29

Neuroimaging Analysis II: Magnetic Resonance Imaging

Estela Camara, Josep Marco-Pallarés, Thomas F. Münte and Antoni Rodríguez-Fornells

ANALYSIS OF MAGNETIC RESONANCE IMAGING

Cognitive processes are widely distributed across the whole brain, involving interacting and overlapping brain regions. Magnetic resonance imaging (MRI) provides a non-invasive, *in vivo*, quantitative measurement of psycho-physiologically relevant parameters that are related to cognitive operations in the normal and abnormal brain. The combination of sophisticated experimental designs and powerful statistical analysis of MRI signals has become a powerful remarkably useful tool for cognitive neuroscience and psychology. The present chapter tries to give an overview of the main statistical tools used in structural and functional MRI analysis. In addition, we will also consider the analysis (preprocessing) and treatment of magnetic resonance images. While the following sections are restricted to MRI, certain points discussed below can also be applied to other neuroimaging techniques like positron emission tomography (PET).

While the electroencephalography (EEG) section in Chapter 28 is focused on the temporal properties of the signals, this chapter emphasizes the spatial aspects. This is a reflection of the two major approaches of MRI in cognitive neuroscience: (1) *structural imaging* is performed to obtain an accurate description of the morphological characteristics of the studied brain; and (2) *functional imaging* provides information about the average hemodynamic response in each part of the brain (which is compartmentalized into many small volume units, 'voxels') when a subject performs a specific task. In both approaches, a detailed (structural or functional) spatial image of the brain is obtained. As the temporal resolution of the hemodynamic response is relatively slow, this has led to a preference for MR analysis to describe spatial aspects of the response.

[16:37 21/4/2009 5283-Millsap-Ch29.tex] Job No: 5283 Millsap: The SAGE Handbook of Quantitative Methods in Psychology Page: 697 697–715

STRUCTURAL MAGNETIC RESONANCE IMAGING

Based on the magnetic properties of different tissues, high-resolution structural MRI can generate three-dimensional images with detailed structural definition because of its high sensitivity to soft-tissue contrasts. Additionally, diffusion tensor imaging (DTI) measures the diffusion of water in the brain; at the microscopic level, neural tissue structures have distinct boundaries, including axon membranes and myelin sheaths, which constrain the diffusional propagation of water molecules and force the latter in certain preferential directions. This allows researchers to characterize the micro-structure of the medium studied from differences in diffusional properties in various physiological and pathological states. Different statistical tools have been developed in order to test for specific regional structural changes at different spatial scales.

Region-of-interest analysis

Traditionally, the simplest approach used to compare local anatomical differences is commonly referred to as *region-of-interest* (ROI) analysis, in which a defined region is identified and statistical comparisons are made relating its size or intensity value with a particular task under study. Obviously, the crucial point is to delimit the studied region. Then, a mean value of this region can be extracted and compared between groups. For example, Figure 29.1 (left panel) shows a typical ROI analysis in which the relative

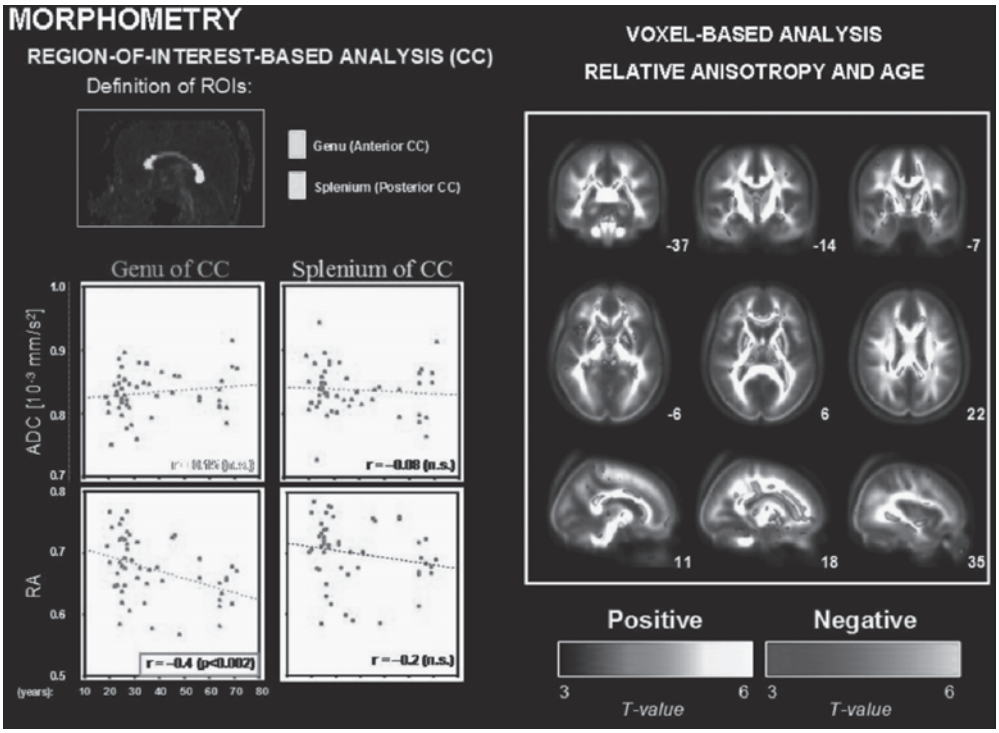

Figure 29.1 (A) Region of interest (ROI)-based analysis of the anterior and posterior corpus callosum (CC). The scatter plots depict the relationship between relative anisotropy (RA) and age, and apparent diffusion coefficient (ADC) and age. (B) Normalized and averaged axial RA map overlays from a sample of 54 healthy volunteers (range 19–71 years) with RA-related t-scores. Figures show positive (red) and negative (blue) correlations. Refer to color plates at the end of this volume for a colored version of this figure.

anisotropy (RA) parameter (a micro-structural index that reflects the integrity of white matter fibers) correlates negatively with age in the anterior part of the corpus callosum (CC). ROI analyses are restricted to the (few) regions selected for analysis, and these regions are usually derived from a priori hypotheses. Possible bias might be introduced due to manual or semi-automated definitions of the ROIs. The additional averaging over a brain region also reduces spatial resolution, and some biologically meaningful differences that might be detected at the voxel-level might be missed (Virta et al., 1999). Given these concerns, a voxel by voxel comparison between groups of subjects might be an attractive method to investigate local changes.

Voxel-based morphometry

Voxel-based morphometry (VBM) permits researchers to make statistical inferences at the voxel-level for the whole brain by estimating changes in local tissue concentrations and volumes. This procedure is relatively straightforward and can be broadly divided into three steps (Ashburner and Friston, 2000): (1) normalization; (2) segmentation; and (3) smoothing. In addition, some studies introduce a modulation step. Figure 29.2 (upper panel) summarizes the main steps of this process. At this point, statistical analysis from the subsequent voxel-based analysis is directly comparable to the ROI approach, since each voxel in the smoothed image contains the average value from the surrounding voxels (see Figure 29.2B). Thus, regions that differ significantly with respect to a particular effect are commonly revealed based on the general linear model (GLM) framework. Figure 29.1 shows an example in which RA values are correlated with age in the whole brain by applying a voxel-based approach.

