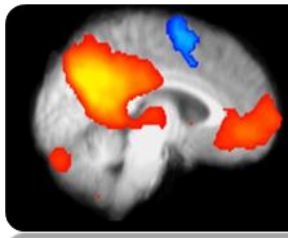
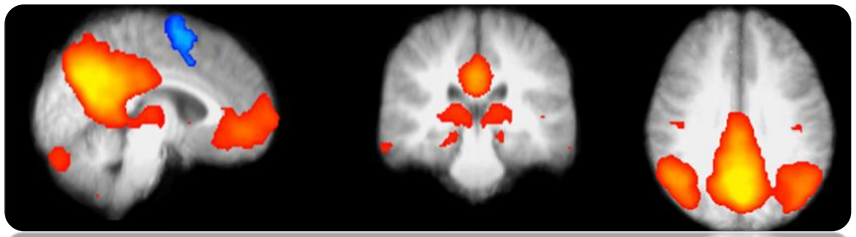


ICA Applications to resting- and event-related fMRI: theme & variations

Andrea Greve
Timothy Rittman
Simon Davis

Application of ICA-based methodology

Resting-state Networks



resting-state data

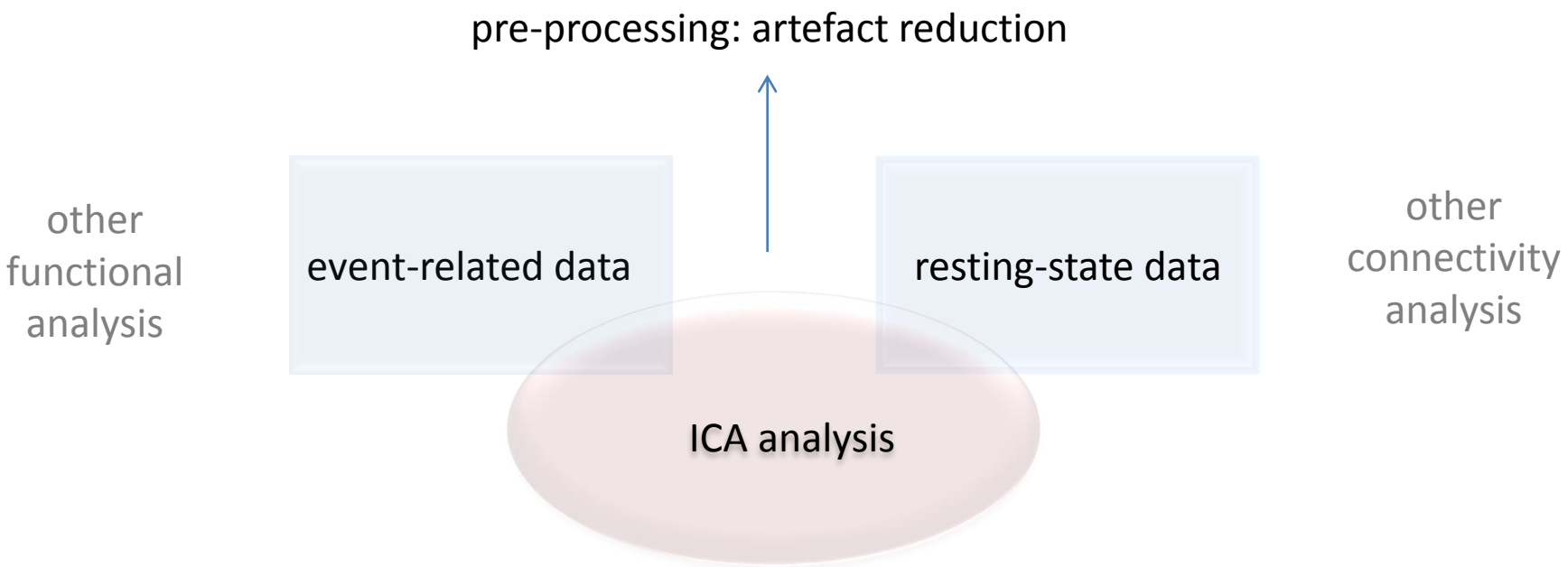
other connectivity analysis

low temporal frequency (<0.1 Hz)
correlation between functionally related brain regions
fluctuations can occur during rest (i.e. no explicit stimulus or task)

The hypothesis is that these low frequency fMRI signal fluctuations reflect correlated neuronal fluctuations in a network of task-related brain regions

Also referred to as: 'low-frequency correlations', 'default activity', 'default mode', 'spontaneous network correlations', 'intrinsic connectivity networks' ...

Application of ICA-based methodology

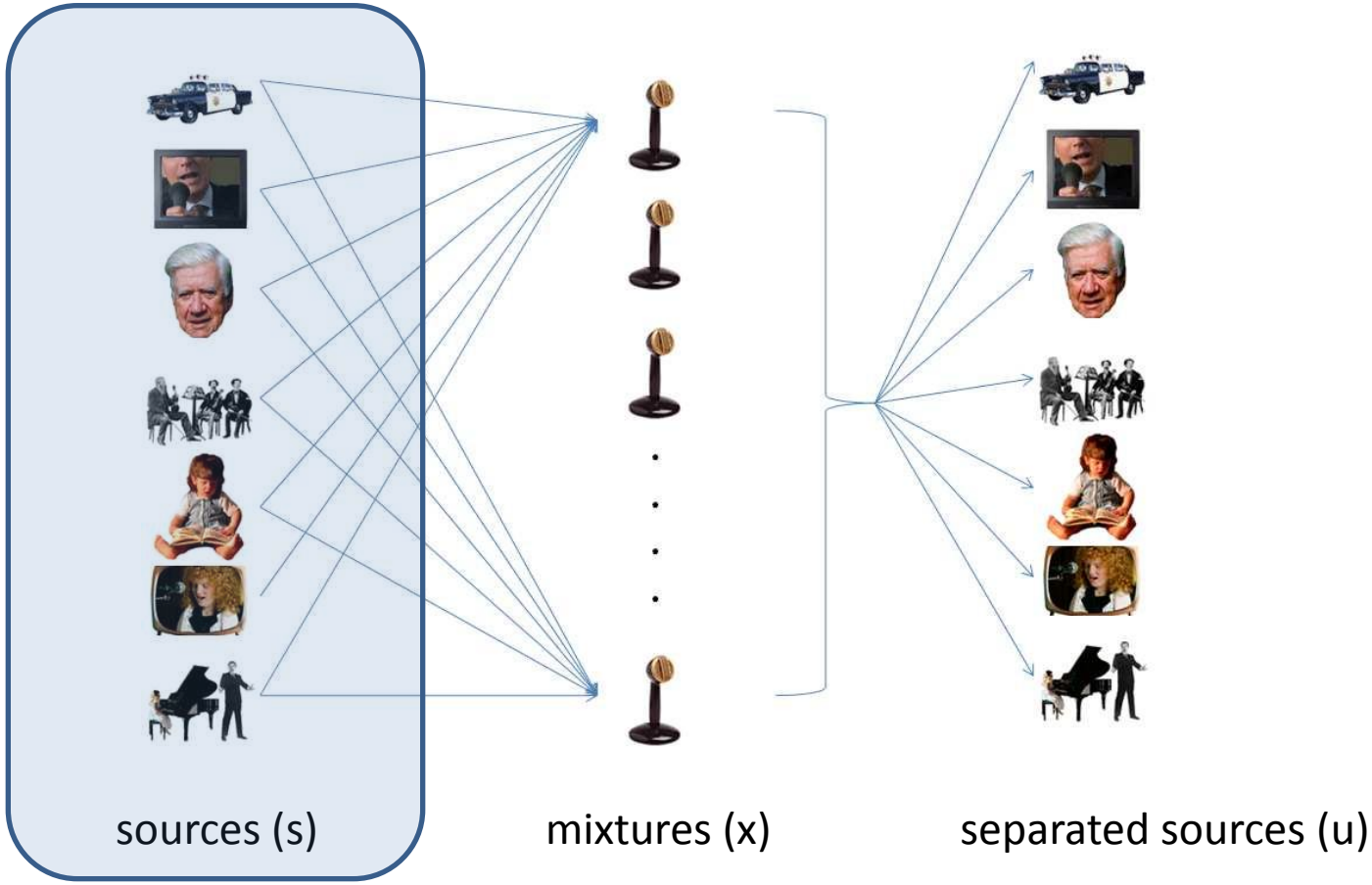


Spatial ICA, tensor based ICA
Single subject ICA vs. Group ICA
Dimensionality reduction (# comp.)
Selection and reliability of components
Statistical analysis, establishing group difference

ICA and blind source separation: a simple introduction

linear mixture by unknown matrix A

un-mixing matrix W



sources (s)

mixtures (x)

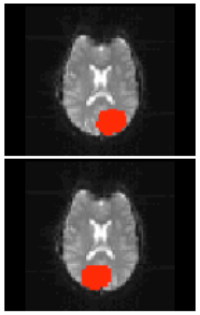
separated sources (u)

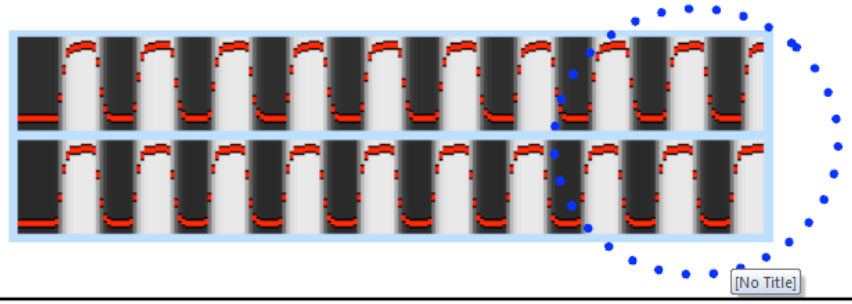
$$x_1 = a_{11}s_1 + a_{12}s_2$$
$$x_2 = a_{21}s_1 + a_{22}s_2$$
$$\ldots$$
$$x = As$$

PCA vs ICA

Simulated Data

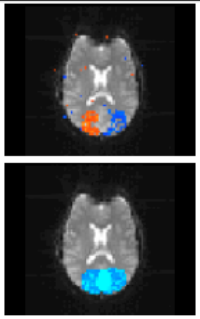
(2 components, slightly different timecourses)

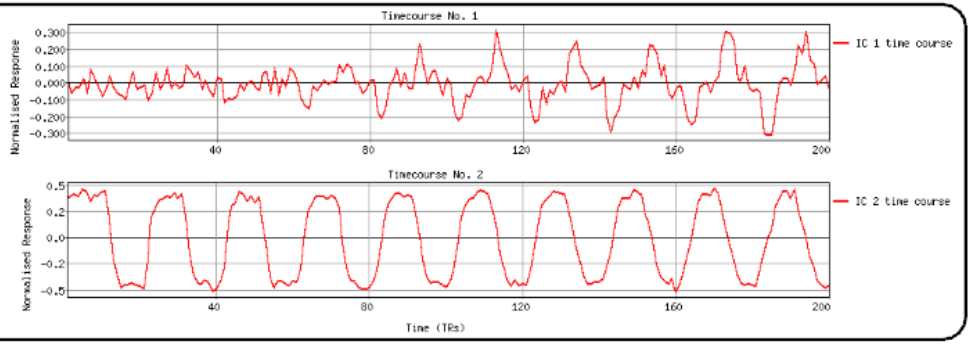




PCA

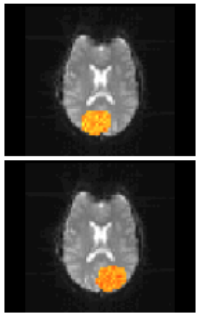
- Timecourses orthogonal
- Spatial maps and timecourses "wrong"

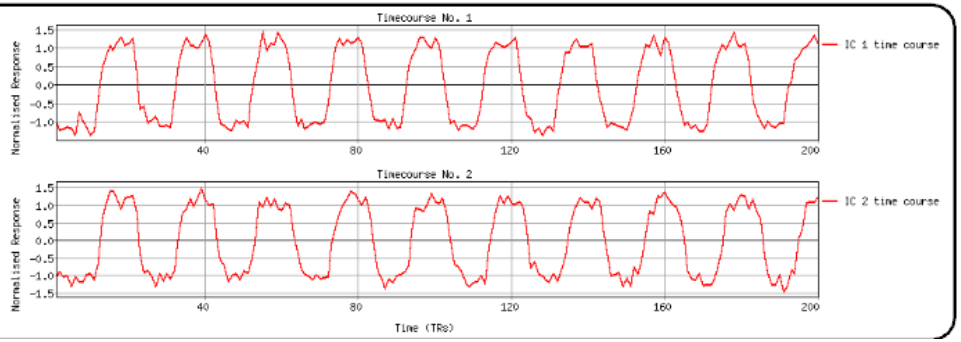




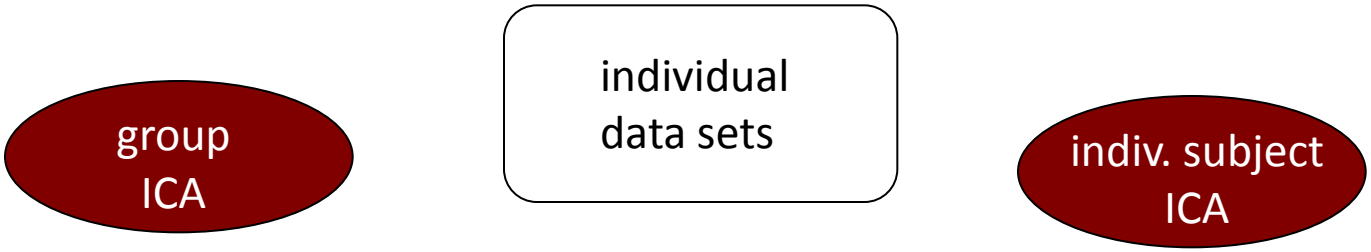
ICA

- Timecourses non-co-linear
- Spatial maps and timecourses "right"





ICA-based methodology for multi-subject RSN analysis



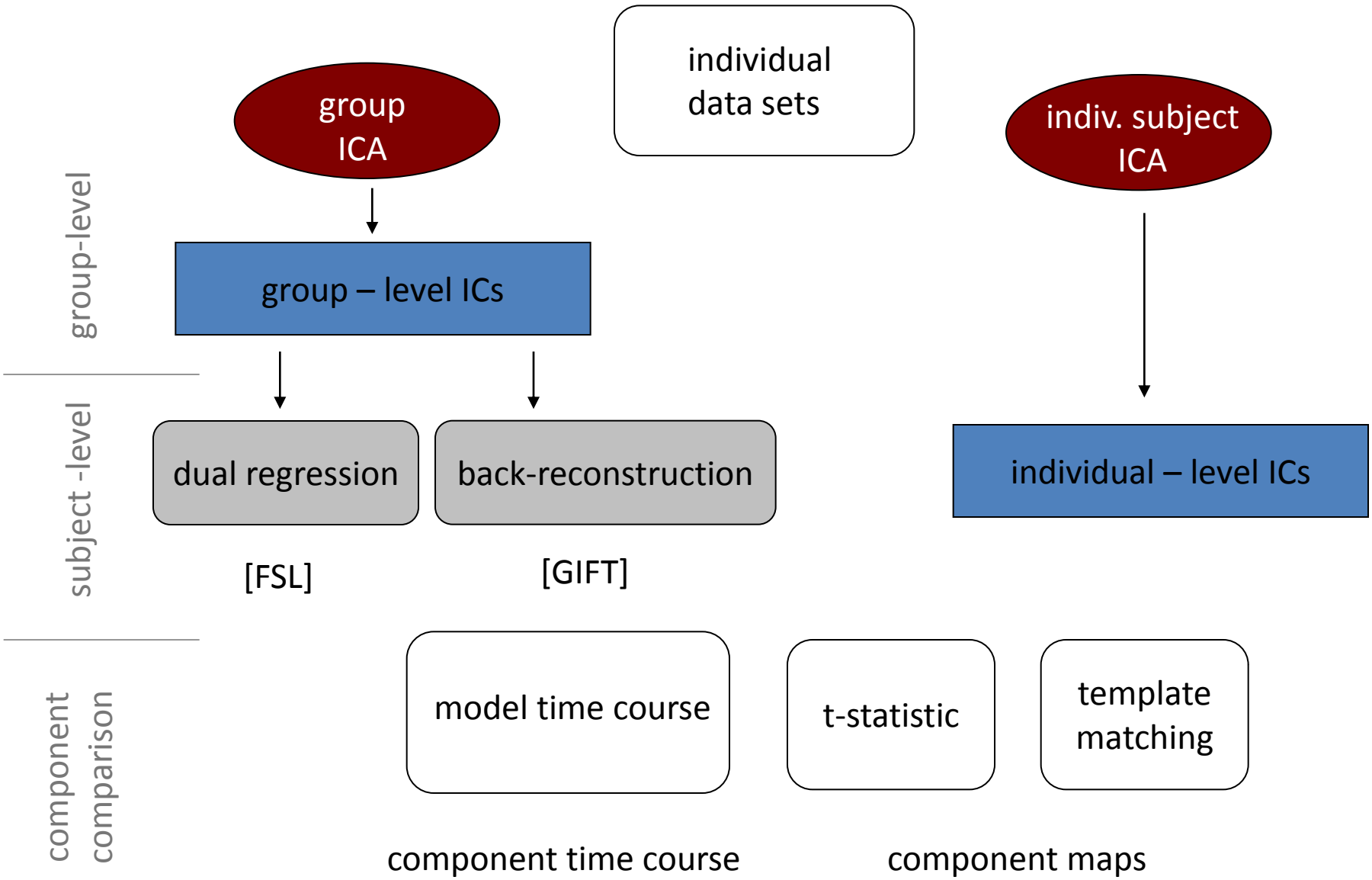
Why not just run ICA on each subject separately?

