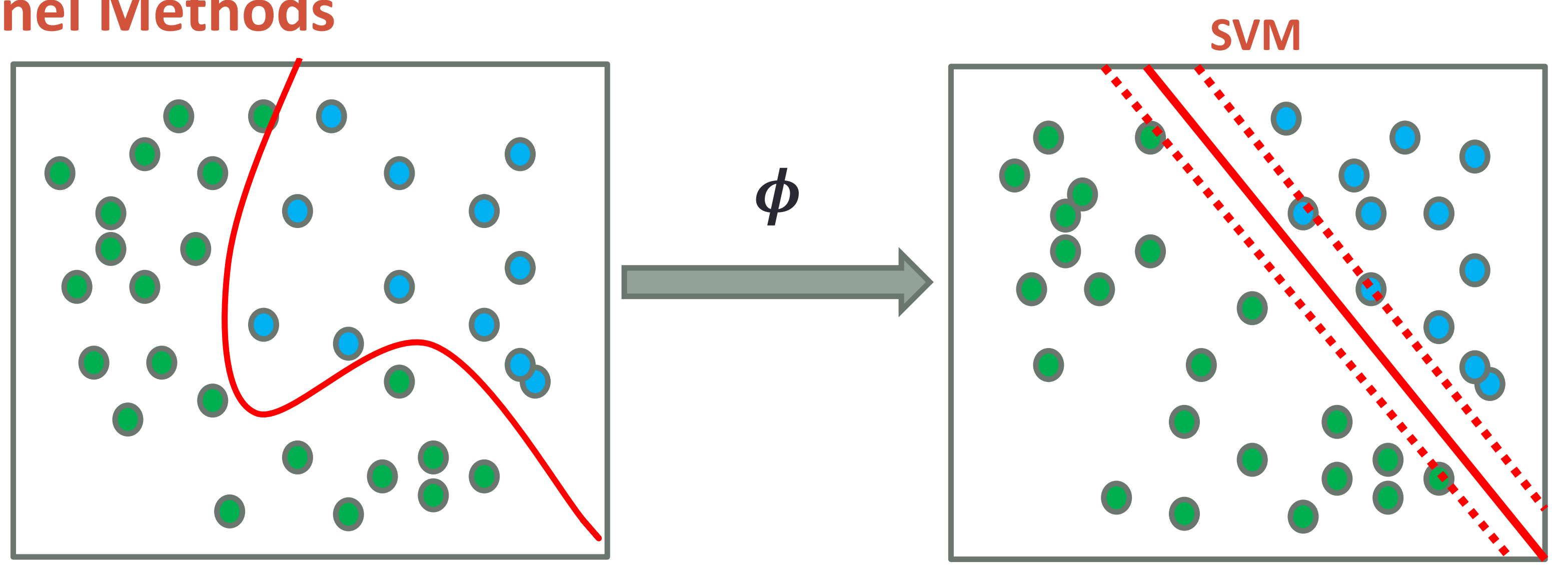


Classification

- Generative classifier
- Discriminative classifier



Kernel Methods

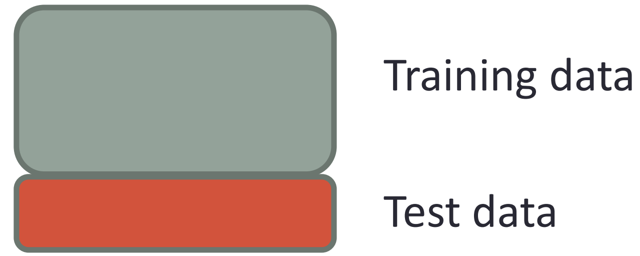


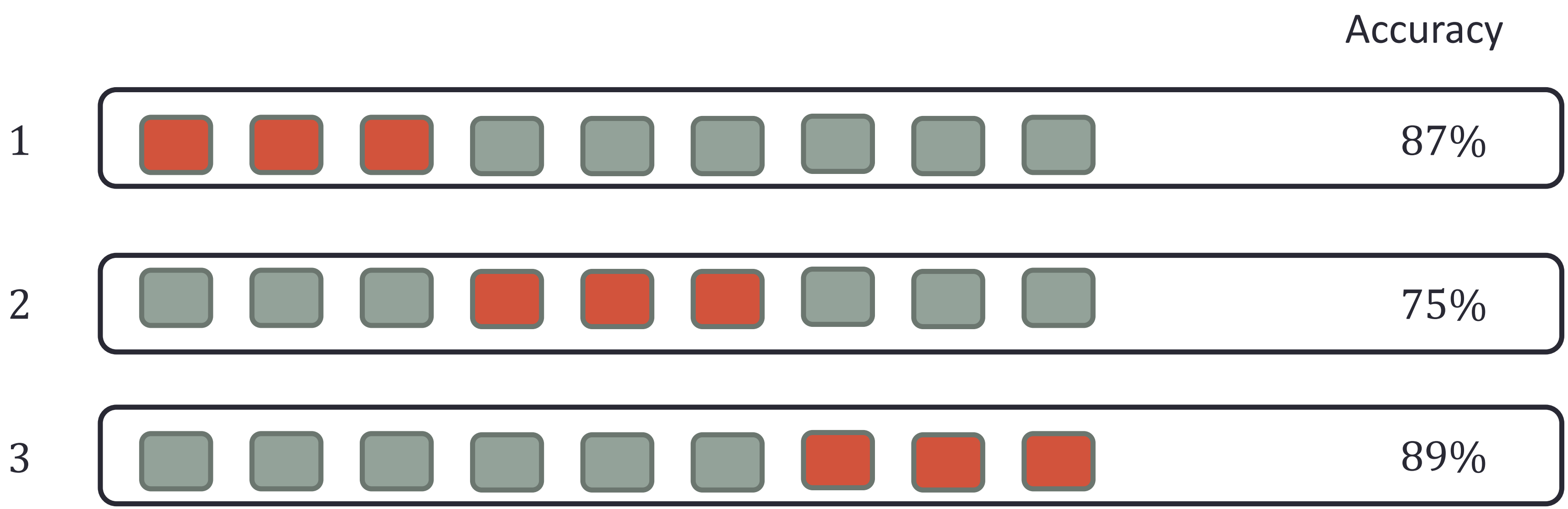
- Kernel Function – $K(x_i, x_j) = \phi(x_i).\phi(x_j)$