Finally, in particular with structural diffusion data, it should be noted that some difficulties may arise in this procedure if the intention is to examine tissue properties at a voxel level. In this case, identical brain co-ordinates have to be compared across the whole study population, and even mesoscopic-structural differences are discarded. Thus, the main challenge facing voxel-based diffusion parameter analysis involves meeting the requirement for an optimal matching of the brains being compared; this is usually quite difficult to achieve. Indeed, traditionally, voxel-based analyses of diffusion data have only been treated as an exploratory tool. Despite the possible artefacts derived from these methodological constraints, the development of complementary tools for voxel-based analysis in diffusion data sets has created very powerful approaches for the analysis of diffusion information (Camara et al., 2007).

FUNCTIONAL MAGNETIC RESONANCE IMAGING

The capacity to map brain functions non-invasively *in vivo* with functional MRI has been critical for the success of cognitive neuroscience in the past decade. Blood oxygenation level dependent (BOLD) contrast is the main mechanism measured by functional MRI (Ogawa et al., 1990). Activity dependent changes in local deoxyhemoglobin levels are theorized to result from changes in oxygen extraction, blood flow, and blood volume regulation within the brain. All of these parameters change during neural activity (Buxton et al., 1998). While the vascular changes underlying neurovascular coupling are highly correlated with neural changes, the different time constants of the neural and vascular (and by extension BOLD) phenomena need to be taken into account. Whereas neural activity changes within milliseconds, the hemodynamic response has a long time constant and therefore low temporal resolution (see Appendix 1 for the mathematical characterization of the hemodynamic response function). The hemodynamic response begins with an initial short dip of the BOLD signal and then shows a steep increase, with a maximum between 4–6 seconds after the onset of neural activity. The hemodynamic response to a given stimulus can last between 20 and 30 seconds until complete return to

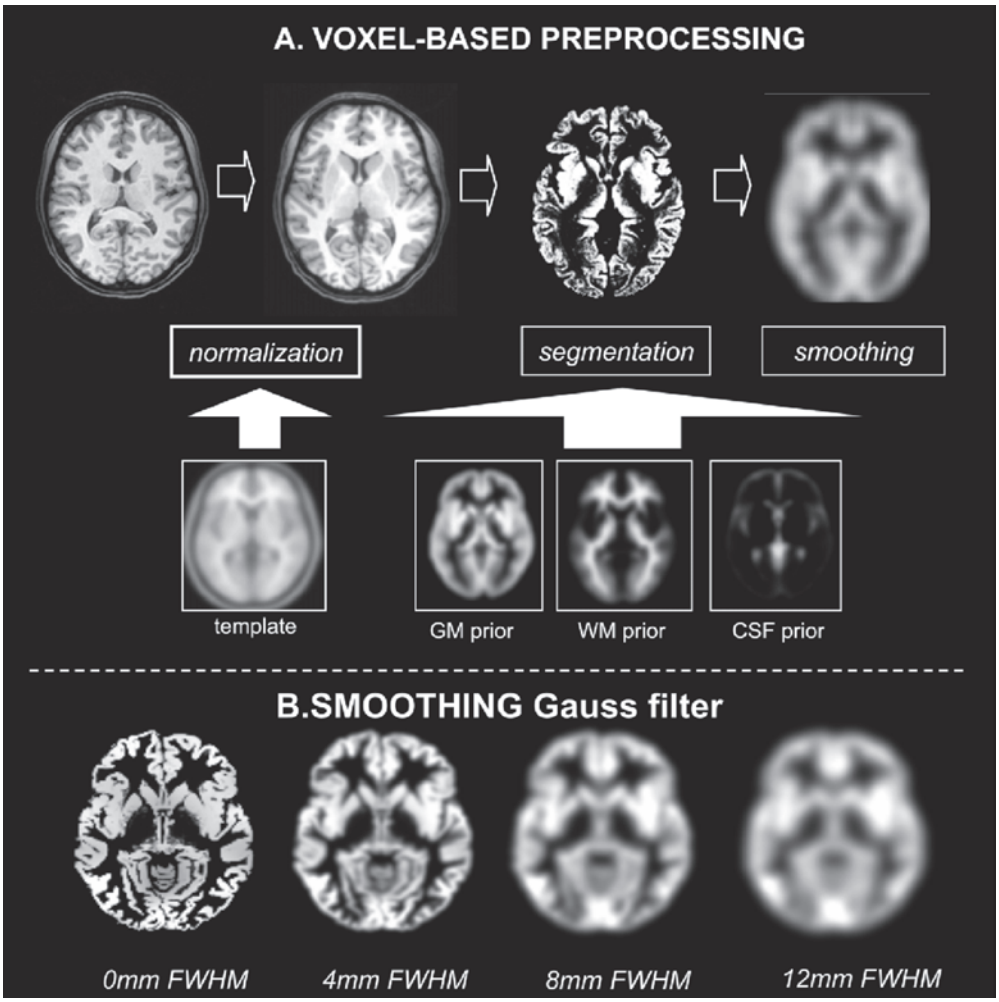

Figure 29.2 (A) Voxel-based morphometry main preprocessing steps. (B) Four different Gaussian filters are applied to a normalized gray matter segmented image. The effects of using different degrees of Gaussian smoothing appear as a blur of the image. FWHM, full-width half maximum. Refer to color plates at the end of this volume for a colored version of this figure.

a baseline level, but this pattern can vary between regions and subjects (Aguirre et al., 1998; Figure 29.3).

The initial dip in the BOLD response is thought to reflect a local increase in oxygen consumption that likely reflects an increase in neural activity. Consequently, this effect should be spatially highly specific, but it is unfortunately quite inconsistent across studies. Such a discrepancy might be due to the fact that the amplitude of the initial dip is much smaller than the main BOLD signal. It is necessary to use high magnetic fields in order to enhance the signal-to-noise ratio and observe the effect (Yacoub et al., 2001). The subsequent increase of the BOLD response has been related to the adaptive behavior of neural activity.

