A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic, and ERP data

- **Journal**: Neuroimage (2009)
- **Author**: Vince D. Calhoun, Jingyu Liu, and Tülay Adalı
- **Link**: https://www.ncbi.nlm.nih.gov/pubmed/19059344

Contents

- Abstract
- Introduction and background
- Theory and implementation
- Examples

Abstract

- **General linear model (GLM)** that requires the user to parameterize the data.
- **Independent component analysis (ICA)** is intrinsically a multivariate approach.
- ICA is a powerful and versatile data-driven approach for studying the brain.

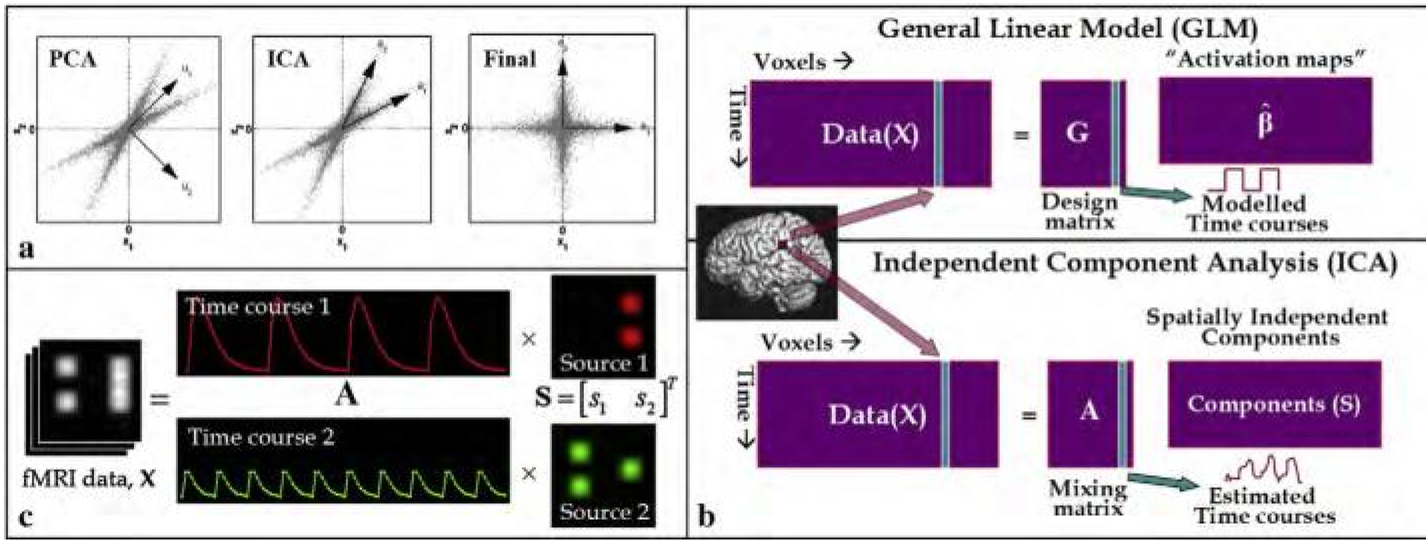
Introduction and background

- ICA of fMRI data
- Group ICA of fMRI data
- ICA for data fusion

- Uncorrelatedness is only partway to independence.
- A typical ICA model assumes that the source signals are not observable, statistically independent and non-Gaussian, with an unknown, but linear, mixing process.
- Since to achieve ICA, statistical information higher than second order is needed, it can either be generated using nonlinear functions or can be explicitly calculated.
- Suppose there are m observed signals and n sources, ICA model is expressed as $x(t) = As(t)$, where $x(t) = [x_1(t), x_2(t), \ldots, x_m(t)]^{\mathrm{T}}$, $s(t) = [s_1(t), s_2(t), \ldots, s_m(t)]^{\mathrm{T}}$, and A is called the **mixing matrix**.

- A number of algorithms derived based on:
    - maximum likelihood estimation
    - maximization of information transfer
    - mutual information minimization
    - maximization of non-Gaussianity
- Commonly used ICA algorithms:
    - **Infomax** (Bell and Sejnowski, 1995; Lee et al., 1999)
    - **FastICA** (Hyvarinen and Oja, 1997)
    - **Joint approximate diagonalization of eigenmatrices (JADE)** (Cardoso and Souloumiac, 1993)

- **Fig. 1**


PCA
ICA
Final
a
General Linear Model (GLM)
Voxels →
"Activation maps"
Time →
Data(X)
=
G
β̂
Design matrix
Modelled Time courses
Independent Component Analysis (ICA)
Spatially Independent Components
Voxels →
Time →
Data(X)
=
A
Components (S)
Mixing matrix
Estimated Time courses
b
Time course 1
Source 1
A
S = [s1 s2]T
fMRI data, X
Time course 2
Source 2
c

