Available at: http://www.ictp.it/~pub_off IC/2004/69

United Nations Educational Scientific and Cultural Organization
and
International Atomic Energy Agency

THE ABDUS SALAM INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS

EXPLORATORY ANALYSIS OF FUNCTIONAL MAGNETIC RESONANCE IMAGING DATA

R. Mutihac[1]
F.M. Kirby Research Center for Functional Brain Imaging, Kennedy Krieger Institute, Baltimore, MD 21205, USA,
Department of Radiology, Johns Hopkins University, Baltimore, MD 21205, USA,
Department of Electricity and Biophysics, University of Bucharest, Bucharest 79600, Romania,
and
The Abdus Salam International Centre for Theoretical Physics, Trieste, Italy,

J. Gillen
Department of Radiology, Johns Hopkins University, Baltimore, MD 21205, USA,

Peter C.M. Van Zijl and J.J. Pekar
F.M. Kirby Research Center for Functional Brain Imaging, Kennedy Krieger Institute, Baltimore, MD 21205, USA
and
Department of Radiology, Johns Hopkins University, Baltimore, MD 21205, USA.

MIRAMARE – TRIESTE

August 2004

[1] Regular Associate of ICTP.

Abstract

The present contribution briefly reviews the state-of-the-art in functional magnetic resonance imaging (fMRI) data exploratory analysis. Hypothesis-driven methods are discussed in contrast with data-driven model-free techniques in the context of processing and interpretation of brain imaging data. The specificity of applying various exploratory methods to fMRI time series is highlighted and, consequently, their benefits and limitations are comparatively pointed out and typified by experimental investigations reported in literature and by our own research as well. The emphasis is put on the independent component analysis (ICA) considered as a promising data-driven multivariate approach to neuroimaging data analysis. ICA is based on a minimum of statistical assumptions on the latent sources, namely non-Gaussianity and mutual independence, and it allows to discover in data feature reporting on the organization of the nervous system. In this respect, we reconsidered some previously published data on neural substrates of response inhibition in a visuo-motor task and which, subject of ICA, revealed activity predominantly localized to auditory regions, with time courses consistent with the experimental paradigm. Our results concluded that exploratory approaches can discover activity in fMRI data beyond that predicted in advance and modeled as regressors in a linear model, so that new independent regressors may be added in a linear model with potential benefits in model accuracy and physiological interpretation.

INTRODUCTION

Biomedical signals are a rich source of information about physiological processes, but they are typically mixtures of unknown combinations of sources summing differently at different loci, and often contaminated with artifacts and/or noise. Further, for many data sets even the nature of the sources is an open question. In recent years, advanced non-invasive medical imaging techniques such as positron emission tomography (PET), dynamic computer tomography (CT), and magnetic resonance imaging (MRI) has been introduced into biomedical practice. Beyond the plain imaging of morphological structure, the analysis and visualization of biomedical image time-series data is a challenge with growing importance for both basic research and clinical application. In this respect, functional MRI (fMRI), as a non-invasive technique in localizing dynamic brain processes in intact living brain, is by far the most complex and informative approach in neuroscience imaging. Specifically, fMRI is used to track brain function, by visualizing changes in chemical composition of brain areas or changes in the flow of fluids that occur over time spans of seconds to minutes. It is based on the magnetic susceptibilities of oxygenated hemoglobin (HbO_2) and deoxygenated hemoglobin (HbR) and is used to track blood-flow-related phenomena accompanying or following neuronal activations. The classical principle behind detecting activations using fMRI is essentially a voxel-by-voxel *t*-test on a series of images acquired under different conditions [Friston *et al.,* 1995; Worseley and Friston, 1995]. The index of neuronal activity used generally in brain imaging data analysis is the *blood oxygenation level dependent* (BOLD) contrast [Ogawa *et al*., 1990]. The basic assumption is that an increase in neuronal activity within a brain region entails an increase in local blood flow, leading to reduced concentrations of deoxyhemoglobin in the blood vessels. Unlike oxyhemoglobin that is paramagnetic $(\chi > 0)$, the deoxyhemoglobin is diamagnetic $(\chi < 0)$, which means a different magnetic susceptibility, χ, in relation to the surrounding tissue. Therefore, relative decreases in deoxyhemoglobin concentration attract a reduction in local field unhomogeneity and a slower decay of the MR signal, resulting in higher intensities in T_2* (spin-spin relaxation time) weighted images (Fig. 1).

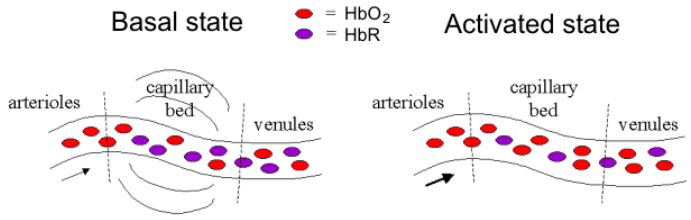


Fig. 1 – *Level of HbR and CBV are basal in normal blood flow (left), whereas an increase in flow (right) entails increased CBV and decreased HbR due to lower field gradients around vessels, which increases the MRI signal.*

The main benefits of fMRI as a technique to image brain activity related to a specific task or sensory process include: (*i*) the signal does not require injections of radioactive isotopes, (*ii*) the total scan time required is short, such as on the order of *1.5* to *2.0* min per run (depending on the paradigm), and (*iii*) the in-plane resolution of the functional image is generally about

$1.5 \times 1.5\,\mathrm{mm}^2$, although resolutions less than *1* mm are possible.

The analysis of fMRI brain data is a complex task, since the fMRI signals have varied, unpredictable time courses that represent the summation of signals from hemodynamic changes as a result of neural activities, from subject motion and machine artifacts, and from physiological cardiac, respiratory and other pulsations. The relative contribution and exact form of each of these components are largely unknown, suggesting analysis methods of blind source separation (BSS) type. A general trend in data analysis consists in finding an adequate representation of multivariate data, which is expected to provide the underlying factors describing their origin and essential structure. Linear transformations are often envisaged to perform such a task due to their computational and conceptual simplicity. Some common linear transformation methods are principal component analysis (PCA), factor analysis (FA), projection pursuit (PP) [Huber, 1985], and more recently the independent component analysis (ICA). ICA [Comon, 1994; Cardoso, 1998] has emerged as a useful extension of nonlinear PCA and developed in context with BSS [Bell and Sejnowski, 1995; Cardoso, 1998] and digital signal processing (DSP). ICA is also related to recent theories of the visual brain, which assume that consecutive processing steps lead to a progressive reduction in the redundancy of the representation [Barlow, 1961; Olshausen and Field, 1996]. It is also related to work on sparse [Field, 1994] and low entropy coding [Atick, 1992].

The analysis of large and complex data sets is efficiently performed by two overlapping classes of techniques known as *exploratory data analysis* (EDA) and *data mining*. EDA helps to cope with data in a fairly informal way and reveal structure relatively quickly and easy. The emphasis is on flexible probing of data, often before comparing them to any probabilistic model. EDA is used to identify systematic relations between variables when there are no (or not complete) *a priori* expectations as to the nature of those relations. In a typical EDA process, many variables are taken into account and compared, using a variety of techniques in search for systematic patterns. The methods of EDA are *best* compromises for a broad range of situations and, quite often, are close to *best* solution for each situation alone. However, the exploration of data can only serve as the first stage of data analysis and its results can be treated as tentative at best as long as they are not confirmed (e.g., crossvalidated) using a different data set or an independent subset. If the result of the exploratory stage suggests a particular model, then its validity can be verified by applying it to a new data set and testing its fit such as testing its predictive validity. Case selection conditions can be used to quickly define subsets of data (e.g., for estimation and verification), and for testing the robustness of results [StatSoft].

Data mining is an analytic process designed to explore usually large amounts of data in search of consistent patterns and/or systematic relationships between variables, and then to validate the findings by applying the detected patterns to new subsets of data. The ultimate goal of *data mining* is prediction, hence *predictive data mining* is the most common type of *data mining*. Basically, the process of *data mining* consists of three stages: (*i*) the initial exploration, (*ii*) model building or pattern identification with validation/verification, and (*iii*)

deployment, which means the application of the model to new data in order to generate predictions. *Data mining* is, nevertheless, based on the conceptual principles of statistics including the traditional EDA and modeling, and it shares with them both some components of its general approaches and specific techniques. An important general difference in the focus and purpose between *data mining* and EDA is that *data mining* is more oriented towards applications rather than the basic nature of the underlying phenomena. In this respect, *data mining* is relatively less concerned with identifying the specific relations between the involved variables, rather its focus is on producing a solution that can generate useful predictions. Therefore, *data mining* accepts a *black box* approach to data exploration or knowledge discovery and uses not only EDA techniques, but also complex analytic techniques like *neural networks* (NNs), which can generate valid predictions but are not capable of identifying the specific nature of the interrelations between the variables on which the predictions are based.

EDA and *data mining* emphasize flexible searching for clues and evidence in data, whereas confirmatory data analysis (CDA) stresses evaluating the available evidence. ICA typifies the exploratory methods searching for a suitable linear transformation of a random vector onto a basis that minimizes the statistical dependence between its components. Spatial transformations are ubiquitous and essential for the proper characterization of evoked hemodynamics changes in fMRI data analysis, both in terms of removing unwanted variance from data and in terms of anatomical localization. The applications of ICA to human brain electromagnetic data have proved meaningful applied to a set of average responses in separating the observed spatially labile activity into spatially fixed components that account for the responses in all the conditions [Makeig *et al.,* 1997].

II. NEED FOR EXPLORATORY ANALYSIS OF FUNCTIONAL MRI DATA

The current techniques (Fig. 2) for analyzing fMRI data can be loosely dichotomized into either data-driven (model-free) methods, such as PCA, ICA, and clustering analysis, or hypothesis-led (model-driven) methods, like the general linear model (GLM) [Friston, 1996]. These two approaches are complementary and mirror the exploratory and confirmatory aspects of scientific investigation. Imaging studies driven by hypotheses derived from cognitive psychology and related disciplines can at best support or refute currently formulated psychological models. Unanticipated time courses of activation of localized brain areas are less likely to be discovered with such analysis methods [McKeown *et al*., 1998a].