> correspondence problem of ICs across subjects
> different splitting *sometimes caused by small* changes in the data

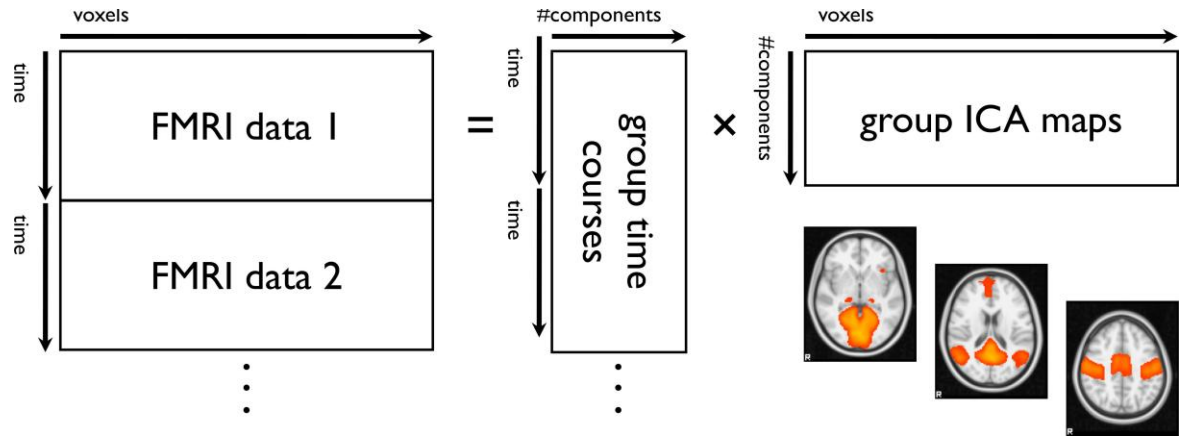
Instead - start with a "group-average" ICA

> but then need to relate group maps back to the individual subjects

ICA-based methodology for multi-subject RSN analysis

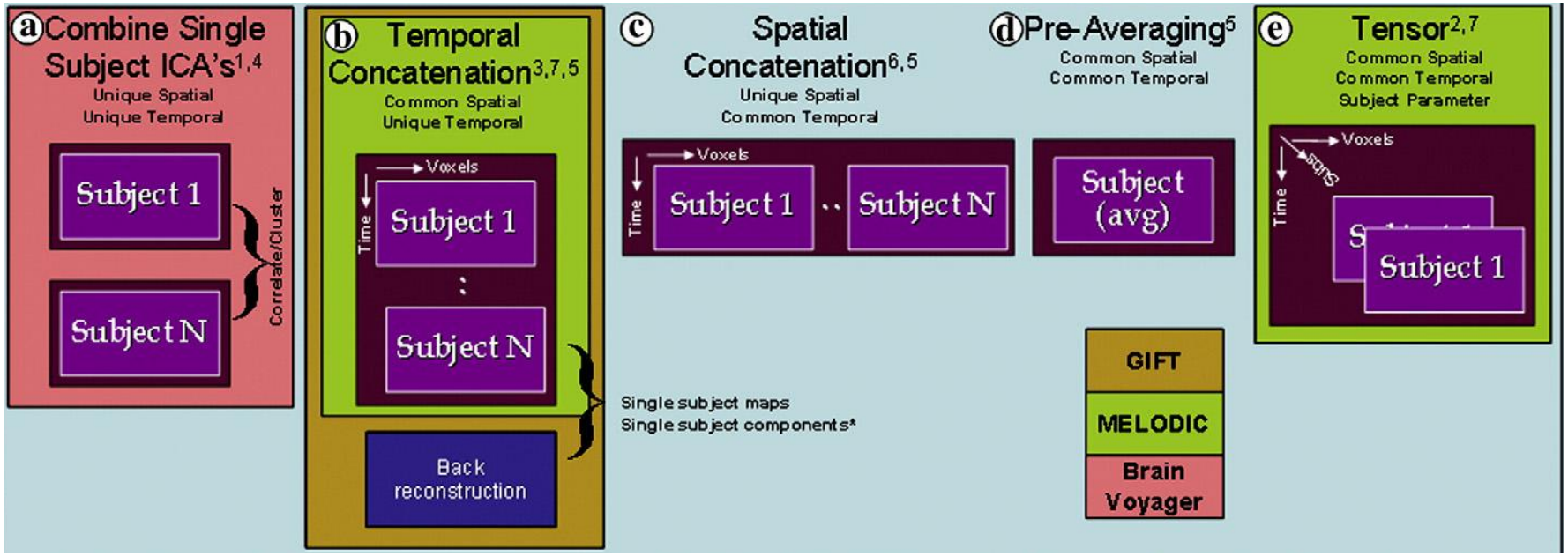


Temporal concatenation group-ICA



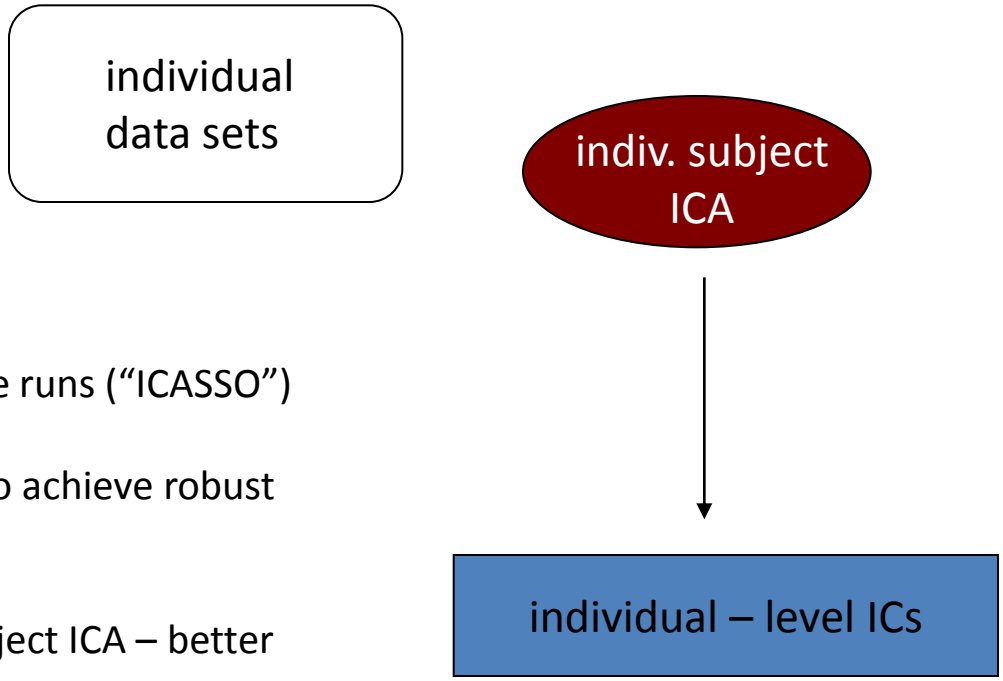
concatenate all subjects' data temporally
group-based PCA reduction on each subject

Several group ICA approaches



[Calhoun et al. 2009]

ICA-based methodology for multi-subject RSN analysis



Separate ICA for each subject

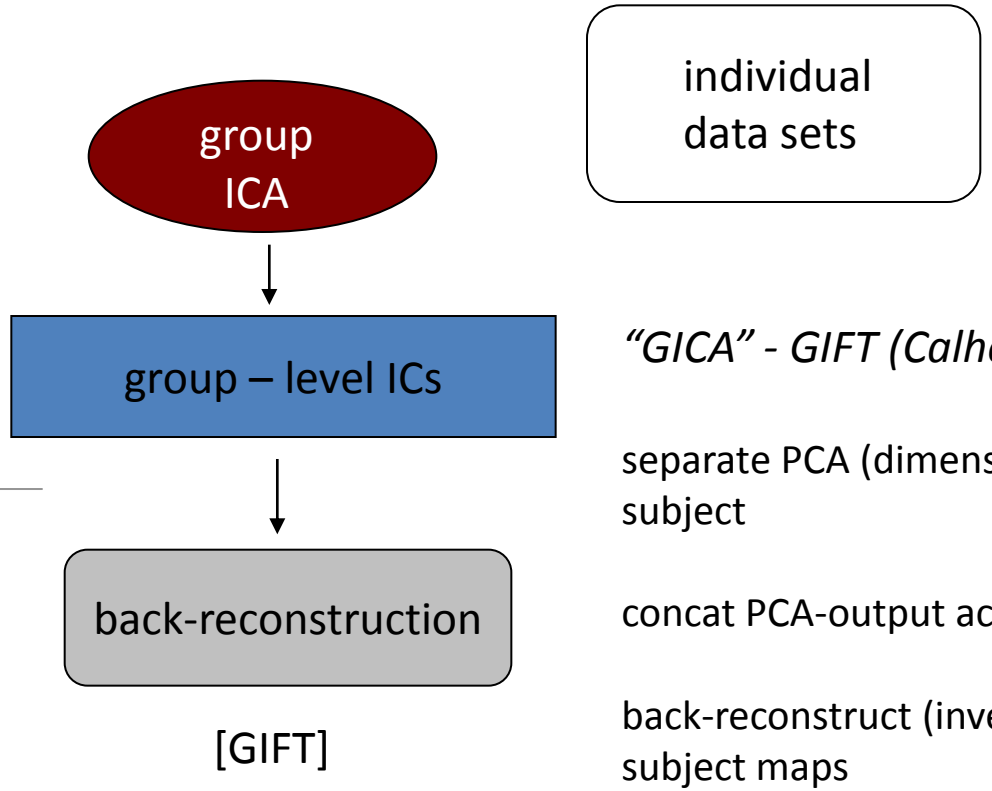
Robustness can be improved via multiple runs (“ICASSO”)

Compare components across subjects, to achieve robust matching of any given RSN

Advantage: Keeps benefits of single-subject ICA – better modelling of structured noise in data

Disadvantage: Correspondence problem, in particular different splitting in different subjects caused by even very small changes in the data

ICA-based methodology for multi-subject RSN analysis



"GICA" - GIFT (Calhoun, HBM, 2001)

separate PCA (dimensionality reduction) for each subject

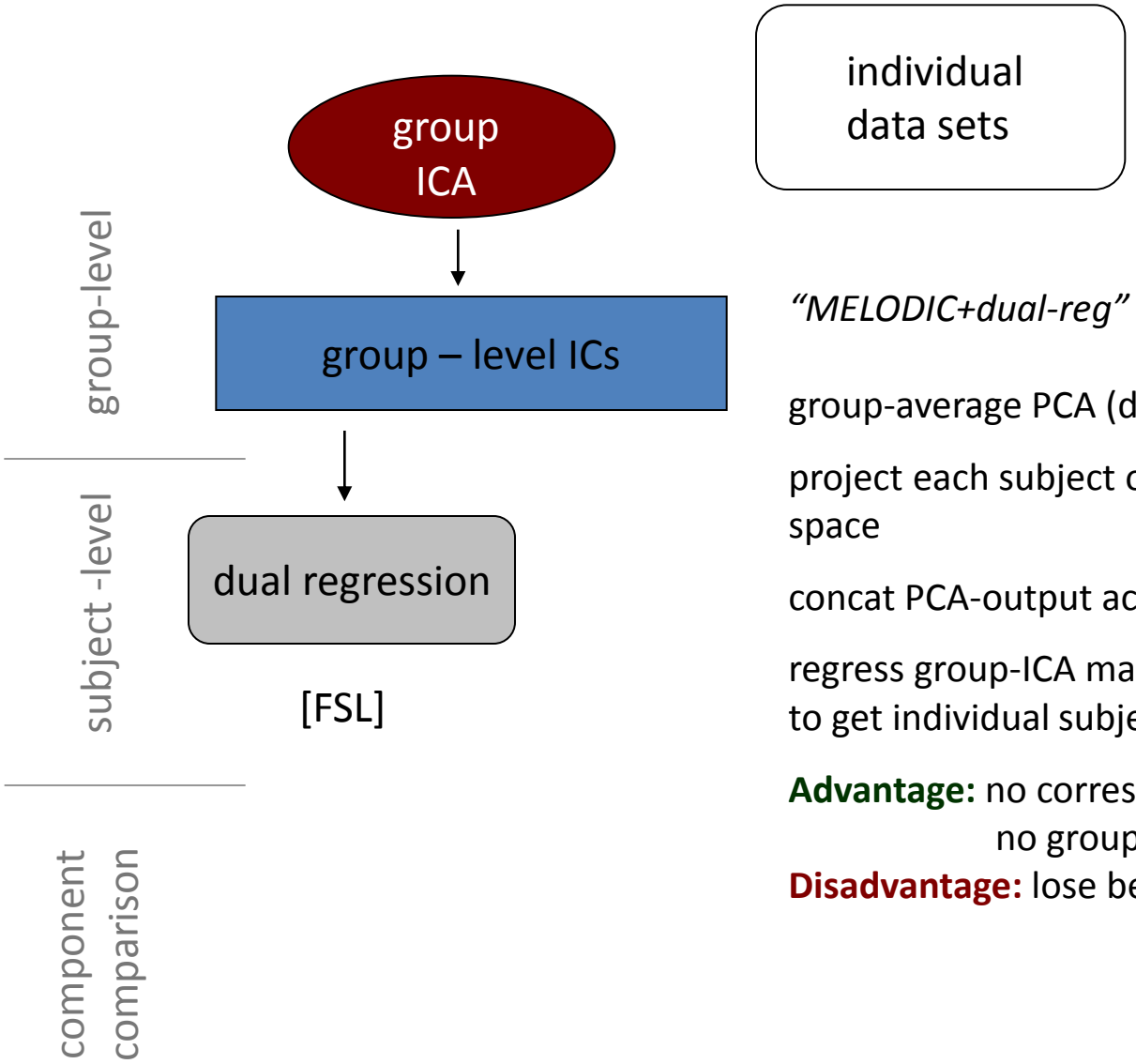
concat PCA-output across subjects and do group-ICA

back-reconstruct (invert) ICA results to get individual subject maps

Advantage: No correspondence problem

Disadvantage: lose benefits of single-subject ICA, PCA-bias

ICA-based methodology for multi-subject RSN analysis



"MELODIC+dual-reg" - FSL (Beckmann, OHBM, 2009)

group-average PCA (dimensionality reduction)

project each subject onto reduced group-average-PCA-space

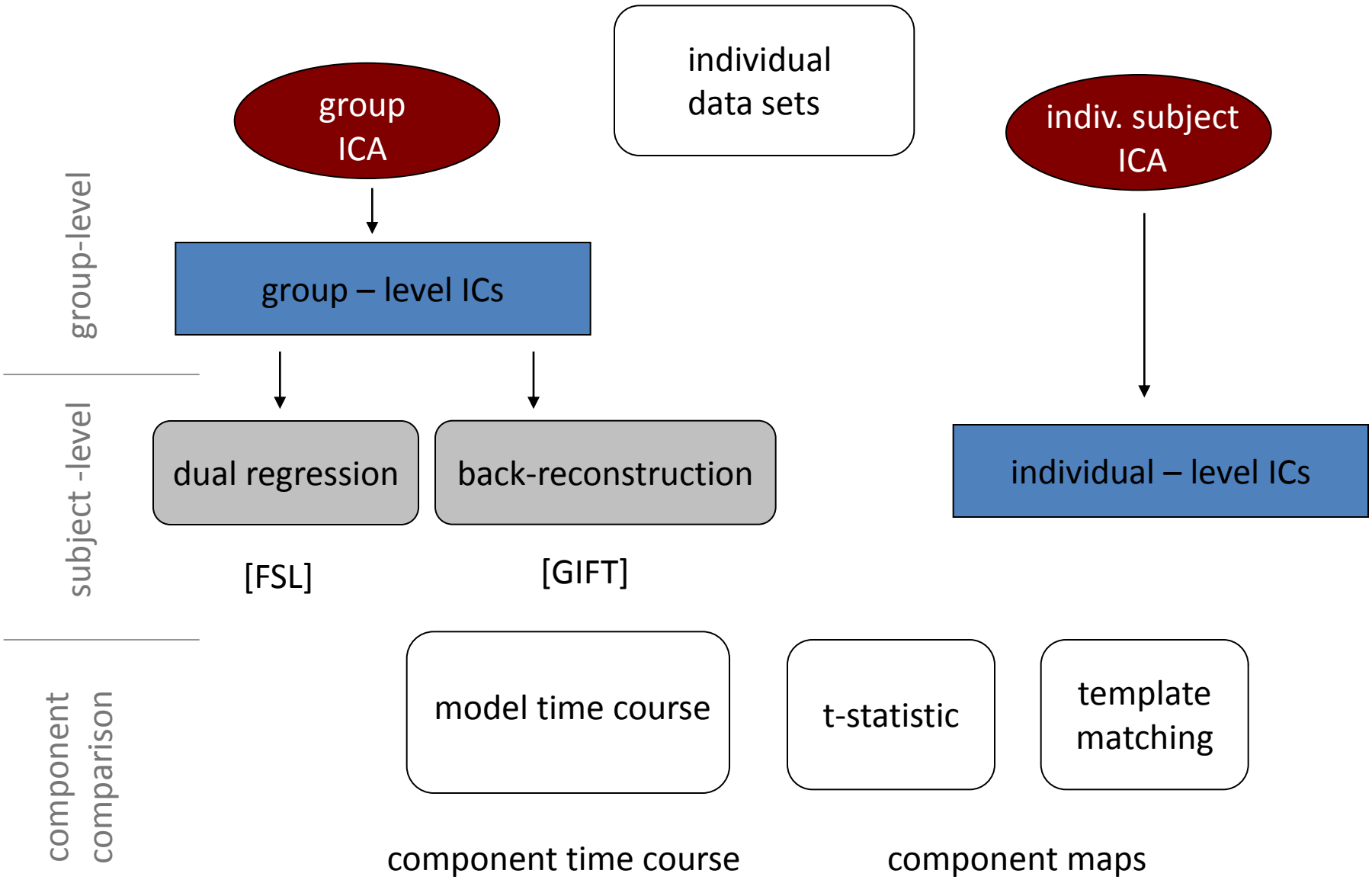
concat PCA-output across subjects and do group-ICA

regress group-ICA maps onto individual subject datasets to get individual subject maps