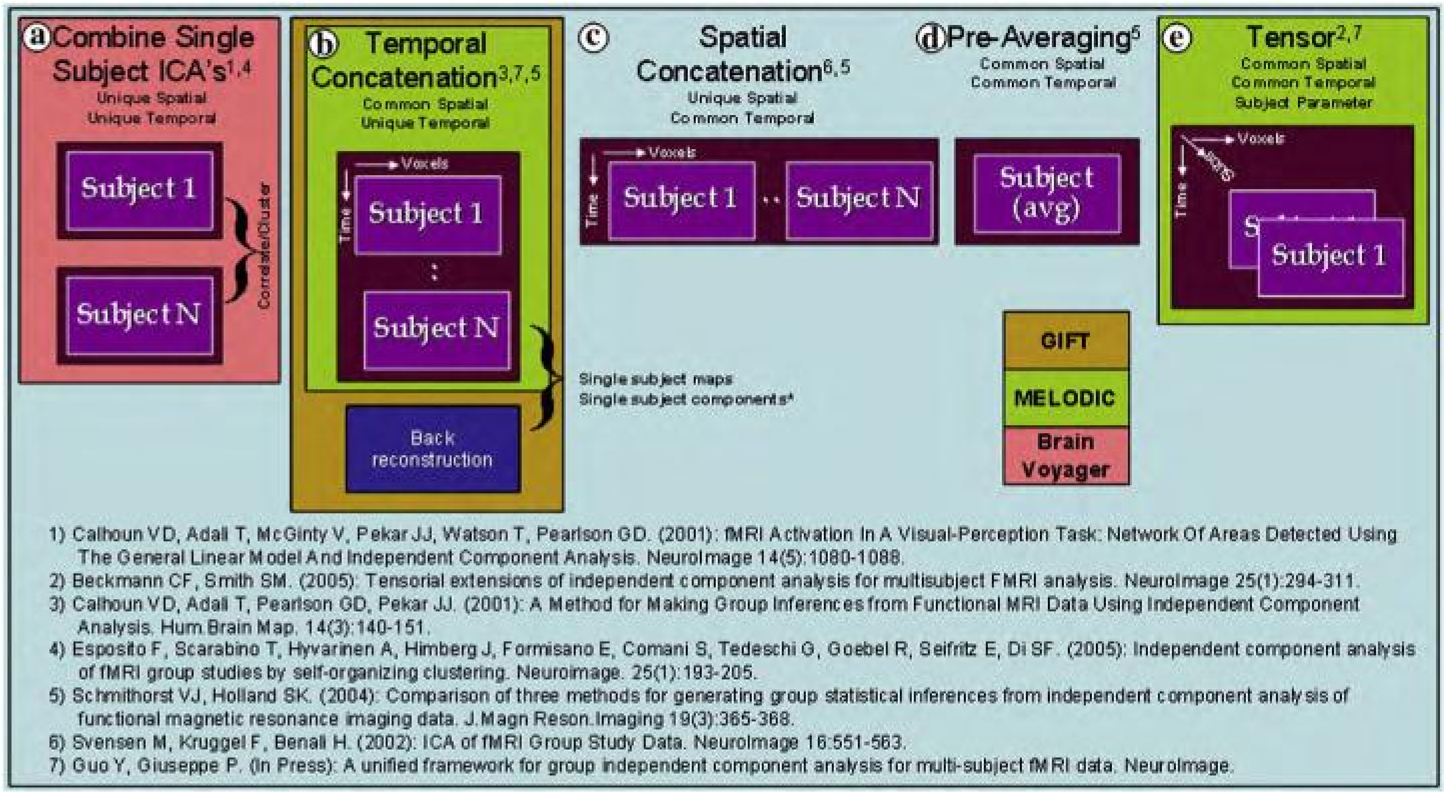
ICA of fMRI data

- **Spatial ICA** finds systematically non-overlapping, temporally coherent brain regions without constraining the shape of the temporal response.
- A strength of ICA is its ability to reveal dynamics for which a temporal model is not available.
- McKeown et al. argued that the sparse distributed nature of the spatial pattern for typical cognitive activation paradigms would work well with **spatial ICA (sICA)**.
- Infomax algorithm with a sparse prior is very well suited for spatial analysis and has also been used for **temporal ICA**.