Kernel methods for pattern analysis, Taylor , Cristianini, 2004

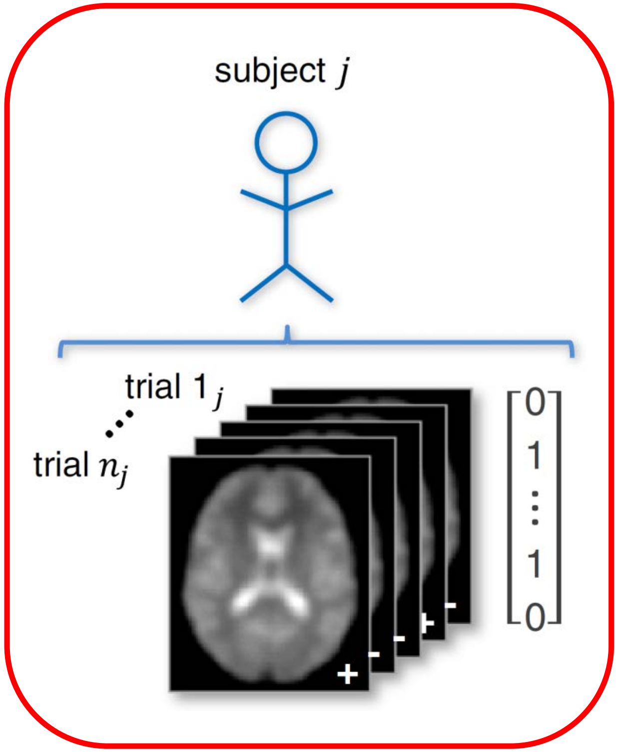
K-fold Cross Validation

- Model Selection
- Performance evaluation





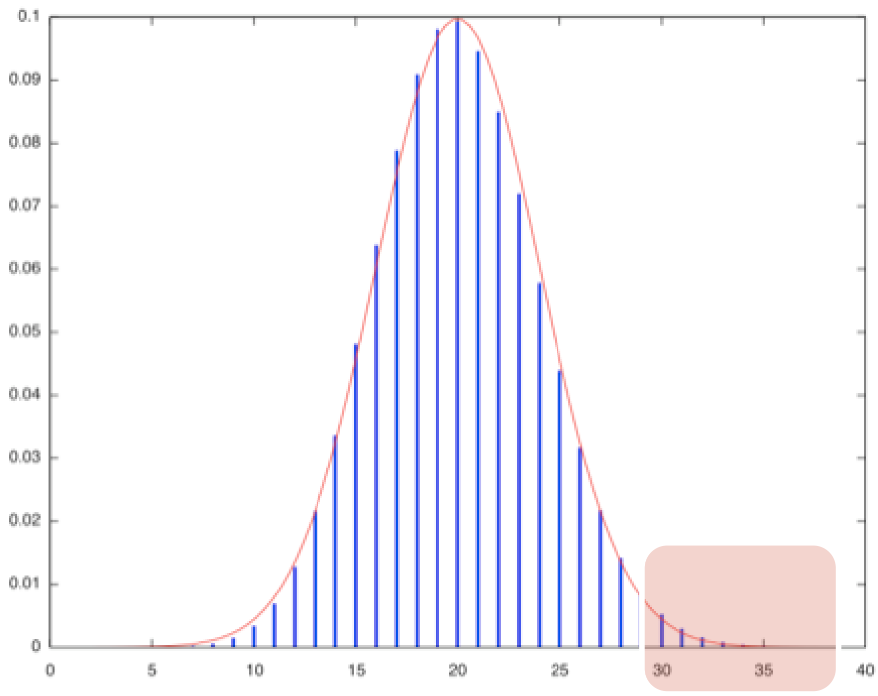
Performance – Single Subject



Brodersen et al. 2013, *NeuroImage*

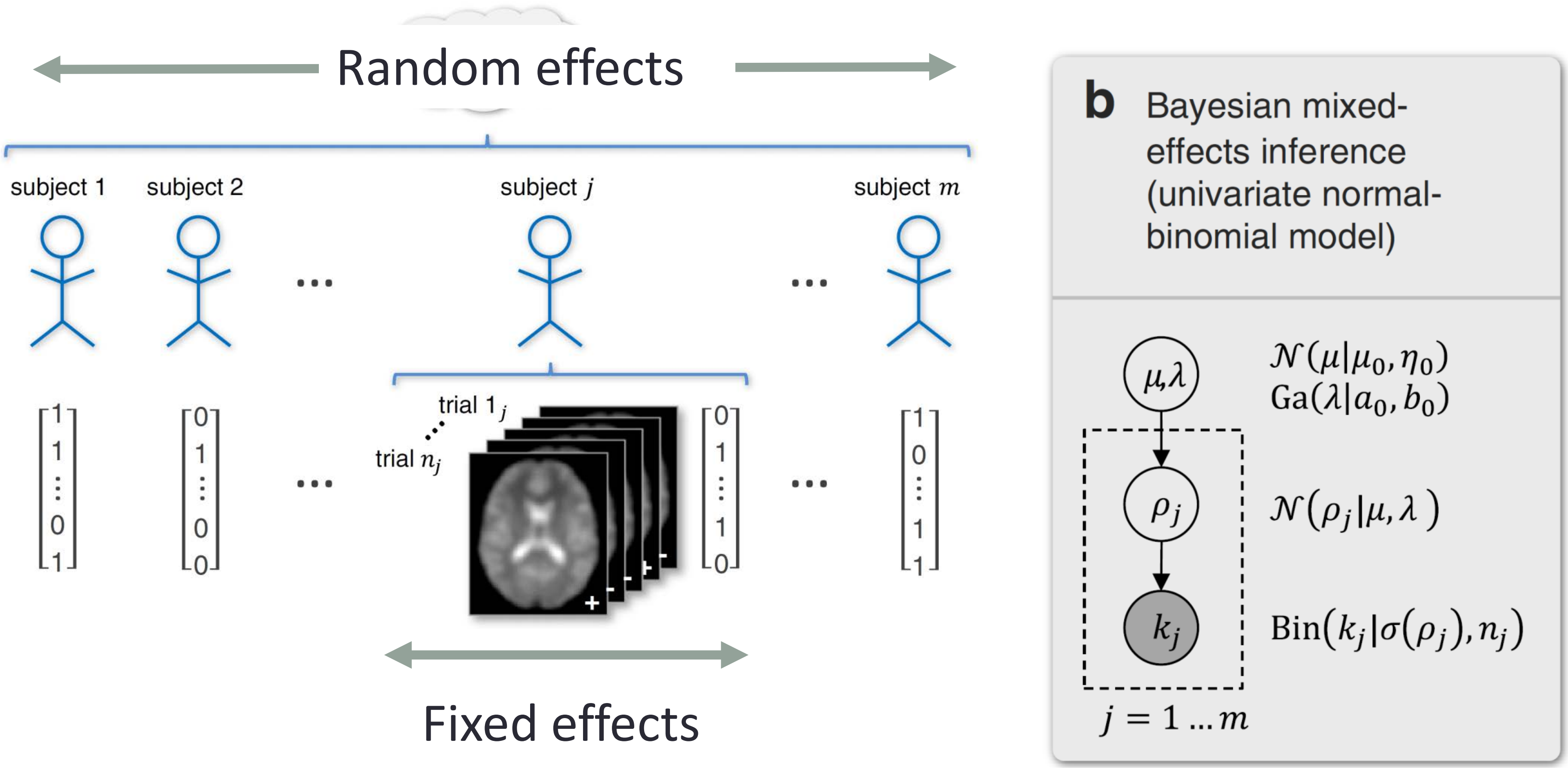
Binomial Test

$$p = P(X \geq k|H_0) = 1 - B(k|n, \pi_0)$$



k=30

Performance – Multiple Subjects

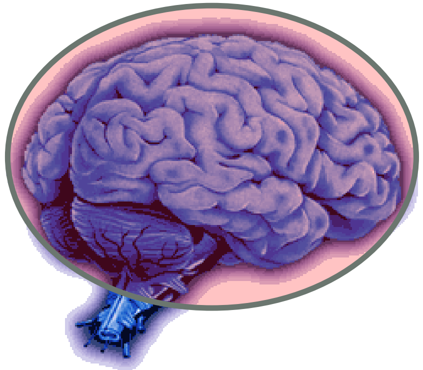


http://www.translationalneuromodeling.org/tapas/

Brodersen et al. 2013, *NeuroImage*

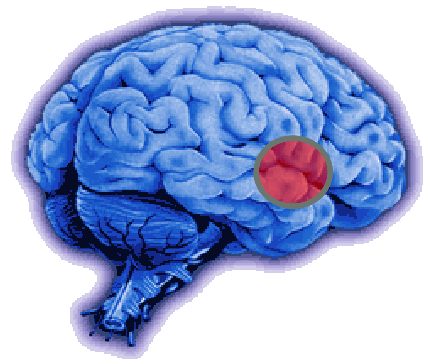
Using classification for fMRI data

Whole brain



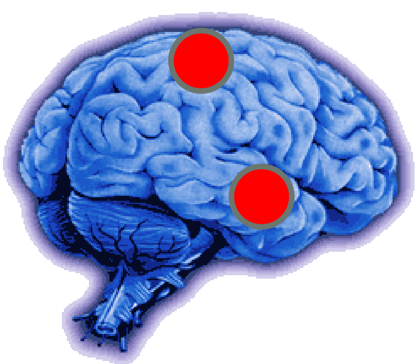
Mourao-Miranda et al. (2005) *NeuroImage*