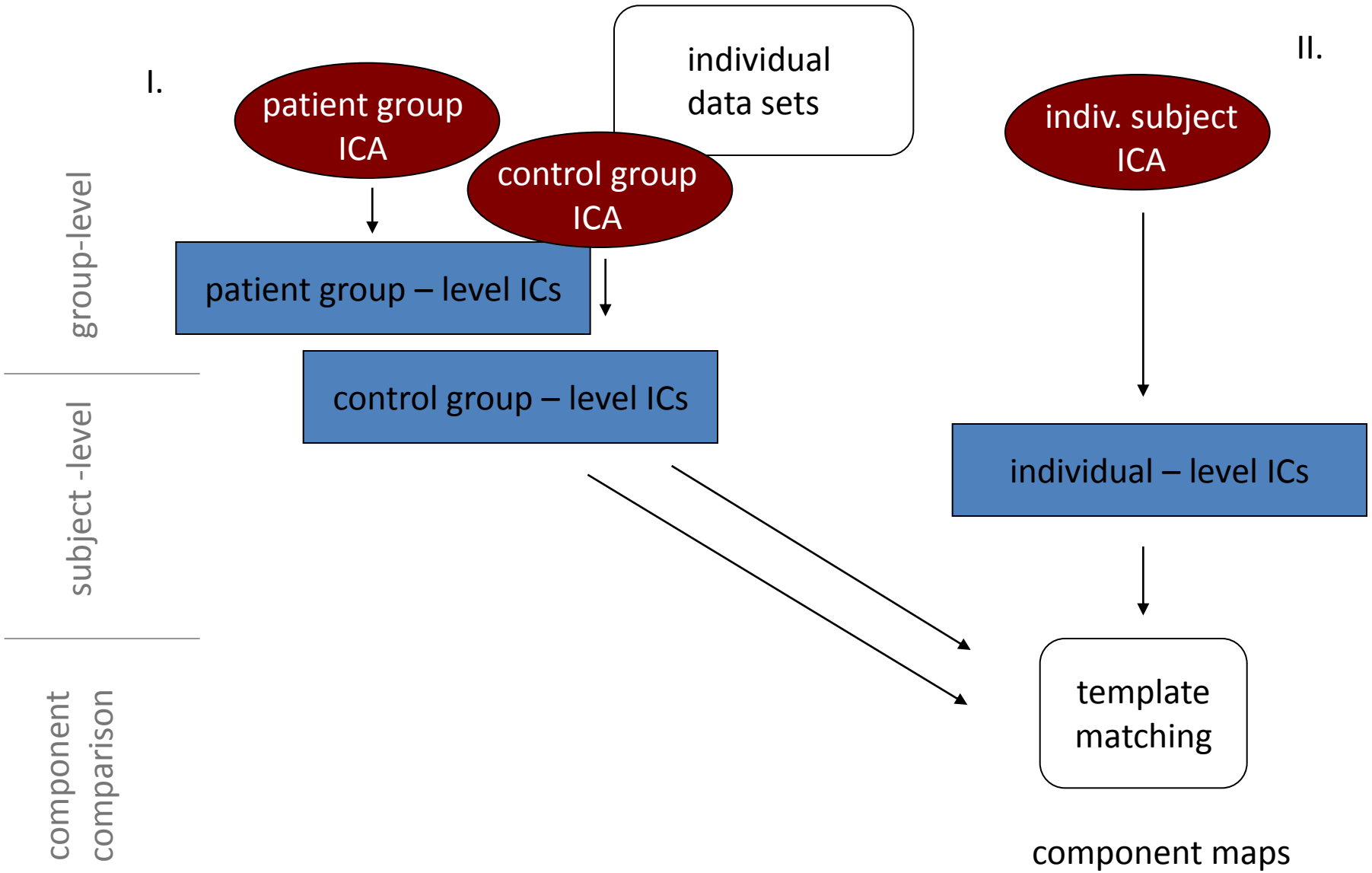
Advantage: no correspondence problem,
no group/PCA bias

Disadvantage: lose benefits of single-subject ICA

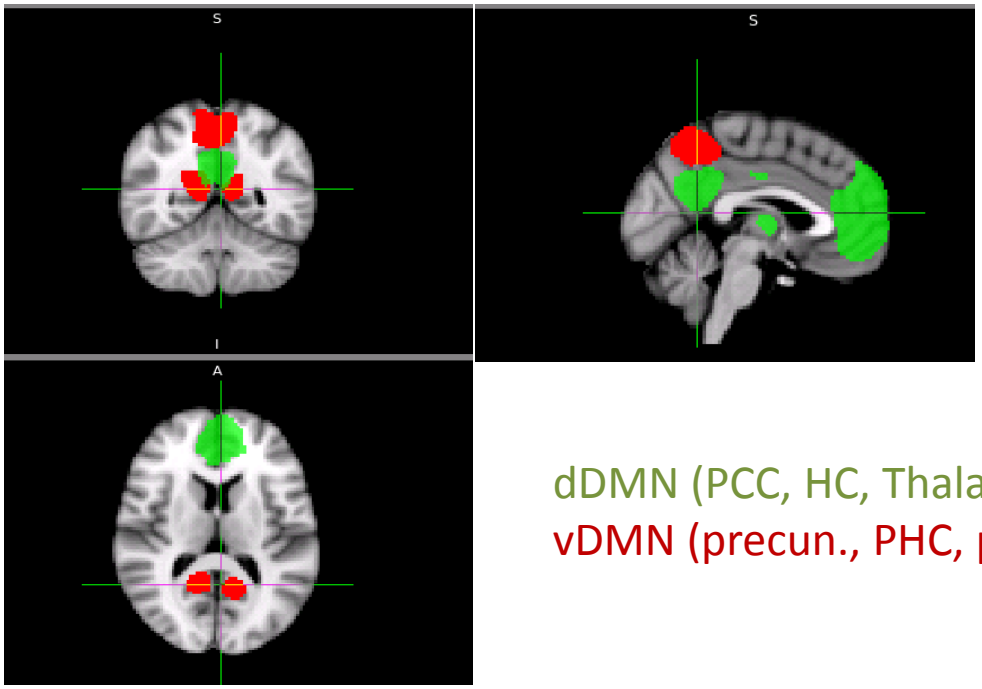
ICA-based analysis of 4 amnesic patients and 29 matched controls



ICA-based analysis of 4 amnesic patients and 29 matched controls



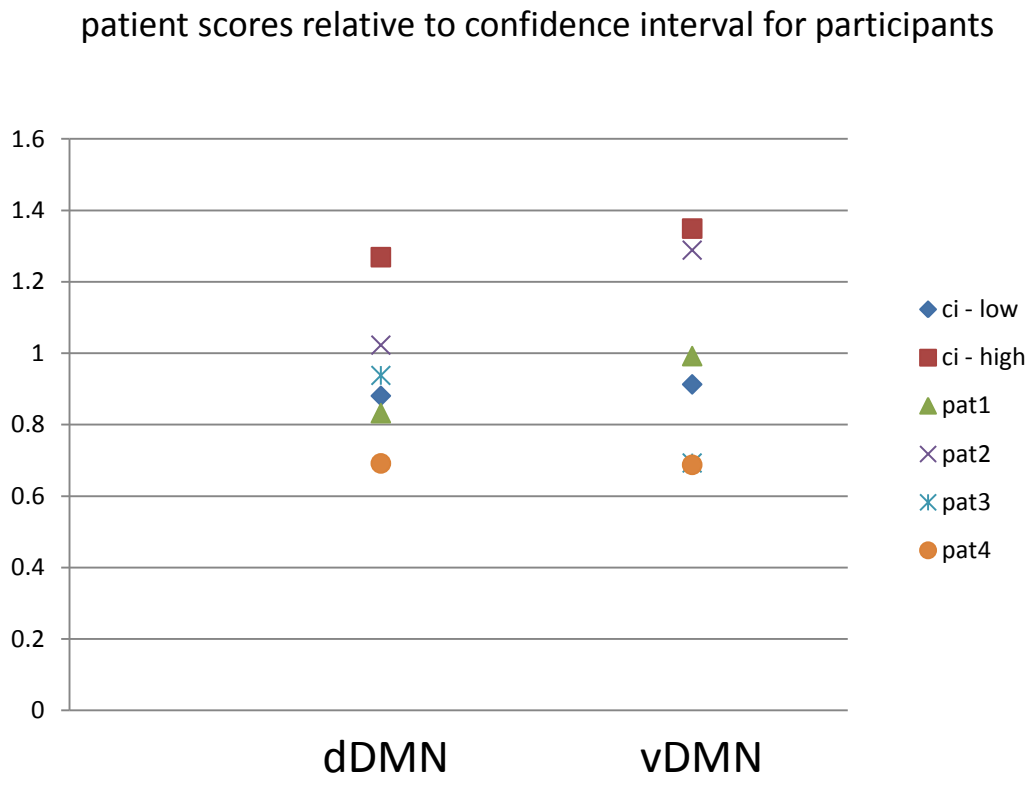
Template matching: functional templates from Mike Greicius' lab



dDMN (PCC, HC, Thalamus, mPFC)
vDMN (precun., PHC, parietal /occipit. L)

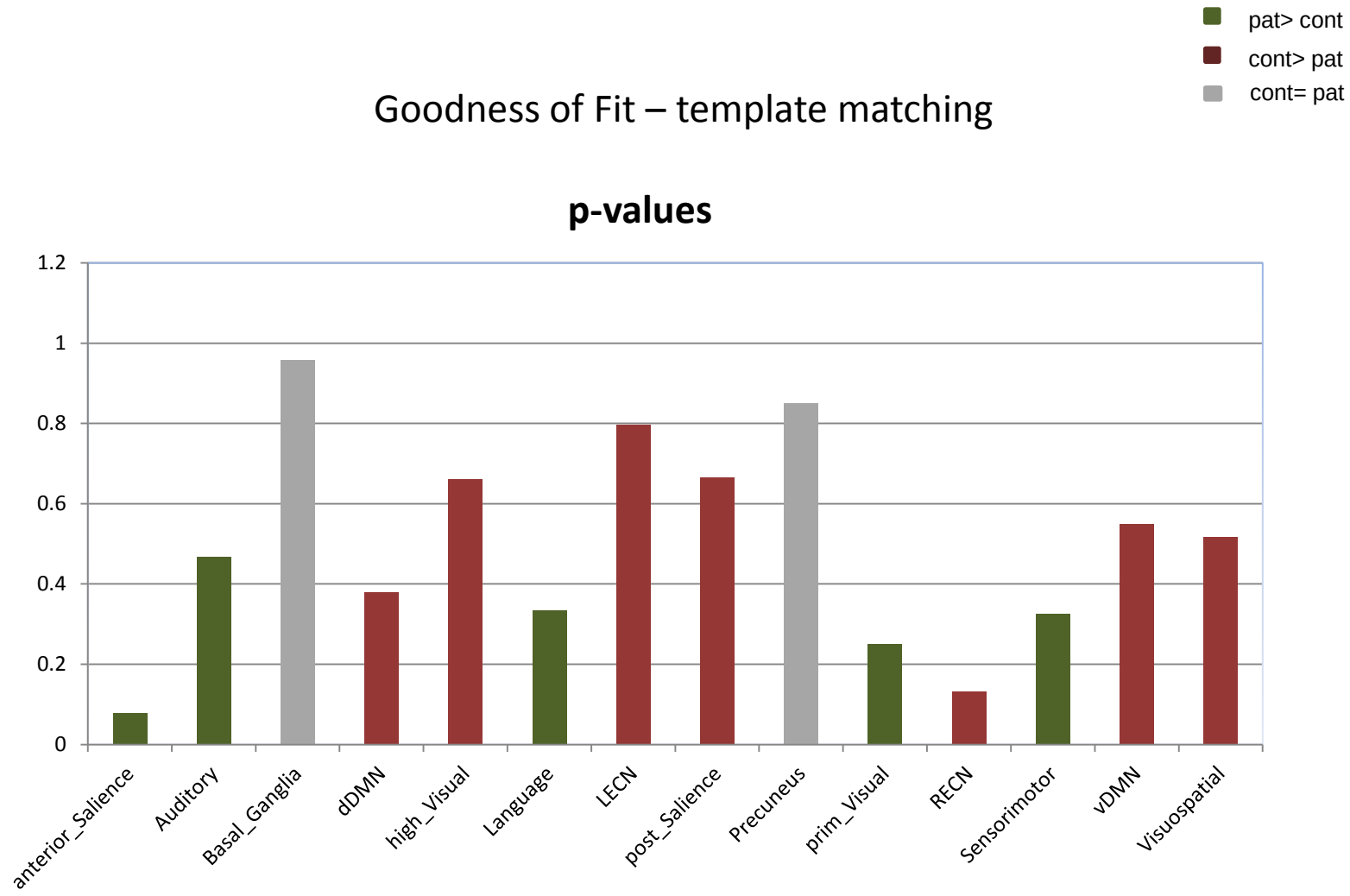
(**http://findlab.stanford.edu/research.html**) 14 binary templates in MNI space

Goodness of Fit – template matching



dDMN t(31) = -0.891, p=0.380
vDMN t(31) = -0.607, p=0.548

Goodness of Fit – template matching



	ant_Sal	Auditory	Basal_Gan	dDMN	high_Vis	Language	LECN	post_Sal	Precuneus	prim_Vis	RECN	Sensorim	vDMN	Visuospat
p-values	0.0775	0.4681	0.957	0.3797	0.6613	0.3337	0.796	0.664	0.8497	0.2503	0.1317	0.3259	0.5485	0.5165
t-values	1.8257	0.7346	0.0543	-0.8912	-0.4424	0.9819	-0.2607	-0.4386	-0.1911	1.1715	-1.5482	0.9982	-0.6067	-0.6563

Matched templates

Prob. template matching using correlation measure

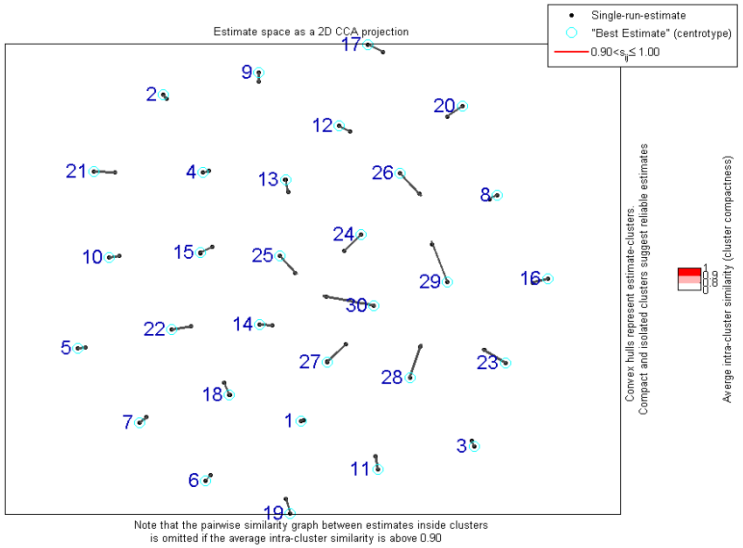
DMN_ICA

Control

17
17
17
23
17
17
23
17
23
23
17
23
17
17
17
17
17
17
17
17
17
17
17
17
23
17
17

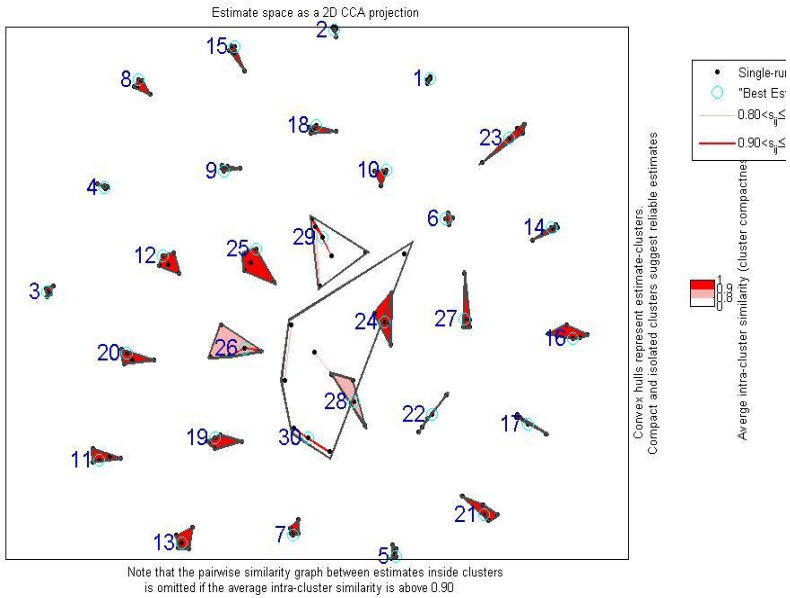
Patients

3
3
3
3



GIFT component sorting

ICmap	score
17	0.43
23	0.39
22	0.18
1	0.14

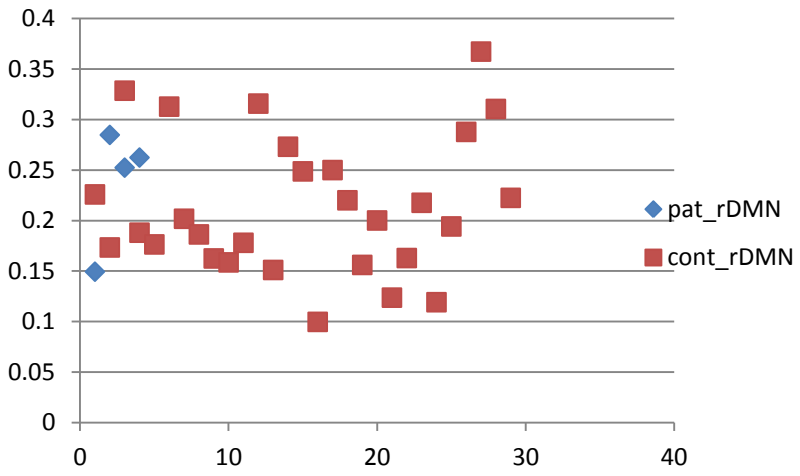


GIFT component sorting

ICmap	score
3	0.33
23	0.13
18	0.12
28	0.09

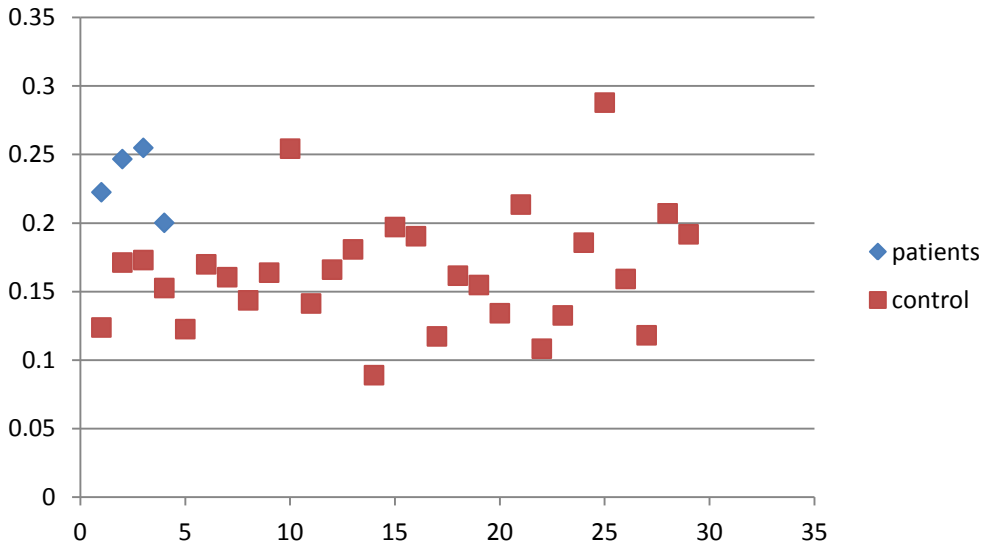
Prob. template matching using correlation measure