Group ICA of fMRI data

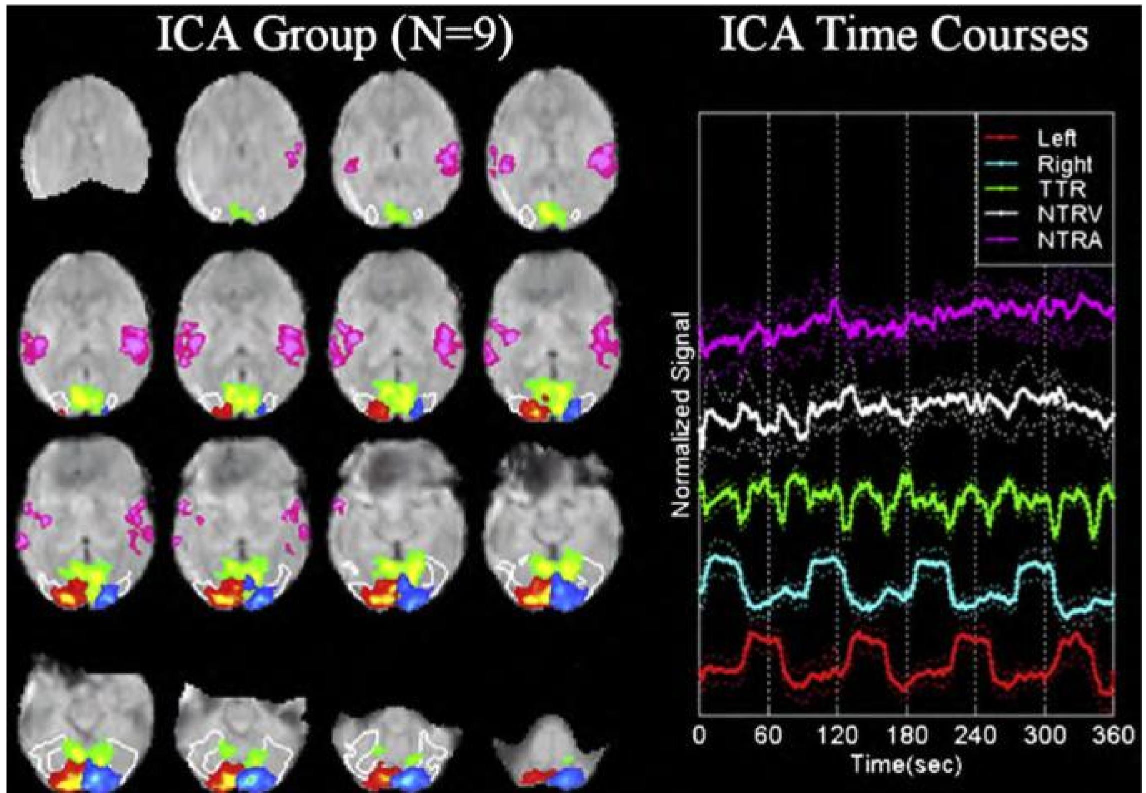
- Fig. 2a illustrates approaches that perform single-subject ICA and then attempt to combine the output into a group post hoc by using approaches such as self-organized clustering or spatial correlation of the components.
- The other four approaches involve an ICA computed on the group data directly. They perform one ICA, which can then be divided into subject specific parts, hence the comparison of subject differences within a component is straightforward.
- ICA with **temporal concatenation** plus **back-reconstruction** can capture variations in subject specific images.
- The tensorial approach in Fig. 2e involves estimating a common time course and a common image for each component but allows for a subject specific parameter to be estimated.

- **Fig. 2**


a Combine Single Subject ICA's1,4
Unique Spatial
Unique Temporal
Subject 1
Subject N
Correlate/Cluster
b Temporal Concatenation3,7,5
Common Spatial
Unique Temporal
Voxels
Time
Subject 1
Subject N
Back reconstruction
c Spatial Concatenation6,5
Unique Spatial
Common Temporal
Voxels
Time
Subject 1
Subject N
d Pre-Averaging5
Common Spatial
Common Temporal
Subject (avg)
e Tensor2,7
Common Spatial
Common Temporal
Subject Parameter
Voxels
Time
Subject 1
Single subject maps
Single subject components*
GIFT
MELODIC
Brain Voyager
1) Calhoun VD, Adali T, McGinty V, Pekar JJ, Watson T, Pearlson GD. (2001): fMRI Activation In A Visual-Perception Task: Network Of Areas Detected Using The General Linear Model And Independent Component Analysis. NeuroImage 14(5):1080-1088.
2) Beckmann CF, Smith SM. (2005): Tensorial extensions of independent component analysis for multisubject FMRI analysis. NeuroImage 25(1):294-311.
3) Calhoun VD, Adali T, Pearlson GD, Pekar JJ. (2001): A Method for Making Group Inferences from Functional MRI Data Using Independent Component Analysis. Hum.Brain Map. 14(3):140-151.
4) Esposito F, Scarabino T, Hyvarinen A, Himberg J, Formisano E, Comani S, Tedeschi G, Goebel R, Seifritz E, Di SF. (2005): Independent component analysis of fMRI group studies by self-organizing clustering. Neuroimage. 25(1):193-205.
5) Schmithorst VJ, Holland SK. (2004): Comparison of three methods for generating group statistical inferences from independent component analysis of functional magnetic resonance imaging data. J.Magn Reson.Imaging 19(3):365-368.
6) Svensen M, Kruggel F, Benali H. (2002): ICA of fMRI Group Study Data. NeuroImage 16:551-563.
7) Guo Y, Giuseppe P. (In Press): A unified framework for group independent component analysis for multi-subject fMRI data. NeuroImage.