Searchlight classifier



Nandy & Cordes (2003) *MRM* , Kriegeskorte et al. (2006) *PNAS*

Pattern localization

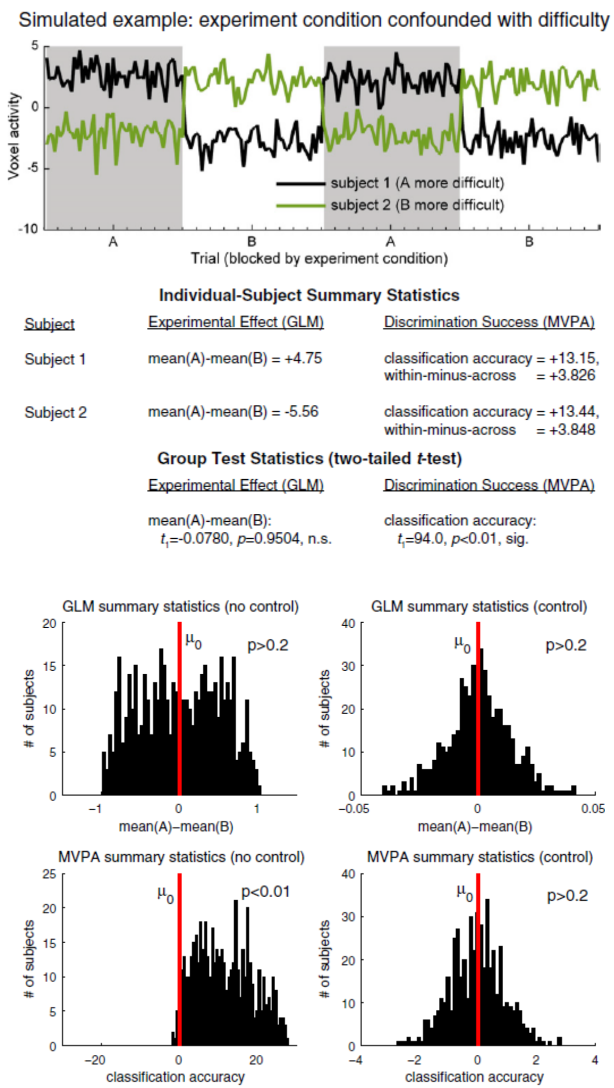


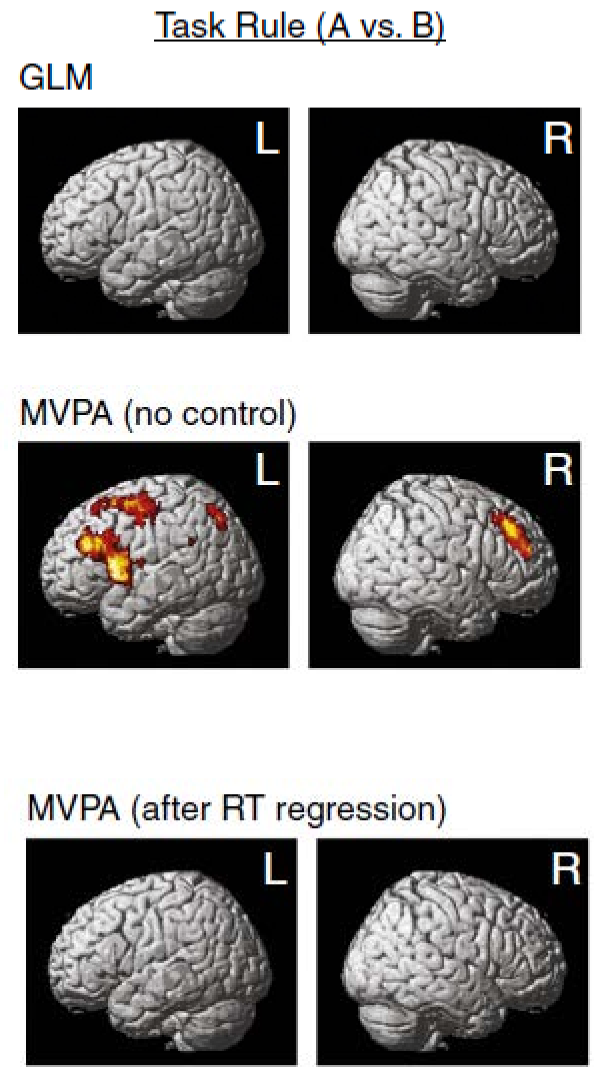
Word Category

Food
Building
People

Pereira et al. (2009) *NeuroImage*, Mitchell et al. (2004) *Machine Learning*

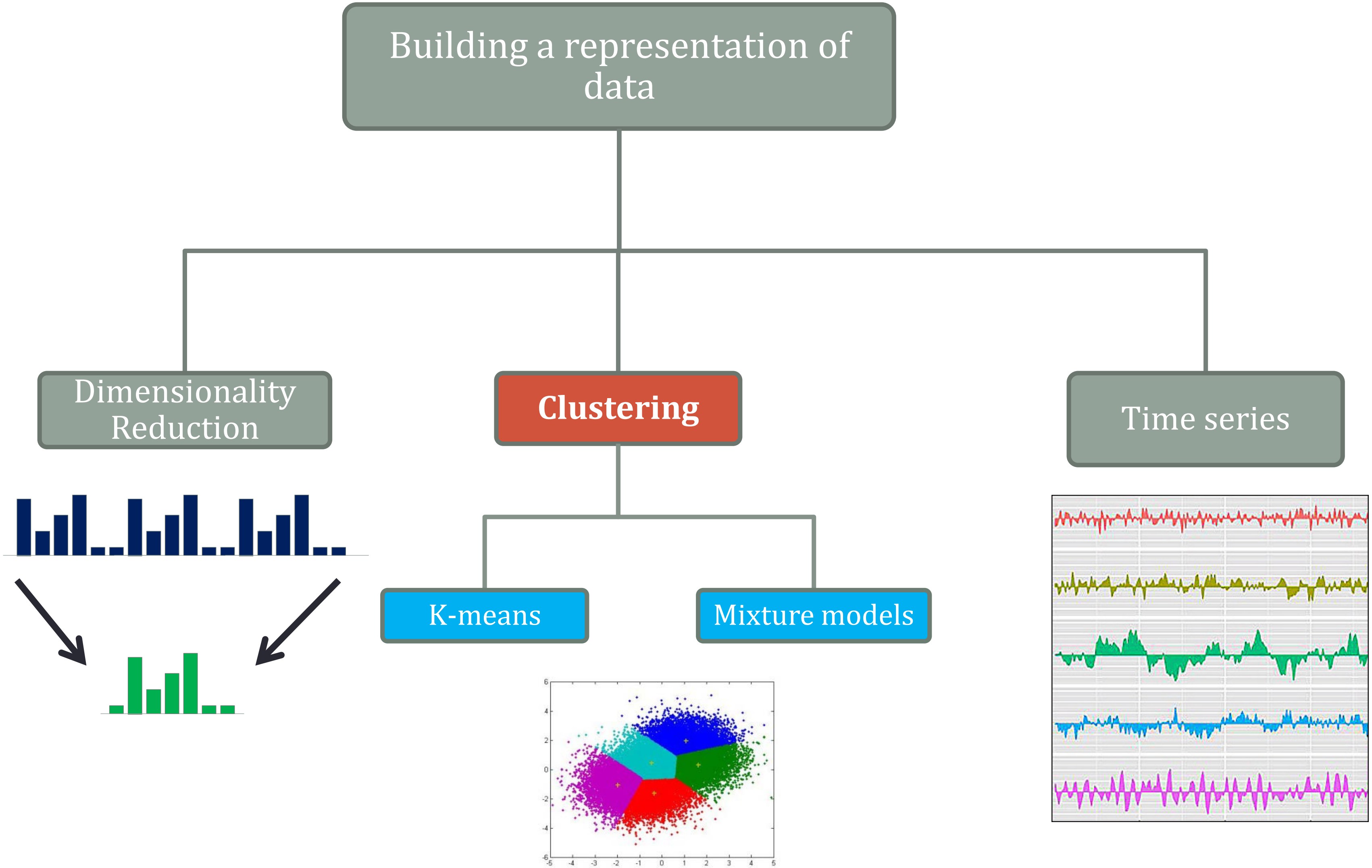
Confounds – GLM vs MVPA





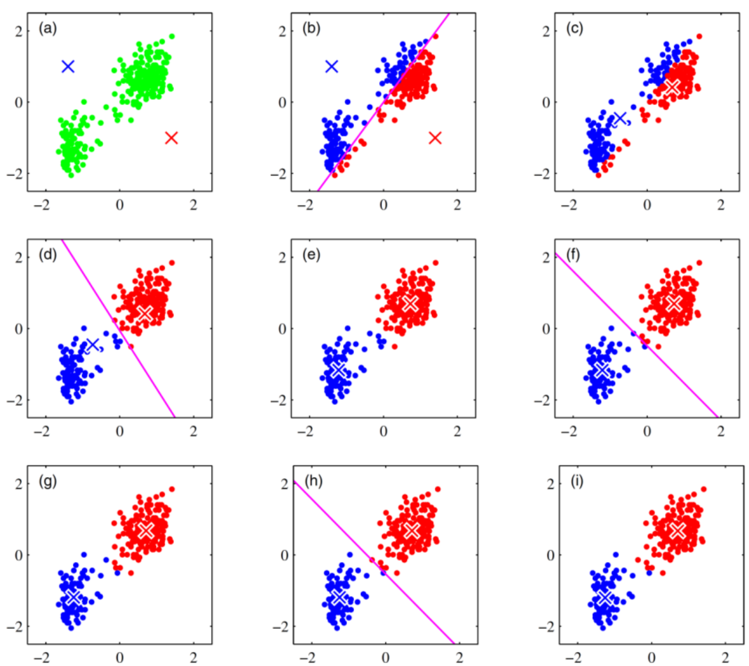
Todd et al. 2013, *NeuroImage*

Unsupervised Learning

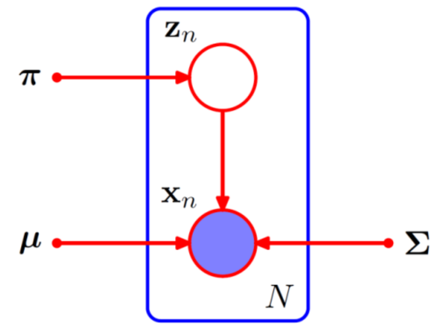


K-means

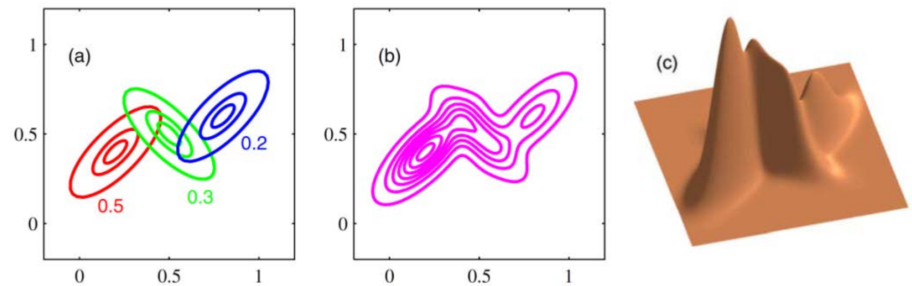
$$\widetilde{J} = \sum_{n=1}^{N} \sum_{k=1}^{K} r_{nk} \mathcal{V}(\mathbf{x}_n, \boldsymbol{\mu}_k)$$



Mixture of Gaussians



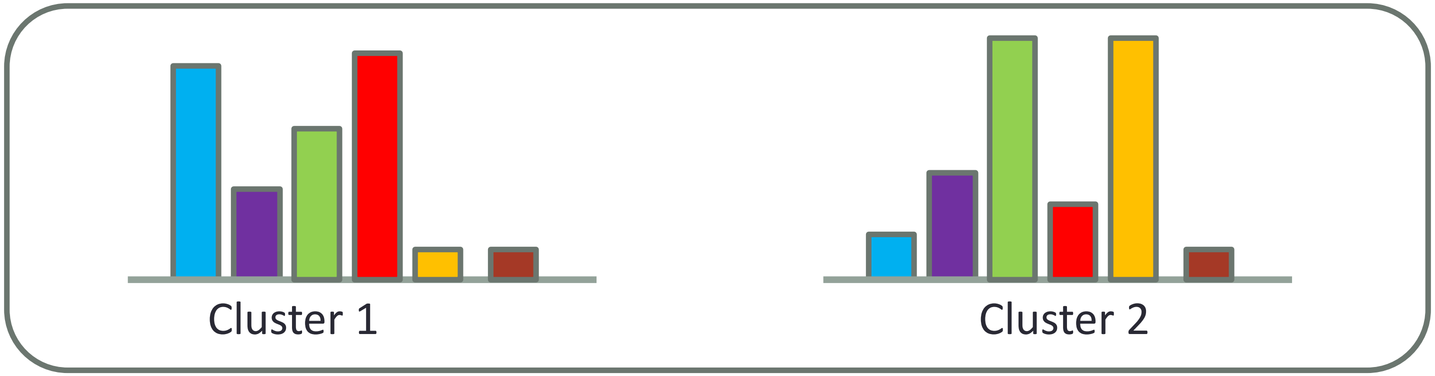
$$p(\mathbf{x}) = \sum_{k=1}^{K} \pi_k \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$$



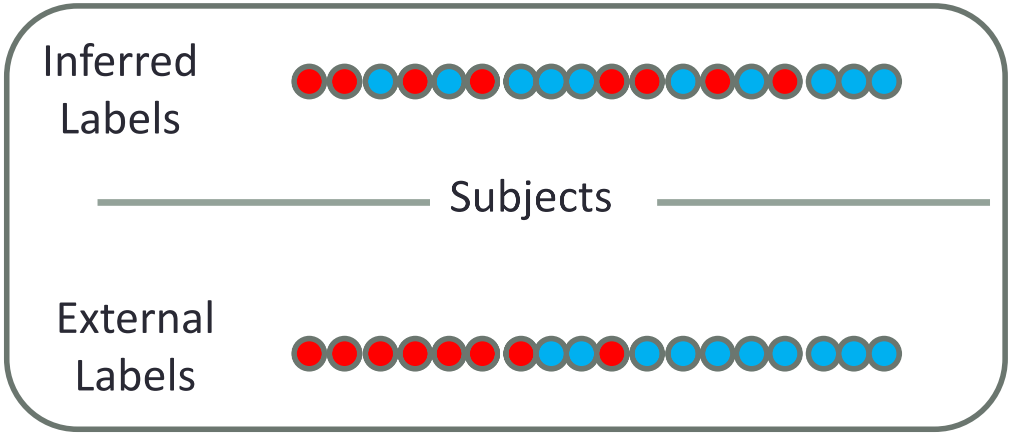
Bishop (2006)

Interpretation

- Cluster parameters



- Purity



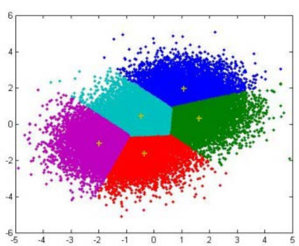
Normalized Mutual Information (NMI)