A major goal of functional MRI analysis is to mathematically estimate a hemodynamic response function that captures the BOLD signal associated with task-related neural activity changes. Nevertheless, differences between hemodynamic response

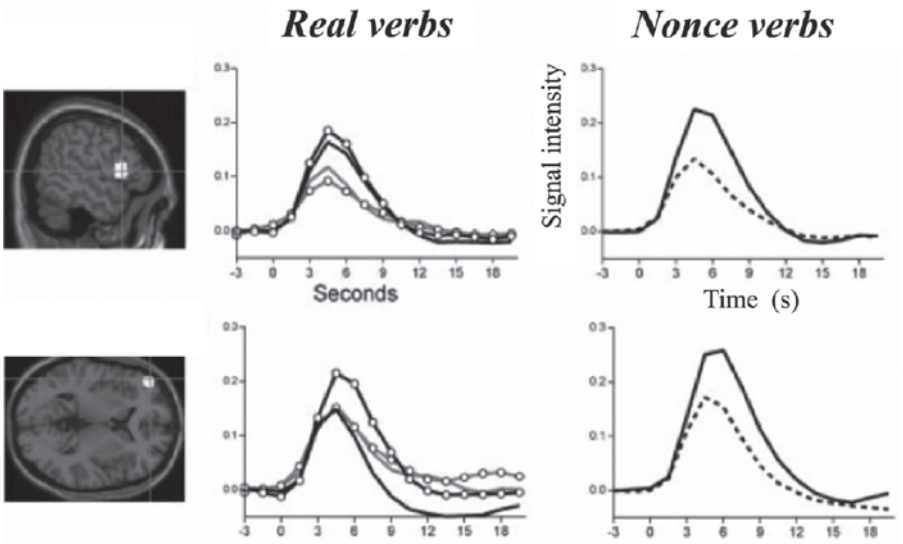

Figure 29.3 Time-course reconstruction of the hemodynamic response of two different regions of interest, the inferior frontal gyrus and the left middle frontal gyrus, evoked by different task-related conditions (de Diego et al., 2006). Refer to color plates at the end of this volume for a colored version of this figure.

patterns exist from one brain region to another and across subjects. The differences arise primarily because of variation in the vasculature, although non-linearities at the neuronal level (e.g., the adaptive behavior of neural activity) can induce hemodynamic differences as well (Logothetis et al., 1999; Logothetis, 2003). Figure 29.3 shows different BOLD responses evoked by different task-related conditions (de Diego et al., 2006). Notice that different time course responses are observed both between regions and between conditions.

Experimental designs

Block and event-related designs (Amaro and Barker, 2006) are the main types of experimental designs in functional MRI.

Block designs

Following the tradition of block designs used in PET, the first studies using functional MRI employed block designs because of their statistical power and simplicity. A series of trials of one condition are grouped together during a period of time (a block), usually about 30 seconds. Therefore, it is possible to pinpoint brain activity that is related to cognitive processing differences between experimental and control blocks. Key aspects of a block design are the number of conditions, the block length, and the spacing between blocks (Bandettini and Cox, 2000). Introducing more conditions in a block design leads to a decrease of the signal-to-noise ratio (SNR), because the duration of the experiment is limited. In a standard block design, the number of experimental conditions is therefore limited to three or four. The block length depends more on the demands of the experiment itself, given that some processes cannot be modulated over short or long periods. In principle, the shape and timing of the hemodynamic response is insensitive to the change of the length of the blocks. Hence, long periods can be robustly modulated, but one might encounter variations in task performance or non-stationarity effects.

Event-related designs

Because of the problems of block designs mentioned above, experimental designs that

use random sequences of events are desirable. Moreover, some cognitive processes could not be studied by block designs. For example, in action monitoring tasks aiming at the delineation of error-related activity or in memory tasks trying to differentiate successfully and unsuccessfully remembered items, it is not possible to predict when an error or memory hit will be produced; hence, block-designs are not feasible. Event-related designs generally consist of rapid trial sequences comprised of different event types of brief duration presented in a random order. The development of rapid data acquisition sequences has been crucial, since the reconstruction of the time course of the hemodynamic response associated with an individual event requires frequent sampling of the measured signal. It is important to apply optimized paradigms in which it is possible to reduce the effective sampling rate (Dale, 1999). By randomizing the stimulus onset-asynchrony from one event to the next, the time course of the BOLD response to a particular class of events can be extracted from the BOLD signal (Dale, 1999; Miezin et al., 2000). Nevertheless, the BOLD response associated with a specific event class in a rapid event-related design is so weak that the number of repetitions of each event has to be sufficiently large to permit estimation of the response. The cost associated with these designs is the increase in the duration of the experiment. Additionally, it is important to note that every possible combination of trial sequences should be presented to allow an optimal deconvolution. In sum, a trade off between the number of conditions and the number of trials per condition and an optimal trial spacing are crucial factors in event-related designs (Birn et al., 2002; Dale, 1999).

Mixed-models designs

The mixed designs represent an integration of classical block and event-related designs. In the mixed design, control blocks are alternated with task blocks, during which trials are presented with different intervals between them. In block designs, the overall activity evoked during a task block is estimated, thereby confounding sustained and transient activity. Event-related designs, on the other hand, ignore sustained activity and reveal only activity that is transient in nature and associated with a concrete event of interest. Indeed, one of the main advantages of mixed designs is the ability to dissociate sustained task-related processes and transient trial-related processes (Visscher et al., 2003). Mixed designs might therefore appear optimal for many applications, but they require sophisticated analysis to separate transient and sustained activities, particularly for cases in which a brain region shows transient and sustained activation.

General linear model

One of the simplest statistical approaches used to infer differences between two conditions is to compare their mean intensity value, typically by applying a simple t-test. Because of the variability between block transitions and the low number of blocks per condition, a direct comparison between the respective means is not feasible. Therefore, comparisons between conditions are performed by averaging the sustained activity between conditions and excluding the transitions between blocks. In these cases, a normal distribution of the data can be assumed and the statistical significance of the comparison is enhanced. The major shortcoming of using only direct differences between means is that neither the variability in the shape of the hemodynamic response nor possible temporal condition-related variations are captured by steady-state averaging.

The most widely used mathematical approach is the GLM framework (Figure 29.4). It models each experimental session as an independent single time series decomposed into the sum of separate factors (conditions) and additive noise. This approach allows researchers to estimate and evaluate a predicted model whose parameter estimates are adjusted based on their best fit with the experimental data.

Formally, time series of signal intensities at each voxel are modeled independently as a linear combination of explanatory variables

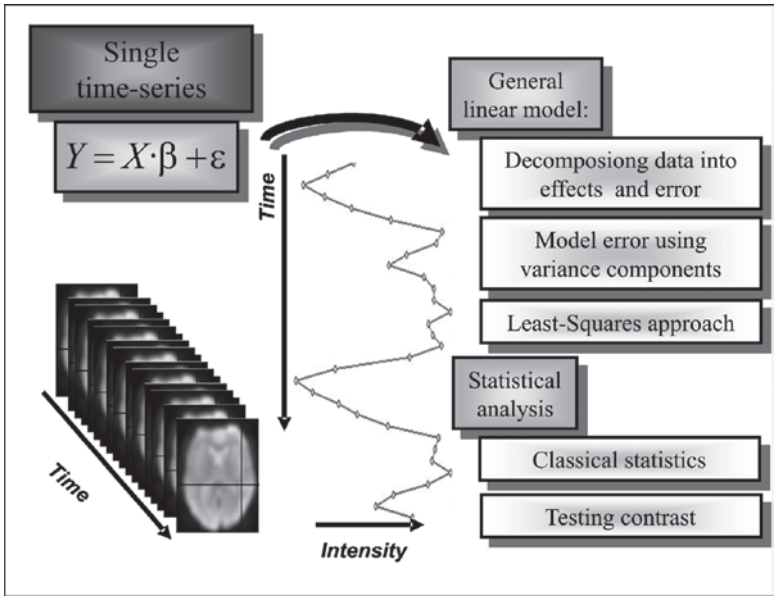

Figure 29.4 Flowchart describing the main steps used in the study of functional time series embedded in the General lineal model. Refer to color plates at the end of this volume for a colored version of this figure.