Frequently, fMRI experiments reveal coactivation of spatially disparate brain regions, which cannot be rigorously investigated with univariate techniques, because they examine each voxel individually to determine if a give voxel is deemed task-related by a specified criterion and ignore the relationships between voxels. As for instance, two voxels may both be individually correlated with the task reference function (an estimate of the expected task-

related changes seen in a voxel) above a certain threshold, yet be uncorrelated with one another. The voxels found task-related on the base of some statistical significance test (e.g., exceeding a predefined level of significance of a *t*-statistic under the null hypothesis that the distribution of voxel values are identical during control and experimental conditions) are subsequently assembled to form a spatially distributed map of task-related activation. Contrarily, multivariate techniques separate data into a set of spatial patterns of activity (maps), enabling the analysis of co-activation in spatially divergent areas within a given map.

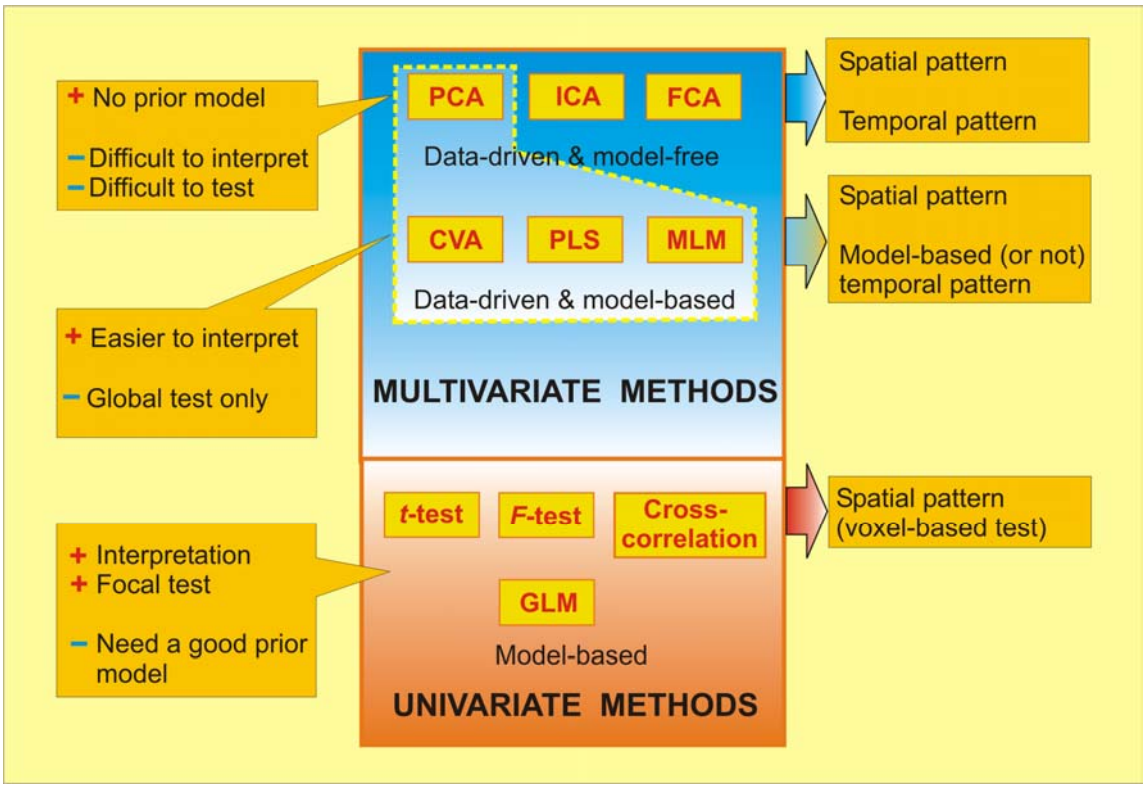


Fig. 2 – *Some typical univariate/multivariate hypothesis/data-driven analysis methods specified along with their main strong points (+) and pitfalls (–).*

In data-driven analysis no statistical model about what inferences are made need to specify. Multivariate data analysis relies on the covariance paradigm and is free of prior assumptions on activation functions. Contrary to inferential approach, exploratory data analysis is capable to detect the functional activity without reference to the experimental protocol and can also reveal new components in the data. Generally, neuromorphic and alike methods of unsupervised learning fall in the class of data-driven analysis [Barlow, 1987], such as eigenimage analysis, self-organizing artificial neural networks, temporal clustering analysis (CA) and fuzzy clustering analysis (FCA), factor analysis (FA), projection pursuit (PP), PCA, and ICA.

Tukey [1962] argued that classical statistics leaning on analyzing small, homogeneous, stationary data by means of known distributional models and assumptions will prove inappropriate to deal with the problems raised by the analysis of large and complex data. A typical *3D* multislice volume fMRI brain dataset for a single subject ranges from 10^8 to 10^{10} bytes, and the values increase further in simultaneous inter-subject studies. Two features of fMRI data, which are characteristics for massive datasets, are *nonstationarity* and *distributional heterogeneity*. It is claimed that the difference between real-life large datasets and smaller ones consist not only in size but in qualitative terms as well [Huber, 1994]. Consequently, the investigations of functional brain imaging data should primarily rely on critical consideration of methods that belong to data mining and exploratory data analysis. We consider that the analysis of typically large, complex, and rather heterogeneous fMRI data ought to start with an exploratory method that will reveal the intrinsic structure in data with no need for prior models and assumptions. A critical evaluation and comparison of the data-driven methods used in fMRI data analysis has not been published to date. Besides, as briefly stated by Huber [1994], "… there are no panaceas in data analysis" either. In other words, all methods have highs and lows, so that an educated choice appears problem domain-dependent.

The constraints imposed by PCA and FA, which segregate data by partitioning its total variance into uncorrelated components, appear unrealistic in fMRI since may lead to ambiguous separation of time courses corresponding to activation, noise, and artifacts. This is due to the relative small amplitude of the task-related components and non-specificity of variance partitioning. Variance partitioning in ICA is based on mutual information (MI), though constraints of spatial and/or temporal statistical independence, as well as non-Gaussianity are imposed, which may only partly be true in real data. Most applications of ICA include PCA as a preprocessing step for whitening data, dimension reduction, and/or filtering out some noise, though by removing the many smallest principal components one runs the risk of potentially removing small details of interest. Individual PCA components are necessarily both spatially and temporally uncorrelated, making them unlikely to represent functionally distinct brain systems. Rotation methods such as *Varimax* and *Promax* (Hendrickson and White, 1994; Makeig *et al.*, 2000) might be used to relax the orthogonality constraint, but their utility for fMRI data analysis has not yet been explored and the relevance of their underlying assumptions to fMRI data may also be questionable [Duann *et al*., 2002]. Additional possible limitations of ICA refer to linearity and the global characterization of data (i.e., even if data sets are statistically heterogeneous, ICA attempts to describe them using the same global features as if the data were spatially homogeneous) [Somorjai and Jarmasz, 1999]. Theoretically, nonlinear ICA might circumvent the distributional heterogeneity [Parra *et al*., 1996], though its implementation becomes rather computationally intensive. A fast deflation-type fixed-point based ICA algorithm introduced by Hyvärinen and Oja [1997] relaxed considerably the computational demand of ICA. In contrast, clustering, and particularly FCA, is local in the sense that the cluster centroids do not consist of linear combinations of the time courses of activations, hence does not get confounded by global

heterogeneity. Moreover, algorithmic implementation of FCA can be made fast, which is important in processing large and complex data sets. FCA and ICA are complementary in the sense that spatial ICA could be used subsequently to FCA if it appears that the centroids are linear mixtures of well-defined temporal shapes. This view advanced by Somorjai and Jarmasz [1999] was supported by Karhunen *et al.* [2000], who proposed a preprocessing step by *k*-means clustering of data. The idea behind *k*-means clustering is to classify individual voxels in the volume with respect to their activation time courses. A *k*-means algorithm needs *k* cluster centroids to be chosen of the same dimensionality as the time series, then each voxel is assigned to the cluster centroid with the best match. Subsequently, local ICA performs demixing of the *k* clusters (or their centroids). In conditions where the time courses of activation change significantly in amplitude during experiments, *wavelet analysis* (WA) can more accurately detect activations than most commonly used data analysis methods invoked so far [Brammer, 1998].

Whatever method or combination of methods would be used in exploratory analysis to discover new hypotheses (models) extracted directly from data, they have to be subsequently tested and verified by some more conventional statistical inferential methods of analysis. The combined information gathered from two or more methods may reveal structure in data that any single method could not have provided. The more an approach embeds prior knowledge that we are aware of about the structure to be discovered, the higher the chance of its detection. This suggests starting the analysis in an adequate Bayesian framework that incorporates all available information on the data and continuously updates the state of knowledge when new data are presented [Mutihac *et al*., 2000]. We limited our comments to mostly compelling exploratory techniques that were reported to have been employed so far in fMRI data analysis and pinpointed their ability to cope with large datasets. Due to low level signals generated in blood-flow techniques, the most of the functional neuroimaging approaches have primarily been developed to rely upon averaging data across time and/or space, which cancels out random processes including spurious noise. Most often, fMRI data preprocessing includes PCA dimension reduction, which may discard some small but significant components. ICA yields components as independent as possible even if the assumption of source independence holds only loosely, and no intrinsic meaning is associated with them either. It is therefore compulsory to thoroughly check the validity of the assumptions on which ICA decomposition is based in order to evaluate the reliability and infer the functional significance of the resulting components. In this respect, a correct estimation of the data model complexity (i.e., the number of relevant estimated components) makes the interpretation task easier and more meaningful [Mutihac *et al*., 2004] along with reducing the computational demand.