Balanced purity

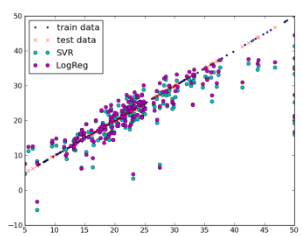
Rand Index

Motivation

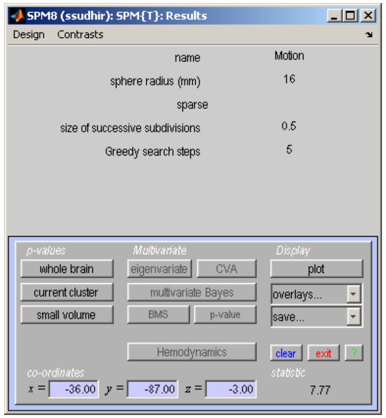
Modelling



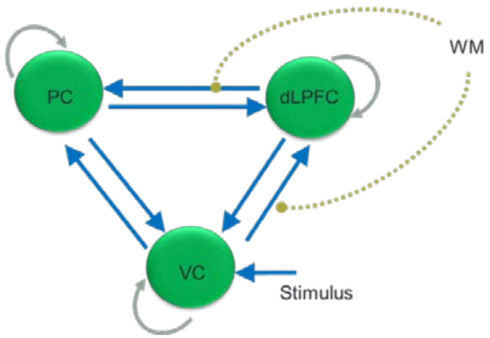
Learning from Data



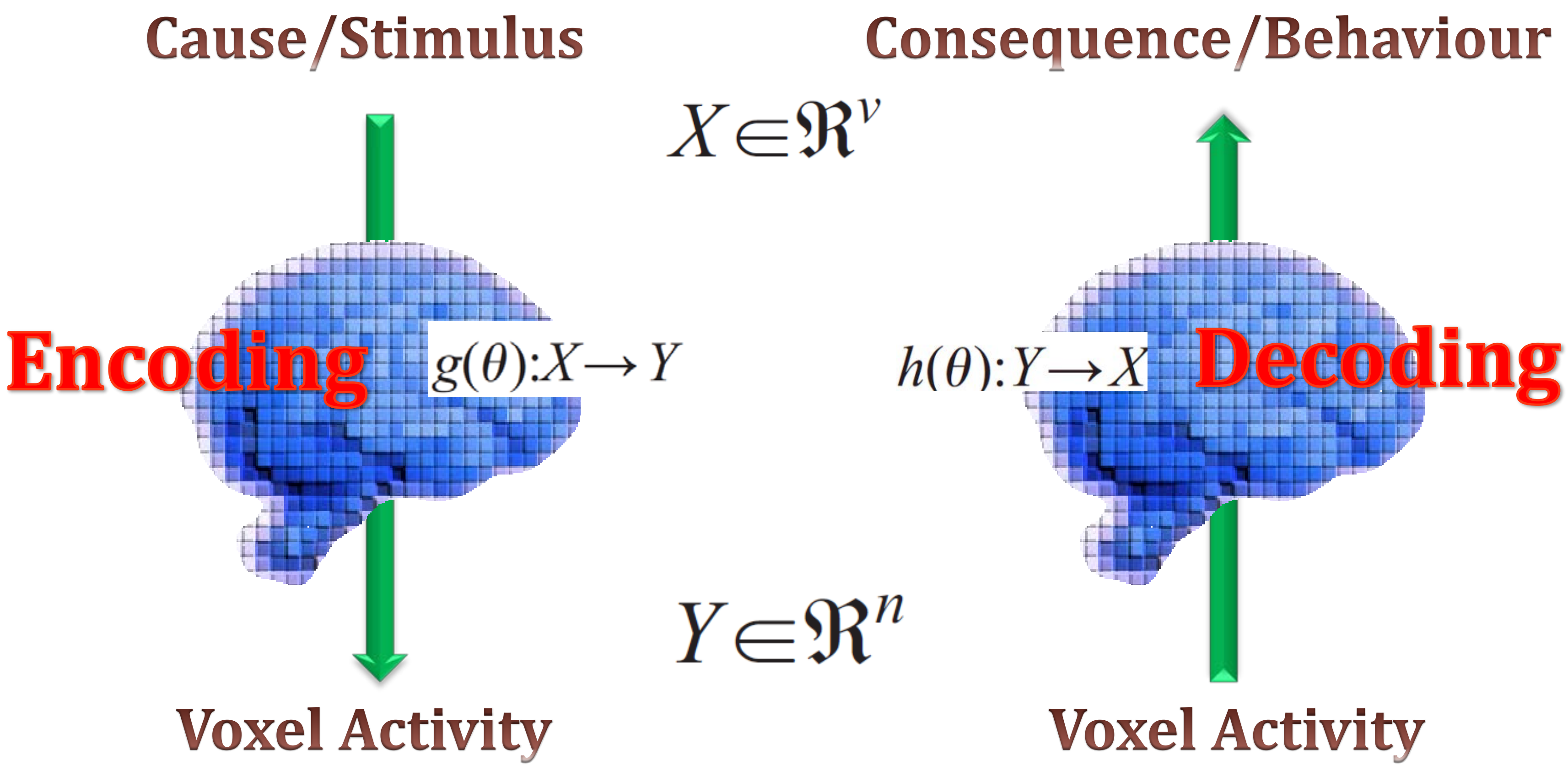
Multivariate Bayes in SPM



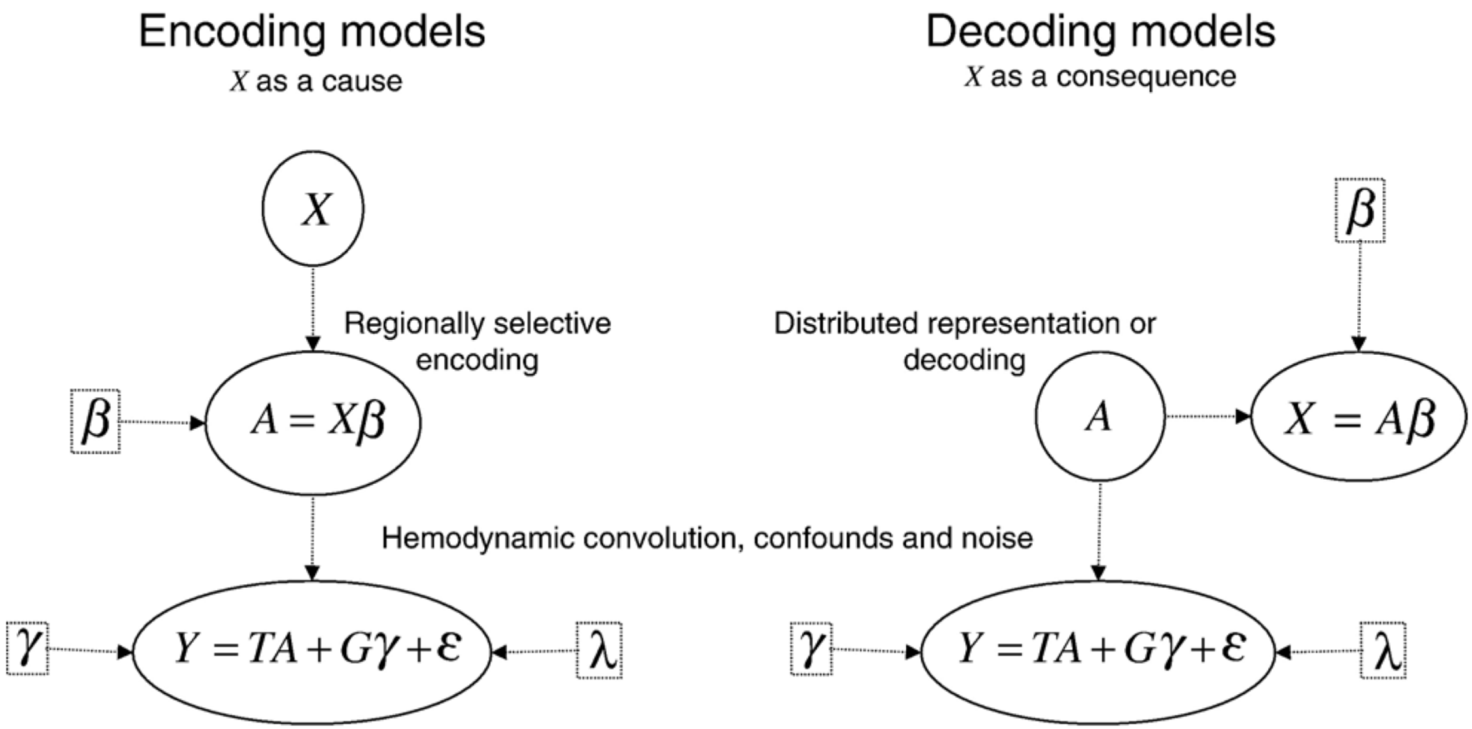
Generative Embedding



Encoding Vs Decoding Models



Encoding Vs Decoding



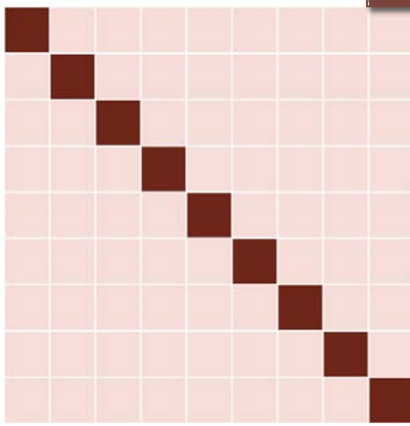
$$g(\theta) : X \rightarrow Y$$
$$Y = TX\beta + G\gamma + \varepsilon$$

$$g(\theta) : Y \rightarrow X$$
$$TX = Y\beta - G\gamma\beta - \varepsilon\beta$$

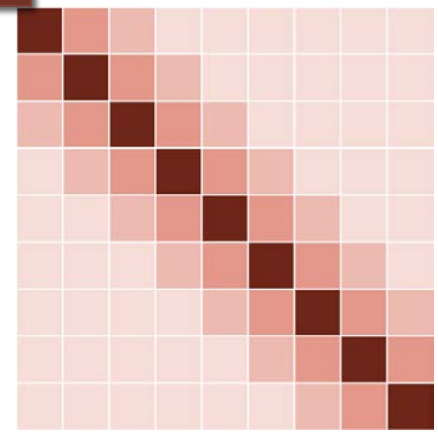
Friston et al. 2008 NeuroImage

Coding hypotheses

Sparse vectors



Spatial vectors

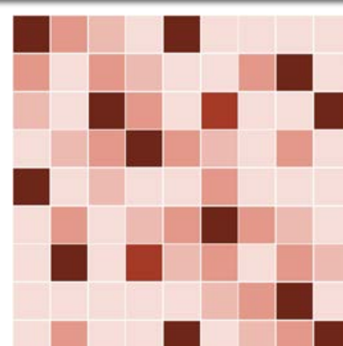


Smooth vectors

Distributed vectors

Singular vectors
of data

$$UDV^T = RY^T$$



Support vectors

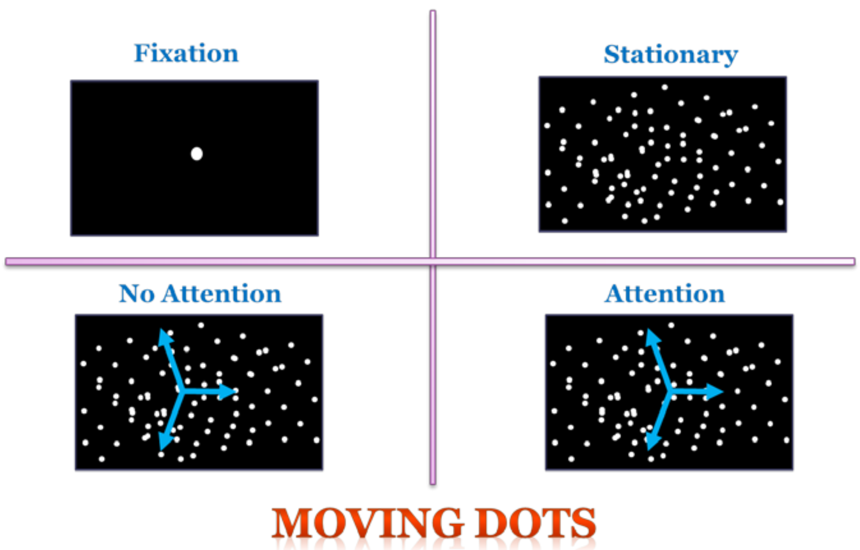
$$U = RY^T$$

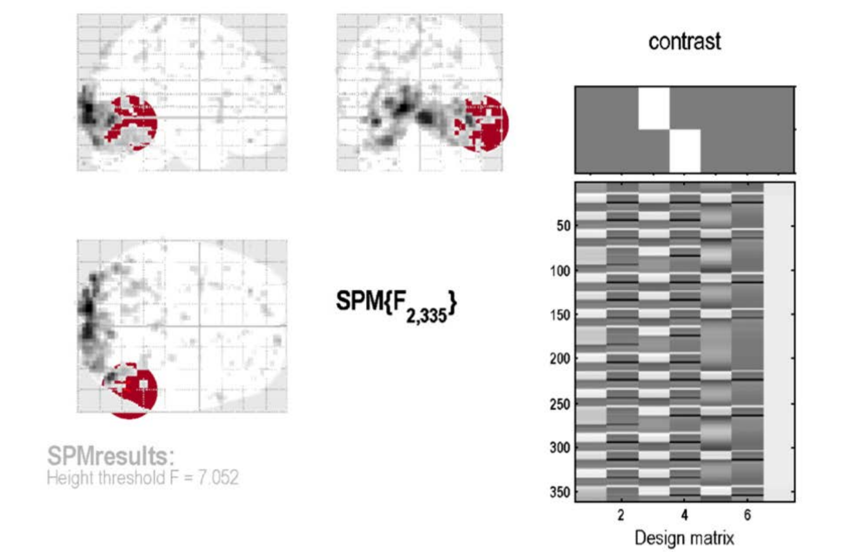
Bayesian decoding of motion

Attention to motion dataset - Büchel & Friston 1999 Cerebral Cortex

Experimental factors:

1. Photic
2. Motion
3. Attention





Statistics: ***search volume: 16.0mm sphere at [48,-63,0]***

set-level		cluster-level			voxel-level					mm mm mm		
p	c	$p_{corrected}$	k_E	$p_{uncorrected}$	$p_{FWE\text{-}corr}$	$p_{FDR\text{-}corr}$	F	$(Z_{\equiv})$	$p_{uncorrected}$			
0.006	4		39		0.000	0.000	40.66	Inf	0.000	39	-72	0
					0.000	0.000	40.09	Inf	0.000	42	-75	-3
			3		0.002	0.000	13.35	4.55	0.000	39	-63	-12
			1		0.124	0.003	8.67	3.52	0.000	51	-66	-3
			1		0.436	0.010	7.36	3.18	0.001	48	-63	0

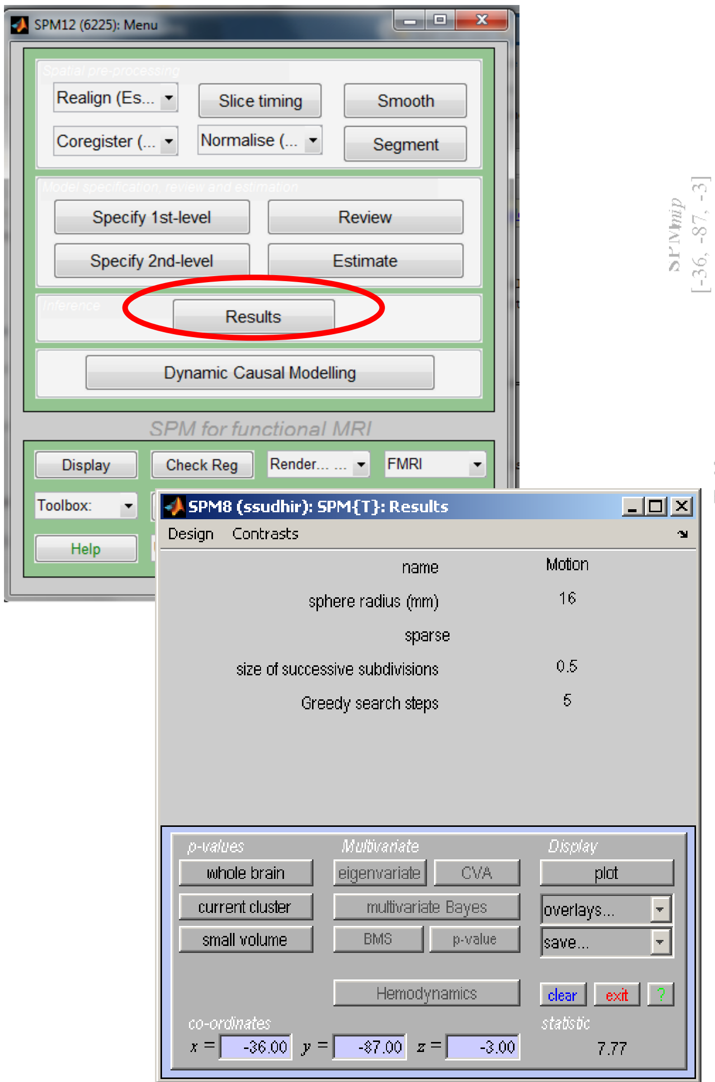
table shows 16 local maxima more than 4.0mm apart

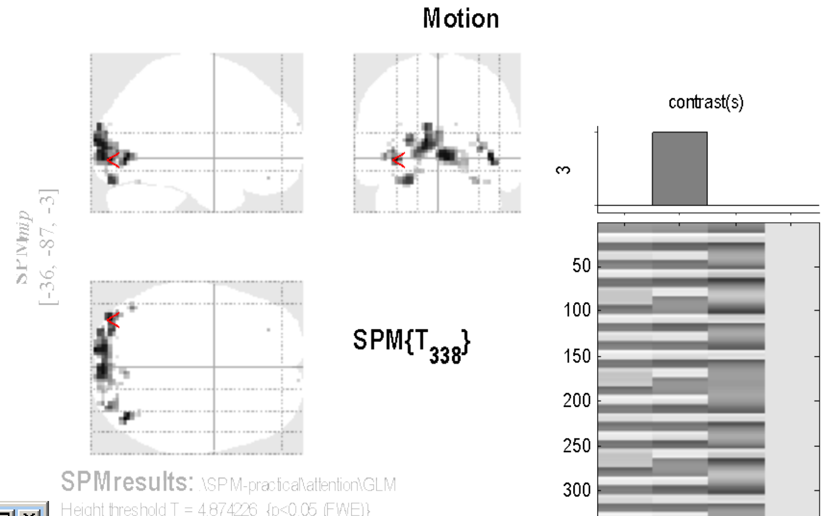
Height threshold: F = 7.05, p = 0.001 (0.513) {p<0.001 (unc.)}
Extent threshold: k = 0 voxels, p = 1.000 (0.513)
Expected voxels per cluster, <k> = 1.212
Expected number of clusters, <c> = 0.72
Expected false discovery rate, <= 0.01