Melodic indiv. Subj ICA
prob. template (from GIFT) used in template matching (correlation measure)

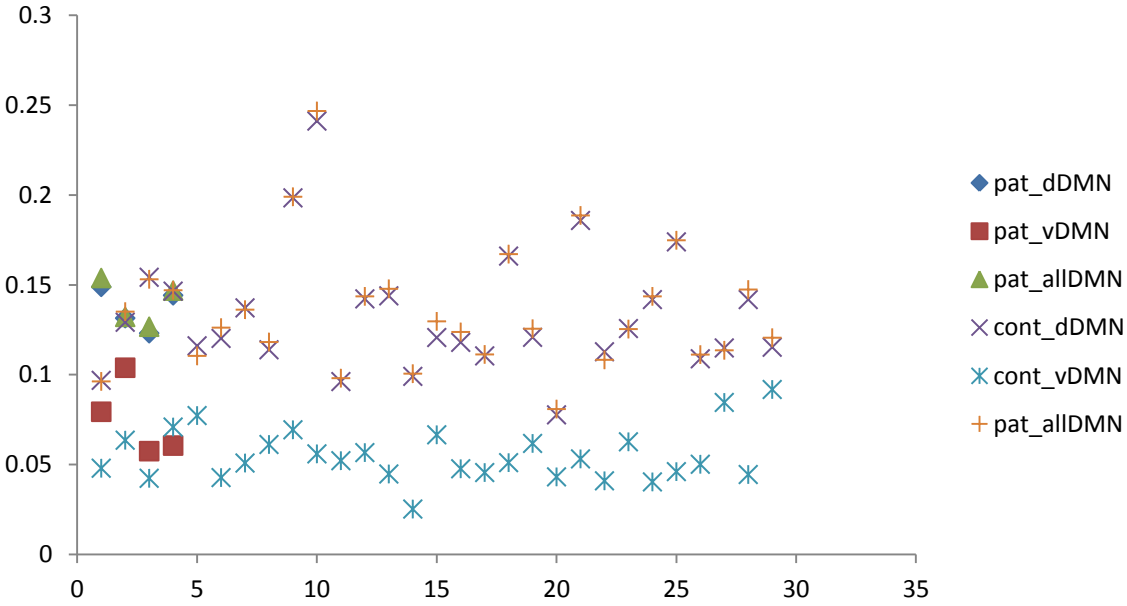


Prob. template matching using correlation measure

GIFT: Group ICA [pat] temporal concatenation
GFIT: Group ICA [contr] temporal concatenation
indiv. spatial template: back projection into subject space
prob. template matching using correlation measure



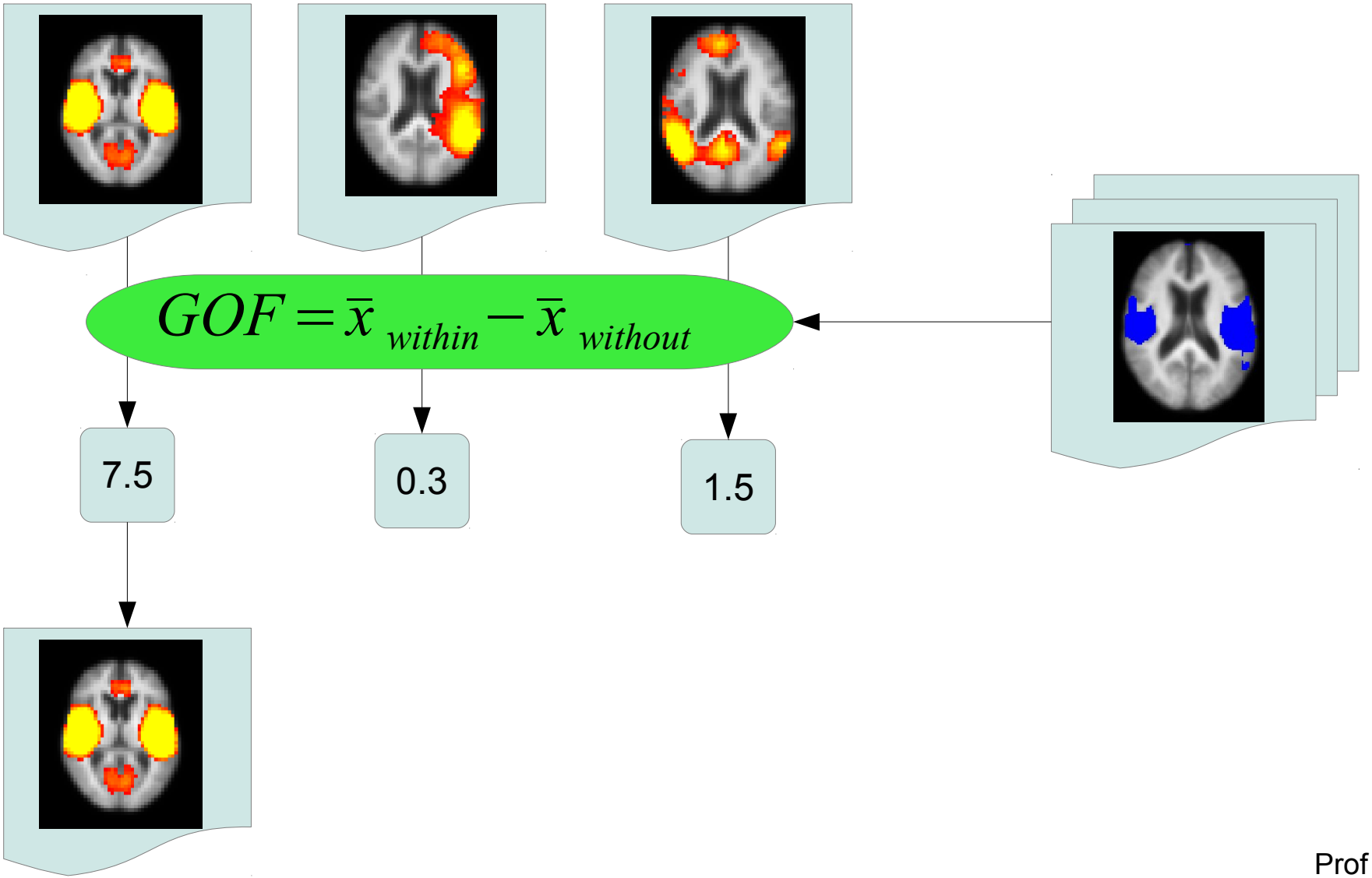
Prob. template matching using correlation measure



Study outline

- Disease groups with dementia
  - Progressive supranuclear palsy
  - Corticobasal degeneration
- Wider context
  - development of network related biomarkers
  - correlation between disease pathology, macroscopic networks and clinical measures
- Challenge
  - group comparison of ICA data