- Separate components for primary visual areas on the left and the right visual cortex were consistently task-related with respect to the appropriate stimulus.
- Group inference or comparison of groups can be performed by performing statistics on either the ICA images or the time courses.

- **Fig. 3**


ICA Group (N=9)
ICA Time Courses
Left
Right
TTR
NTRV
NTRA
Normalized Signal
0
60
120
180
240
300
360
Time(sec)

ICA for data fusion

- One promising data fusion approach is to first process each image type and extract features from different modalities.
- An advantage of ICA over variance-based approaches like SVD or PCA is the use of higher-order statistics to reveal hidden structure.

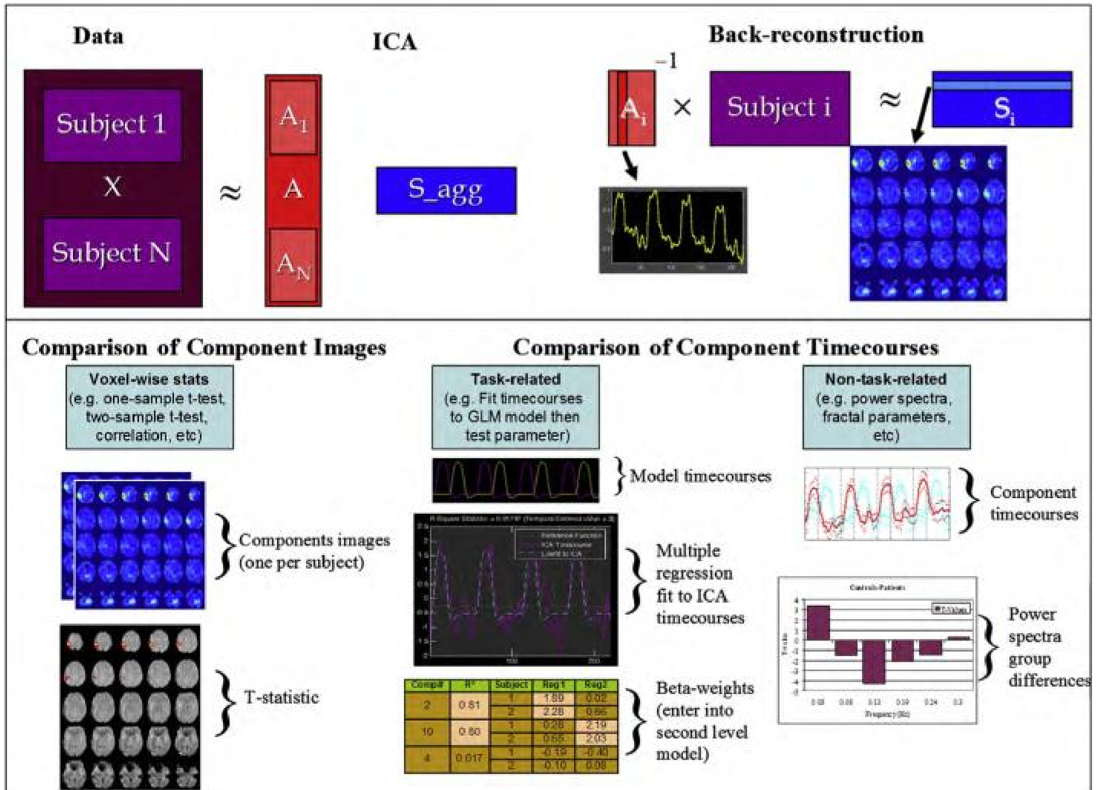
Theory and implementation

- Group ICA of fMRI
- Joint ICA
- Parallel ICA

Group ICA of fMRI

- The group ICA approach implemented in GIFT incorporates temporal concatenation plus back-reconstruction.
- Once the mixing matrix is estimated, the component maps for each subject can be computed by projecting the single subject data onto the inverse of the partition of the mixing matrix that corresponds to that subject.
- Group inferences can be made by analyzing the subject specific time courses and spatial maps.
- The time courses can be analyzed by fitting to a GLM, but instead of fitting to the voxel-wise data the ICA time courses are the dependent variable.