and Gaussian noise. In this way, the experimental data for each voxel (Y_k) is defined as a linear combination of weighted dissociable factors (x_i) plus an additive error term (ε):

$$Y_k = x_{k1}\beta_1 + \ldots + x_{ks}\beta_s + \ldots + x_{kS}\beta_S + \varepsilon_k \quad (1)$$

In a matrix representation:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (2)$$

Indeed, the $K \times S$ matrix $\mathbf{X}$ (where K is the number of time points and S the number of explanatory factors) represents the design matrix, and its complexity depends on the experimental constraints. Given that the BOLD response is known, the shape of the estimated response at each sample time point for each condition is achieved by convolving the onset times of every condition with the shape of the predicted hemodynamic response function (HRF). Therefore, effects of interest in a block design are modeled by convolving the canonical HRF with a constant value in the blocked intervals and zeros in the rest of the experiment (the boxcar regressor function). In contrast, event-related designs are modeled by convolving each trial onset with the canonical HRF. In addition, parametrical modulations (i.e., reaction time or the accuracy of a trial) or specific regressors that reflect movement-correlated patterns can be included in the design matrix as a covariate of the model factors. It is also common to introduce other known sources of variability not related to the experiment (nuisance factors), such as linear drifts of intensity as a result of small deviations in the scanner magnetic field or physiological parameters like small-scale pulsations in the brain derived from the heartbeat or respiration, as additional model factors. The inclusion of nuisance factors in the experimental design allows researchers to reduce the amount of residual variance explained by the error term and consequently increases the significance of the contrasts. Nevertheless, it is important to bear in mind that more degrees of freedom are available if more factors are involved, and thereby the statistical corrections applied become more difficult.

Given the acquired functional volumes **Y** and defined design matrix **X**, it is possible to estimate the combination of $\hat{\boldsymbol{\beta}}$ values that best fits the experimental data **Y** with the predicted model by using a least-squares approach:

$$\hat{\boldsymbol{\beta}} = (X'X)^{-1}\mathbf{X}'\mathbf{Y} \tag{3}$$

$\hat{\boldsymbol{\beta}}$ values adjust the weight of each factor model and define the parameter estimates, indicating how much each factor contributes to the overall signal. Alternative approaches have to be taken into account when the inverted ($\mathbf{X}'\mathbf{X}$) does not exist.

Once the parameter estimates have been extracted, researchers must evaluate whether these parameters can explain the experimental data. Squared residuals give a measure of how well the model fits the data based on the estimation of its variance ($\hat{\sigma}$). The GLM assumes that residuals are independent and normally distributed after adjusting the model, but extensions of the generalized linear model account for non-normally distributed errors. The residual variance can be assessed as the residual sum-of-squares divided by the corresponding degrees of freedom.

$$\hat{\sigma}^2 = \frac{\hat{\varepsilon}^t\hat{\varepsilon}}{K-p} \tag{4}$$

$$p = \text{rank}(\mathbf{X}) \tag{5}$$

At this point, it is possible to quantify the significance of an effect of interest in a given voxel for a particular factor by comparing the amplitude of the parameter estimates with the distribution of the error measure. Thus, by applying a t-statistic, differences in the means between one-dimensional contrasts and the null hypothesis ($\beta_1 = 0$) can be tested. Afterwards, the significance of the effect (*P*-value) can be computed by comparing the t-value to a Student's t-distribution as a function of the available K-p degrees of freedom.

It is important to note that we have been treating each time series as independent, but time series are temporally correlated. Therefore, scans are not independent measures and their associated error at a given scan is correlated with its temporal neighbors. Accordingly, the number of degrees of freedom available is lower than the number of scans considered. Different approaches have been developed in order to deal with short serial correlations. As a specific example, statistical parametric mapping (SPM) uses an autoregressive plus white noise model (AR(1) + wn) to track the temporal covariance, making the assumption that the pattern of error correlations is the same over all the voxels but that its amplitude differs. Serial correlations are mainly caused by cardiac, respiratory, or vasomotor sources (Mitra et al., 1997).

Additionally, the combination of several beta parameters (a subset of the model) is of interest in some experiments. Here, inferences can be made based on an *F*-statistic approach by assuming identically distributed errors. Thus, under the null hypothesis, no linear combination of the effects accounts for significant variance.

Evidence for an effect is provided by significance thresholding. Classical thresholding methods are based on the rejection of the null hypothesis, which states that the distribution is the same for both conditions apart from differences due to random noise. However, testing for significant differences over the whole brain involves a large number of simultaneous statistical comparisons (i.e., one for every voxel). In this sense, significance values need to be lowered to account for possible false positive effects. Several approaches have been described for adjusting the statistical threshold to multiple comparisons.

Multiple-comparison correction

The simplest and more conservative approach used in order to circumvent the problem of the large number of simultaneous statistical tests is the *Bonferroni correction*. Information from nearby voxels tends to be correlated, since typically data has been spatially smoothed and it is difficult to determine the number of independent measures that exist. In this regard, this correction is too restrictive and

introduces the risk of increasing the false negative rate.

As a response to this problem of the Bonferroni-style correction, *random field theory* has been applied to functional data (Birn et al., 2002; Dale, 1999; Worsley et al., 1992) in order to generate an accurate theoretical significance threshold for spatially smoothed data. Smoothness can be assessed from the observed spatial correlation in the image, but is usually known from the width of the smoothing kernel (full-width half maximum; FWHM) previously applied. Based on this value, the number of independent comparisons, called *resells* (r), can be approximated by the quotient between the total number of voxels that belong to the brain and the number of voxels contained in the volume defined by the FWHM spatial filter applied. Additionally, due to the intrinsic spatial correlation, clustered blobs are expected instead of individual significant voxels. The *Euler characteristic* function predicts at high thresholds the number of blobs that should be found by chance in a random image at a given statistical threshold.

$$\mathrm{EC} = r(4\ln 2)(2\pi)^{-\frac{3}{2}} Z_t e^{-\frac{z_t^2}{2}} \qquad (6)$$

Thus, given a z-score (Z_t), the Euler characteristic function returns the significance threshold from which it is possible to conclude that any cluster that remains after thresholding has occurred by chance. This approach only accounts for the number of clusters, however, and it does not consider the size of the blob. Usually, regional differences (or activated voxels) are not expected to be found in isolated voxels. Thus, it is less likely to find a set of contiguous voxels that are all active by chance than it is to find a single voxel active by chance. Cluster-size thresholding or Family-wise approaches face this issue.