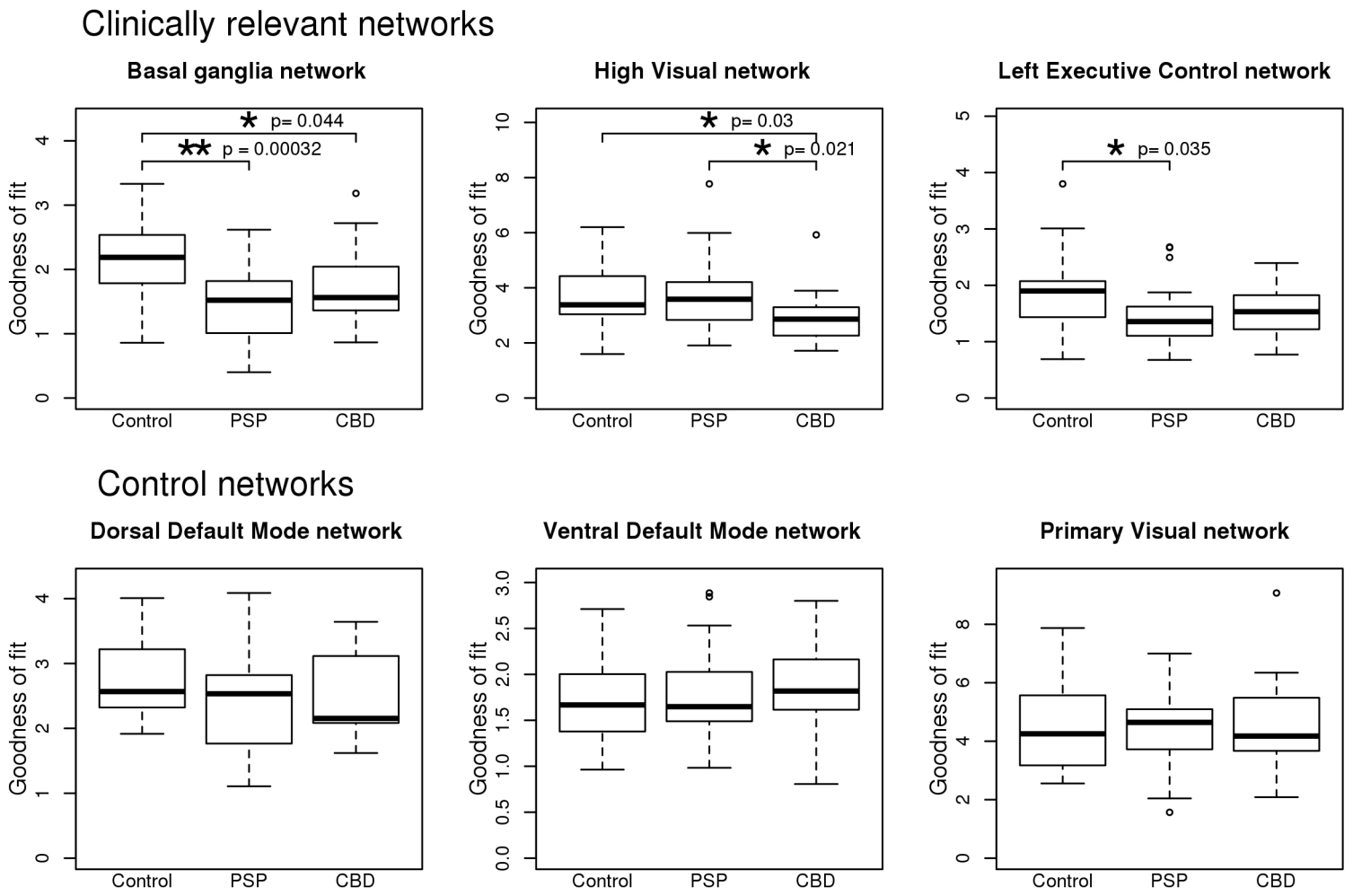
Goodness of fit



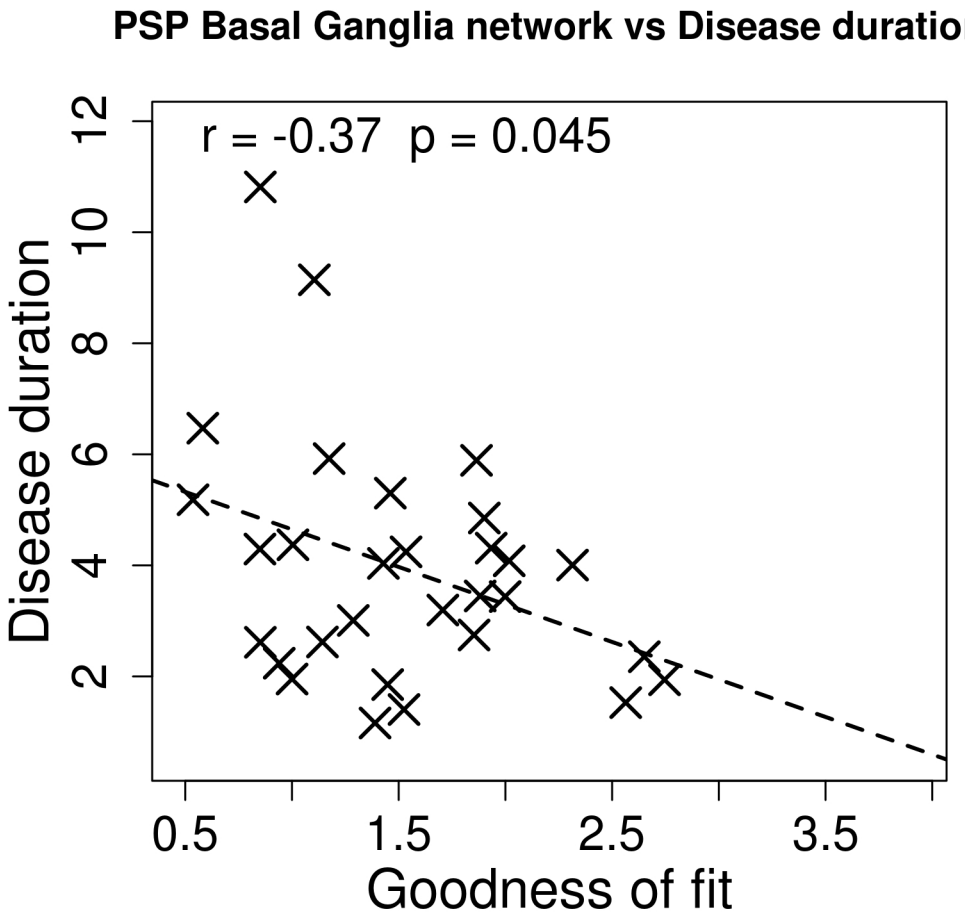
Prof Mike Greicius

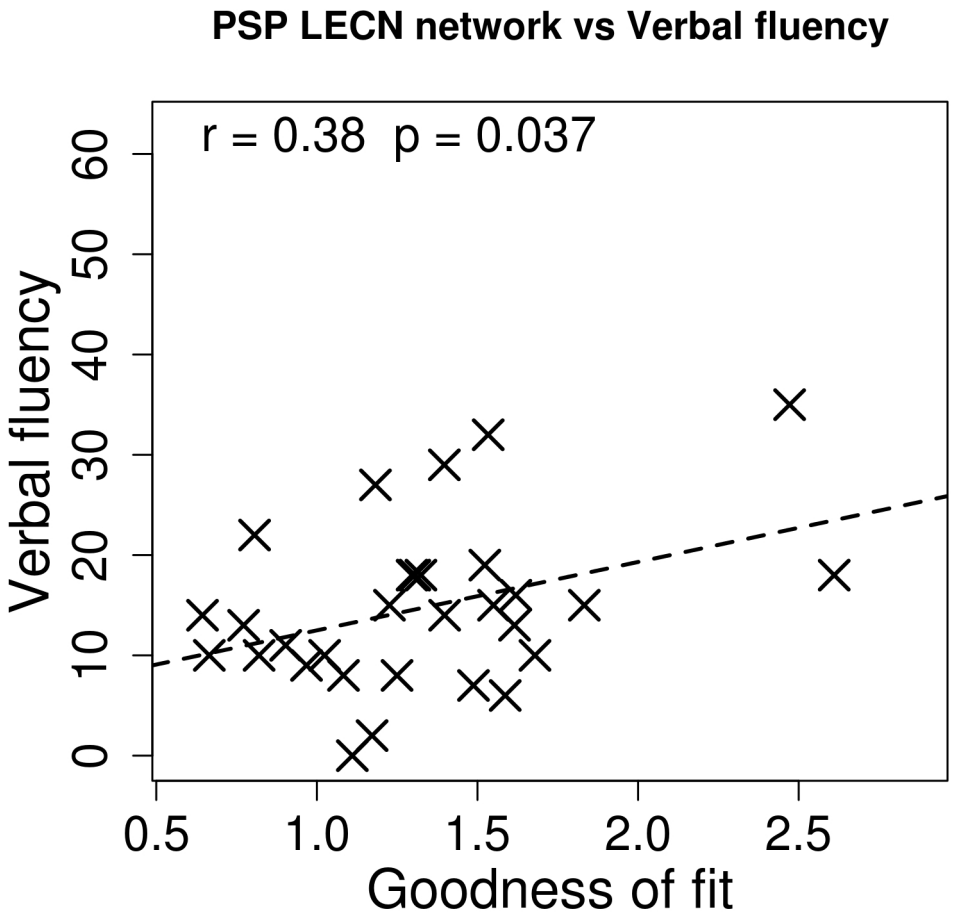
Greicius et al 2004, PNAS

Clinically relevant networks differ between patients and controls

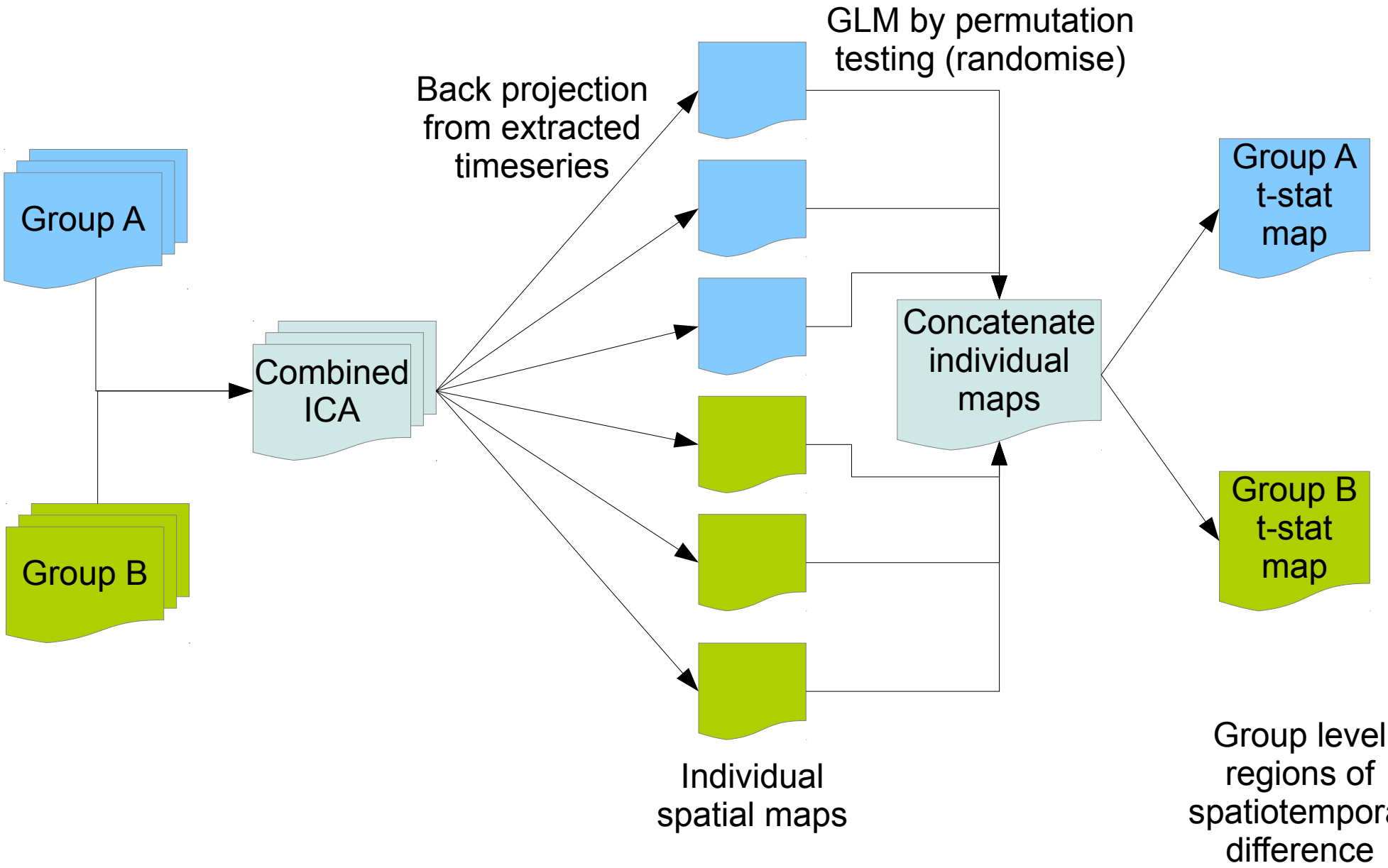


Correlation with clinical measures



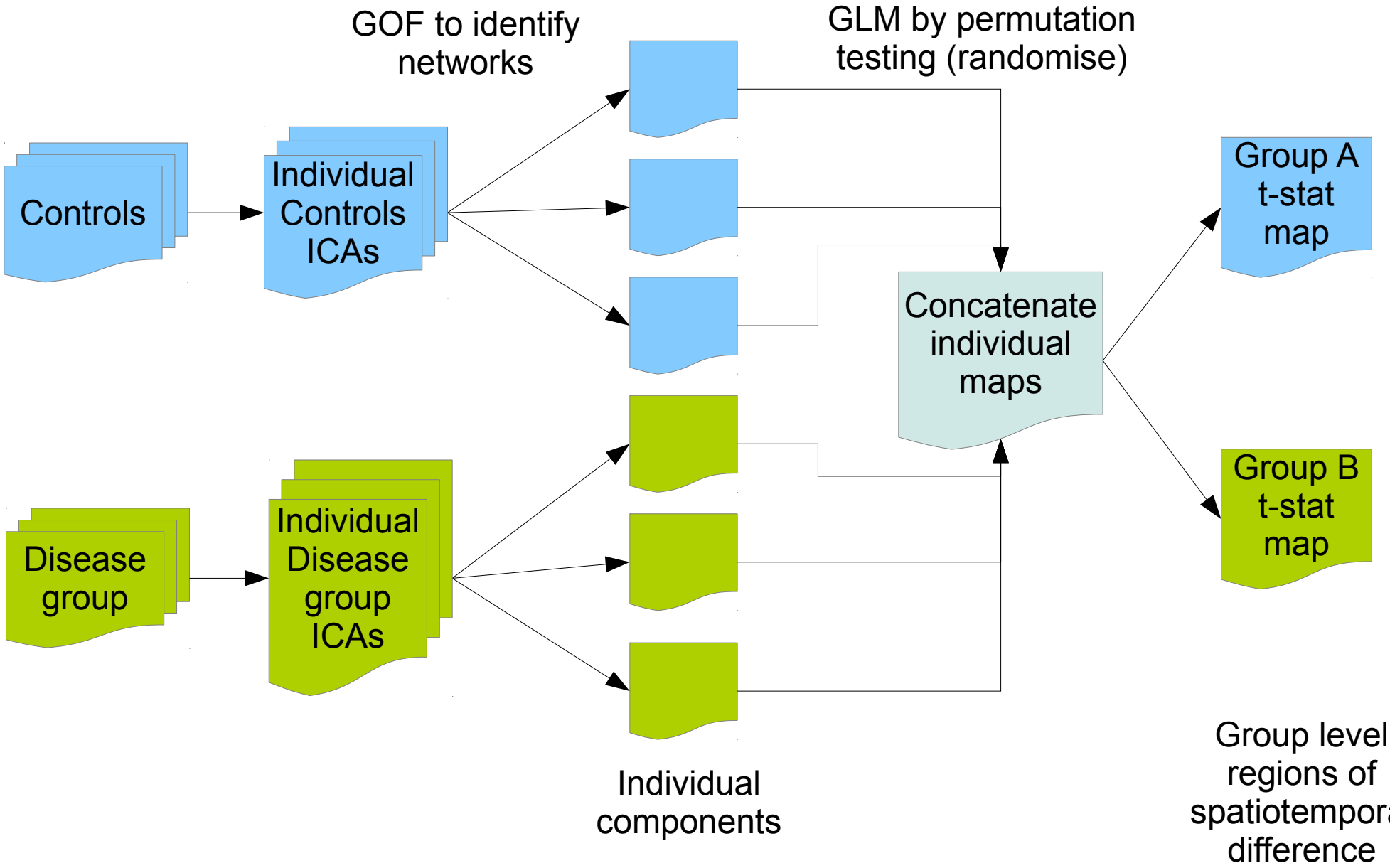


Dual regression



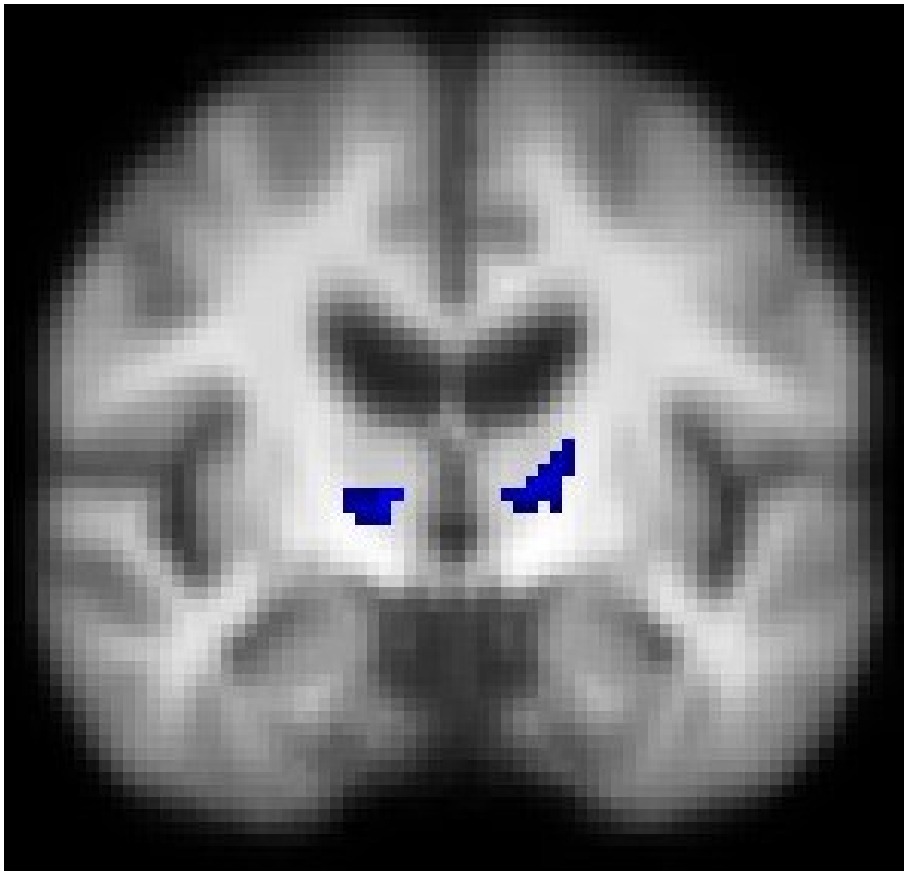
Beckmann 2009, Proceedings of the 15th Annual Meeting of Organization for Human Brain Mapping

Modified regression



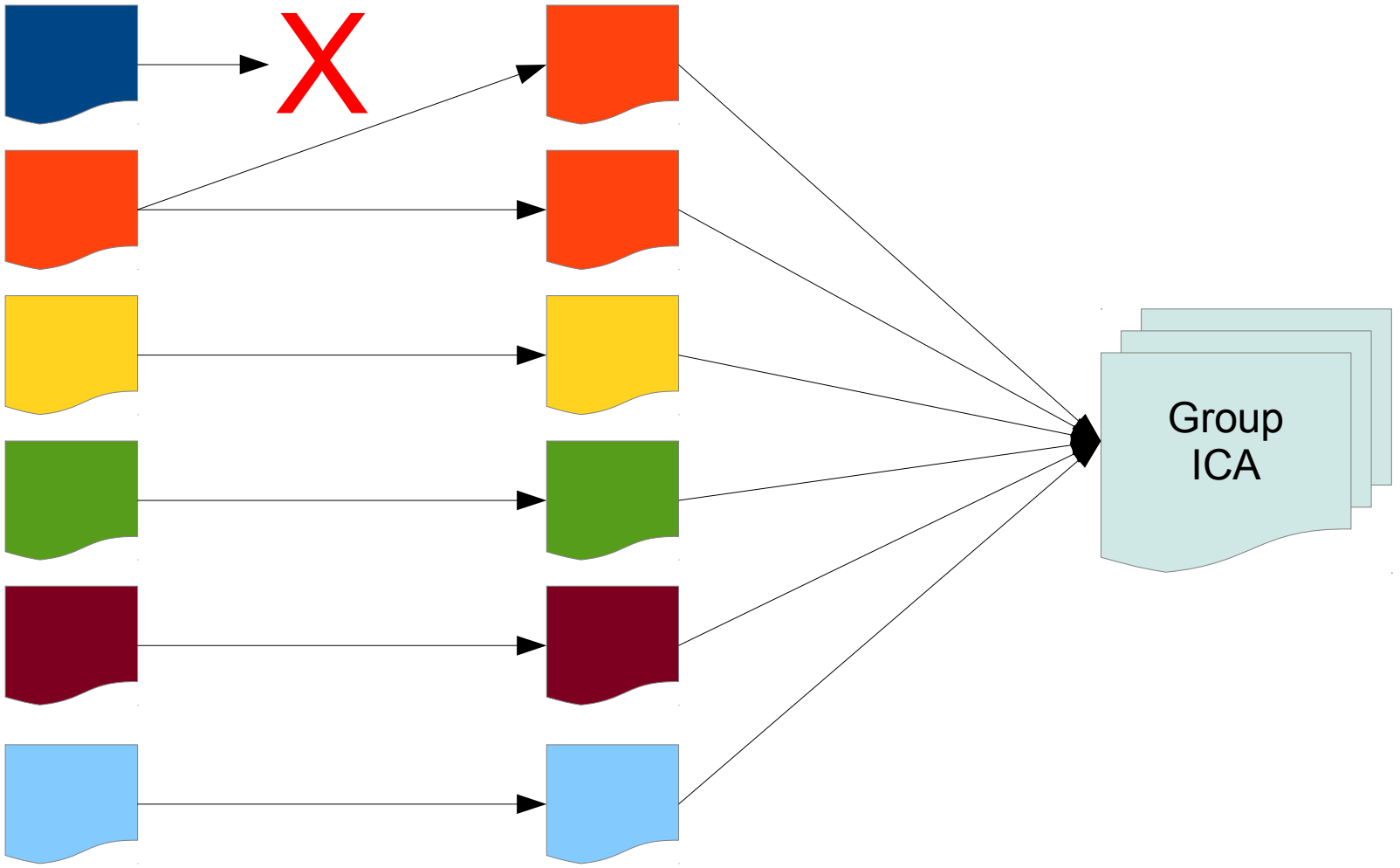
Beckmann 2009, Proceedings of the 15th Annual Meeting of Organization for Human Brain Mapping