- **Fig. 4**


Data
ICA
Back-reconstruction
Subject 1
X
Subject N
≈
A1
A
AN
S_agg
Ai
−1
×
Subject i
≈
Si
Comparison of Component Images
Comparison of Component Timecourses
Voxel-wise stats
(e.g. one-sample t-test, two-sample t-test, correlation, etc)
Task-related
(e.g. Fit timecourses to GLM model then test parameter)
Non-task-related
(e.g. power spectra, fractal parameters, etc)
Model timecourses
Component timecourses
Components images
(one per subject)
Multiple regression fit to ICA timecourses
Power spectra group differences
T-statistic
Beta-weights
(enter into second level model)

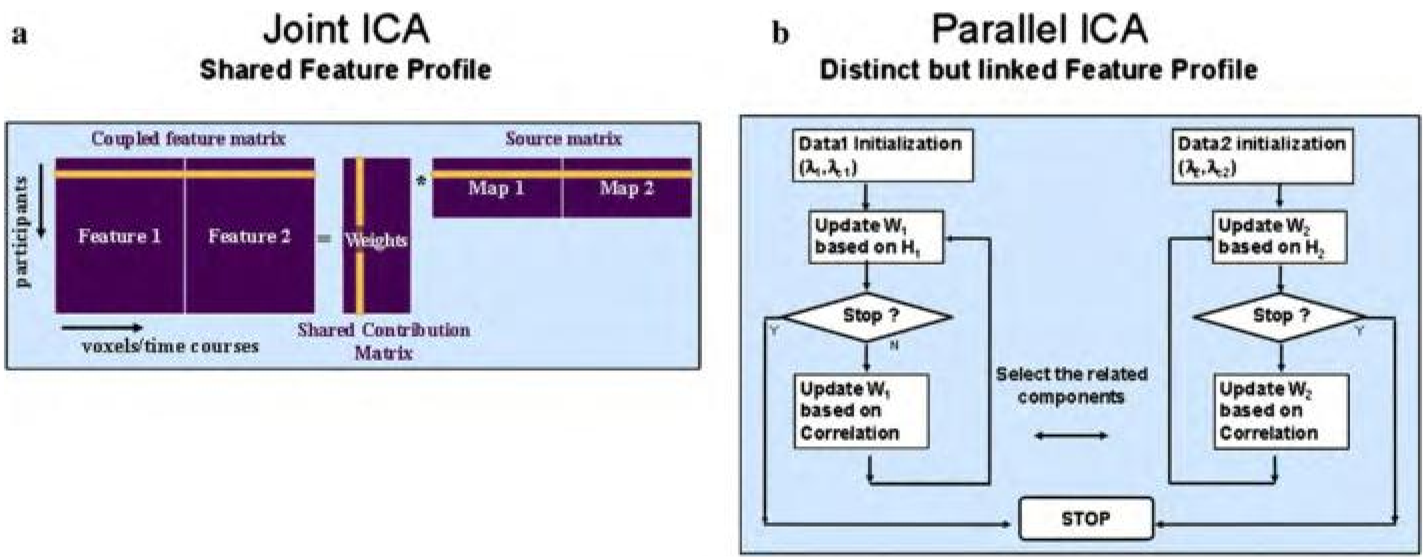
Joint ICA

- **Joint ICA (jICA)** is an approach that enables us to jointly analyze multiple modalities which have all been collected in the same set of subjects.
- When posed as a maximum likelihood problem, we estimate a **joint unmixing matrix** W such that the likelihood $L(W)$ is maximized.
- This formulation characterizes the basic jICA approach and assumes that the sources associated with the two data types (F and G) modulate the same way across N subjects (see Fig. 5a).
- We can require that the form of modulation across samples for the sources from two data types to be correlated but not necessarily the same.

Parallel ICA

- **Parallel ICA (paraICA)** identifies components of both modalities and connections between them through enhancing intrinsic interrelationships (see Fig. 5b).
- The relationship between two data types is calculated as the correlation between the columns of mixing matrices A.
- Thus, we have the correlation term and the maximization function based upon entropy.
- Two demixing matrices W are updated separately during which the component with highest correlation from each modality is selected and used to modify the update of the demixing matrix based on the correlation value using appropriate stopping criteria.

- **Fig. 5**


a
Joint ICA
Shared Feature Profile
Coupled feature matrix
Source matrix
participants
Feature 1
Feature 2
=
Weights
*
Map 1
Map 2
voxels/time courses
Shared Contribution Matrix
b
Parallel ICA
Distinct but linked Feature Profile
Data1 Initialization
Data2 initialization
Update W1 based on H1
Update W2 based on H2
Stop ?
Stop ?
Y
N
Y
Update W1 based on Correlation
Update W2 based on Correlation
Select the related components
STOP