A *family-wise* null hypothesis can be tested in several ways, such as by selecting statistical values larger than the expected in the case of a null distribution or considering significant effects that survive at the cluster level after thresholding for the spatial extension of the cluster. Typically, cluster-size thresholds for imaging data are set around 10–20 voxels. However, these values depend on many parameters, such as the value of the threshold selected or the number of voxels in the imaging data.

Given these concerns, multiple comparisons approaches allow researchers to select the proper significance threshold for a voxel-based (functional) magnetic resonance imaging analysis. While an increase of sensibility is obtained applying such corrections, the localizing power is reduced since only clusters or sets of clusters can become significant. Such approaches might not be sensitive enough to detect biologically meaningful differences, which could be otherwise detected at the voxel level. Additionally, when a particular hypothesis exists about an anatomical region involved in a concrete process, the number of voxels tested is largely reduced, minimizing the need of multiple comparison correction. ROI approaches have several advantages when a priori expectations exist for a specific region.

Fixed-effect analysis versus random-effect analysis

Up to this point, we have focused on single subject analysis. The combination of data across subjects allows researchers to test experimental hypotheses in a more traditional way. In early studies of neuroimaging, a fixed-effect approach was employed. *Fixed-effects analysis* assumes that experimental manipulations have the same effect in terms of functional activation across all subjects aside from random noise. Based on this assumption, data from different subjects is concatenated and treated as if it came from a single-subject analysis. Fixed-effect analysis enhances the sensitivity of the statistical analysis due to the averaging across subjects. Inter-subject variability is not taken into account, however, restricting the analysis to the sample collected. Nevertheless, in neuroimaging studies, we usually want to make inferences at a more general level about the population from which the subjects are drawn. Random-effect approaches deal with inter-subject differences

by considering that the experimental manipulation has a distribution of the effect of interest across subjects. *Random-effect analysis* is implemented by a two-level model. At the first level, standard statistical maps are estimated for each subject for the comparison of interest. Then, in a second level, the distribution of individual statistical patterns is tested for significance. Therefore, both within-subjects and between-subjects variance sources are controlled. Random-effect analysis is the most common approach implemented in functional analysis to make interferences about the population from which the subjects are drawn.

We have been describing the established standards in fMRI analysis, however, several limitations of this approach should be considered. Classical approaches are mainly mass-univariate analyses, meaning the analysis is constrained to the voxel-based level. Such analyses basically result in a static picture of the brain in action. By contrast, there is increasing evidence that brain functions are supported by integrated brain networks flexibly co-operating during task execution. Voxel-based inferences might not be the most appropriate approach to assess such dynamic interactions, since the whole brain information is not exploited. Multivariate approaches can be considered as an alternative. Pattern recognition methods, for instance, have been used to analyze fMRI data with the goal of decoding the information represented in the whole brain at a particular time (Haynes and Rees, 2006) by using sophisticated classifiers, such as support vector machine (SVM) approaches or artificial neural networks among many others. Principal component analysis (PCA) or single value decomposition (SVD) approaches have also been incorporated in functional analysis, but their main power still remains in de-noising the data before preprocessing.

Additionally, classical inferences are defined in a standard parametric domain. However, an increasing number of studies have applied non-parametric tests, such as permutation tests and Bayesian frameworks, in order to enhance the sensitivity of the analyses. Nevertheless, classical parametric models combined with brain connectivity analyses still persist as the most common approaches used to study brain functional specialization and dynamics.

BRAIN CONNECTIVITY: EFFECTIVE AND FUNCTIONAL CONNECTIVITY

In many brain imaging experiments, multiple areas are found to be activated in a given task. A pressing question in cognitive neuroscience, therefore, is to determine how such areas act in concert as functional networks. If different regions are involved in a network, they should have strongly correlated activity patterns. Indeed, the inherently multivariate nature of functional MRI allows researchers to investigate how anatomically distant brain regions interact during a specific cognitive task. The concepts of functional and effective connectivity were introduced by Friston et al. (1993) to identify functional networks. Both concepts assume that temporal correlations in the BOLD signal reflect synchronous neural firing in the interacting regions. In particular, functional connectivity measures the temporal correlation between spatially remote neurophysiological events. In contrast, effective connectivity is defined as the influence of one neural system on another. Therefore, functional connectivity is operationally defined, whereas effective connectivity depends on a specific model.

Functional connectivity

Correlational analysis

Functional connectivity basically assesses the correlation between a small number of preselected regions or between voxels. Then, the correlation field can be introduced as a matrix of all pair-wise correlation coefficients that indicating those regions coactivated with a particular activation pattern. Notice that voxels within the FWHM should not be taken into account, because high correlations are introduced by the smoothing process.

The autocorrelation matrix can also be transformed and thresholded by applying a *t*-test.

Coherence analysis

Coupling between neural components might result in phase locking of their BOLD response (i.e., increased coherence). Coherence provides a measure of the frequency-specific association between two time series, x and y, at a frequency λ (Sun et al., 2004). It is defined as the cross-spectrum of two time series, $f_{xy}(\lambda)$, normalized by the power spectrum $f_{xx}(\lambda)$ of each time series:

$$\mathrm{Coh}_{xy}(\lambda) = \frac{\left|f_{xy}(\lambda)\right|^2}{f_{xx}(\lambda)\, f_{yy}(\lambda)} \tag{7}$$

where:

$$f_{xy}(\lambda) = \sum_{u=1}^{n} \mathrm{Cov}_{xy}[u] e^{-j\lambda \cdot u} \tag{8}$$

$$f_{xx}(\lambda) = \sum_{u=1}^{n} \mathrm{Cov}_{xx}[u] e^{-j\lambda \cdot u} \tag{9}$$

where the cross-covariance function is defined below for a stationary time series x and y as:

$$\mathrm{Cov}_{xy}[u] = E\{(x[t] - \mu_x)(y[t+u] - \mu_y)\} \tag{10}$$

$E\{\}$ denotes the expected value and μ the mean time series value.

Eigenimages-singular value decomposition

SVD seeks to reduce the dimensions of the correlation structure to a small number of weighted orthogonal modes (principal components) (Worsley et al., 2005). SVD transforms the original time-series (**X**), where each column represents a voxel and each row a scan of mean-corrected data, into two sets of unitary orthogonal matrices (**V**, **U**) and a **S** diagonal matrix of decreasing singular values:

$$\mathbf{X} = \mathbf{USV}' \tag{11}$$

Columns of **V** define the spatial components, representing a distributed brain system that can be displayed as an image (eigenimage). Columns of **U** represent temporal or scan-order patterns accounting for the time-dependent profiles associated with each eigenimage (eigenvariate). Eigenvalues are the singular values squared of the S matrix, estimating the relative amount of variance accounted for by each principal component. They permit a qualitative assessment of the importance of each eigenimage/eigenvariate. Thus, the first eigenimage expresses the pattern over voxels that accounts for the greatest variability across all the scans, whereas the first eigenvariate is the temporal pattern that reflects the greatest variability across all voxels.