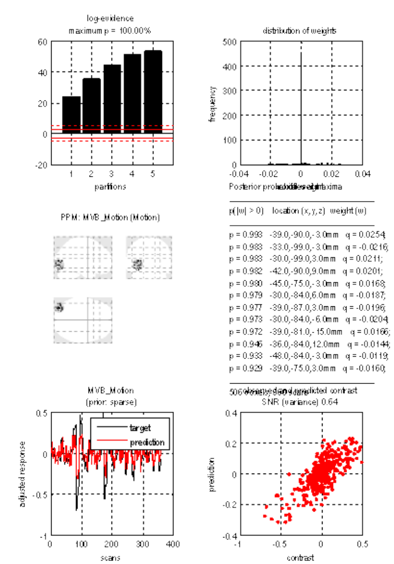
Degrees of freedom = [2.0, 335.0]
FWHM = 7.5 7.4 7.1 mm mm mm; 2.5 2.5 2.4 {voxels};
Volume: 15741; 583 voxels; 43.3 resels
Voxel size: 3.0 3.0 3.0 mm mm mm; (resel = 14.69 voxels)

Friston et al. 2008 NeuroImage

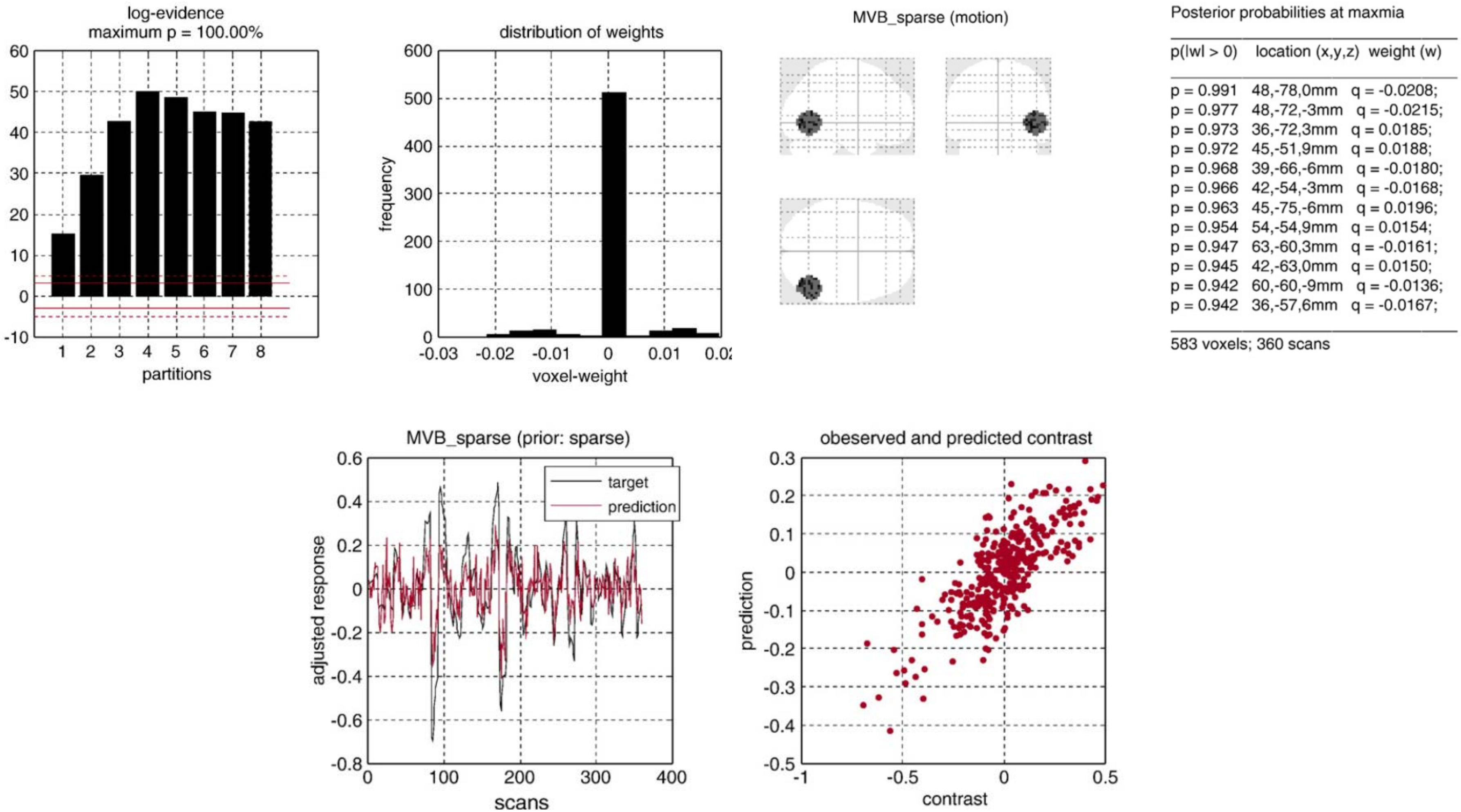
Multivariate Bayes in SPM





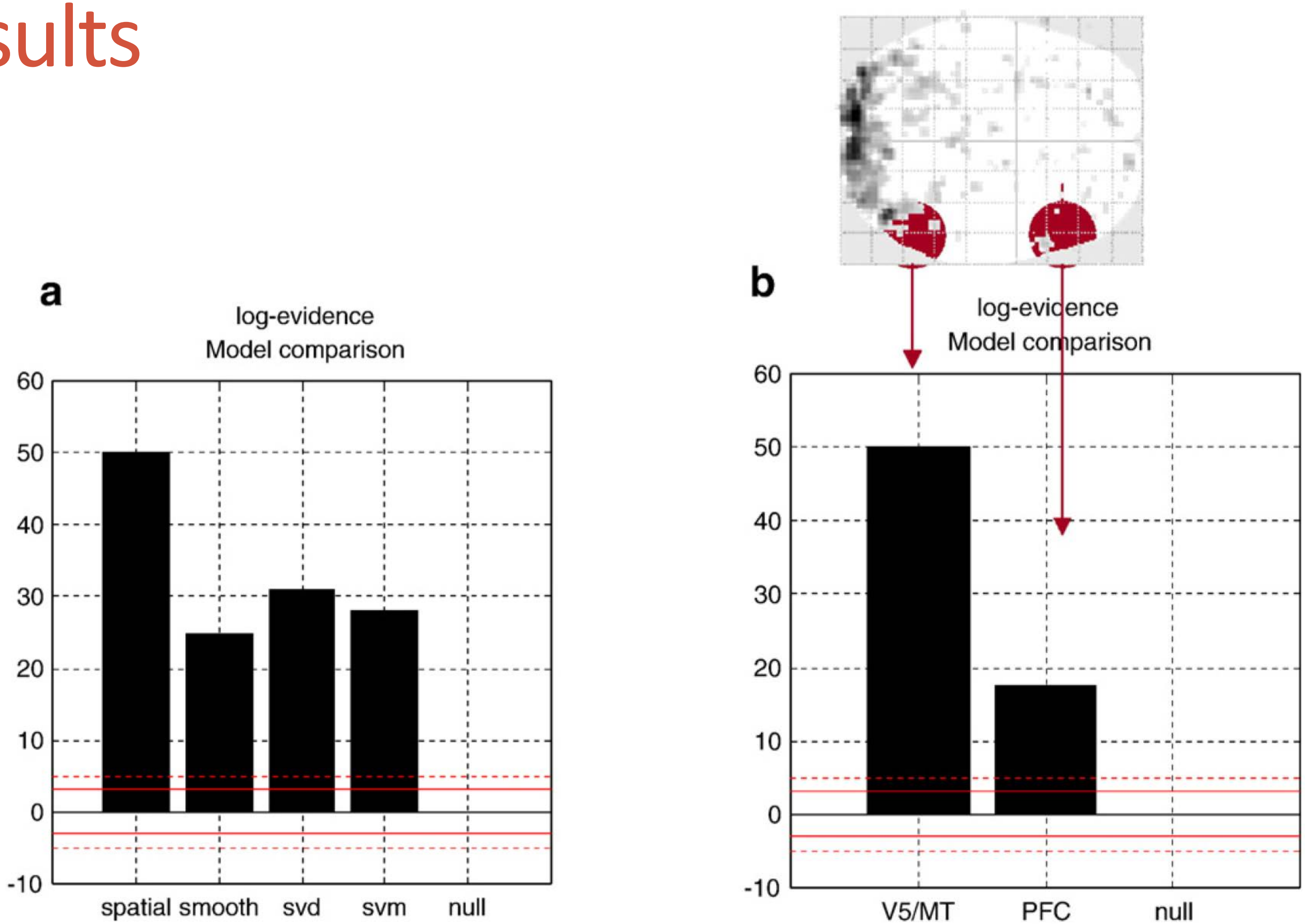


Results



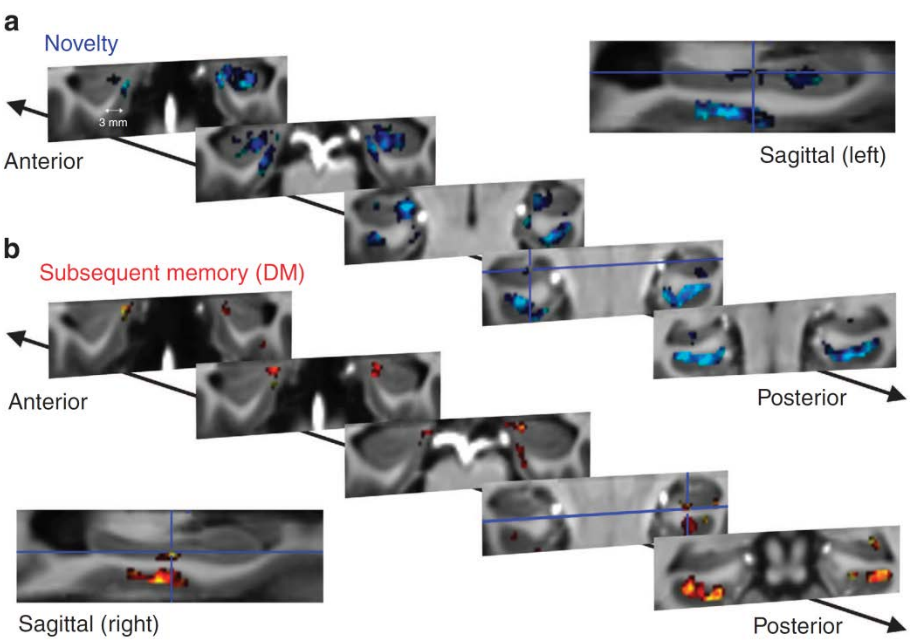
Friston et al. 2008 NeuroImage

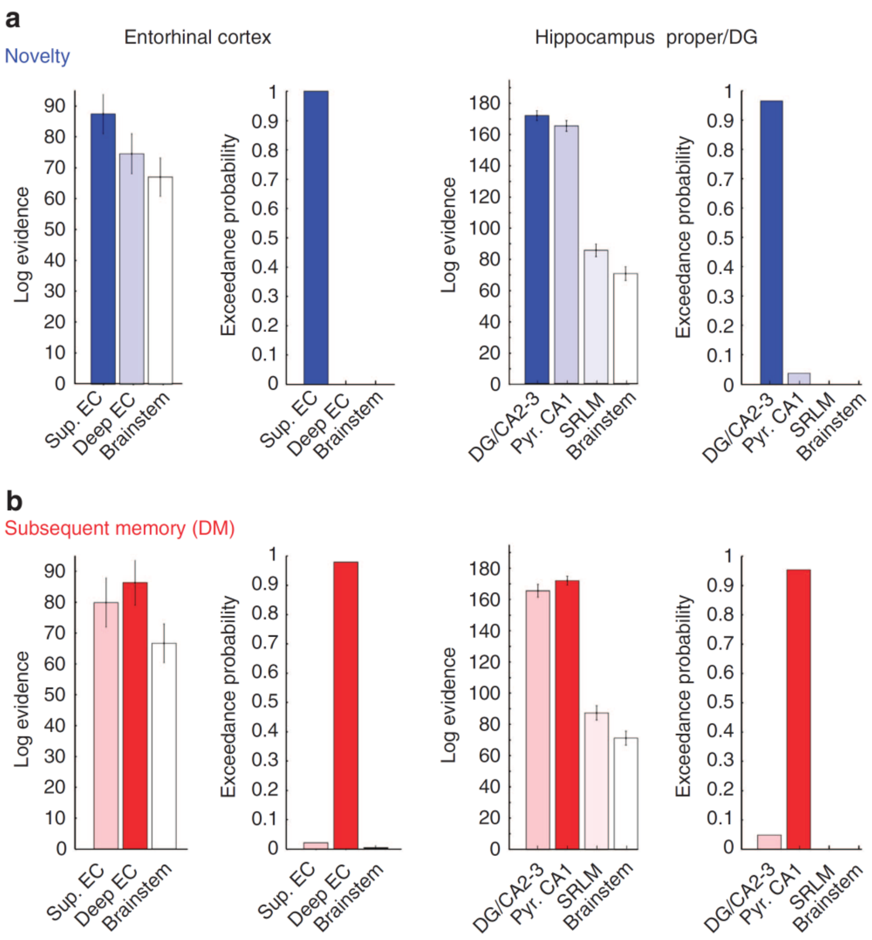
Results



a
log-evidence
Model comparison
60
50
40
30
20
10
0
-10
spatial smooth
svd
svm
null
b
log-evidence
Model comparison
60
50
40
30
20
10
0
-10
V5/MT
PFC
null