PSP vs Controls, Basal Ganglia network



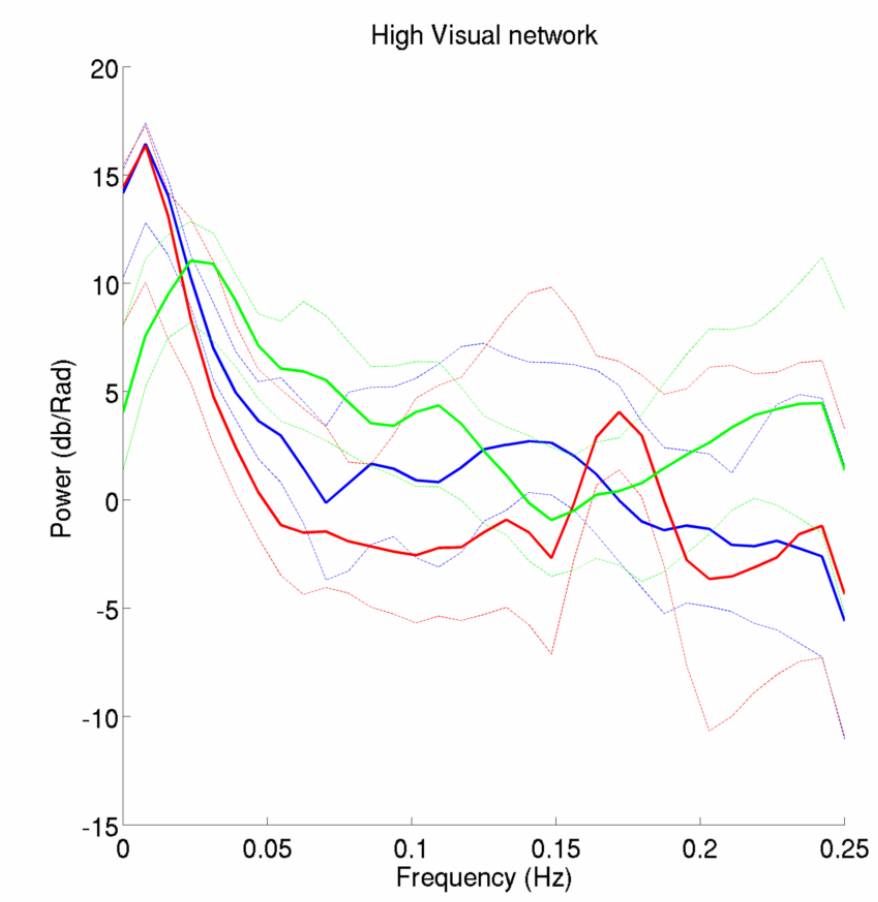
FWE p<0.05

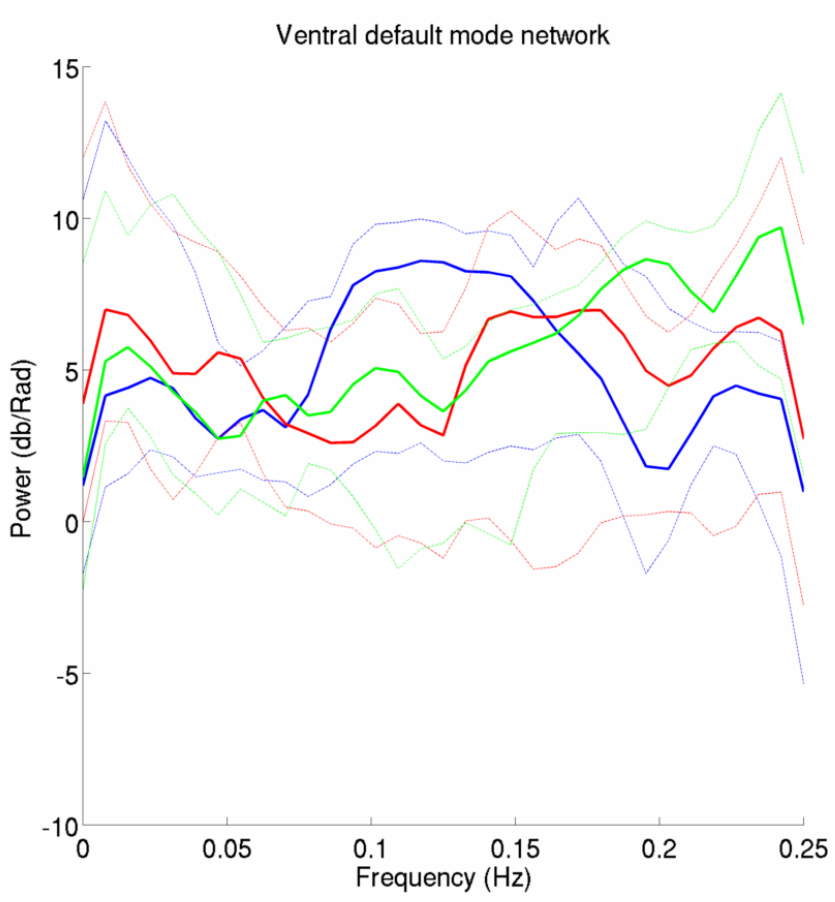
Robustness – within groups



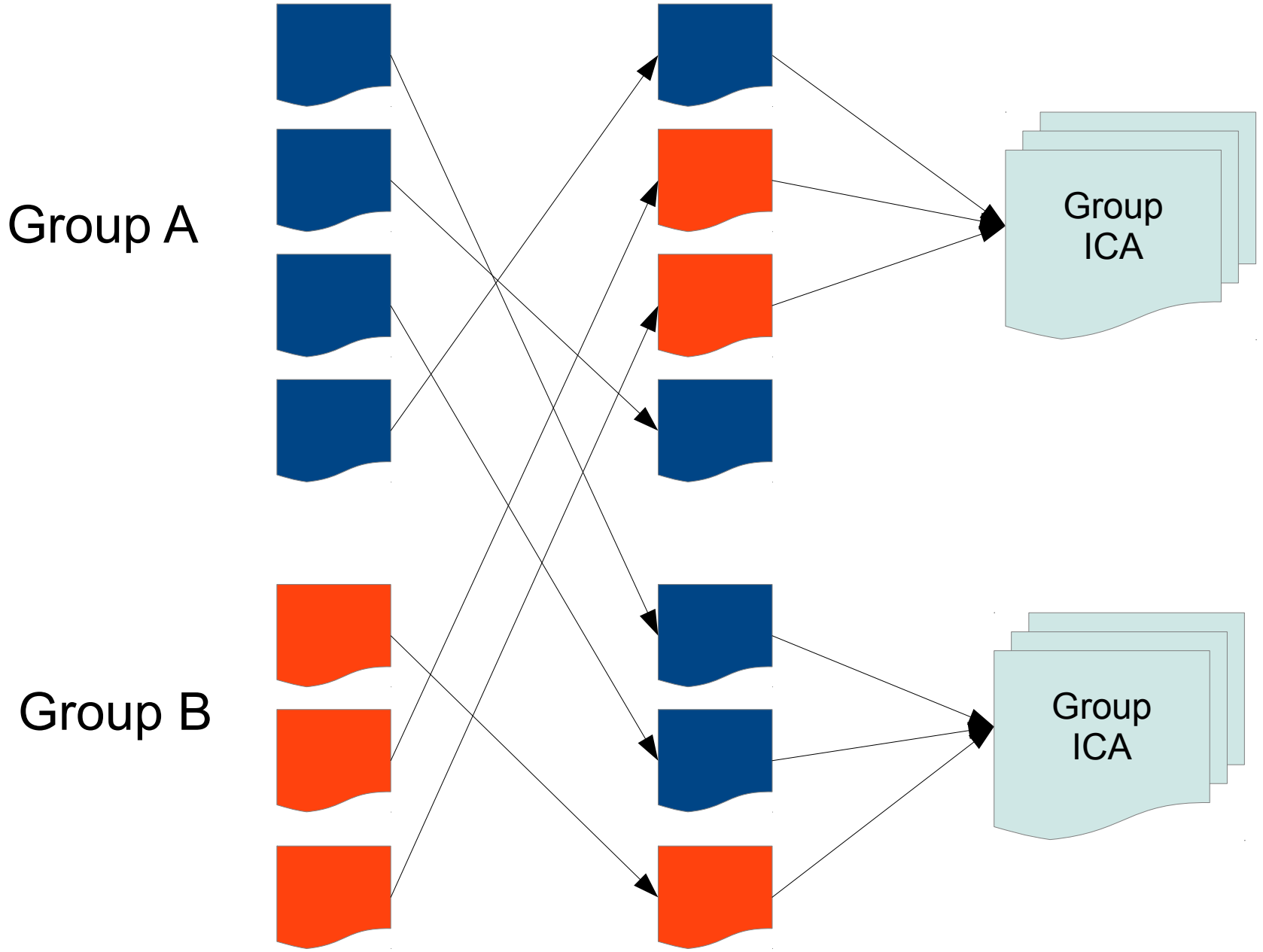
Robustness – within groups

PSP
CBD
Control



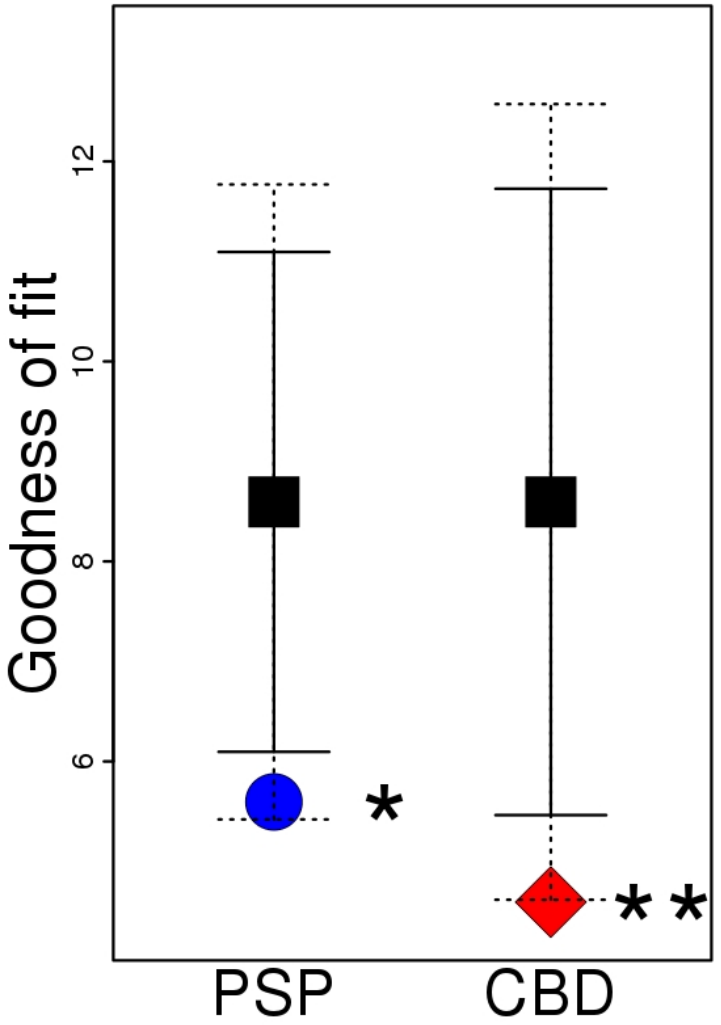


Robustness – across groups

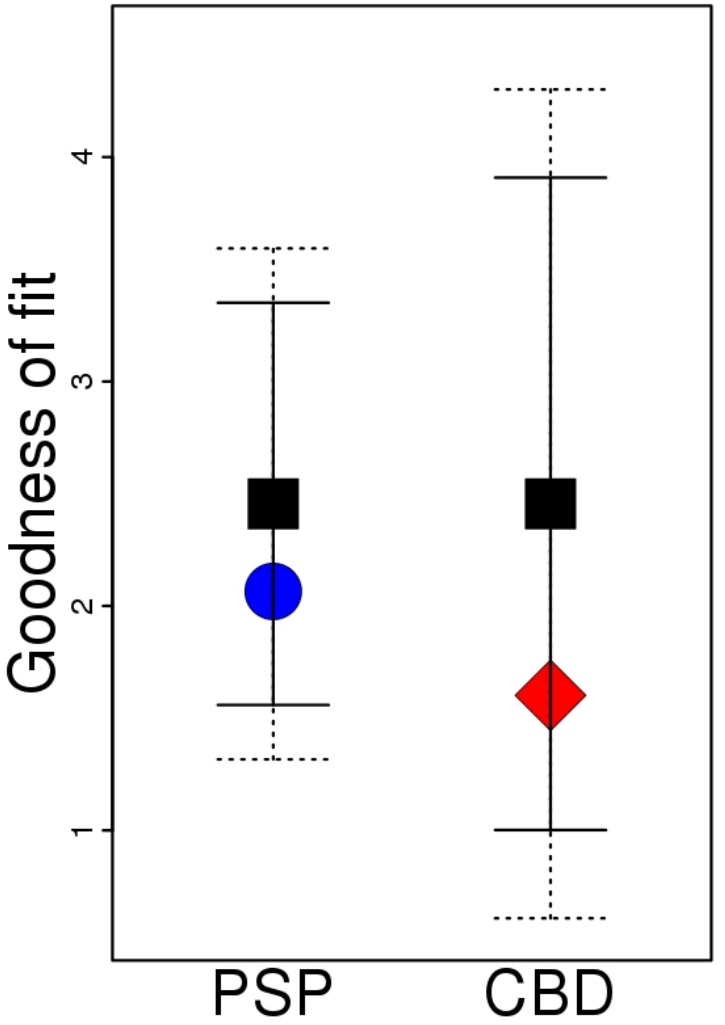


Robustness – across groups

Basal Ganglia network



Left Executive Control network



Conclusions

- Goodness of fit score
  - identify networks
  - compare groups
- Regression analysis
  - identify regions of spatiotemporal difference
- Robustness using bootstrapping
  - identify variance within data

task-related ICA

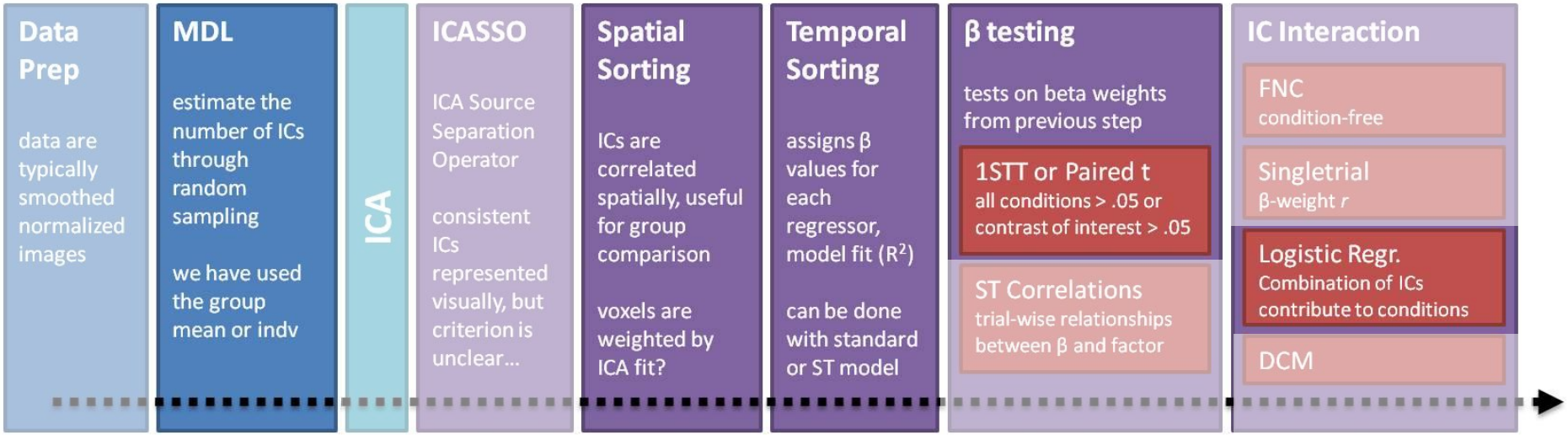
how many components do I use?

what threshold do I use to display my data?

how do I compare within/across groups?

selection before

selection after



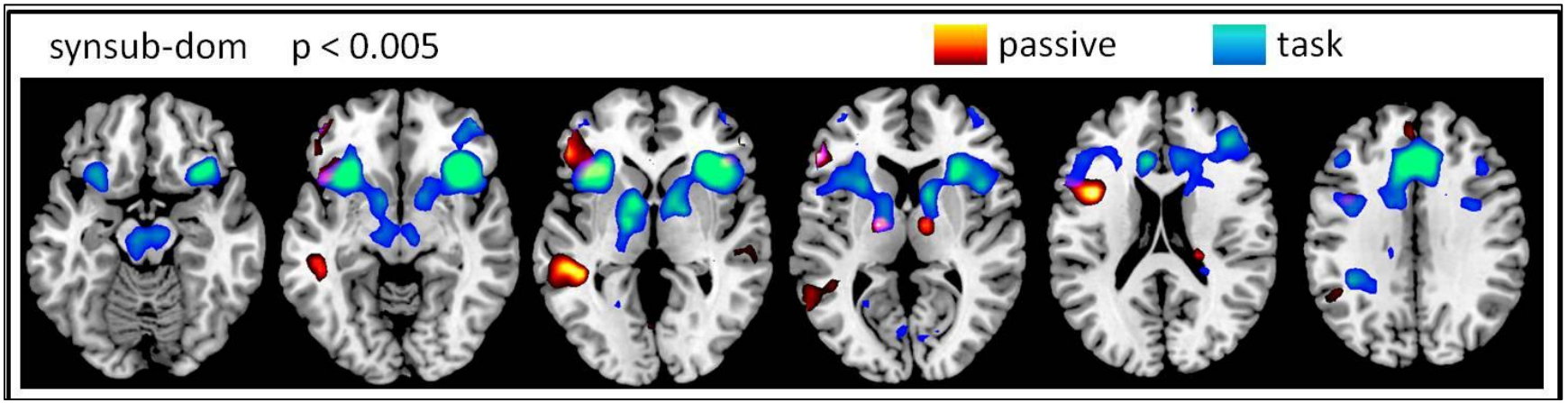
motivation

	syntax	Semantic	Ss
passive	"... boring colleagues were approaching ..."	"... wet palms swayed in the cool..."	12 YAs, 19 matures
task	"... boring colleagues **were**"	"...wet palms **swayed**"	12 YAs, 19 matures

subordinate: "... boring colleagues **was** damaging his career"

Subjects make acceptability judgments either during scanning (**task**) or in a post-scan questionnaire (**no-task**)

conventional multivariate SPM analysis



lack of LMTG involvement in the subordinate condition motivated an analysis which was able to discover hidden sources

how many components do I use?

what threshold do I use to display my data?

how do I compare within/across groups?

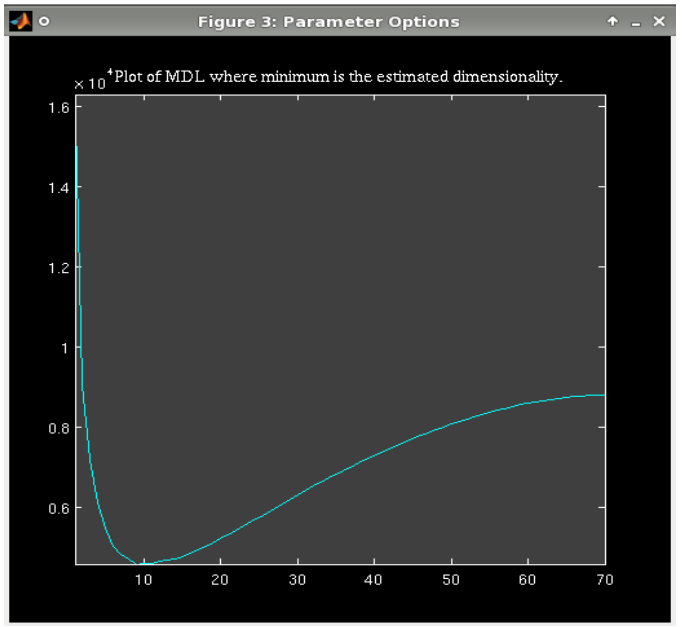
selection before

MDL – Minimum Description Length
used to estimate the likely number of sources

essentially a statistical instantiation of Occam's Razor

V is the number of voxels, **M** is the number of subjects
£ (Q ˆ N) is the log of the maximum likelihood estimate of the model parameters (estimated from the data, e.g., fMRI data)
ML is the number of time points following the first reduction
N is the number of sources.

$$MDL(N) = -V(ML - N)\mathcal{L}(\hat{\theta}_N) + \frac{1}{2}\left(1 + NL + \frac{1}{2}(N-1)\right)\ln V \quad (6)$$



alternatively, a single subject approach can be used to estimate both MDL and subsequent ICA

Spatial sorting is then used to compare or cluster individual components

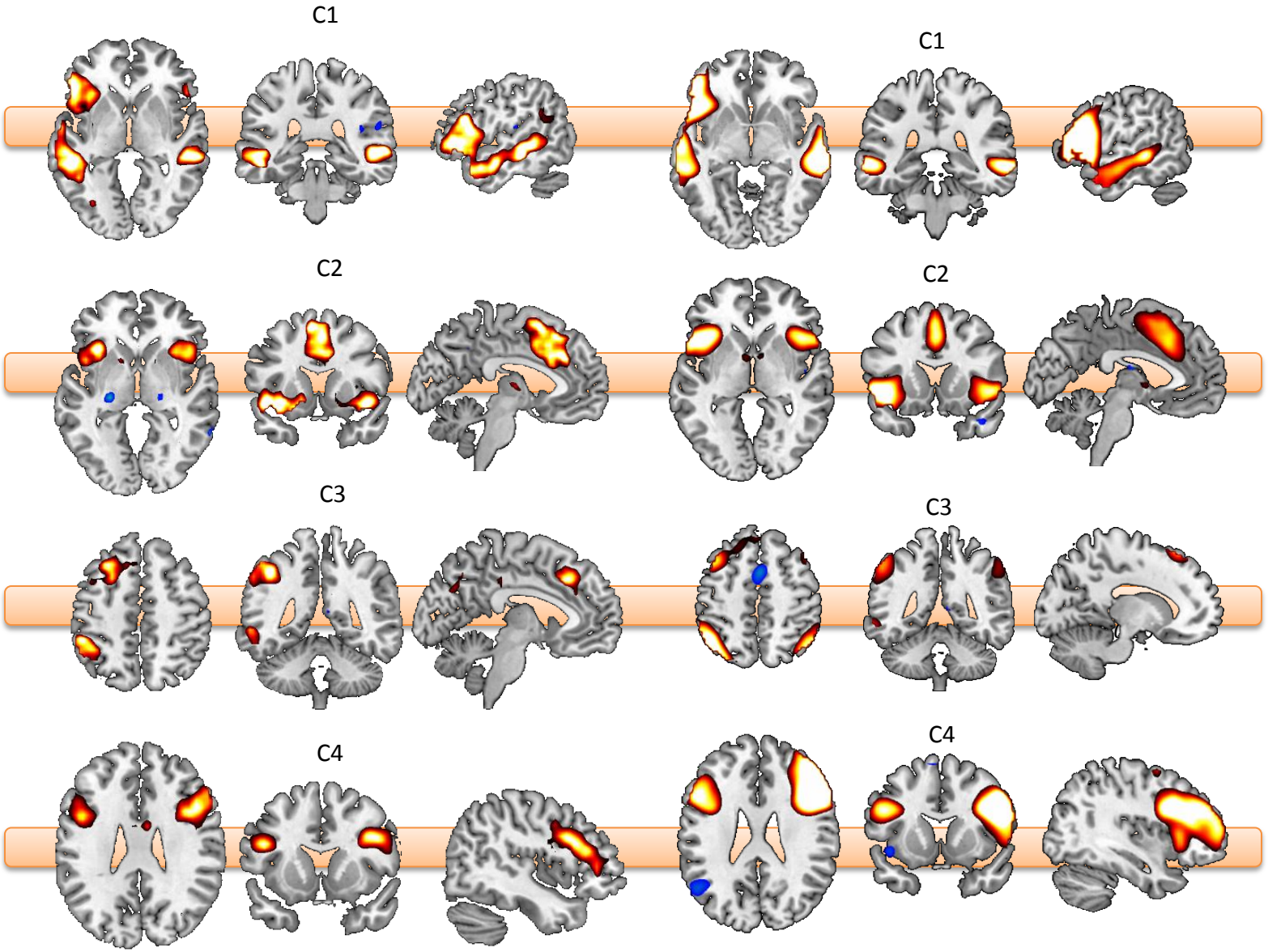
allows for unique spatial and temporal features