A comparison between correlation-based and SVD approaches shows that thresholding correlations directly yield those voxel pairs that are highly correlated whereas SVD detects connectivity patterns in a more qualitative way. Worsley et al. (2005) showed that correlations are highly sensitive to detect focal interactions in practice, whereas SVD are more prone to detecting connections between more extensive regions. Additionally, the main drawback of using SVD lies in selecting the appropriate number of components that should correspond to the number of networks under study in a particular experiment. More components will lead to a more complete but more complex description of the system. Therefore, only few principal components are typically selected, and those components that only contribute minimally to the explained variance are neglected.

Effective connectivity

The parameters introduced so far allow us to identify those regions involved in a particular brain system by demonstrating high correlations between them. The concept of effective connectivity takes this idea a step further, because it seeks to describe and quantify the influence of one brain area upon another. In the following section, we will describe three approaches to capture effective connectivity: (1) psycho- physiological interactions

(PPI); (2) structural equation modeling; and (3) dynamic causal modeling (DCM).

Psycho-physiological interactions

PPIs can be understood as the modulation that one cerebral region exerts over another in a specific experimental context (Friston et al., 1997). In other words, if the activity of one region is regressed in terms of another, the slope of this regression reflects the influence of the second area over the first one. Variations in the slope of such regressions are related to changes in experimental cognitive conditions. Thus, regressions are computed for every voxel separately for each condition, and then inferences are made between the different experimental conditions.

Structural equation modeling

SEM is model-dependent, requiring the *a priori* specification of an anatomical model that graphically defines those anatomical connections considered to be functionally relevant (Goncalves and Hall, 2003). The connectivity model states not only which regions are connected to each other but also the direction of the connection. Note that only a small number of regions can be included in such a model; otherwise, computational problems arise. Thus, SEM estimates the connection strengths (path coefficients) that best predict the inter-regional covariances of the functional imaging data under the given model. Further information about SEM can be found in Chapters 21 to 25.

Dynamic causal modeling

DCM shifts the focus from regionally-specific activations to inter-regional path-specific activations using a dynamic deterministic non-linear model. The basic idea is to model the brain as a dynamic system that is subject to inputs and produces outputs in terms of parameters that represent the coupling between unobserved brain states. Thus, inputs modulated by the experimental conditions can induce neuronal responses in specific anatomical regions, but they also might change the effective connectivity by influencing the coupling between nodes.

The main idea is to submit the system to different controlled experimental inputs that directly generate variation in the outputs. By measuring the responses after perturbing the system, free model parameters can be estimated. Effective connectivity can be expressed following any non-linear function characterizing the neurophysiological inputs related to a brain region relative to other regions. DCM is also supported with a forward model, which transforms the neural responses to a measurable hemodynamic response (i.e., the output). In general, DCM does not restrict the number of connections that can be modeled, and consequently a large number of free parameters have to be estimated. Several constraints have to be imposed (for example, the neural activity cannot diverge exponentially to infinite values). A Bayesian framework is an appropriate approach for tackling such an analysis. A complete mathematical explanation of DCM can be obtained from Friston et al. (2003).

Additional approaches for connectivity mapping include multidimensional scaling (Welchew et al., 2002) and hierarchical clustering (Stanberry et al., 2003). These use dissimilarities rather than partial correlations, and thus they permit the use of complementary multivariate techniques.

COMMON STATISTICAL PROCEDURES IN ELECTROENCEPHALOGRAPHY AND FMRI: INDEPENDENT COMPONENT ANALYSIS

One of the classical assumptions in the study of the physiological correlates of behavior is that the physiological responses (i.e., cerebral electrical signals registered using EEG or the fMRI BOLD signal) are stationary with properties that are constant over time. For example, the mean of all electrical signals time-locked to a certain event will give an evoked potential as a result, and this potential is supposed to be constant over time. However, several

studies have shown that EEG (Blanco, Garcia, Quiroga, Romanelli, and Rosso, 1995) and fMRI signals (Turner and Twieg, 2005) show non-stationary behavior. Although traditional analysis approaches of brain signals have provided relevant information about the implementation of cognitive functions, other analytical techniques not based on stationary *a priori* are needed. One such technique that has become increasingly popular over the past few years is independent component analysis (ICA). ICA had been developed to solve technical problems, such as the separation of multiple human voices in a cocktail party setting recorded with several microphones. It turns out that many other problems, including the denoising of images, face or speech recognition, or extracting the components of brain activity related to an event, lend themselves to treatment in a similar way.

Independent component analysis applied to the study of brain signals

Brain signals recorded at a scalp-electrode (EEG) or from a voxel in the brain (fMRI) can be thought of as representing a mixture of several signals coming from different sources (see Figure 29.5 and 29.6). Three sources located at different places in the brain, each with its own temporal evolution, generate a signal at the scalp that is the sum of the three attenuated signals. Decomposing the signals that generate the recorded response is not trivial. Importantly, each source is associated with a certain scalp map (see Figure 29.5C), and temporal evolution of each source occurs independently. The goal of ICA applied to EEG data is to find sources (or components) with independent temporal evolutions (and their associated scalp maps) based on the time course of the activity at the different scalp electrodes (see Figure 29.5B and D). The components can then be examined using the standard logic of ERP research (see Makeig et al., 2002) and applied to fMRI (Esposito et al., 2005).

Another application of ICA to EEG and fMRI signals involves separating the brain signals related to the performance of a given task from the signals related to other factors (e.g., noise and movements). Signals produced by movements usually comprise lower frequencies than signals evoked by a visual simulus and are independent of the presentation of a stimulus. Given that the statistical properties of these signals are different, ICA is able to separate brain signals from other signals, such as ocular movements, muscular artifacts, and electrical noise. It can thus be used to denoise EEG data. Moreover, ICA can also be useful in denoising fMRI data (Figure 29.6) and studying BOLD activity related to the events presented to the subject (McKeown et al., 1998). For a mathematical specification of ICA, see Appendix 2.