III. ICA MODEL

3.1 Data modeling

Basically, the linear, stationary, noise-free ICA model (Fig. 3) transforms an *M*-dimensional random vector $\mathbf{X} = (X_1(t),...,X_M(t))$ into an estimated vector $\mathbf{Y} = (Y_1(t),...,Y_M(t))$ of the original source signals $\mathbf{S} = (S_1(t),...,S_N(t))$ that are assumed mutually statistically independent and non-Gaussian distributed. The observed signals **X** are linear mixtures of the original signals $\mathbf{X} = \mathbf{AS}$ at any sample index *t* (e.g., time), where **A** is a fixed *mixing matrix* of size $M \times N$. Mixing is assumed to be instantaneous, so there is no time delay between the source signals S_i, $i = 1,2,...,N$ mixing into an observable signal X_j, $j = 1,2,...,M$. Both the original (source) signals and the mixing matrix are generally unknown, but the number of observed mixtures *M* should exceed or equal the number of source signals *N* and the matrix **A** should be full-column rank. The ICA task is to find the best estimators for the original source signals **S**, that is $\mathbf{Y} = \hat{\mathbf{S}} = \mathbf{BX}$, where the *separating matrix* **B** is the (pseudo)inverse of the mixing matrix **A**, in such a way that the estimates are as statistically independent as possible. It can be shown [Hyvärinen *et al*., 2001] that the problem is well-defined, that is, the model can be estimated if and only if the components $\{S_i\}$ are *non-Gaussian*. This is a fundamental requirement setting the main difference between ICA and *factor analysis*, in which the non-Gaussianity of data is not taken into account [Spearman, 1927]. In fact, ICA could be considered as non-Gaussian FA, since data are also modeled in FA as linear mixtures of some underlying factors.

It is remarkable that the latent sources can be estimated from linear mixtures on the basis of their independence and non-Gaussianity assumption only. Uncorrelatedness is not enough to separate the components, which is the reason why PCA and FA yield uncorrelated components but fail to separate the original source signals. We can transform any linear mixture of independent components (ICs) into uncorrelated components by using common decorrelation methods, in which case the mixture is orthogonal [Hyvärynen *et al*., 2001]. The goal in ICA is to estimate the orthogonal transformation that is left after decorrelating the observed data, a goal that cannot be achieved by classical methods since they are based on essentially the same covariance information as decorrelation. For random variables that are merely uncorrelated, nonlinear transformations do not have zero covariance in general. Thus, ICA makes use of a stronger form of decorrelation, by finding a representation where the estimates are uncorrelated even after some nonlinear transformations.

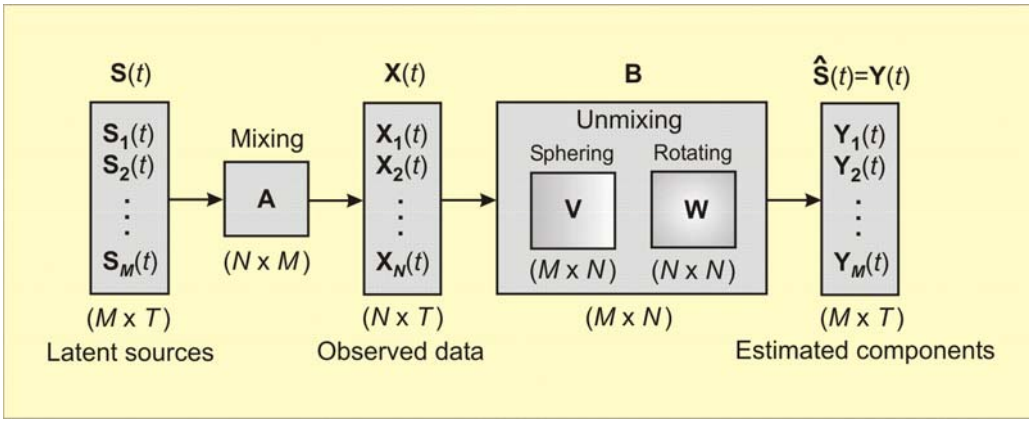


Fig. 3 – *Standard (stationary, linear, noise-free) ICA model.*

3.2 ICA Estimation Principle

There are many methods for estimating the ICA model. Their common feature is that they consider some form of *higher-order statistics*, which specifically means information not contained in the covariance matrix of the observed data. Using the covariance matrix, we can decorrelate the components in the ordinary linear sense, but not any stronger. Two kinds of higher-order information are currently used for ICA modeling, though some others may be used as well: the nonlinear correlations and kurtosis.

Nonlinear decorrelation is probably the basic ICA method. It may be stated as finding a matrix **B,** so that for any $i \neq j$, the components Y_i and Y_j are uncorrelated, *and* the transformed components $g(Y_i)$ and $h(Y_j)$ are uncorrelated, where g and h are some suitable nonlinear functions. If the nonlinearities are adequately chosen, then this method comes up with the independent components. The principles for proper choice of the nonlinearities are recommended by the estimation theory (which provides methods like *maximum likelihood* of estimating any statistical model) and information theory (which provides exact measures of independence like *mutual information*).

Maximum non-Gaussianity of the estimated components is an intuitive and practical method for ICA. Its formulation is to find the local maxima of non-Gaussianity of a linear combination of the observed variables $Y = \sum_i b_i X_i$ under the constraint that the variance of Y is constant. Each local maximum must give one independent component, otherwise, if it were a mixture of two or components, it would be closer to a Gaussian distribution due to the Central Limit Theorem.

3.3 ICA Algorithms

Apart from the estimation principle, an algorithm is needed as part of ICA in order to implementing the computations required. Since nonquadratic functions are generally involved by the estimation methods, numerical algorithms are needed, which are quite computationally demanding. The current algorithms for ICA can be loosely classified in two categories. One category contains *adaptive algorithms* generally based on stochastic gradient methods and implemented in neural networks [Juten and Herault, 1991; Moreau and Macchi, 1993; Bell and Sejnowski, 1995; Delfosse and Loubaton, 1995; Karhunen *et al.,* 1995; Amari, Cichocki, and Yang, 1996; Hyvärinen and Oja, 1996]. These algorithms exhibit slow convergence and their convergence depends crucially on the correct choice of the learning rate parameters. The other category relies on *batch computation* minimizing or maximizing some relevant criterion (objective) functions [Cardoso, 1992; Comon, 1994]. Their drawback is that require complex matrix or tensorial operations.

IV. FUNCTIONAL MRI TIME SERIES

4.1 Hemodynamic response-based signals

Functional MRI studies rely upon the detection of small intensity changes of BOLD signal over time, often with a contrast-to-noise ratio of less than 1.0. Virtually all fMRI studies analyze the magnitude images from the MRI scanner. A standard approach is to correlate the time-series data with an assumed reference signal [Bandettini *et a.l*, 1993]. Many generalizations have been proposed, usually involving linear modeling approaches utilizing an estimate of the hemodynamic response [Worsley and Frinston, 1995]. The information contained in the phase images is ignored in such analyses.

There are several types of signals that can be encoded within the hemodynamic signals measured by fMRI [McKeown *et al.*, 1998c]. The separated signals are commonly classified as *signals of interest* and *signals of no interest*. The signals of interest include *task-related*, *function-related*, and *transiently task-related*. The *task-related* signal is the easiest to model. A reference waveform, based upon the paradigm, is correlated with the data. The responses of the brain to a given task may not be regular however, for example the signal may fade out before the stimulation is turned off or change over time as repeated stimuli are applied, leading to a *transiently task-related* signal. It is also conceivable that there are several different types of transiently task-related signals coming from different regions of the brain. The *function-related* signal manifests as similarities between voxels within a particular functional domain (*e.g.*, the motor cortex on one side of the brain will correlate most highly with voxels in the motor cortex on the opposite side of the brain) [Biswal *et al*., 1995]. The

signals not of interest include *physiology-related*, *motion related*, and *scanner-related* signals. *Physiology-related* signals like breathing and heart rate tend to come from the brain ventricles (fluid filled regions of the brain) and areas with large blood vessels present, respectively. *Motion-related* signals can also be present and tend to be changes across large regions of the image (particularly at the edges of images) [Molgedey and Schuster, 1994]. The *scanner-related* signals are generally varying in time (e.g., scanner drift and system noise) or varying in space (e.g., susceptibility and radio-frequency artifacts) [Beckmann *et al*., 2000]. Finally, there are also various types of noise involved in an fMRI experiment, which may be considered as signals of no interest. Specifically, there is noise due to the magnetic resonance acquisition which can be discussed in terms of (*i*) object variability due to quantum thermodynamics and (*ii*) thermal noise. It was shown that the thermal noise will result in white noise with a constant variance in the image dimension [Wang, 1995]. Additionally there is noise due to subject movement, brain movement, and physiologic noise. In principle, ICA need not explicitly model these sources of noise, rather they pop out as separate components [McKeown *et al*., 1998a,c]). Each ICA component map is described by a distribution of values, one for each voxel. These values represent the relative amount a given voxel is modulated by the activation of that component. To find and display voxels contributing significantly to a particular component map, the map values may be scaled to *z*-scores (the number of standard deviations from the map mean). A certain threshold may be set, so that voxels whose absolute *z*-scores are larger than the threshold are considered *active* voxels for that component. Negative *z*-scores indicate voxels whose fMRI signals are modulated *opposite* to the time course of activation for that component. ICA implementation fully characterizes the observed fMRI data by decomposing them into sparse maps, or spatial modes, that are spatially as independent as possible, and their associated time courses, both of which provide a unique representation of the data (up to scaling and permutation).

It is important to understand the properties of signals in fMRI when developing methods for analyzing data. If the activations do not have a systematic overlap in time and/or space then the distributions can be considered independent [Calhoun *et al.*, 2001, McKeown *et al*., 1998a,b]. The temporal distribution of a task-related waveform is often nearly bimodal (off/on) and thus the algorithm needs to incorporate this fact. The assumption that components are spatially independent and add linearly was evaluated and it was concluded that the fMRI signals and noise are non-Gaussian and the accuracy of the ICA model may vary in specific regions of the brain. The signals of interest in fMRI are typically focal and thus have a sub-Gaussian spatial distribution [Calhoun *et al*., 2003]. However, the artifactual signals will be more varied and potentially super-Gaussian.