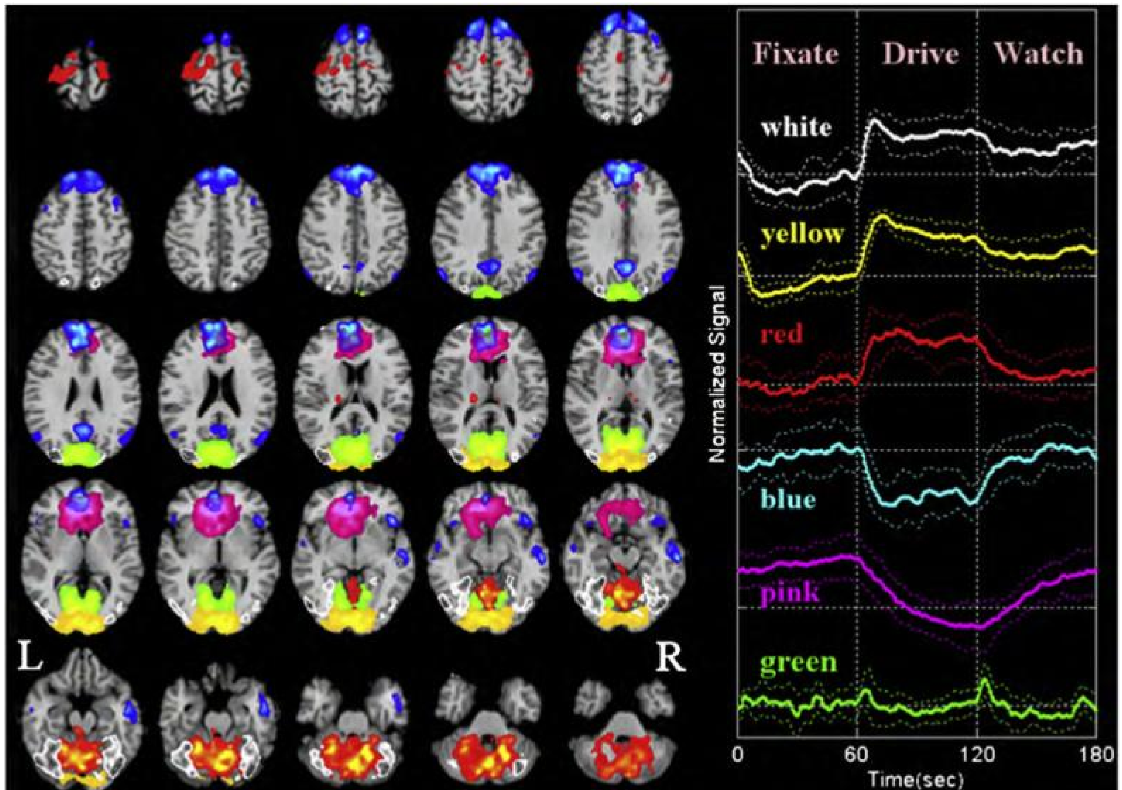
Examples

- Example 1
- Example 2
- Example 3
- Example 4

Example 1

- fMRI data from 15 subjects were collected during a 10 min paradigm with alternating 1 min blocks of fixation, simulated driving, and watching.
- ICA time courses were first analyzed to evaluate task-relatedness. Six components were identified and entered into a voxelwise one-sample t-test.