Limitations and problems in the use of independent component analysis

The ICA approach presents some limitations that should be carefully considered before applying this method. The most important problems are:

1. There exists an ambiguity with regard to the sign of the maps (unmixing matrix). Given that the original data is recovered by multiplying the unmixing matrix by the activations, this may lead to ambiguous results.
2. Contrary to PCA, there is no equivalent to 'variance explained' in ICA, and hence it is difficult to establish an 'order' of the different components. In other words, the specification of the 'component that explains the most of the variance' is delicate in the ICA case. To solve this problem, it has been suggested to define the activations x(t) as unit variance and suppose that the columns of the mixing matrix reflect power of each component in the space. Although this can help in the determination of the 'importance' of each component in the decomposition, however, the ambiguity in the determination of an order between components still persists (James and Hesse, 2005).
3. ICA can decompose the data maximally into as many components as there are sensors. This number usually ranges from 19 to 256 in EEG situations, but it can be increased by a factor of two in MEG cases and 1000 in fMRI situations. Thus, a method for determining the

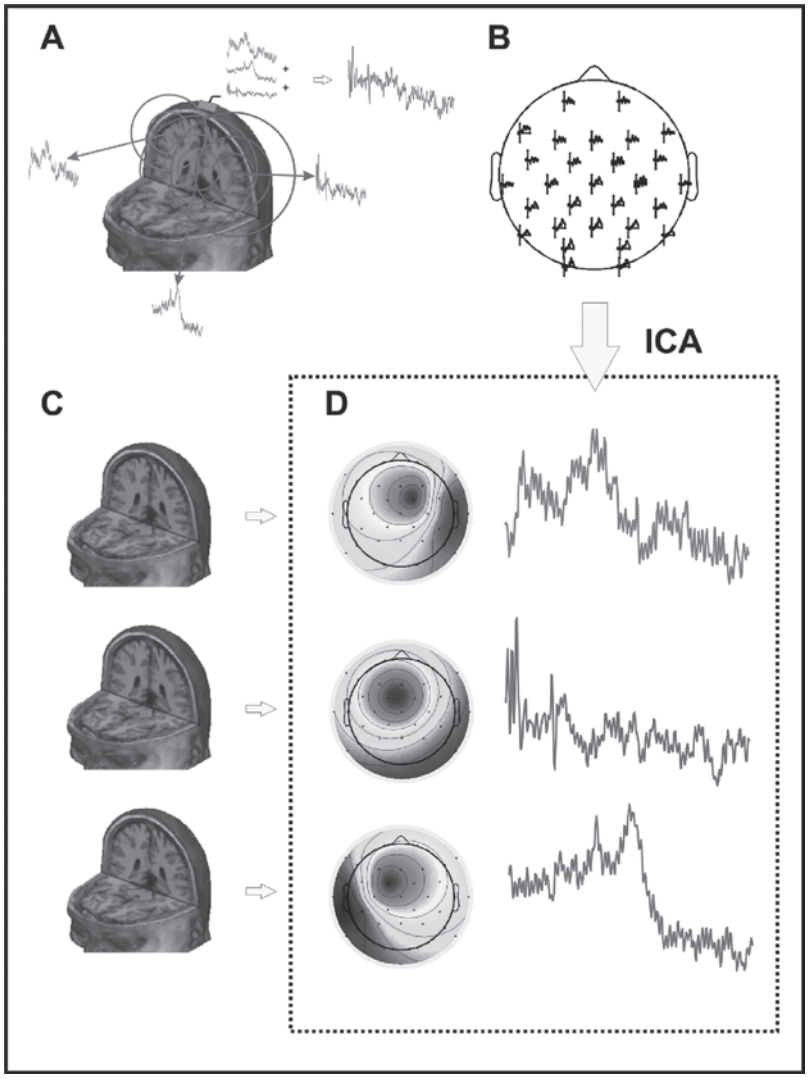

Figure 29.5 Application of independent component analysis (ICA) to electroencephalography (EEG) data. Three independent sources in the brain (A) generate a scalp signal that is the sum of their respective contributions (B). The ICA procedure finds three independent components (D), each displaying a specific scalp distribution and a temporal evolution that coincides with the original temporal evolution of the sources. In addition, the application of an inverse solution to the scalp maps of independent components permits the localization of the original sources (B). Refer to color plates at the end of this volume for a colored version of this figure.

number of components in the solution would be desirable. Although some methods have proposed to solve this problem (i.e., PCA based, Hyvarinen, Karhunen, and Oja, 2001), however, an ideal solution for this problem has not yet been found. In general, the most common method involves including a relatively high number of components and selecting those that either make 'physiological' sense or can be related to some kind of noise.

CONCLUSIONS

In the preceding pages, we have discussed some of the most important analysis

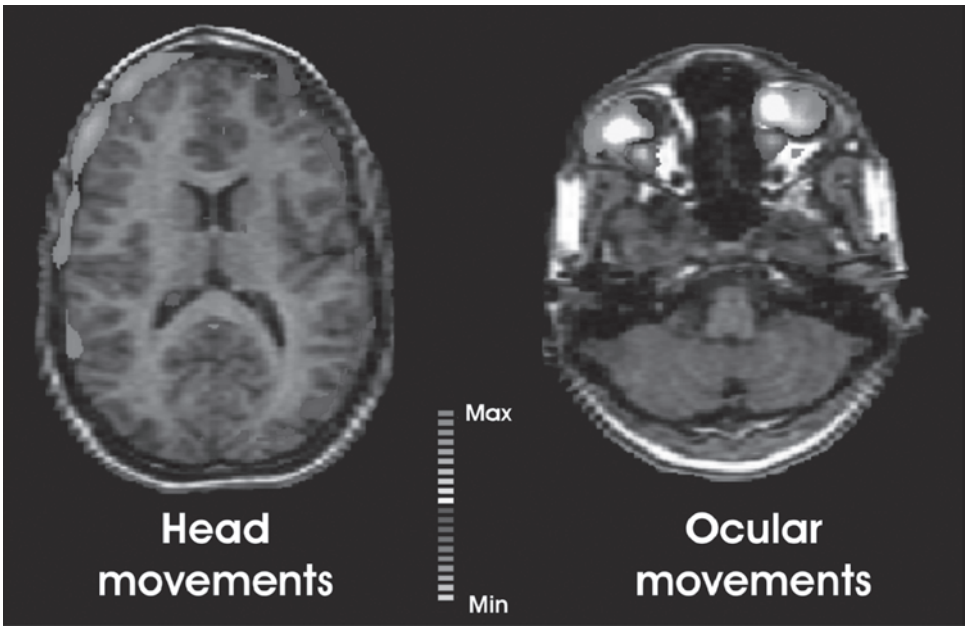

Figure 29.6 Application of independent component analysis (ICA) in the denoising of fMRI data. Two independent components account for head movements (left) and ocular movements (right). Refer to color plates at the end of this volume for a colored version of this figure.

techniques in (f)MRI research. It is important to bear in mind that these technical developments are not self-serving but are necessary prerequisites for the successful application of temporal and spatial neuroimaging techniques in cognitive neuroscience. In fact, the progress of this research field has resulted largely from the rapid development of new analysis tools. Recent years have seen an increase in the speed of such new developments owing to huge investments in neuroimaging centers in many countries throughout the world. Some of the more advanced analysis techniques (such as dynamic causal modeling and related methods) have just begun to penetrate the field and may significantly change the way cognitive neuroscientists think about the brain in action.