Many aspects of fMRI signals are well known and could be incorporated into an ICA analysis. First, local spatial correlation exists in MR images due only to the acquisition process. It is often the case that partial *k*-space acquisitions involve sampling fewer frequency samples than the desired number of spatial samples. One can use the fact that the matrix of frequency data is Hermitian-symmetric to reconstruct the image using a partially acquired

frequency matrix (with the trade-off being a decrease in signal-to-noise ratio). Another well-known method involves sampling the lower frequencies and padding the high frequencies with zero (with the trade-off being a decrease in spatial resolution). This broadens the well described MRI spatial point spread function in one direction, although it has been suggested that there is a real gain in resolution when zero padding is up to as much as twice the original number of samples [Constable *et al.*, 1989]. This results in spatial correlation of the MR signal. In addition, spatial correlation is induced by the process being measured. The hemodynamic sources to be estimated have a spatial hemodynamic (vascular) point spread function. This is partially due to the hemodynamics, but is also a function of the pulse sequence and the parameters used. Differing degrees of sensitivity to blood flow and blood oxygenation as well as differences between low and high field magnets will measure different hemodynamics. The pulse sequence, parameters, and magnetic field strength are considered as constant to enable discussion of the hemodynamic point spread function without introducing the complexities of these parameters. A typical ICA model of fMRI data in notations that facilitate a direct comparison with the GLM is presented in Fig. 4.

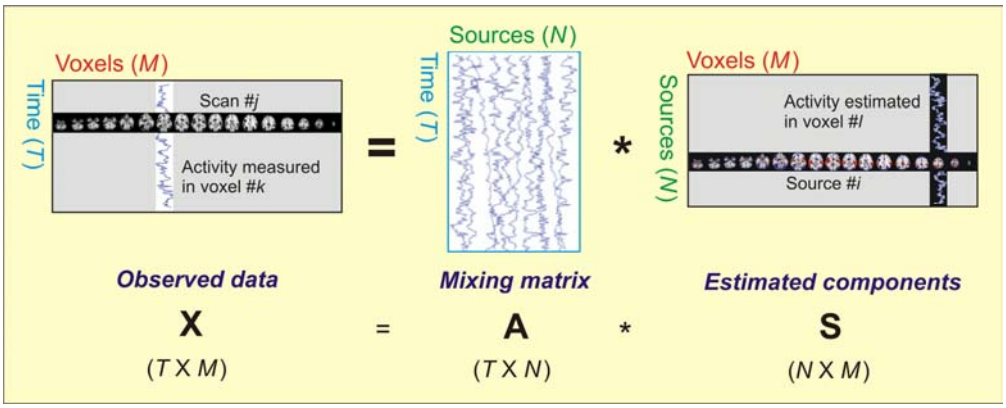


Fig. 4 – *ICA model for fMRI data. The estimation of the latent sources is carried out in such a way as to maximize the mutual independence of the rows in the matrix* **S** *by means of an appropriate mixing matrix* **A.**

There may also be some degree of temporal correlation. Temporal correlation is introduced by: *i*) rapid sampling (a scanner parameter) on the time scale of the magnetic equilibrium constant, *ii*) the temporal hemodynamic (vascular) point spread function (a physiologic variable), and *iii*) poorly understood temporal autocorrelations in the data.

4.2 Functional organization of the brain

Functional organization of the brain is based on two complementary principles, namely *localization* and *connectionism* [Phillips *et al.*, 1984]. Localization implies that each

psychomotor function is mainly performed in a small set of brain areas. Occasionally, some lesions interrupt anatomical connections between macroscopic loci required to perform a specific psychomotor task [Gardner, 1975; Takahashi *et al*., 1992; Duffield *et al*., 1994], yet the brain recovers its normal functionality. This implies the complementary principle of connectionism meaning that the brain regions involved in a given psychomotor function may be widely distributed so that the brain activity required to perform a given task may be the functional integration of activity in multiple microscopic loci or distinct brain subsystems [McKeown *et al*., 1998a]. Consequently, the multifocal brain areas activated by performance of a psychomotor task should be unrelated to the brain areas whose activity is affected by artifacts, such as head movements, physiological pulsations, and machine noise which may dominate fMRI experiments. Any of these separate processes may be represented by one or more *spatially-independent* components, each associated with a *single time course* of enhancement and/or suppression and *a component map*. It is assumed [McKeown *et al*., 1998b] that each component map, which is specified by a spatial distribution of fixed values for each voxel, represents possibly overlapping, multifocal brain areas of statistically dependent fMRI signal influence. The maps will be independent if active voxels in the maps are sparse and mostly nonoverlapping though some overlap usually exists.

4.3 ICA of fMRI time-series

The general framework for the analysis of imaging time-series was established in positron emission tomography (PET) neuroimaging and extended thereafter for fMRI. Basically, two approaches are employed: hypothesis-driven (i.e., inferential) or data-driven (i.e. exploratory). Most of imaging neuroscience relies on inferential hypothesis-led analysis, which is based on some *statistical parametric mapping* (SPM) [Friston, 1995]. Testing a hypothesis reduces operationally to specifying what one expects to see. In SPM the image reconstruction is carried out using a voxel-specific statistic that tests hypotheses on dynamics. The statistic is usually derived under parametric assumptions using the *general linear model* (GLM).

ICA, as a model-free approach, has traditionally been used in two complementary ways to decompose an image sequence into a set of images and a corresponding set of time-varying image amplitudes: *spatial* ICA (sICA) which finds a set of mutually independent images (i.e., maps) and a corresponding dual set of unconstrained time courses, and *temporal* ICA (tICA), which finds a set of independent time courses and a corresponding (dual) set of unconstrained images. ICA can find statistically independent signals out of linear mixtures by making use of the extra degrees of freedom implicit in the unconstrained dual signals, even if the underlying sources are not statistically independent. Therefore, ICA may come out with independent signals, which are not the underlying sources. A natural modification of ICA is to enforce constraints on dual signals as well by maximizing the degree of independence of the dual signals. This approach called *spatiotemporal* ICA (stICA) [Stone *et al*., 1999] maximizes the degree of independence over space *and* time simultaneously, without necessarily producing

independence in either space or time, but searching for the best compromise. The assumption that both spatial and temporal signals are independent is rarely valid in practice, so stICA yields solutions in which the degree of spatial independence is maximized subject to the constraint that the degree of temporal independence is maximized (and reciprocally). It turns out that stICA is based on a more physically plausible assumption that both spatial and temporal sources are *almost* independent. Sources of this type are likely to be present in a spatio-temporal medium like brain tissue, in which correlations in activity between nearby points in space and time are the norm. Such a relaxed assumption over the approximate independence of both the estimated components and their dual signals permits the recovery of sources that are correlated over time and space.

The assumptions of ICA apply to fMRI data in a different way than to other time series analysis. Here the principle of brain modularity suggests that, as different brain regions perform distinct functions, their time courses of activity should be separable (though not necessarily independent, particularly when, typically, only a few hundred or fewer time points are available). Spatial modularity, plus the relatively high *3D* spatial resolution of fMRI, allows the use of ICA to identify maximally *spatially* independent regions with distinguishable time courses. Decreases as well as increases in brain activity are observed, which allows components to have overlapping spatial regions and still be approximately independent. However, the spatial independence of active brain areas is not perfect, and therefore the nature and functional significance of independent fMRI components must be validated by convergent physiological and behavioral evidence (e.g., simultaneously running EEG).

Mathematical formalism

Assume that we have T components (i.e., volumes) and each component consists of V elements (i.e., voxels). The component map distribution is considered spatially independent, and hence uniquely specified. In mathematical terms, if $p_k(\mathbf{C}_k)$ is the probability density distribution (pdf) of the voxel values $\mathbf{C}_k$ in the k-th component map, then the joint probability distribution of all N components factorizes:

$$p(\mathbf{C}_1,\mathbf{C}_2,...,\mathbf{C}_N)=\prod_{k=1}^{T} p_k(\mathbf{C}_k) \tag{1}$$

where each of the component maps $\mathbf{C}_k$ is a vector with components C_{ki}, $i = 1,2,...,V$. The (spatial) independence of the component maps is much stronger than the decorrelation condition imposed by PCA, that is, the voxel values between pair of components are merely uncorrelated:

$$\mathbf{C}_i \cdot \mathbf{C}_j = \sum_{k=1}^{T} C_{ik}C_{jk} = 0, \quad \text{for } i \neq j \tag{2}$$

since (1) implies that the higher-order correlations are also zero. Basically, independence is a much stronger property than uncorrelatedness.