- **Fig. 6**

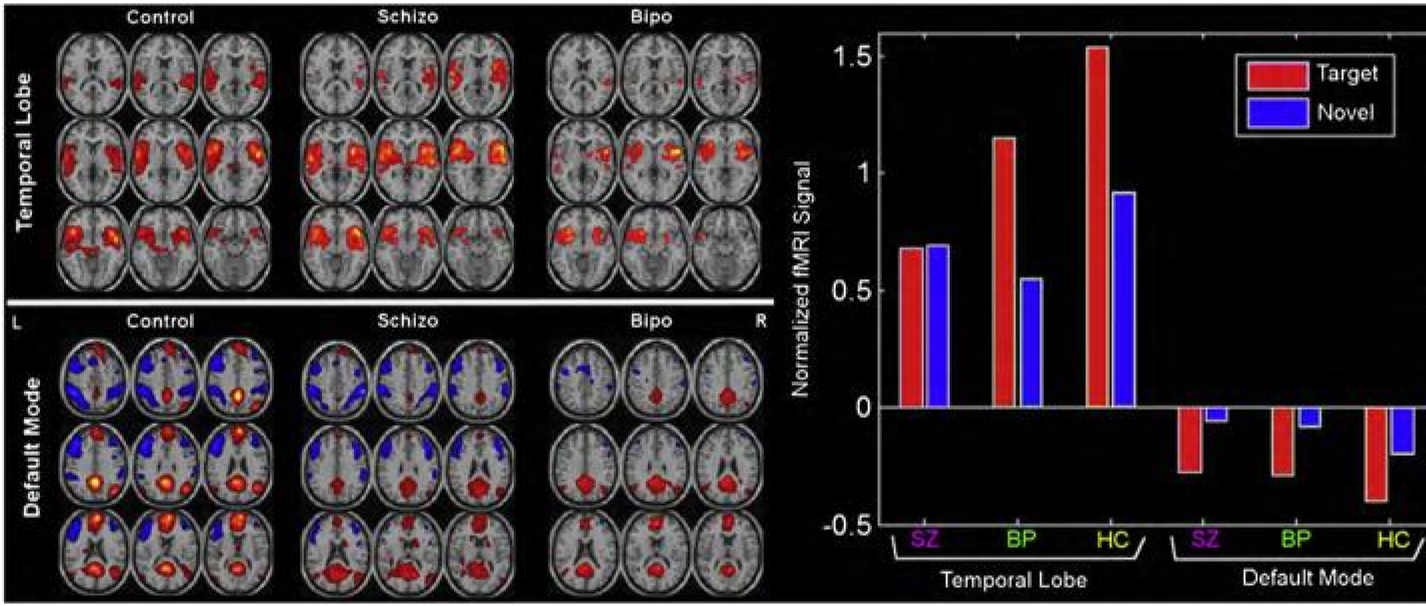


Fixate
Drive
Watch
white
yellow
red
blue
pink
green
Normalized Signal
0
60
120
180
Time(sec)
L
R

Example 2

- fMRI data were collected from an auditory oddball task for two patients groups as well as healthy controls.
- Back-reconstructed component maps were entered into two sample t-tests to evaluate pair-wise differences between the three groups.
- We performed a multiple regression including the target, novel, and standard stimuli and the mean of the estimated beta parameters is shown in Fig. 7.

- **Fig. 7**

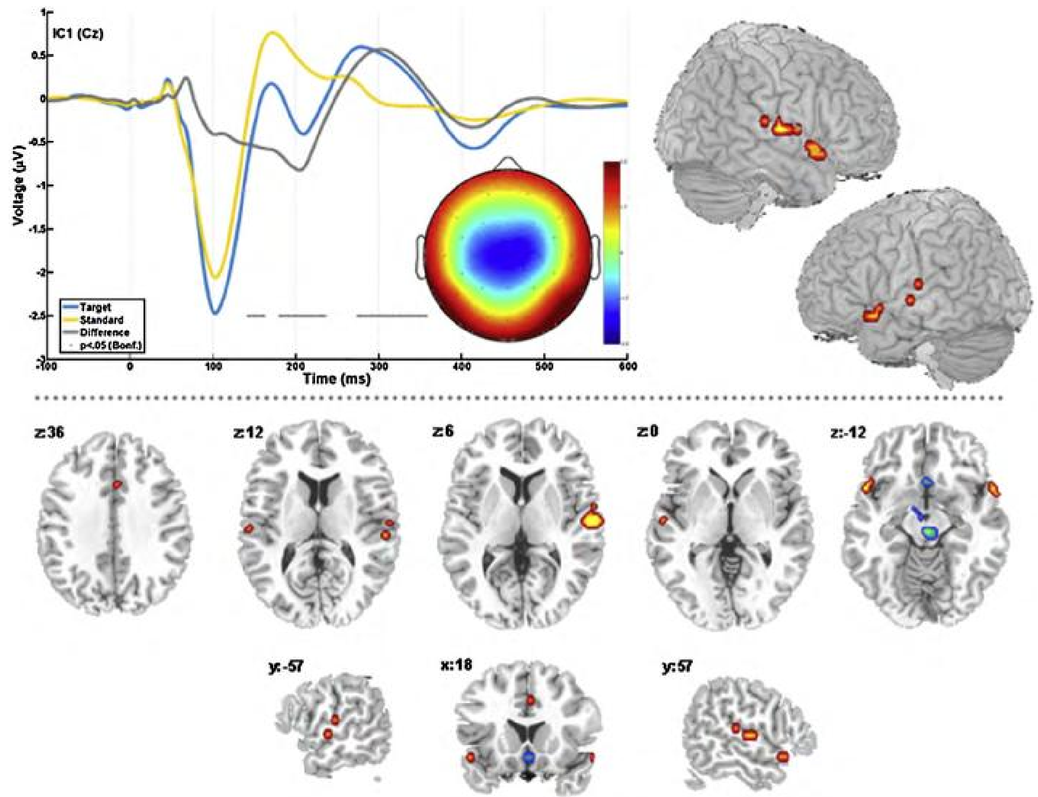


Control
Schizo
Bipo
Temporal Lobe
L
R
Default Mode
Normalized fMRI Signal
1.5
1
0.5
0
-0.5
Target
Novel
SZ
BP
HC
Temporal Lobe
SZ
BP
HC
Default Mode

Example 3

- The fMRI data and the 64 channel ERP data are entered into a joint ICA analysis.
- Parallel spatial and temporal independent component analysis for concurrent multi-subject single-trial EEG-fMRI that addresses the mixing problem in both modalities.
- Integration of the data via correlation of the trial-to-trial modulation of the recovered fMRI maps and EEG time courses.