Laminar activity related to novelty and episodic encoding

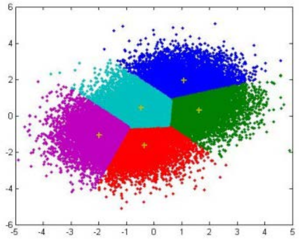


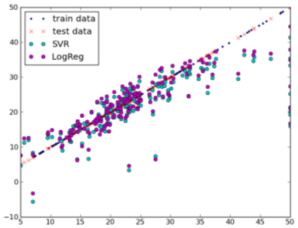


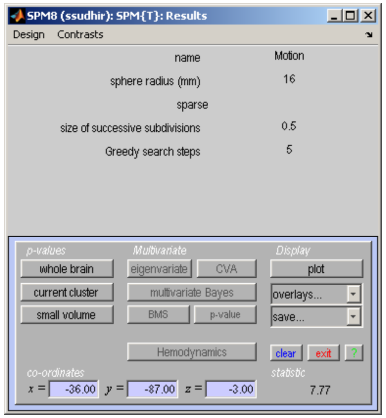
Maas et al. 2014 Nature Communications

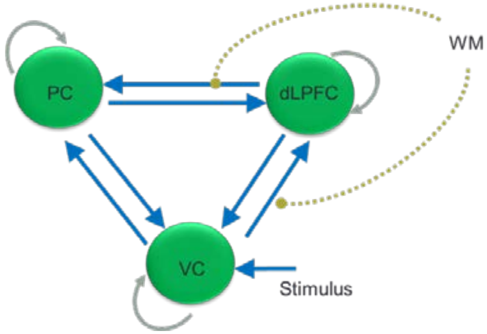
Motivation

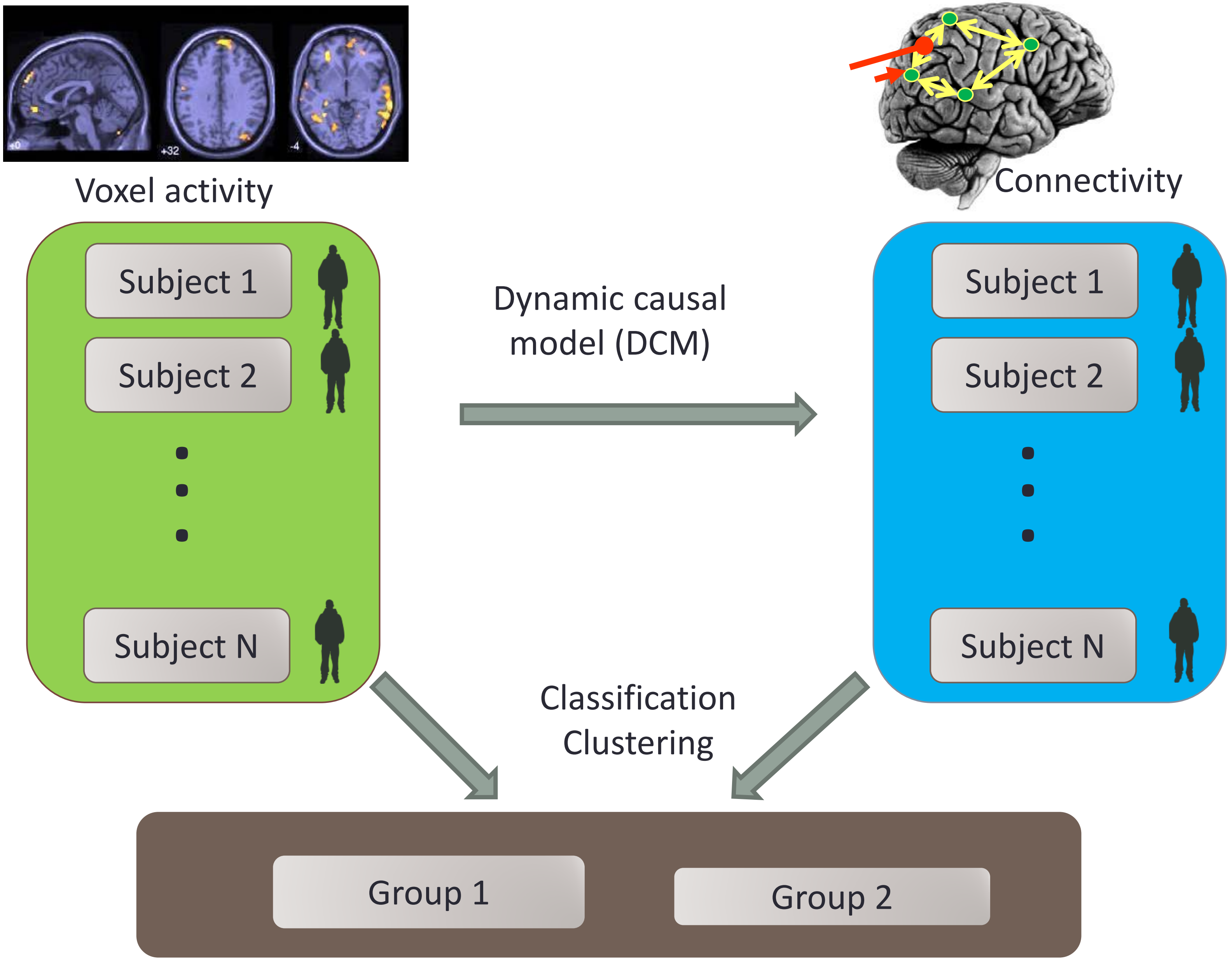
- Modelling Principles
- Learning from Data
- Multivariate Bayes in SPM
- Generative Embedding





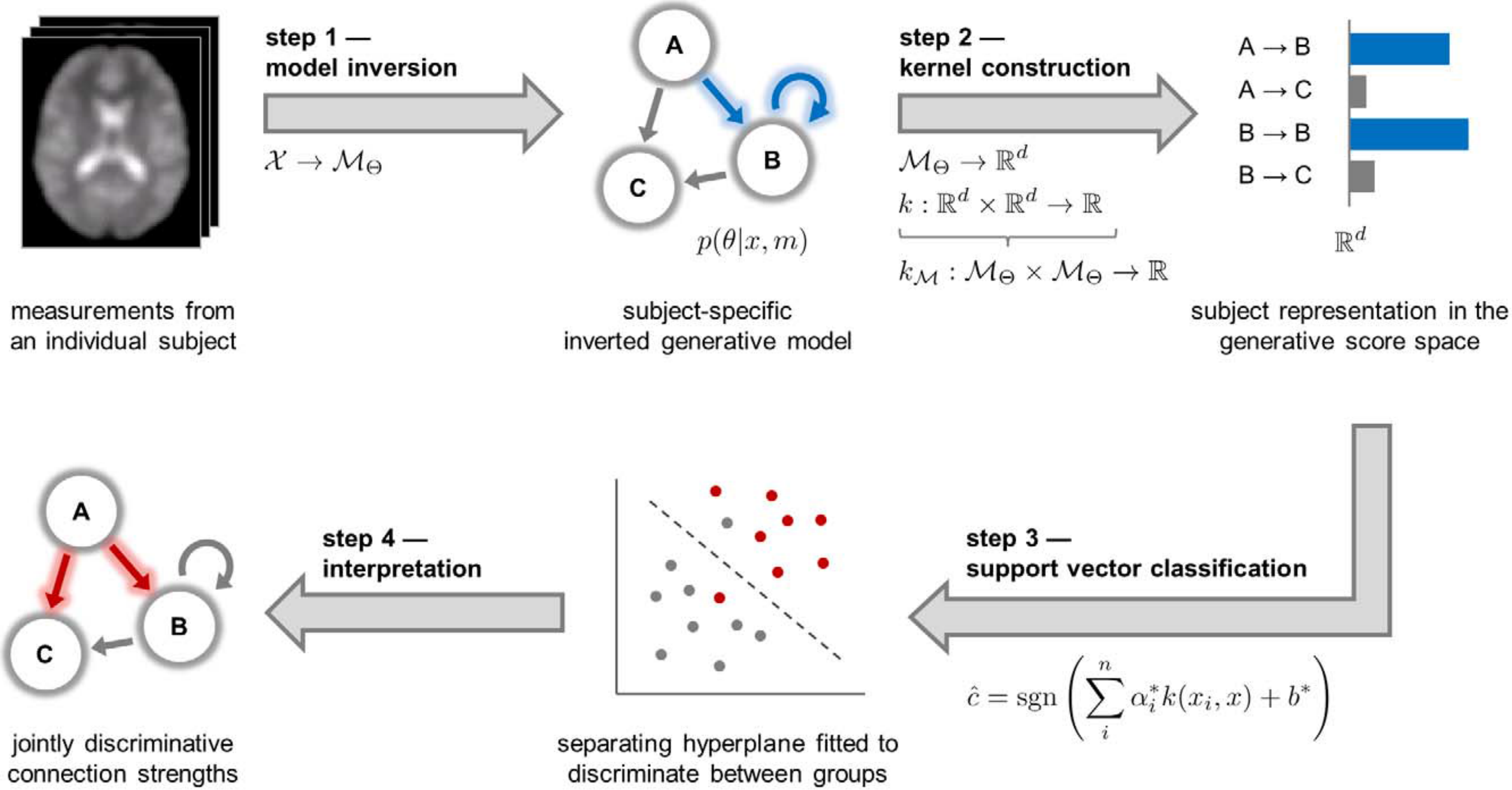






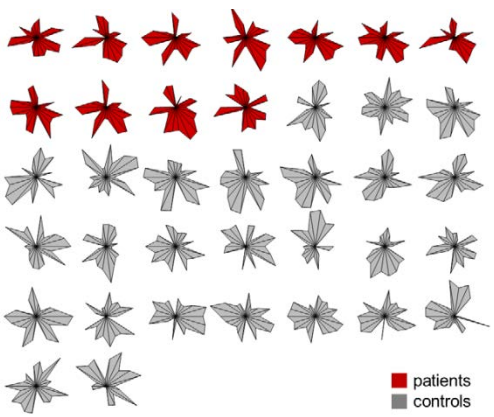
+0
+32
-4
Voxel activity
Connectivity
Subject 1
Subject 2
Subject N
Subject 1
Subject 2
Subject N
Dynamic causal
model (DCM)
Classification
Clustering
Group 1
Group 2