Functional MRI signals can be decomposed into a number of independent component maps and their associated component activation waveforms using an appropriate ICA algorithm. No *a priori* assumptions are necessary on: (*i*) the time courses of activation of different components, (*ii*) whether a given component is activated by specific psychophysiological systems or is related to machine noise or other artifacts. In linear ICA decomposition of fMRI data, the observed data matrix **X** can be transformed into a set of volume maps **C** by taking linear combinations, defined by a $T \times T$ unmixing matrix, **W**, of the volumes recorded at each time point (i.e., sample index) $t = 1,2,...,T$, such as:

$$\mathbf{C} = \mathbf{W}\mathbf{X} \tag{3}$$

where **C** is the matrix of the component maps of size $T \times V$, **X** is the row zero-mean data matrix of size $T \times V$ with each row representing an entire volume (i.e., one full scan of all slices) recorded at each given time point (i.e., sample index) *t* and each column representing the values a certain voxel, and **W** is a $T \times T$ matrix containing the coefficients of linear unmixing of the volumes. An ICA algorithm can perform *blind separation* of input data **X** into a linear sum of time-varying modulations of maximally independent component maps, by iteratively building up an unmixing matrix $\mathbf{W} = \|W_{ik}\|$, which ensures a *linear* decomposition of the data like $C_{ij} = \sum_{k=1}^{T} W_{ik} X_{kj}$, where C_{ij} is the value of the *j*-th voxel of the *i*-th component map, X_{kj} is the value of the *j*-th voxel of the *k*-th data vector, and the summation runs over the *T* time points (i.e., components) of the fMRI input data. If the number of IC's is taken equal to the number of time points, then the unmixing matrix **W** is invertible due to the supposition that the (unknown) mixing matrix has full-column rank. Data reconstruction from the IC's follows straightforward: $X'_{ij} = \sum_{k=1}^{T} W_{ik}^{-1} C_{kj}$,where X'_{ij} is the reconstructed data at the *i*-th time point of the *j*-th voxel. In matrix notation:

$$\mathbf{X}' = \mathbf{W}^{-1}\mathbf{C} \tag{4}$$

The protocol used to appropriately select the separation matrix W depends on the adopted estimation principle. In PCA [Friston *et al*., 1993], **W** in (3) is selected so that the resultant component maps in **C** are uncorrelated and summarize the variability in the data in as few maps as possible. In ICA [Comon, 1994, Bell and Sejnowski, 1995], **W** in (3) is selected such as the rows in **C** are made maximally statistically independent. It appears that the stricter criteria of spatial independence in ICA improves the estimates for the temporal and spatial extent of task-related activity and yields an efficient means for exploratory analysis of fMRI data [McKeown *et al*. 1998 b, c]. The columns of $\mathbf{W}^{-1}$ give the time courses of activation for the spatial maps. Unlike PCA, ICA allows that the time courses to be non-orthogonal.

Assuming stationary linear mixture of brain source signals, then $\mathbf{W}^{-1}$ is the best estimate of the mixing matrix.

ICA specificity

Some basic considerations must be discussed in applying ICA to fMRI data. First, as a general requirement in applying ICA to determine the separating matrix **W** in (3), the fMRI data are assumed to consist of linear sum of spatially independent patterns of activity. Task-related activations which vary in space and time can then be modeled as consistently task-related maps, and as several spatially independent transiently task-related maps, each with unique time courses, so that the sum of all task-related components provides a measure of the full spatiotemporal extent of the task related activity. Then, in the basic ICA model (3) used for fMRI analysis, there was no explicit noise model; rather, the noise was assumed to be distributed among one or more of the components. Secondly, note that equation (3) implies that the recorded data (mixtures), **X**, can be accurately modeled as component maps, **C**, linearly combined as specified in the matrix $\mathbf{W}^{-1}$. Thirdly, since the fMRI data change through time, equation (3) assumes that this is a result of changes in the relative contributions from each of the component maps rather than of changes in the component maps themselves. In other words, the maps are considered to be fixed throughout the fMRI experiment. Finally, equation (3) also implies that the relative contribution from each component map at a given time point in the experiment is the same throughout the head. If any of the above assumptions are not valid, then the ICA algorithm will be less able to separate out statistically independent component maps. The estimated probability of observing the data under the null hypothesis that the ICA assumptions are valid will therefore be reduced and signal separation will be suboptimal

As both a theoretical and a technical issue, the assumption of invertibility of **W** amounts to tacitly admitting that the number of time points equals the number of the independent components to be estimated. This is highly questionable, unless we do have prior information concerning the exact number of independent patterns of brain activity. Standard ICA, as a simple linear regression model, differs from GLM as applied in neuroimaging, in some important respects:

1. The mixing process is assumed to be square, that is, the signal is not constrained to be contained within a lower dimensional signal subspace. However, if we assume that a smaller number of source processes represent the dynamics in data, a mismatch between the best linear model fit and the original data is inevitably introduced;
2. Standard ICA does not include a noise model, instead, data are assumed to be completely characterized by the estimations of the sources and the mixing matrix. This precludes: (*i*) the assessment of statistical significance of the source estimates (i.e., the estimated *independent components*) within the framework of a null-hypotheses testing,

(*ii*) the threshold techniques (like converting the component map values into *Z*-scores) are devoid of statistical meaning and can only be conceived as *ad-hoc* recipes.

The interest in spatial ICA of fMRI data stems from many reasons:

1. ICA implementation fully characterizes the observed fMRI data by decomposing them into *sparse maps*, or *spatial modes*, that are spatially as independent as possible, and their associated time courses, both of which provide a unique representation of the data (up to scaling and permutation);
2. ICA finds systematically non-overlapping, temporally coherent brain regions without constraining the temporal domain;
3. ICA can reveal inter-event and inter-subject differences in the temporal dynamics without resorting to any prior temporal model. The temporal dynamics in fMRI experiments is difficult to study due to the lack of a well-understood and/or incomplete brain-activation model.
4. ICA can be used to investigate the spatially distributed brain networks and functional connectivity of the brain.

4.4 Ranking ICA Separated Components

It is useful to rank order the components by the extent of their contribution to the original data. Unfortunately, different ICA component time courses contained in $\mathbf{W}^{-1}$ are generally nonorthogonal so that, unlike PCA, the variances explained by each component will not sum to the variance of original data. The *contribution* each independent component makes to the magnitude of the original data may be expressed in two alternative ways [McKeown *et al.*, 1998a]:

1. It can be estimated by the root mean square (RMS) of the data set reconstructed solely from this component using $\mathbf{X}' = \mathbf{W}^{-1}\mathbf{C}$ in which $\mathbf{C}$ has only one nonzero row corresponding to the appropriate component;
2. It can be regarded as the RMS error introduced per data point when data are reconstructed without this component:

$$\gamma_i = \frac{1}{T \cdot V}\left[\sum_{j=1}^{T}\sum_{k=1}^{V}\left(A_{jk}^{i}\right)^2\right]^{1/2} \tag{5}$$

where γ_i is the contribution to the data from the *i*-th component and A_{jk}^{i} is an $T \times V$ matrix computed from the outer product of the *i*-th component map and the *i*-th column of $\mathbf{W}^{-1}$, that is $A_{jk}^{i} = W_{ji}^{-1}C_{ik}$

4.5 Pitfalls of ICA Model of fMRI Data

The ICA implementation fully characterizes fMRI data by separating them into sparse maps, or spatial modes, and associated time courses. Employing an ICA algorithm capable of looking for non-sparse as well as sparse maps can find maps that all are sparse [McKeown *et al*., 1998a]. Many of these maps can be identified with known artifacts, such as blood vessel pulsations, head movements and slow drifts. These highly spatially structured signals are not easily modeled by a priori estimates as required by hypothesis-driven methods. A basic assumption in ICA, that the maps are spatially independent, does not preclude the possibility of spatial overlap because maximal independence can be achieved with overlap in high-dimensional spaces. Nevertheless, ICA is not an ideal method of fMRI data processing and particular care should pay to the following issues:

1. Functional MRI signal component processes may exhibit saturation or other nonlinear properties and thus may not be appropriate for analysis using exclusively a whole linear model. However, an assumption of additivity may be reasonable [Boynton *et al*., 1996].
2. ICA model assumes that the distribution of voxel values specifying the map for each signal component is statistically independent of the distributions of voxel values specifying all the other component maps. This assumption leads to an essentially unique decomposition and biases the ICA algorithm towards finding components having relatively sparse as well as discrete active component areas [McKeown *et al*., 1998b]. Nevertheless, the optimum way to describe the varying spatial extent of time-dependent task-related activations detected in fMRI data is still a question of debate.
3. The least energetic ICA components, in the sense of their mean back-projected variance in the data space, particularly those with speckled spatial distributions, appear to be noise of unknown origin. It is not transparent to what extent a given component is physiological signal or identifiable artifact, and what is noise. Moreover, it is not very clear which domain of applicability would benefit from noise modeling included in an extension of the basic ICA model, since noise modeling is itself a rather complex and imprecise task.
4. The assumption that the component maps are spatially stationary makes the method sensitive to the detection of movement artifact. Nevertheless, there is no straightforward correction method of suspected head movements. Moreover, the spatial independence of active brain areas is not perfect, and therefore the nature and functional significance of independent fMRI components must be validated by convergent physiological and behavioral evidence.

5. ICA method yields a linear decomposition of data into as independent as possible components, even if the underlying assumption of their independence is weak. That is, the estimates show up with no particular meaning or significance assigned by the ICA method. Apparently, the meaning of the estimated components is domain-specific dependent and not very clear.
6. Model complexity selection is crucial in terms of allotting a meaning to estimated IC's. Overestimating the dimensionality will introduce many spurious components and make difficult subsequent inference on their meaning. Underestimating the number of IC's will possibly discard valuable information rendering the signal separation sub-optimal.
7. The ICA estimated components show up even when the assumption of their independence is weak and they are not endowed with any particular meaning. Therefore, methods for testing the statistical reliability of ICA component time courses and areas of activation need to be developed.

V. RESULTS AND DISCUSSION

5.1 Methods

Our fMRI data originated from *24* randomly selected healthy subjects, males and females, performing *2* similar sessions of MR EPI scanning. The scanner was a 1.5 T Intera Philips Medical Systems, with SENSE both on and off. When on, the SENSE factors were 2 and 3, respectively. The acquisition matrix was $64 \times 64 \times 41$ and the voxel size was $3.75 \times 3.75 \times 4\,\text{mm}^3$ with *0.5* mm gap between slices. *120* volumes/session were acquired at $TR = 2.5$ seconds, and the *6* firstly acquired volumes were discarded as dummies, aiming to reach a steady-state regime. The data model complexity was set to *68* principal components by probabilistic PCA reduction [Beckman *et al*., 2000]. Data were processed by spatial Group ICA, which uses higher-order statistics to express the fMRI data as a sum of temporally-modulated spatially-independent sources. The algorithm was an in-house implementation of the fast fixed-point ICA algorithm introduced by Hyvärinen and Oja [1997]. The *68* spatially estimated mean IC's were ranked by spatial correlation within a pre-defined region of interest (ROI) of the brain responsible for auditory activity (i.e., Broadmann area 41 and 42).