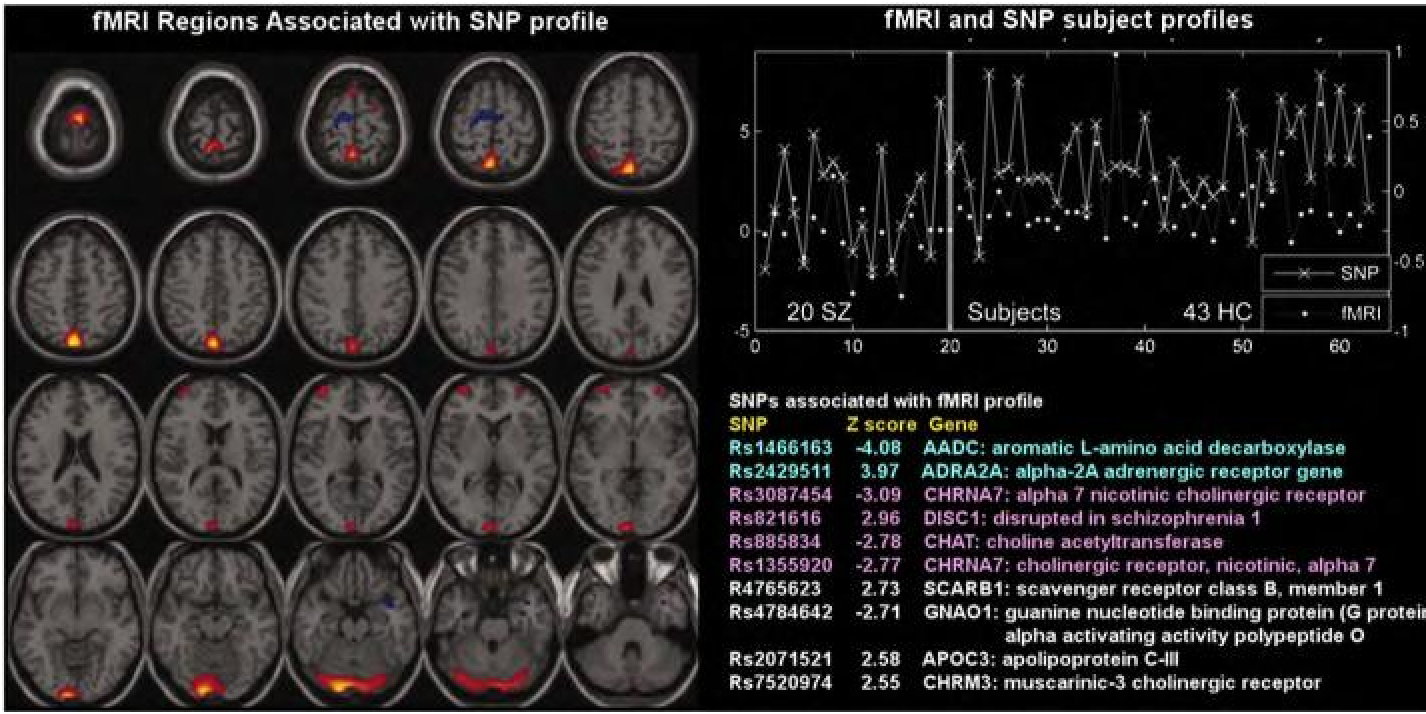
- **Fig. 8**


IC1 (Cz)
Voltage (µV)
Time (ms)
Target
Standard
Difference
p<.05 (Bonf.)
z:36
z:12
z:6
z:0
z:-12
y:-57
x:18
y:57

Example 4

- A parallel ICA analysis of auditory oddball fMRI data and 367 SNPs from schizophrenia patients and healthy controls.
- We found a correlation of 0.38 between one fMRI component and one SNP component.
- This fMRI component consisted of regions in parietal lobe, right temporal lobe, and bilateral frontal lobe.
- Both fMRI and SNP components showed significant differences in loading parameters between the schizophrenia and control groups.

- **Fig. 9**


fMRI Regions Associated with SNP profile
fMRI and SNP subject profiles
20 SZ
Subjects
43 HC
SNP
fMRI
SNPs associated with fMRI profile
SNP Z score Gene
Rs1466163 -4.08 AADC: aromatic L-amino acid decarboxylase
Rs2429511 3.97 ADRA2A: alpha-2A adrenergic receptor gene
Rs3087454 -3.09 CHRNA7: alpha 7 nicotinic cholinergic receptor
Rs821616 2.96 DISC1: disrupted in schizophrenia 1
Rs885834 -2.78 CHAT: choline acetyltransferase
Rs1355920 -2.77 CHRNA7: cholinergic receptor, nicotinic, alpha 7
R4765623 2.73 SCARB1: scavenger receptor class B, member 1
Rs4784642 -2.71 GNAO1: guanine nucleotide binding protein (G protein) alpha activating activity polypeptide O
Rs2071521 2.58 APOC3: apolipoprotein C-III
Rs7520974 2.55 CHRM3: muscarinic-3 cholinergic receptor