however, since fMRI data are noisy the components are not necessarily unmixed in the same way for each subject

spatial sorting

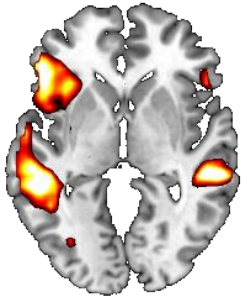
Younger adults

Mature adults



$p < 0.00001$

spatial sorting

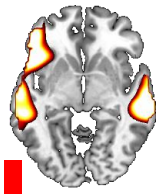


youngs

C1
C2
C3
C4
…
C_N

matures

C1
C2
C3
C4
…
C_N



Network	Overlap (Betas)
Language	.269
RFP	.210
OPER	.152
MFG	.207

OA components (arbitrary order)

selection after

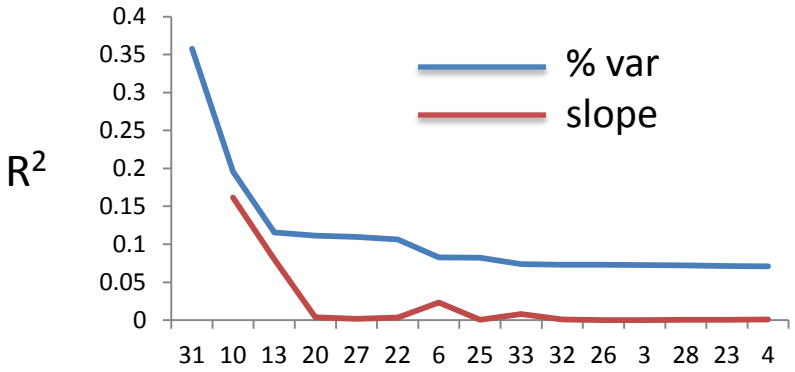
task-independent selection criteria

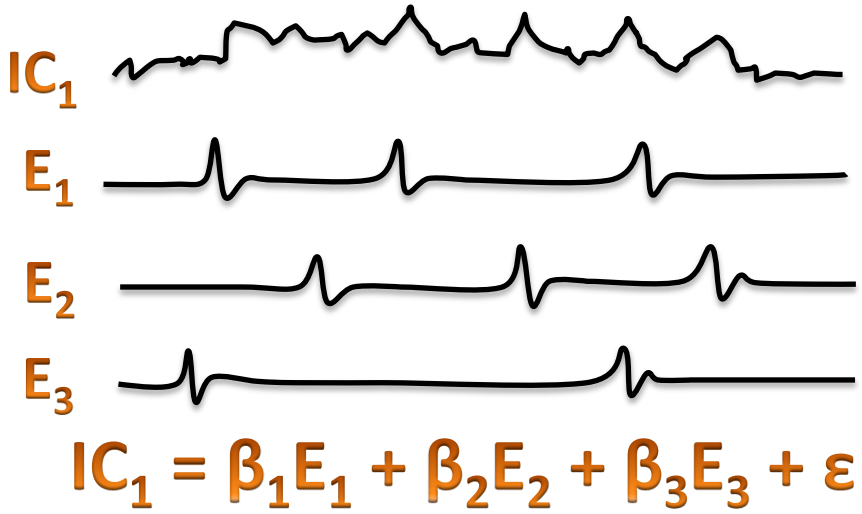
ICASSO (ICA Source Separation Operation) – reliability of extracted timecourses

ICA is run several times and components are clustered based on their absolute value of the correlation between the squared source estimates.

task-dependent selection criteria

"temporal sorting"

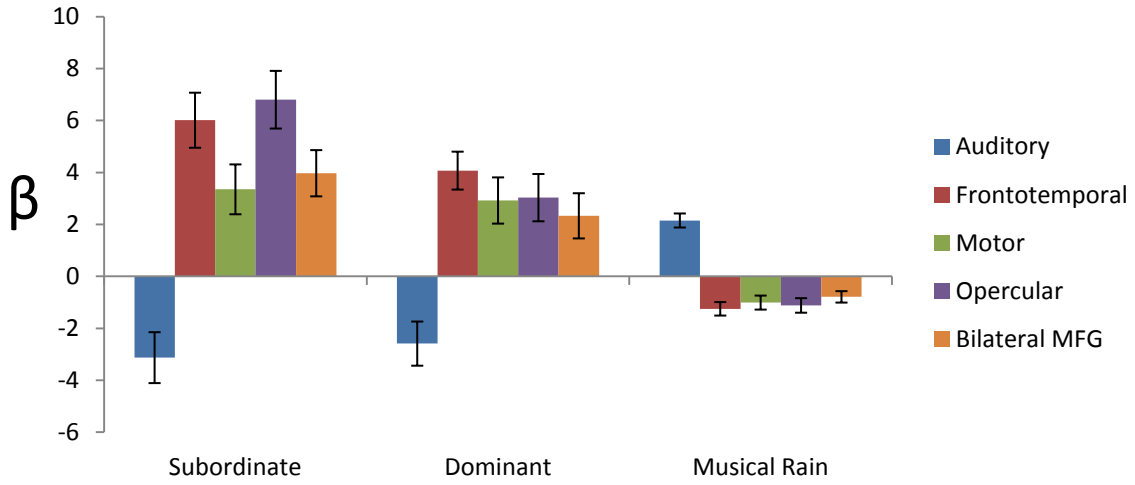




one-sample t-test

of the regression estimates (βs) associated with each trial type

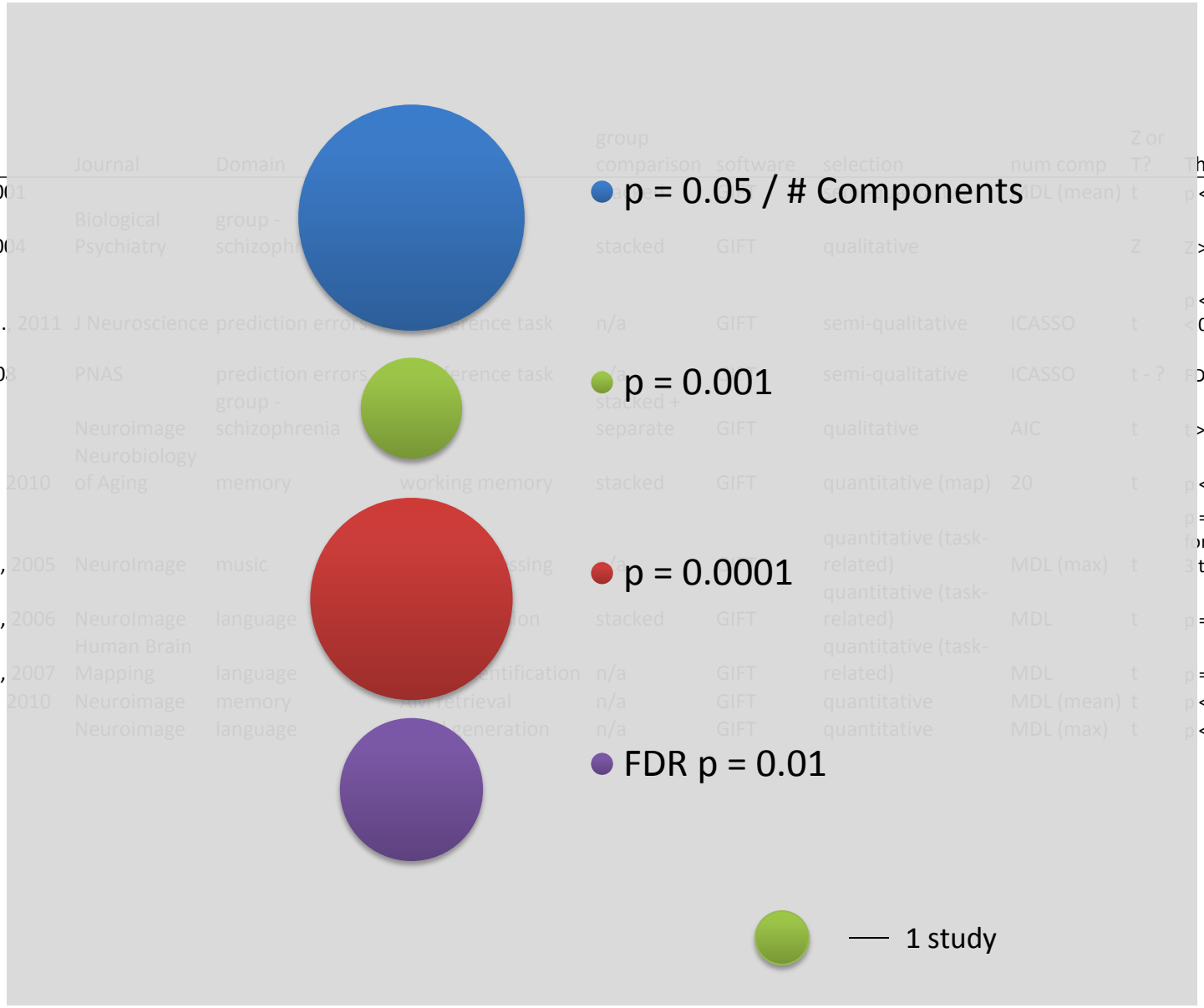
of the components themselves
→ input indv. subject component maps into SPM



how many components do I use?

what threshold do I use to display my data?

how do I compare within/across groups?

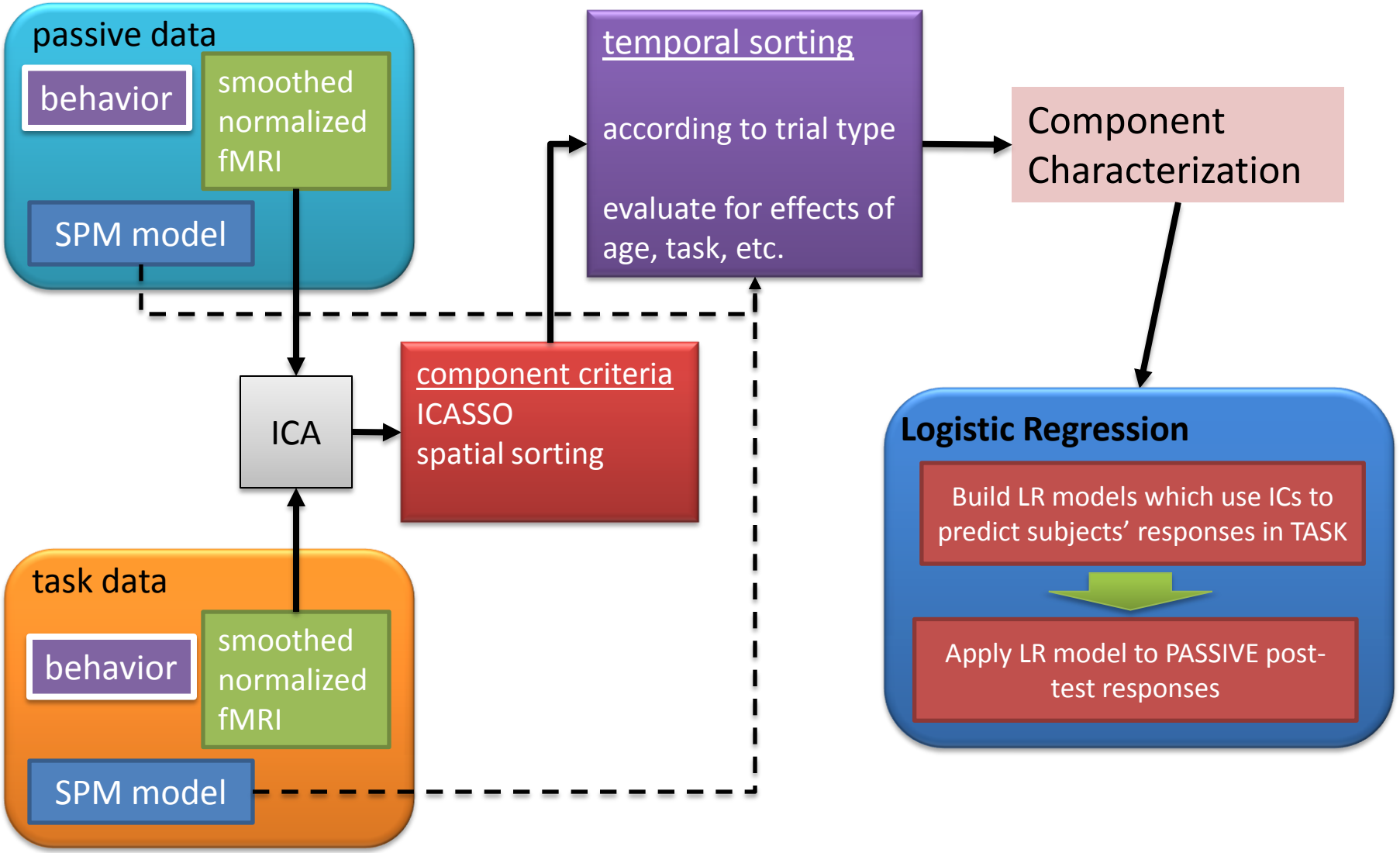


Paper	Journal	Domain	group comparison	software	selection	num comp	Z or T?	Threshold
Calhoun et al., 200						MDL (mean)	t	p < 0.0001
Calhoun et al., 200	Biological Psychiatry	group - schizophr		stacked	GIFT	qualitative	Z	Z > 3.1
Danielmeier et al.,	2011 J Neuroscience	prediction error	ference task	n/a	GIFT	semi-qualitative	ICASSO	t	p < 0.0001, 27 voxels, FDR 0.01
Eichele et al., 200	PNAS	prediction errors	ference task			semi-qualitative	ICASSO	t - ?	DR 0.01 / t = ~4
Jafri et al., 2008	Neuroimage	group - schizophrenia	stacked + separate	GIFT	qualitative	AIC	t	t > 5.6, n = 20
Sambataro et al.,	2010 Neurobiology of Aging	memory	working memory	stacked	GIFT	quantitative (map)	20	t	p < 0.001
Schmithorst et al,	2005 NeuroImage	music	ssing			quantitative (task-related)	MDL (max)	t	p = 0.05/150 (corrected for num components and t tests)
Schmithorst et al,	2006 NeuroImage	language	on	stacked	GIFT	quantitative (task-related)	MDL	t	p = 0.01/52
Schmithorst et al,	2007 Human Brain Mapping	language	ntification	n/a	GIFT	quantitative (task-related)	MDL	t	p = 0.01/52
St. Jacques et al.,	2010 Neuroimage	memory	AM retrieval	n/a	GIFT	quantitative	MDL (mean)	t	p < 0.0001
Tie et al., 2008	Neuroimage	language	eneration	n/a	GIFT	quantitative	MDL (max)	t	p < 0.001

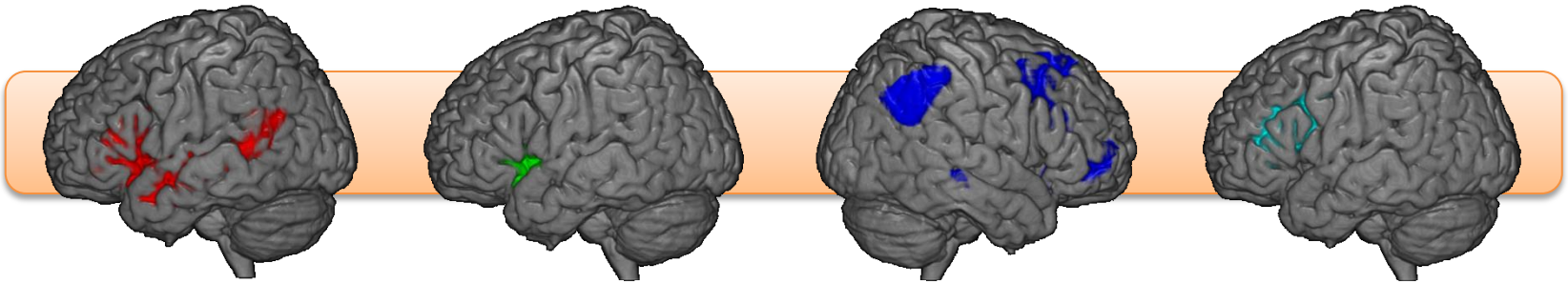
how many components do I use?

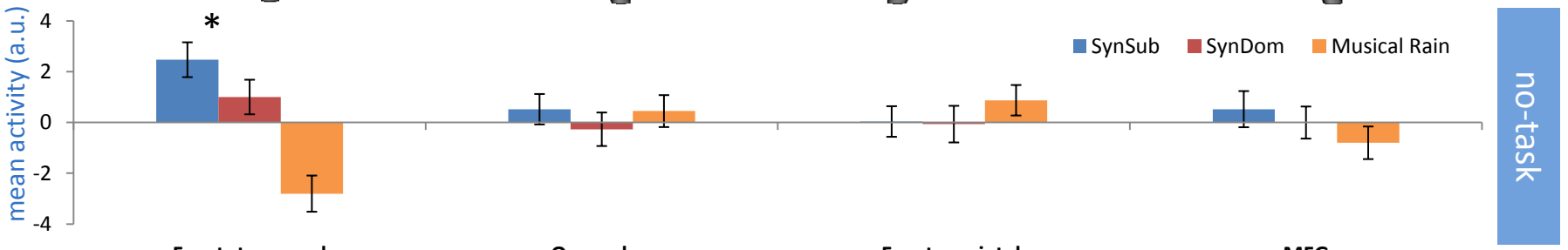
what threshold do I use to display my data?