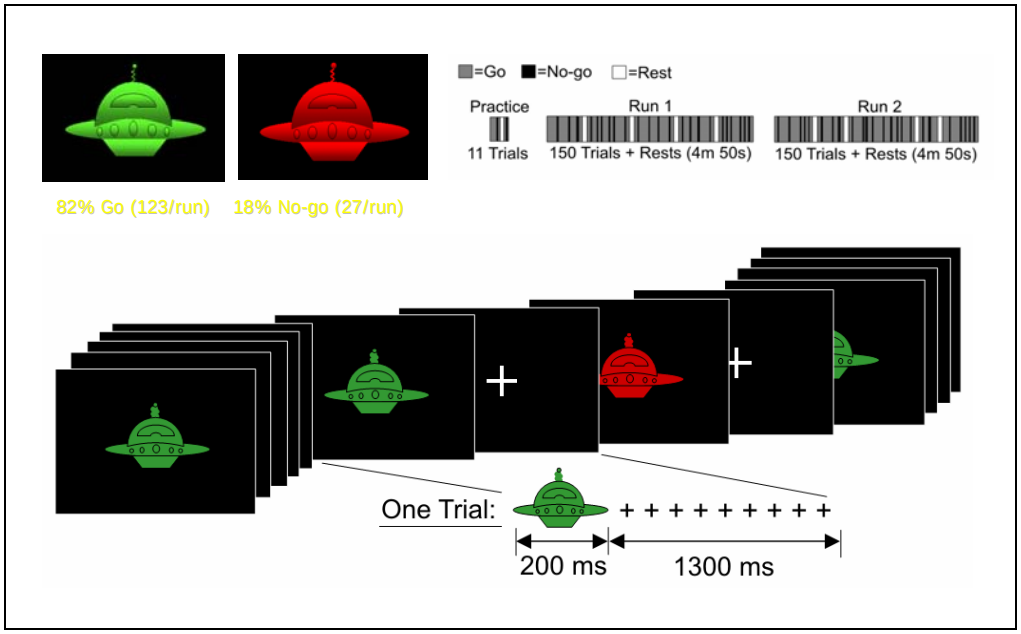
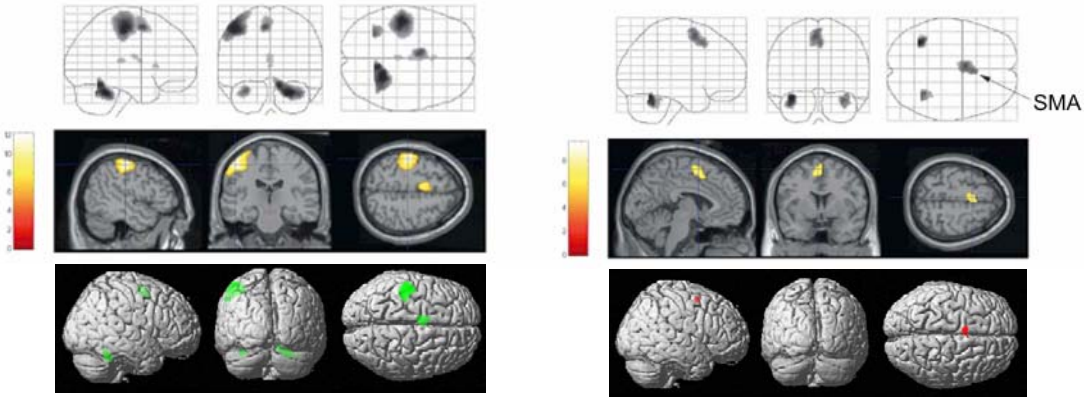


Fig. 5 – *Experimental paradigm for Go/NoGo task.*

We have previously analysed novel "Go/NoGo" tasks to examine neural substrates of response inhibition using fMRI [Mostofsky *et al.*, 2003]. In the simplest task (Fig. 5), subjects saw every *1.5* seconds rapidly presented "spaceships" for *200* ms, and were instructed to press a response button "to rescue" the prevalent (*82*%) green spaceships ("Go") but not the rare (*12*%) red spaceships ("NoGo"); inhibition of the pre-potent response was required. Each "spatial mission" lasted *4*m *50*s, and included *4* brief "rest" periods, apart form the initiating and ending periods of fixation.



(*a*) (*b*)

Fig. 6 – *GLM analysis of Go (a) and NoGo (b) data.*

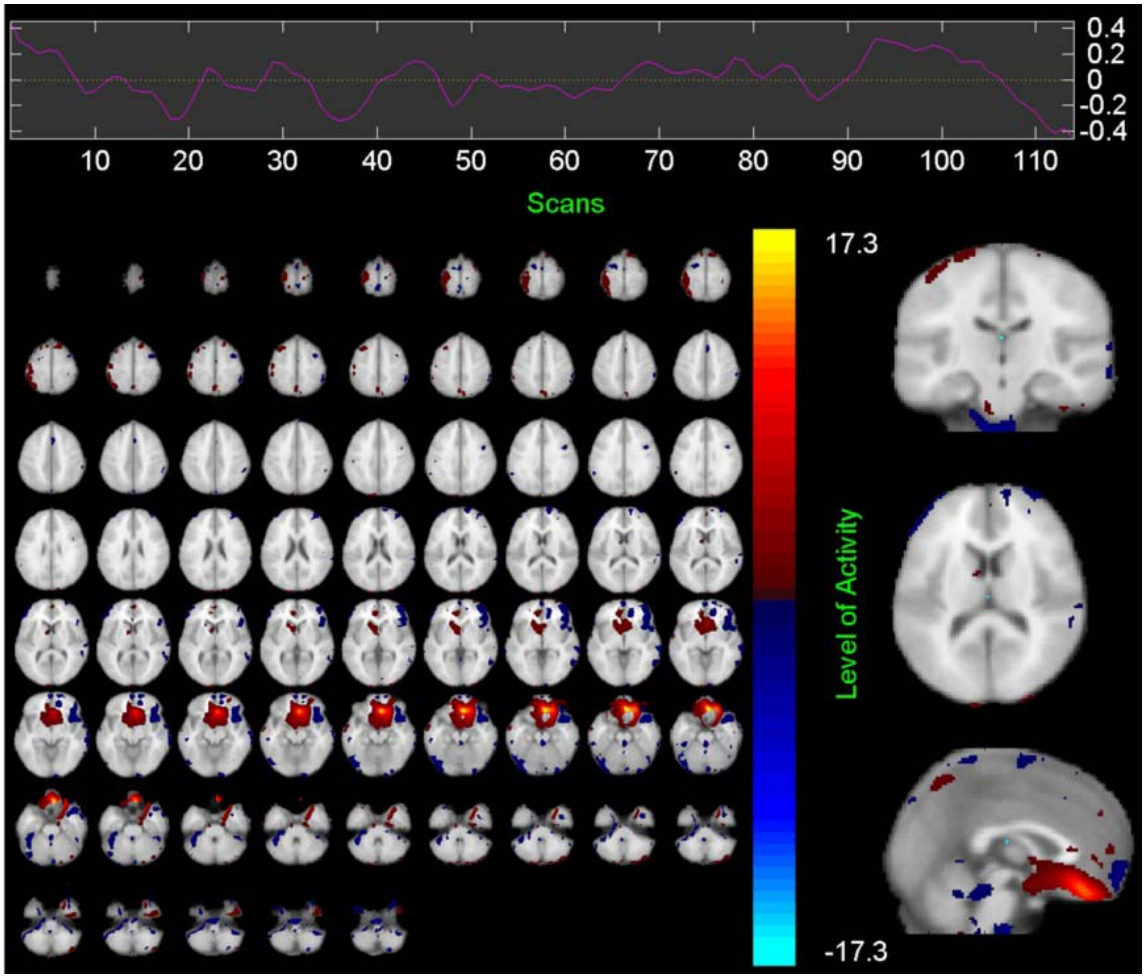


Fig. 7 – *Most energetic mean Group IC for both "Go" and "NoGo" tasks.*

Analysis using the General Linear Model as in SPM2, modeling successful responses as well as errors, revealed activity in contralateral sensorimotor cortex, ipsillateral cerebellum, and supplementary motor area (SMA) for "Go," and in pre-SMA for "NoGo" (Fig. 6).

Our analysis was focused on exploring the Go/NoGo data using spatial ICA, which, in contrast to GLM analysis, makes no specific prior hypotheses on the expected time courses of activation and aims to discover features reporting upon the organization of brain activity. Specifically, we searched to reveal whether the fMRI data from the visuo-motor "Go/NoGo" task reports on attentional modulation of auditory perception. Running Group sICA for all *24* subjects clearly yielded *2* components best correlated with "Go" and "NoGo" regressor, respectively (Figs. 7 and 8).