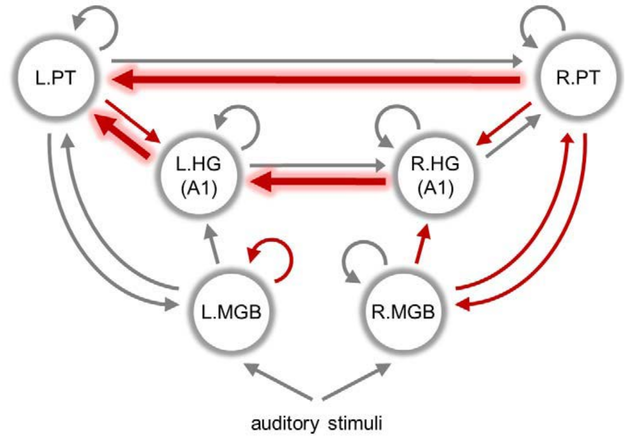
Generative Embedding - Classification

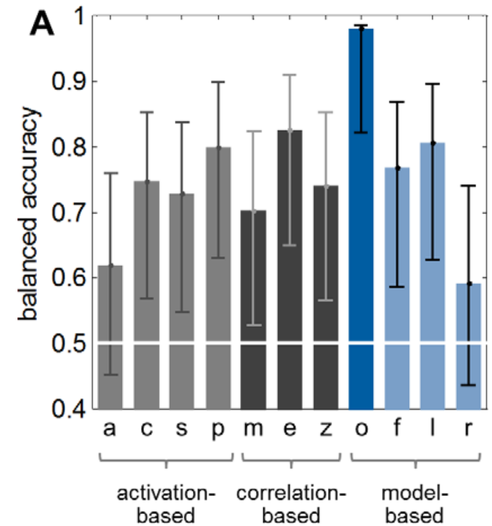


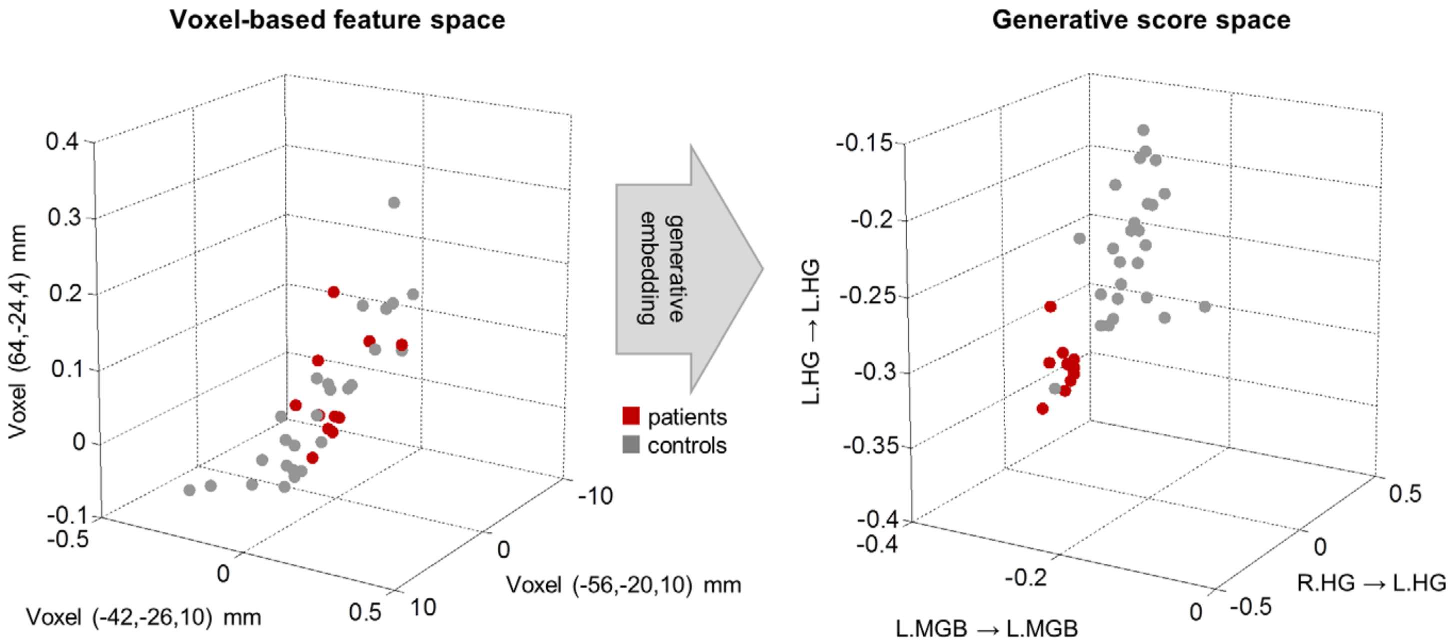
Brodersen et al. pLOS computation biology 2011.

DCM - Speech processing





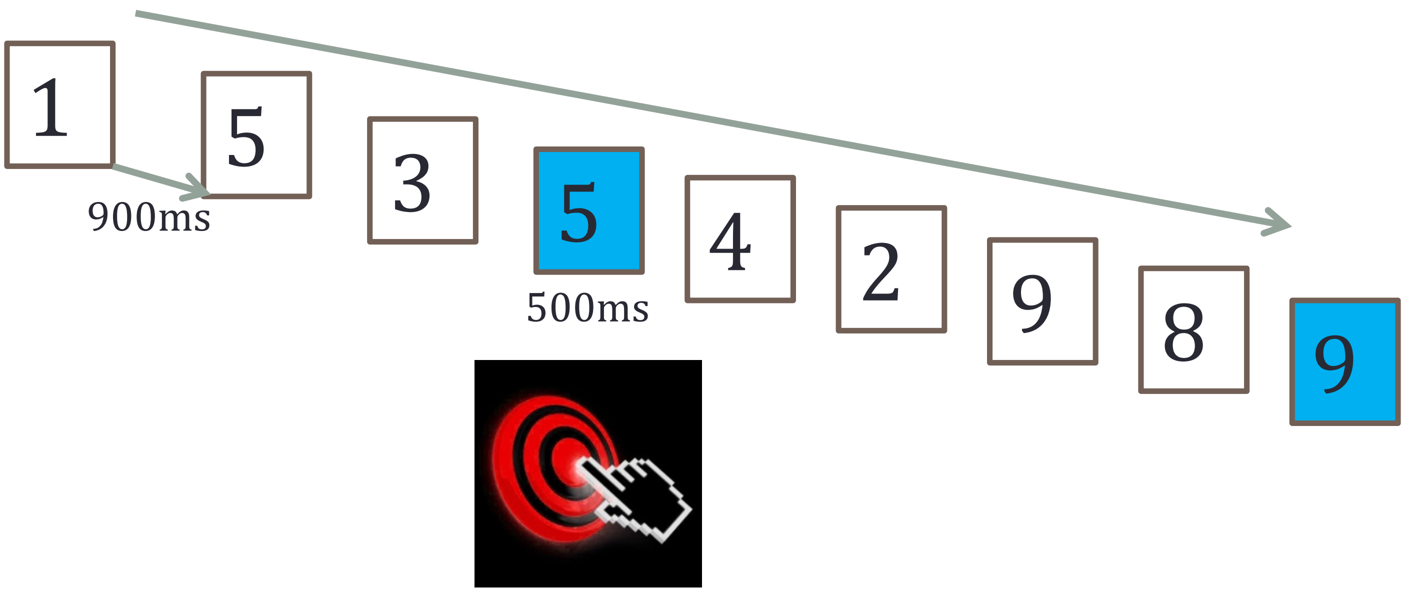




Brodersen et al. pLOS computation biology 2011.

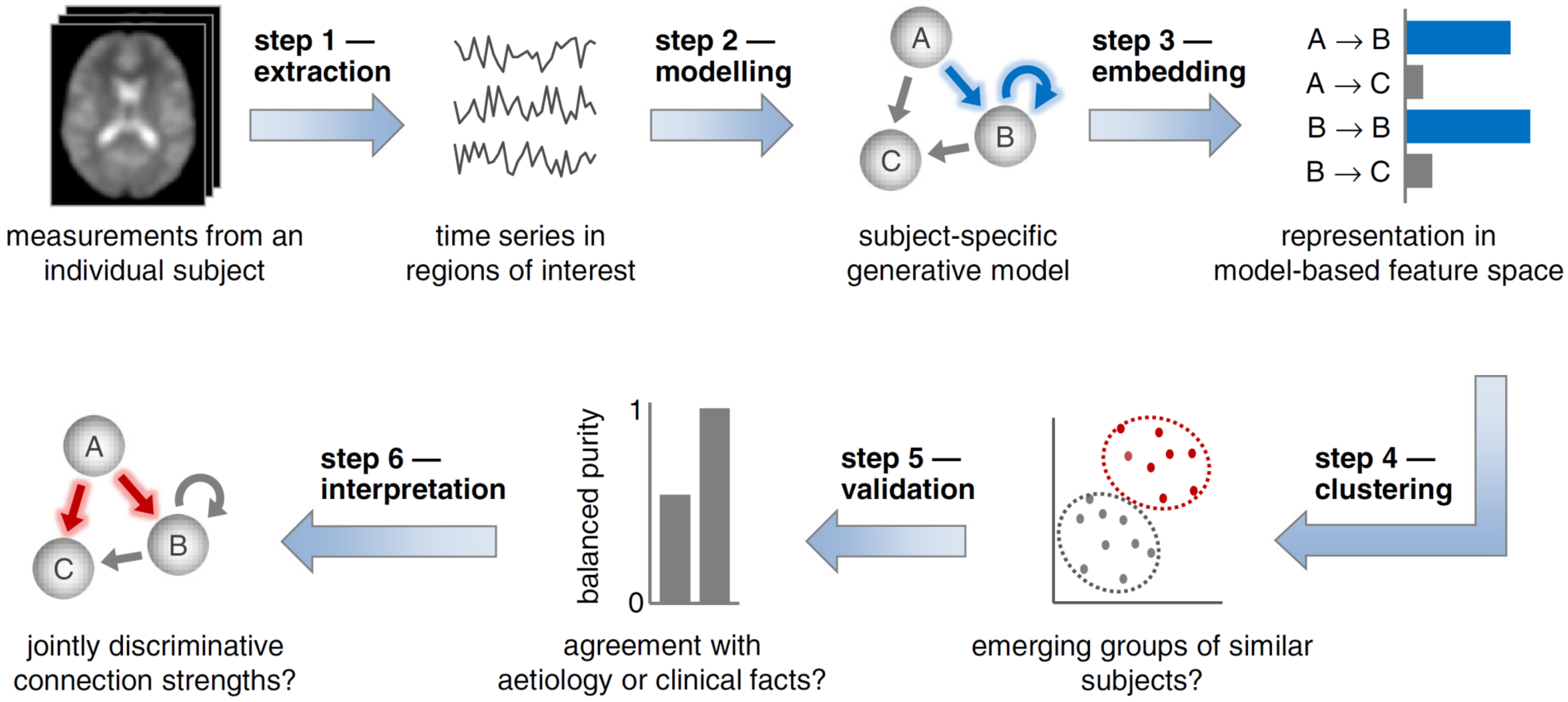
Working memory - fMRI

- 41 Schizophrenia patients (DCM IV,ICD 10), 42 controls
- Visual numeric n-back working memory task



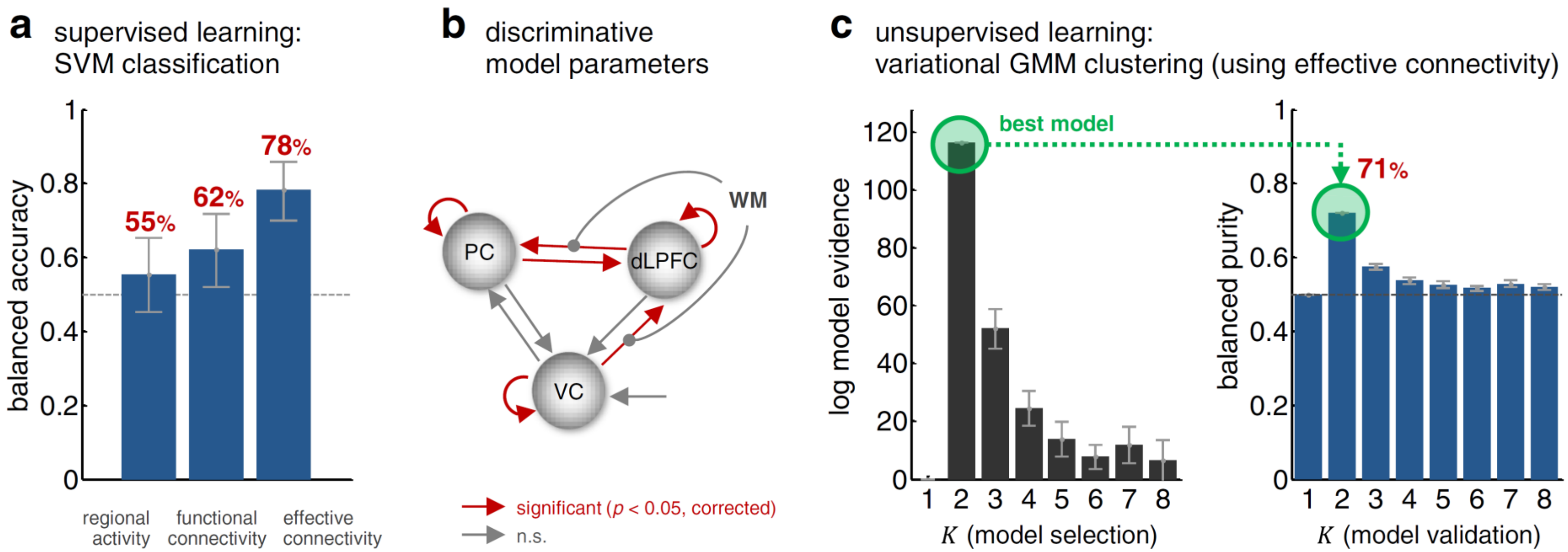
Deserno, Lorenz et al 2012. *The Journal of Neuroscience* 32 (1). Society for Neuroscience: 12–20.