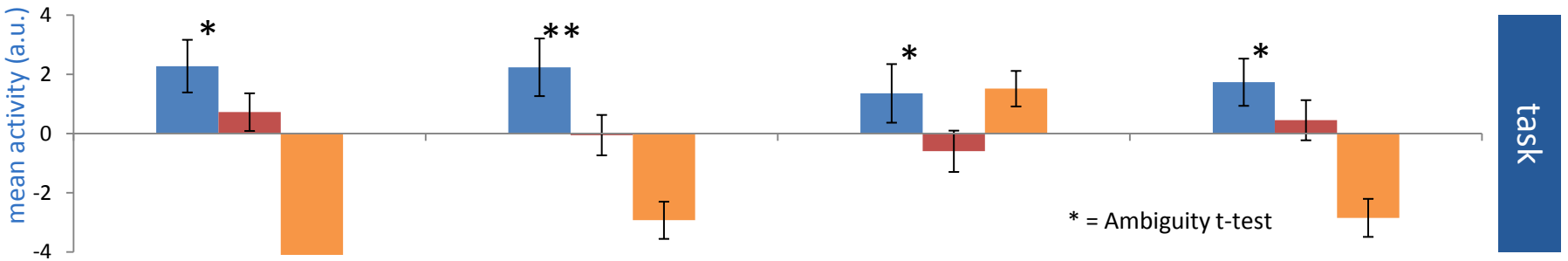
how do I compare within/across groups?

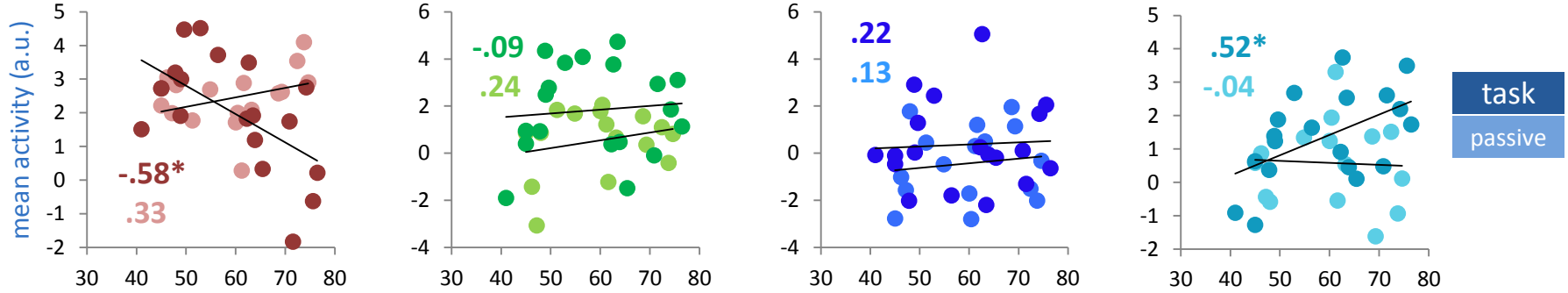

passive data
behavior
smoothed normalized fMRI
SPM model
temporal sorting
according to trial type
evaluate for effects of age, task, etc.
Component Characterization
component criteria
ICASSO
spatial sorting
ICA
Logistic Regression
Build LR models which use ICs to predict subjects' responses in TASK
Apply LR model to PASSIVE post-test responses
task data
behavior
smoothed normalized fMRI
SPM model

component characterization







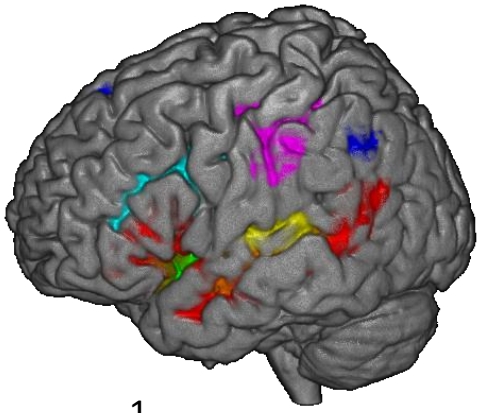


Using Logistic Regression to Predict Subjects Subsequent responses during Passive Listening

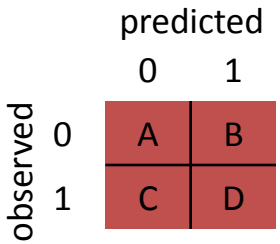
$$P(\text{event}) = \frac{1}{1 + e^{-z}} \qquad \text{where } z = \beta_0 + \beta_1 x\, \beta_1 + \beta_2 x \beta_2 + \ldots \beta_k x \beta_k$$

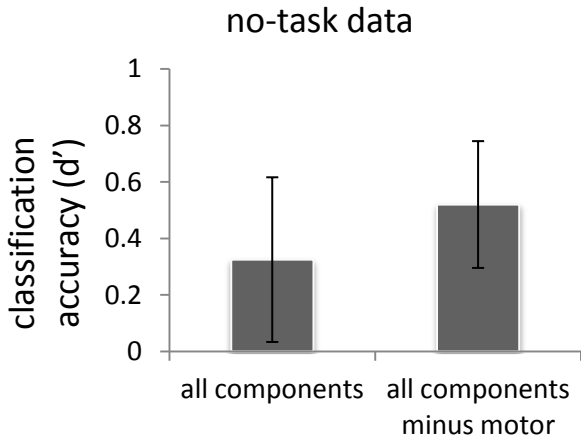
logistic regression is performed on trial-wise information about the loading of each component, with accuracy on each trial (0, 1) as the outcome variable.

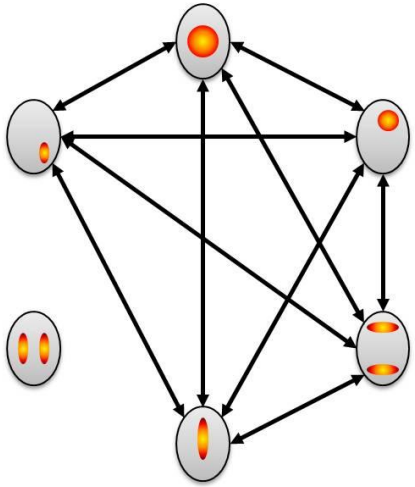
model	%class	χ^2	HosLem GOF	Term	Wald	β
0	52.1			Const	1.07	0.083
1	70.7	33.7	15.7	FRNTMP	**2.18**	0.039
				OPERC	0.47	0.006
				MFG	0.89	0.010
				FP	-1.73	-0.020
				AUD	-0.04	-0.001
				MOT	**-2.64**	-0.032
				Const	2.99	1.313

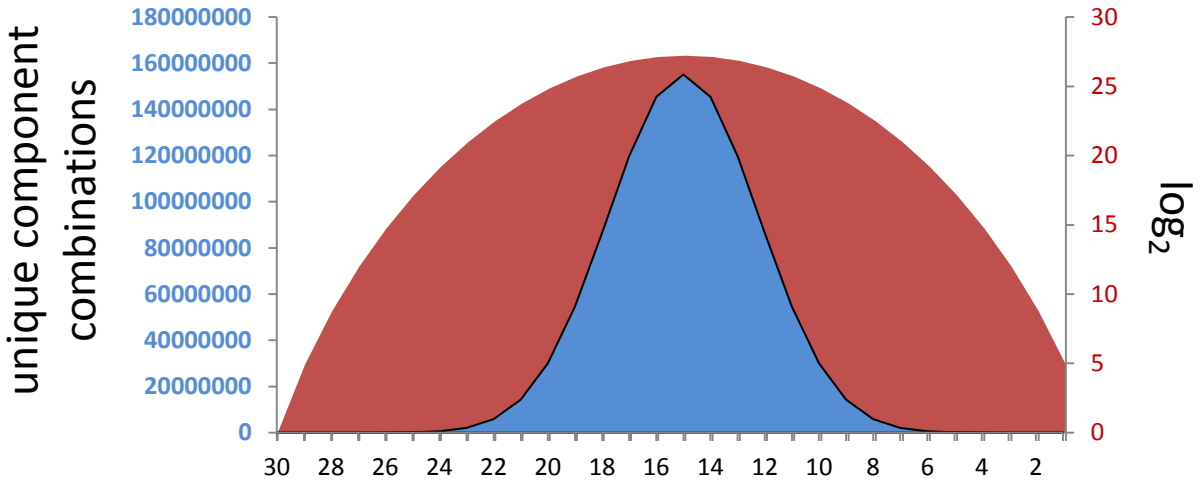


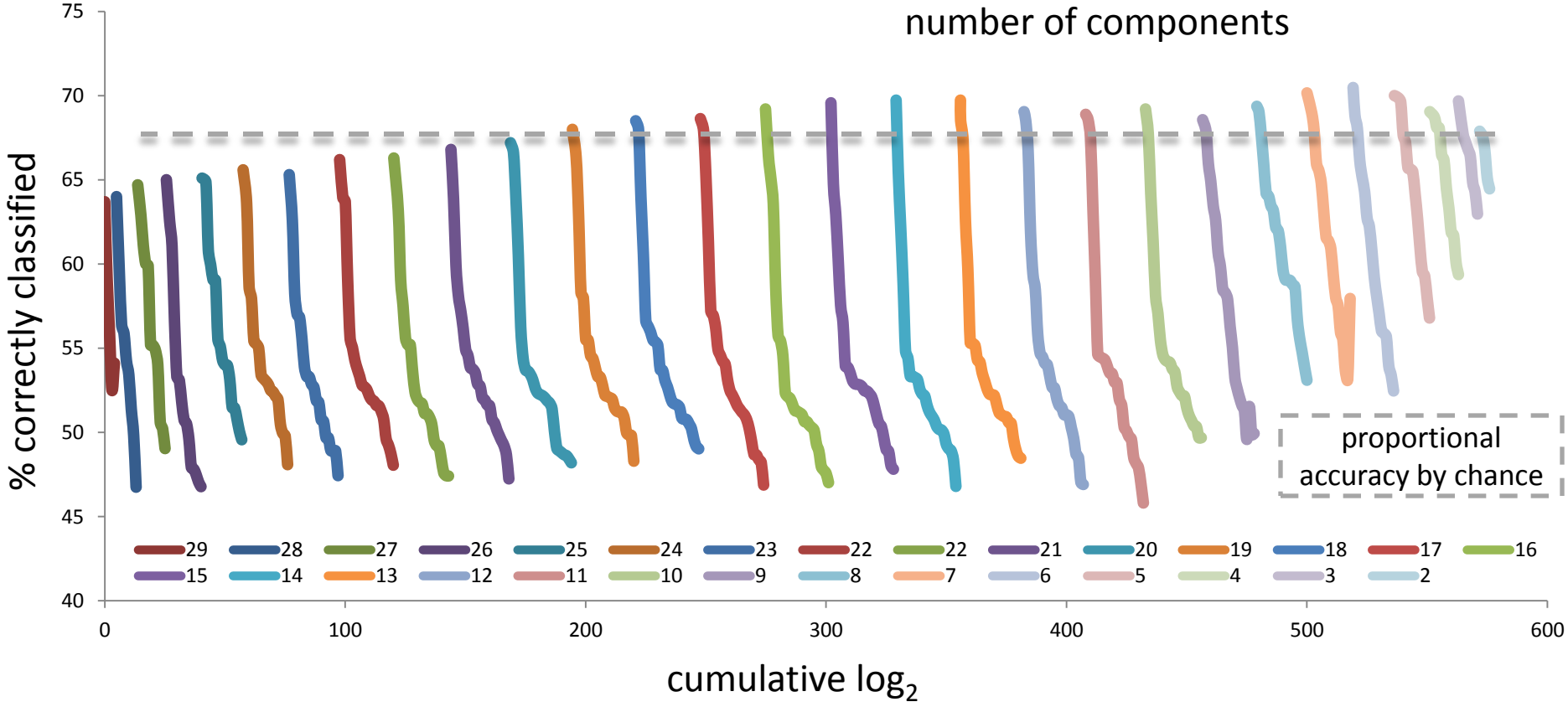
$$\frac{1}{1 + e^{-(.06 + \text{-.032x} + \text{.039x} + \text{.02x} + \text{-.006x} + \text{.001x} + \text{-.01x})}}$$

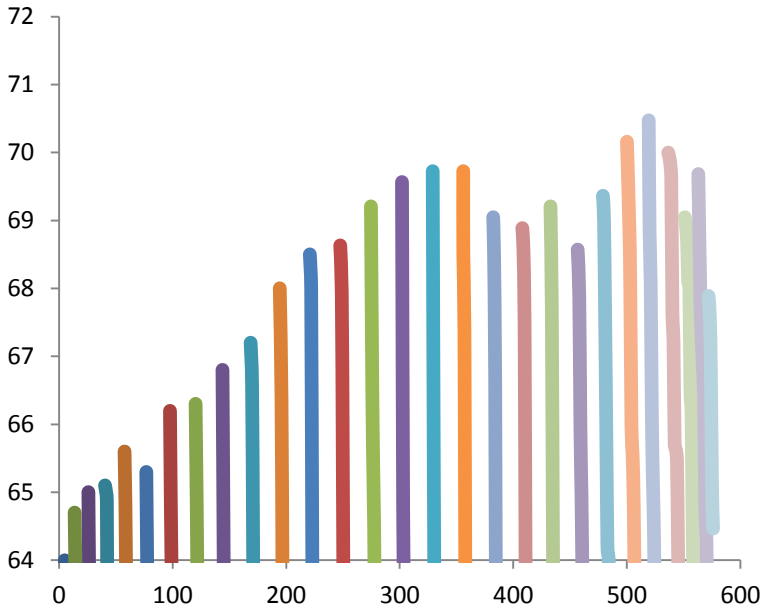




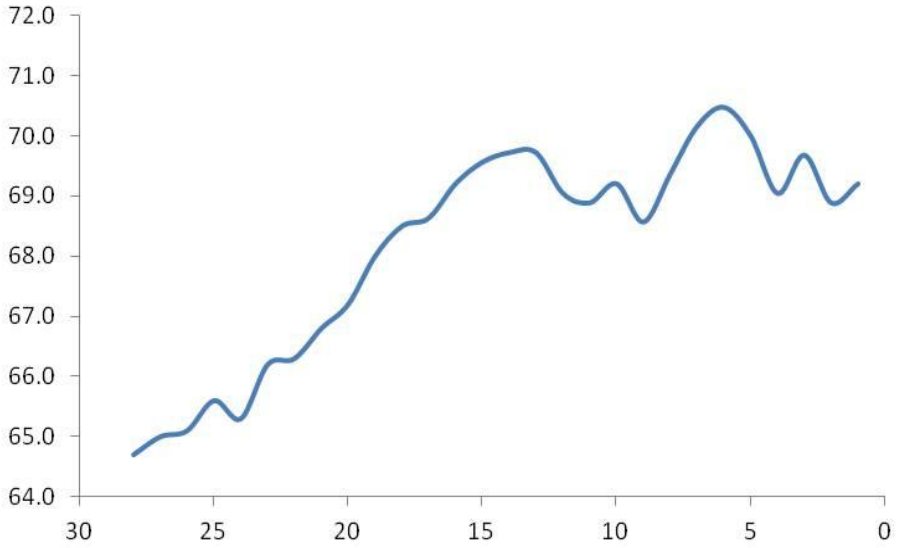



unique component combinations
180000000
160000000
140000000
120000000
100000000
80000000
60000000
40000000
20000000
0
30
25
20
15
10
5
0
log2
30 28 26 24 22 20 18 16 14 12 10 8 6 4 2
number of components

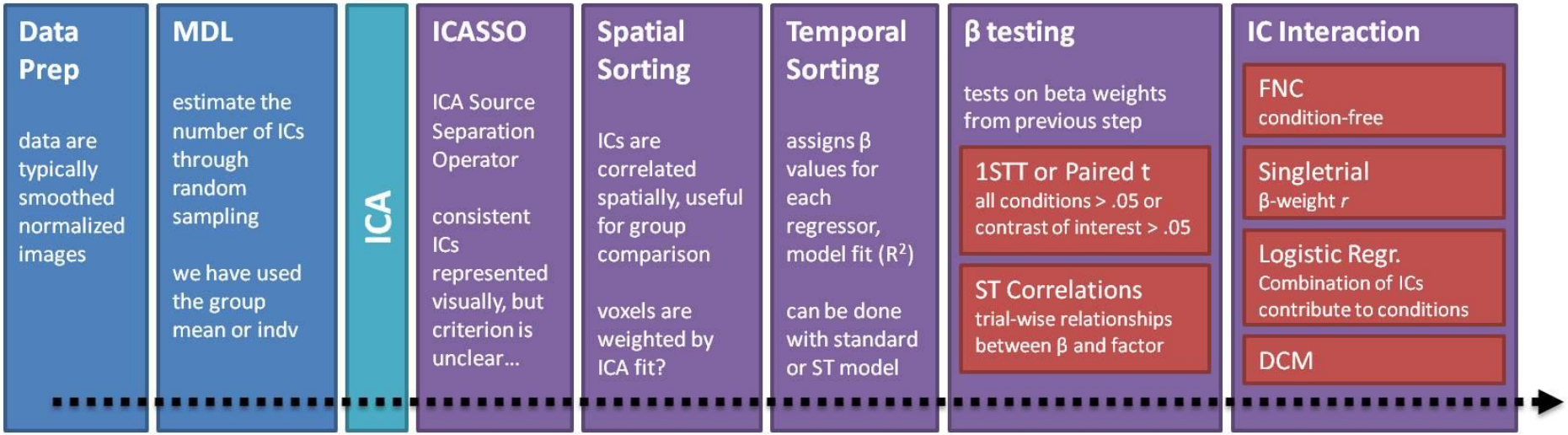

% correctly classified
75
70
65
60
55
50
45
40
0
100
200
300
400
500
600
cumulative log2
proportional accuracy by chance
29 28 27 26 25 24 23 22 22 21 20 19 18 17 16
15 14 13 12 11 10 9 8 7 6 5 4 3 2



Full model logistic



SPSS output –
Backwards Conditional Stepwise Logistic Regression


Data Prep
data are typically smoothed normalized images
MDL
estimate the number of ICs through random sampling
we have used the group mean or indv
ICA
ICASSO
ICA Source Separation Operator
consistent ICs represented visually, but criterion is unclear…
Spatial Sorting
ICs are correlated spatially, useful for group comparison
voxels are weighted by ICA fit?
Temporal Sorting
assigns β values for each regressor, model fit (R^2)
can be done with standard or ST model
β testing
tests on beta weights from previous step
1STT or Paired t
all conditions > .05 or contrast of interest > .05
ST Correlations
trial-wise relationships between β and factor
IC Interaction
FNC
condition-free
Singletrial
β-weight r
Logistic Regr.
Combination of ICs contribute to conditions
DCM

advantages

ICA is a data-driven approach, complementary to hypothesis-driven methods (e.g. GLM) for analyzing fMRI data

Finds reduced dimensionality descriptions of poorly understood, high dimensional spaces

Requires no a-priori knowledge about hemodynamics, noise models, time-courses of subject stimuli,…