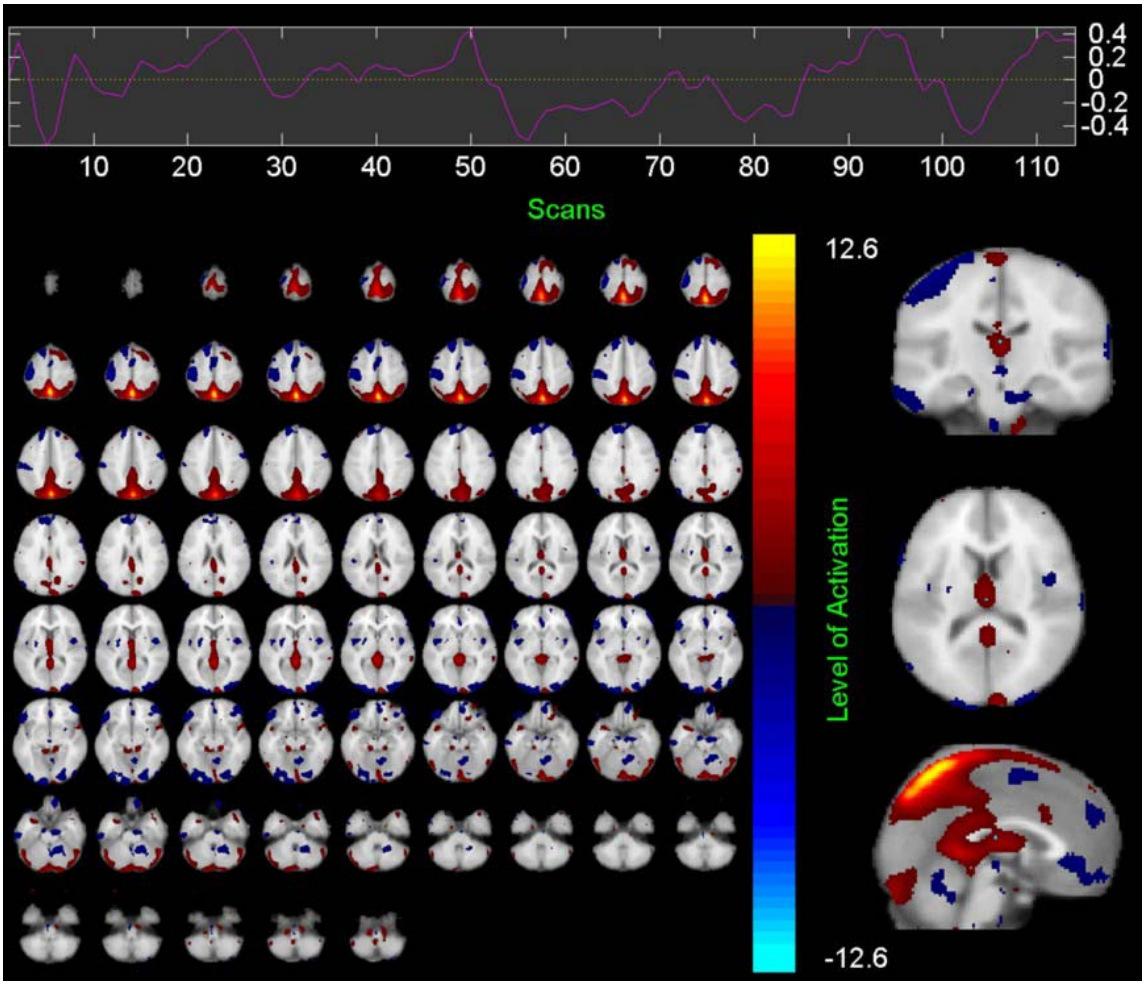


Fig. 8 – *Best correlated mean estimated component by Group ICA with "Go" temporal model, which is consistent with activity in the visual cortex.*

However, Group sICA also yielded a component primarily localized in auditory cortex, with a time course displaying transient deactivation upon initiation of the paradigm, and transient activation during each "rest" period (Fig. 9). This activity, while not predicted *a priori* within the GLM (univariate inferential) analysis and, therefore, not included in the experimental design matrix as a distinct regressor, is easily interpretable, *post hoc*, due to its obvious relationship to the timing of the paradigm events. Such activity predominantly localized to auditory regions, having the time course consistent with attentional modulation of auditory processing is likely due to the acoustically noisy scanner environment, namely to a subjective experience that the scanner's acoustic noise appears "louder" when the participant is not concentrating on task performance. In our opinion, this is a typical example of a brain activation which, because it does not fit an *a priori* hypothesis about changes in brain activity during the task paradigm, was consequently omitted from classical univariate hypothesis-

driven analysis like GLM. Our analysis pointed out that exploratory approaches like ICA can find activity in fMRI data beyond that predicted in advance – in this case, auditory activation during a "non-auditory" task.

The key assumptions that ICA relays on are that the data set consists of *some* spatially independent components, which are linearly mixed and spatially fixed during each run. The number of components extracted can be reduced by initially preprocessing data with PCA. Since higher-order statistics are used to enforce stricter criteria for spatial independence between maps, better estimates for the consistently task-related components are obtained, which supports the assumption of spatial independence. However, spatial dependence between consistently task-related and transiently task-related components can be inferred by the changes in the transiently task-related maps when the consistently task-related component is removed.

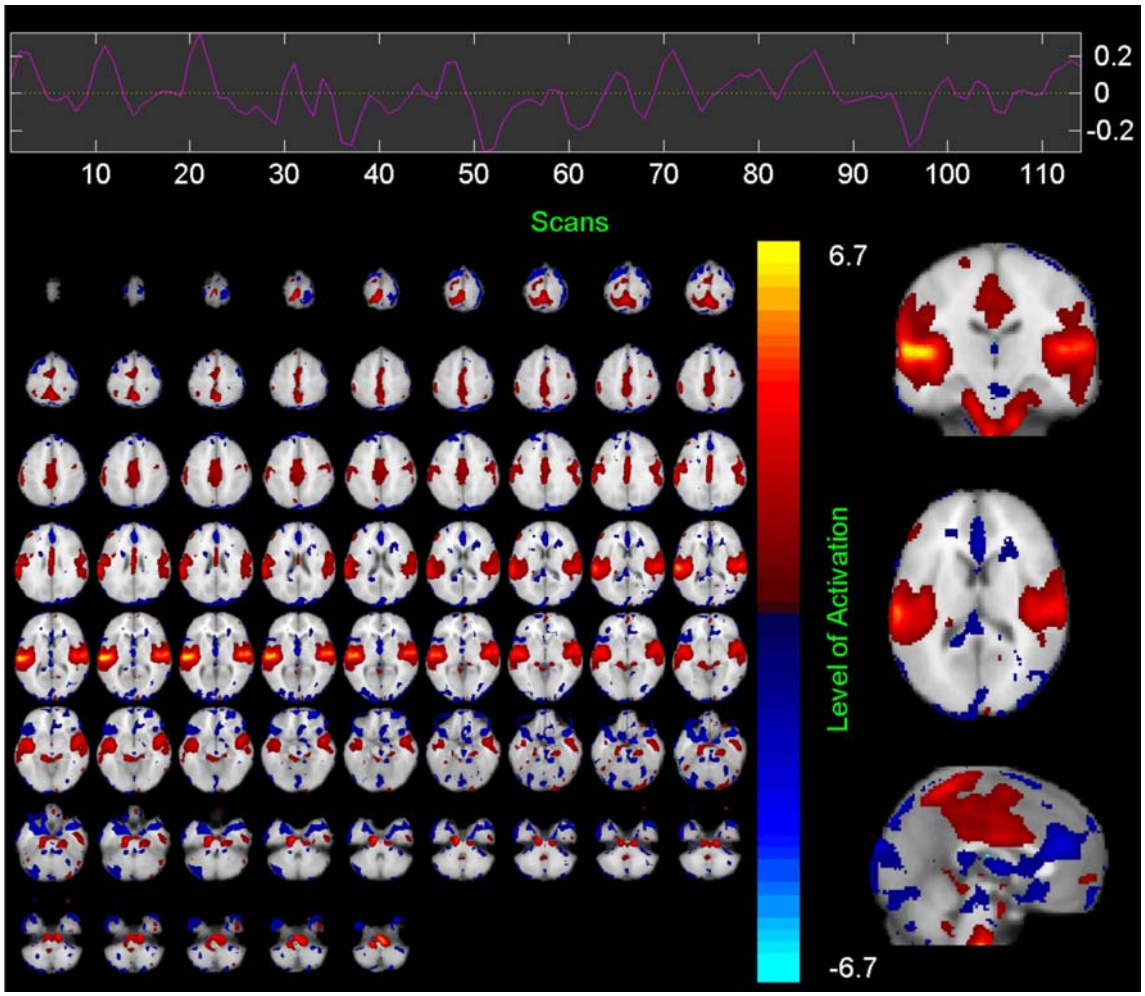


Fig. 9 – *Group ICA separated mean component having the time course consistent with attentional modulation of auditory processing (Broadmann area 41 and 42).*

VI. CONCLUSION

In fMRI, the signal at each voxel consists of a mixture of underlying brain source signals and various uninteresting signals. As a multivariate data-driven statistical analysis method, the ICA model has the ability to discriminate activations that could not be predicted in advance of the experiment, such as transiently task-related ones. Thus ICA may allow straightforward analysis of more complex brain imaging experiments in which unpredictable changes in cognitive activation occur in parallel with changes in arousal or autonomic states for which the exact time courses of activation are not known, too. In this respect, the ICA approach is promising for investigations of patients with pathological conditions that may alter the latencies, amplitudes and brain distributions of their fMRI signals in unpredictable ways. Simulations indicated that the results of ICA are robust in the presence of noise in the data [McKeown *et al*., 1998b].

Applying ICA to brain imaging augments the ongoing re-evaluation of inferential statistics in imaging neuroscience. Inferential approaches, based upon statistical models, allow insights from other fields to be adopted easily and powerfully. ICA does not appeal to a statistical model and no hypotheses are tested, although it might provide insights that allow better models to be designed. The limits of ICA usefulness for fMRI data analysis will ultimately depend on the match between the basic assumptions of the analysis method and the structural composition of fMRI data.

Future techniques in neuroimaging will probably combine data- and hypothesis-driven analysis approaches by initially employing powerful data-driven methods to reveal the underlying nature of the signals and noise, such as stICA, and then testing hypotheses of interest within the previously determined more accurate context, such as multivariate linear models.

Acknowledgments

The ongoing research on applying exploratory analysis to fMRI data is supported by the NIH/NCRR Resource for Quantitative Functional MRI under Grant P41RR015241. This work was done within the framework of the Abdus Salam International Centre for Theoretical Physics, Trieste, Italy. Special thanks goes to Prof. Chela-Flores for his exquisite professionalism provided during many stimulating discussions, as well as to Ms. Grilli and the staff of the Publications and Printing Services who are making their best to have top quality preprints published. Last but not least, R.M. is grateful to M. Iqbal of the Multimedia Group for his full logistic support while at the Centre.