Model based clustering



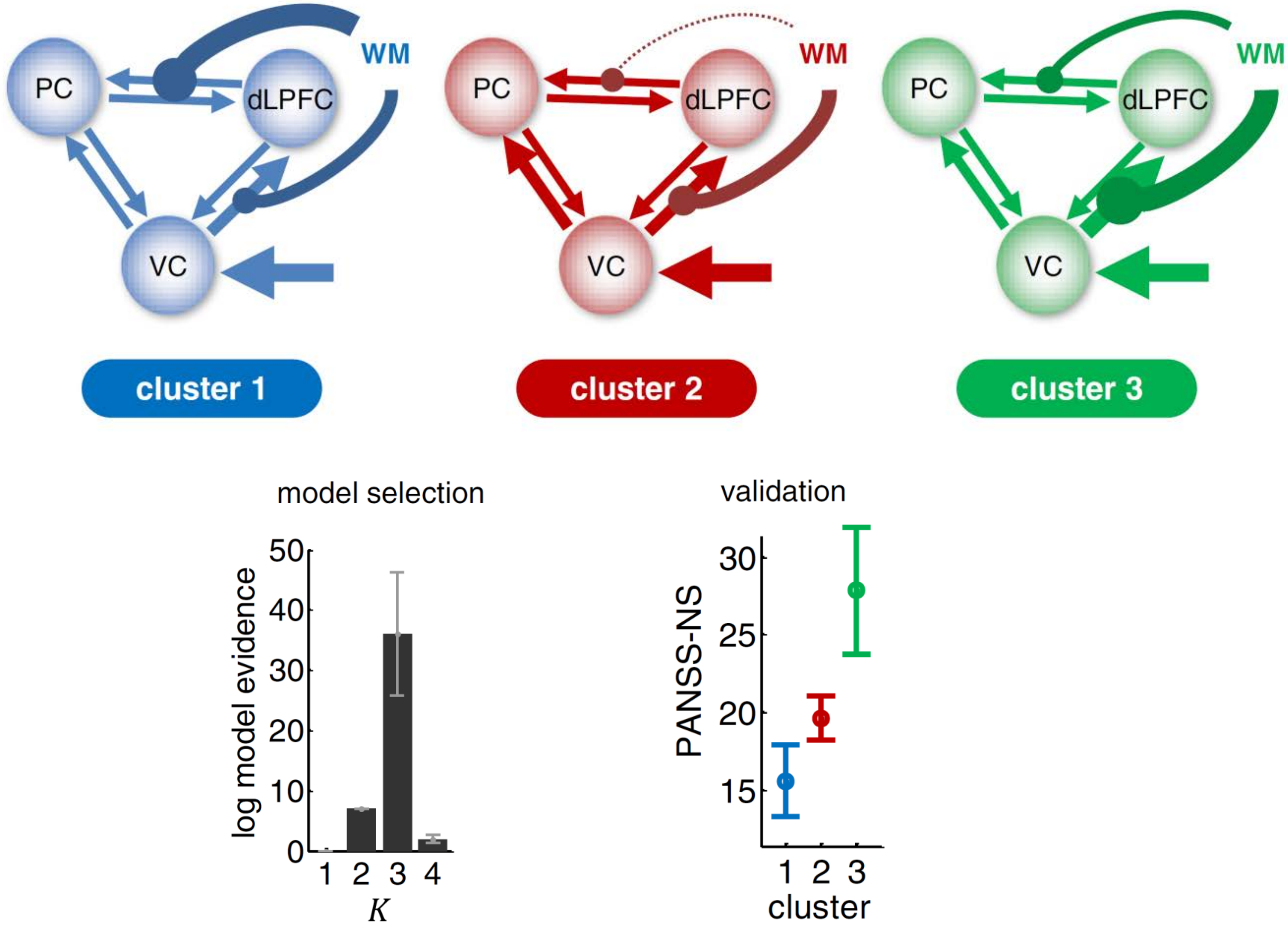
Brodersen et al 2014 Neuroimage

Results – Healthy vs Schizophrenic



Brodersen et al 2014 Neuroimage

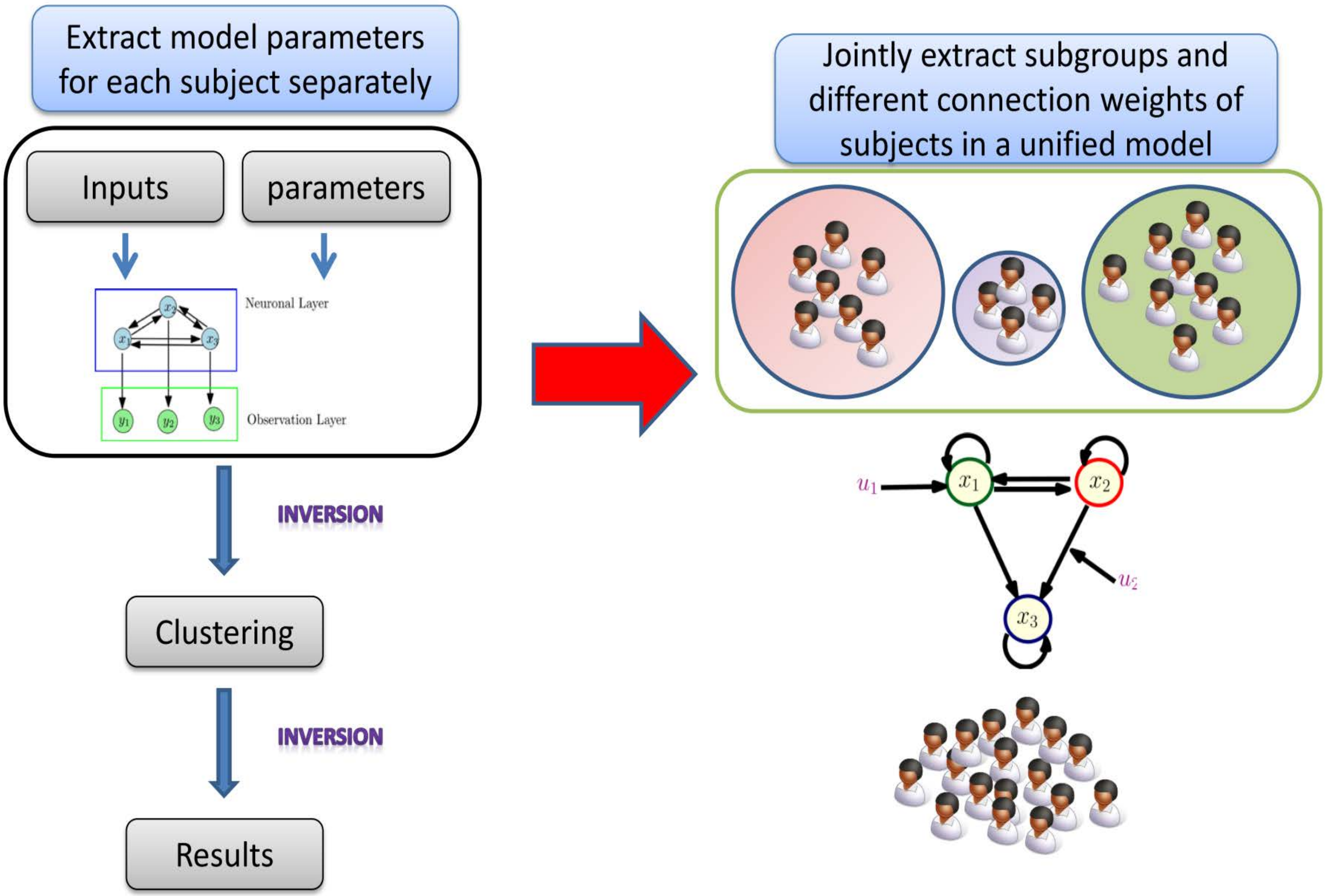
Results – Schizophrenia patients



Brodersen et al 2014 Neuroimage

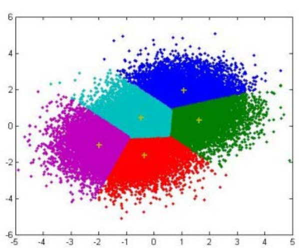
COMING SOON

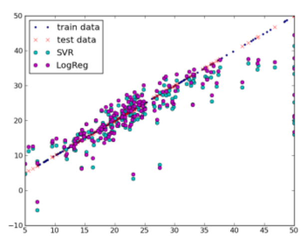
Unified model for identifying subgroups

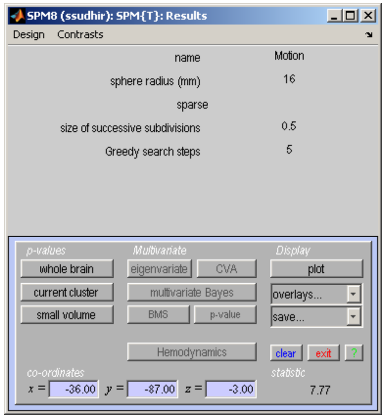


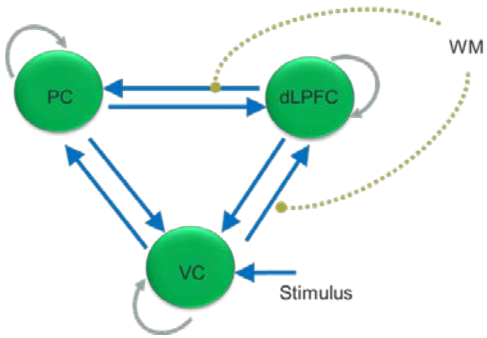
Summary

- Modelling Principles
- Learning from Data
- Multivariate Bayes in SPM
- Generative Embedding



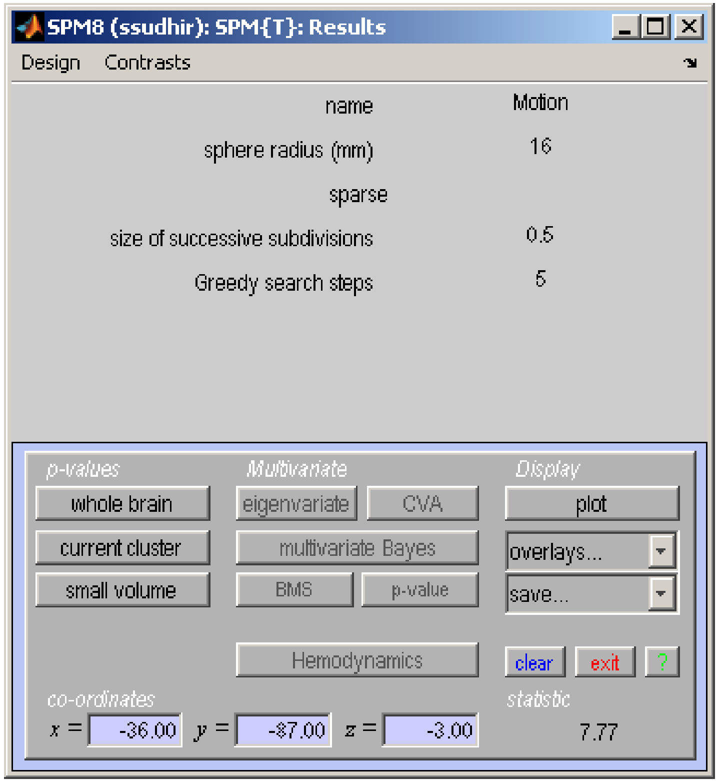




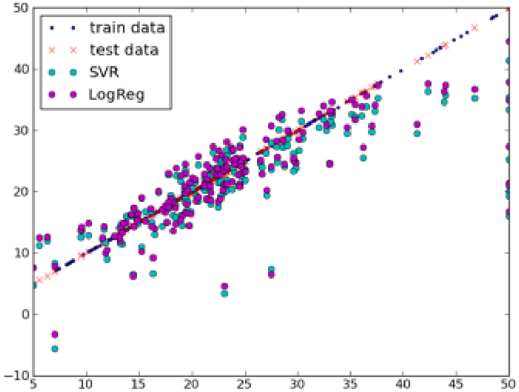


Practicals

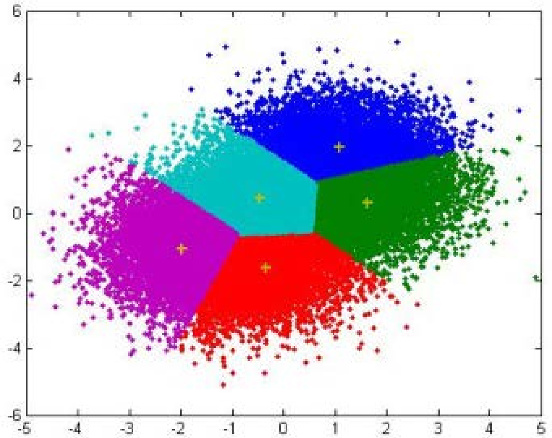
- Multivariate Bayes in SPM



- Classification



- Clustering



Thank you