REFERENCES

Amari S.-I., Cichocki A., and Yang H.H. (1996) A new learning algorithm for blind source separation. *Advances in Neural Information Processing Systems* (Proc. NIPS'95), pp. 752-763, MIT Press, Boston, MA.

Atick J.J., Could information theory provide an ecological theory of sensory processing? *Network* 3:213-251, 1992.

Bandettini P.A., Jesmanowicz A., Wong E.C., and Hyde J.S. (1993) Processing strategies for time-course data sets in functional MRI of the human brain. *Magn. Res. Med.* 30:161-173.

Barlow H.B., Possible principles underlying the transformations of sensory messages. In: Rosenblith W.A. (Ed.) "Sensory Communications." MIT Press, Cambridge MA, 1961, pp. 217-234.

Barlow H.B. (1987) Unsupervised learning. *Neural Comput.* 1:295-311.

Beckmann C.F., Noble J.A., and Smith S.M. (2000) Artefact detection in fMRI data using Independent Component Analysis. *NeuroImage*, 11:S614.

Bell A.J. and Sejnowski T.J. (1995) An information maximization approach to blind separation and blind deconvolution. *Neural Comput.* 7:1129-1159 (code at http://sloan.salk.edu/~tony/ica.html).

Biswal B., Yetkin F.Z., Haughton V.M., and Hyde J.S. (1995) Functional connectivity in the motor cortex of resting human brain using echo-planar MRI, *Magn. Res. Med*. 34:537-541.

Brammer M.J. (1998) Multidimensional wavelet analysis of functional magnetic resonance imaging. *Hum. Brain. Mapp*. 6:378–382.

Boynton G.M., Engel S.A., Glover G.H., and HeegerD.J. (1996) Linear system analysis of functional magnetic resonance imaging in human V1. *J Neurosci*. 16:4207-4221.

Calhoun V.D., Adali, T., Pearlson G.D., and Pekar J.J. (2001) Spatial and temporal Independent Component Analysis of functional MRI: Data containing a pair of task-related waveforms. *Hum. Brain Map*. 13:43-53.

Calhoun V.D., Adali T., Hansen L.K., Larsen J., Pekar J.J. (2003) ICA of functional MRI data: An overview *4th Int. Symposium in Independent Component Analysis and Blind Signal Separation* (ICA2003), Nara, Japan.

Cardoso J.-F. (1992) Iterative techniques for blind source separation using only fourth-order cumulants. *Proc. EUSIPCO*, Brussels, pp. 739-742.

Cardoso J.-F. (1998) Blind signal separation: statistical principles. *Proceedings of the IEEE* 9(10):2009-2025.

Comon P. (1994) Independent component analysis, A new concept? *Signal Processing* 36(3):287-314.

Constable R.T., Kay I., Smith M.R., and Henkelman R.M. (1989) High quality zoomed MR images. *J. Comput. Assist. Tomogr*. 13:179-181.

Delfosse N., Loubaton P. (1995) Adaptive blind separation of independent sources: A deflation approach. *Signal Processing* 45:59-83.

Duann J.-R., Jung T.-P., Kuo W.-J., Yeh T.-C., Makeig S., Hsieh J.-C., and Sejnowski T.J. (2002) Single-trial variability in event-related BOLD signals. *NeuroImage* 15:823–835.

Duffield J.S., de Silva R.N., Grant R. (1994) Pure alexia without agraphia: A classical cortical syndrome revisited. *Scott. Med. J.* 39:178-179.

Field D.J. (1994) What is the goal of sensory coding? *Neural Comput.* 6:559-601.

Friston K.J., Frith C.D., Liddle P.F., and Frackowiak R.S. (1993) Functional connectivity: The principal-component analysis of large (PET) data sets. *J Cereb Blood Flow Metab* 13:5–14.

Friston K.J., Holmes A.P., Worsley K.J., Poline J.-P., Frith C.D., and Frackowiak R.S.J. (1995) Statistical parametric maps in functional imaging: A general linear approach. *Hum. Brain Mapp*. 2:189-210.

Friston K.J. (1996) Statistical parametric mapping and other analyses of functional imaging data. In: "Brain Mapping: The Methods" (Toga A.W. and Mazziotta J.C., Eds.) Academic Press, pp. 363-396.

Gardner E. (1975) Fundamentals of Neurobiology (Saunders W.B., Ed.), 6th ed. Philadephia, USA.

Hendrickson A.E. and White P.O. (1994) Promax: a quick method for rotation to oblique simple structure. *The British Journal of Statistical Psychology*, 17(1):65-70.

Huber P.J. (1985) Projection pursuit. *The Annals of Statistics* 13(2):435-475.

Huber P.J. (1994) Huge data sets. In *Proceedings*, *Compstat* (Dutter R. and Grossman W., Eds), pp. 3-13, *Physica Verlag Heidelberg*.

Hyvärinen A. and Oja E. (1996) A neuron that learns to separate one independent component from linear mixtures. *Proc. IEEE Int. Conf. on Neural Networks*. Washington, D.C., pp. 62-67.

Hyvärinen A. and Oja E. (1997) A fast fixed-point algorithm for independent component analysis. *Neural Computation* 9:1483-1492 (code at http://www.cis.hutfi/projects/ica/fastica/).

Hyvärinen A., Karhunen J., and Oja E. Independent Component Analysis. John Wiley & Sons, 2001.

Juten C. and Herault J. (1991) Blind separation of sources, Part I: An adaptive algorithm based on neuromimetic architecture. *Signal Processing* 24(1):1-10.

Karhunen J., Wang L., and Joutsensalo J. (1995) Neural estimation of basis vectors in independent component analysis. *Proc. Intl. Conf. on Artificial Neural Networks* (ICANN'95). Paris, France, October 1995, pp. 317-322.

Karhunen, J., Malaroiu, S., and Ilmoniemi M. (2000) Local linear independent component analysis using clustering. *Int. Journal of Neural Systems*, 10(6):439-451.

Makeig S., Jung T.-P., Ghahremani D., Bell A., and Sejnowski T.J. (1997) Blind separation of event-related brain responses into independent components, *Proc. Natl. Acad. Sci.* USA 94:10979-10984, 1997.

Makeig, S., Enghoff, S., Jung, T.-P. and Sejnowski, T.J. (2000) Moving-window ICA decomposition of EEG data reveals event-related changes in oscillatory brain activity. *Proceedings of the Second International Workshop on Independent Component Analysis and Signal Separation*, Helsinki, Finland, pp. 627-632.

McKeown M.J., Makeig S., Brown G.G., Jung T.-P., Kindermann S. S., Bell A.J., and Sejnowski T.J. (1998a) Analysis of fMRI data by blind separation into independent spatial components, *Hum. Brain Map.* 6:160-188.

McKeown M.J., and Sejnowski T.J. (1998b) Independent Component Analysis of fMRI Data: Examining the assumptions, *Hum. Brain Map.* 6:368-372.

McKeown M.J., Jung T.-P., Making S., Brown G.G., Kindermann S.S., Lee T.-V., and Sejnowski T.J. (1998b) Spatially independent activity patterns in fMRI data during the Stroop color-namimg task. *Proc. Natl. Acad. Sci*. USA 95:803-810.

Molgedey L. and Schuster H. (1994) Separation of independent signals using time-delayed correlations. *Physical Review Letters*, 72:3634-3637.

Moreau O., Macchi O. (1993) New self-adaptive algorithms for source separation based on contrast functions. *Proc. IEEE Signal Processing Workshop on Higher Order Statistics*. Lake Tahoe, NV, pp.215-219.

Mostofsky S.H., *et al*., *Cogn. Brain Res. 17*:419 (2003).

Mutihac R., Stanciulescu C., Mutihac R.C., Cicuttin A., and Cerdeira A.E. (2000) Topics in Bayesian Maximum Entropy, *Romanian Reports in Physics*, 52(3-4):189-223.

Mutihac R., Gillen J.S., Van Zijl P.M.C., and Pekar J.J. (2004) A Mutual information-based approach to estimate the number of meaningful independent components from ICA decomposition of fMRI data, ISMRM 12th Scientific Meeting and Exhibition, Abstract 492, fMRI Data Analysis, p. 106, Kyoto Kyoto, Japan, May 15-21.

Ogawa S., Lee T.M., Kay A.R., and Tank D.W. (1990) Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proc Natl. Acad. Sci.* 89:9868-9872.

Olshausen, B.A. and Field, D.J. (1996) Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature* 381:607-609.

Parra L., Deco G., and Miesbach S. (1996) Statistical independence and novelty detection with information-preserving nonlinear maps. *Neural Comput.* 8:260-269.

Phillips C.G., Zeki S., and Barlow H.B. (1984) Localization of function in the cerebral cortex. Past, present, and future. *Brain* 107:327-361.

Somorjai, R.L. and Jarmasz, M. (1999) Exploratory data analysis of fMR images: Philosophy, strategies, tools, and implementation. Proceedings of ISMRM 1999, p. 1714.

Spearman C. (1927) The Abilities of Man, Their Nature and Measurement, McMillan, New York.

Stone J.V., Porrill J., Buchel C., and Friston K. (1999) Spatial, temporal, and spatiotemporal Independent Component Analysis of fMRI data. *Proc. of Spatio-temporal Modelling and its Applications*. Dept of Statistics, University of Leeds.

Takahashi N., Kawamura M., Shinotou H., Hyrayama K., Kaga K., and Shindo M. (1992) Pure word deafness due to left hemisphere damage. *Cortex* 28:295-303.

Tukey J.W. (1962) The future of data analysis. *Annals of Statistics*, 33:1-67.

Wang Y. (1995) MR Image Statistics and Model-Based MR Image Analysis, *Ph.D. Dissertation*, UMBC, MD, USA.

Worsley K.J. and Friston K.J. (1995) Analysis of fMRI time-series revisited - Again. *NeuroImage*, 2:173-181.