Because of space limitations, we have not been able to discuss all of the relevant methodological developments in the neuroimaging field. Because of the very different kind of information that EEG-based and MRI-based techniques deliver, it may be desirable in some cases to combine the two techniques. This maybe achieved rather informally by running separate EEG and fMRI experiments with similar or identical stimulus material in different groups of participants (for example, see the complementary results obtained in Bahlmann et al., 2007; Matzke et al., 2002) or very stringently by recording EEG and fMRI signals simultaneously in one session. Recording of EEG signals within the MRI scanner has become possible because of sophisticated methods that denoise EEG from the gradient artifacts introduced by the scanning procedure (see Laufs et al., 2007 for a review). Such simultaneous measurements can provide important constraints for source modeling approaches by seeding dipole sources in those regions activated in the fMRI experiment. It has to be pointed out, however, that such an approach is not without dangers, since one might be inclined to ignore the very different nature of the two signals.

As combined hybrid PET/MRI devices are now available, other important developments will likely result from combining PET and

MRI measurements in the same session. Such devices will allow researchers to investigate task-related molecular (e.g., receptor availability/binding) and blood flow changes simultaneously, thus providing an important link to the knowledge accruing in molecular neuroscience. All of these methodological developments are very promising. It is important to bear in mind, however, that sound experimental designs based on sophisticated models of the cognitive processes under study will remain at the heart of cognitive neuroscience.

NOTE

High resolution color figures in this chapter can be found at www.brainvitage.org/HQMP/

REFERENCES

Aguirre, G.K., Zarahn, E. and D'Esposito, M. (1998) 'The variability of human, BOLD hemodynamic responses.', *Neuroimage*, 8: 360–369.

Amaro, E. Jr. and Barker, G.J. (2006) 'Study design in fMRI: basic principles', *Brain and Cognition*, 60: 220–232.

Ashburner, J. and Friston, K.J. (2000) 'Voxel-based morphometry – the methods', *Neuroimage*, 11: 805–821.

Bahlmann, J., Rodriguez-Fornells, A., Rotte, M. and Münte, T.F. (2007) 'An fMRI study of canonical and noncanonical word order in German', *Human Brain Mapping*, 28: 940–949.

Bandettini, P.A. and Cox, R.W. (2000) 'Event-related fMRI contrast when using constant interstimulus interval: theory and experiment', *Magnetic Resonance Medicine*, 43: 540–548.

Birn, R.M., Cox, R.W. and Bandettini, P.A. (2002) 'Detection versus estimation in event-related fMRI: choosing the optimal stimulus timing', *Neuroimage*, 15: 252–264.

Blanco, S., Garcia, H., Quiroga, R.Q., Romanelli, L., & Rosso, O.A. (1995) 'Stationarity of the EEG series', *IEEE Engineering in Medicine and Biology Magazine*, 14: 395–399.

Buxton, R.B., Wong, E.C. and Frank, L.R. (1998) 'Dynamics of blood flow and oxygenation changes during brain activation: the balloon model', *Magnetic Resonance Medicine*, 39: 855–864.

Camara, E., Bodammer, N., Rodriguez-Fornells, A. and Tempelmann, C. (2007) 'Age-related water diffusion changes in human brain: a voxel-based approach', *Neuroimage*, 34: 1588–1599.

Dale, A.M. (1999) 'Optimal experimental design for event-related fMRI', *Human Brain Mapping*, 8: 109–114.

de Diego, B.R., Rodriguez-Fornells, A., Rotte, M., Bahlmann, J., Heinze, H.J. and Münte, T.F. (2006) 'Neural circuits subserving the retrieval of stems and grammatical features in regular and irregular verbs', *Human Brain Mapping*, 27: 874–888.

D'Esposito, M., Deouell, L.Y., and Gazzaley, A. (2003) 'Alterations in the BOLD fMRI signal with ageing and disease: a challenge for neuroimaging', *Nature Review Neuroscience*, 4: 863–872.

Esposito, F., Scarabino, T., Hyvarinen, A., Himberg, J., Formisano, E., Comani, S., Tedeschi, G., Goebel, R., Seifritz, E., and Di, S.F. (2005) 'Independent component analysis of fMRI group studies by self-organizing clustering', *Neuroimage*, 25: 193–205.

Friston, K.J., Buechel, C., Fink, G.R., Morris, J., Rolls, E. and Dolan, R.J. (1997) 'Psychophysiological and modulatory interactions in neuroimaging', *Neuroimage*, 6: 218–229.

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', *Journal of Cerebral Blood Flow and Metabolism* 13: 5–14.

Friston, K.J., Harrison, L. and Penny, W. (2003) 'Dynamic causal modelling', *Neuroimage*, 19: 1273–1302.

Goncalves, M.S. and Hall, D.A. (2003) 'Connectivity analysis with structural equation modelling: an example of the effects of voxel selection', *Neuroimage*, 20: 1455–1467.

Goutte, C., Nielsen, F.A., and Hansen, L.K. (2000) 'Modeling the haemodynamic response in fMRI using smooth FIR filters', *IEEE Trans. Med. Imaging*, 19: 1188–1201.

Haynes, J.D. and Rees, G. (2006) 'Decoding mental states from brain activity in humans', *Nature. Reviews in Neuroscience,* 7: 523–534.

Hyvarinen, A., Hoyer, P.O., and Inki, M. (2001) 'Topographic independent component analysis', *Neural Comput.*, 13: 1527–1558.

James, C. J., and Hesse, C.W. (2005) 'Independent component analysis for biomedical signals', *Physiol Meas.*, 26: R15–R39.

Laufs, H., Daunizeau, J., Carmichael, D.W. and Kleinschmidt, A. (2007) 'Recent advances in recording electrophysiological data simultaneously with magnetic resonance imaging', *Neuroimage*, 40 (2): 515–528.

Logothetis, N.K. (2003) 'The underpinnings of the BOLD functional magnetic resonance imaging signal', *Journal of Neuroscience*, 23: 3963–3971.

Logothetis, N.K., Guggenberger, H., Peled, S. and Pauls, J. (1999) 'Functional imaging of the monkey brain', *Nature. Neuroscience*, 2: 555–562.

Marco-Pallares, J., Grau, C., and Ruffini, G. (2005) 'Combined ICA-LORETA analysis of mismatch negativity', *Neuroimage*, 25: 471–477.

Makeig, S., Westerfield, M., Jung, T.P., Enghoff, S., Townsend, J., Courchesne, E. and Sejnowsky, T.J. (2002) 'Dynamic brain sources of visual evoked responses', *Science*, 295: 690–694.

Matzke, M., Mai, H., Nager, W., Russeler, J. and Münte, T. (2002) 'The costs of freedom: an ERP: study of non-canonical sentences', *Clinical Neurophysiology*, 113: 844–852.

McKeown, M.J., Makeig, S., Brown, G.G., Jung, T.P., Kindermann, S.S., and Sejnowski, T.J. (1998) 'Analysis of fMRI data by blind separation into independent spatial components', *Human Brain Mapping*, 6: 160–188.

Miezin, F.M., Maccotta, L., Ollinger, J.M., Petersen, S.E. and Buckner, R.L. (2000) 'Characterizing the hemodynamic response: effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing', *Neuroimage*, 11: 735–759.

Mitra, P.P., Ogawa, S., Hu, X. and Ugurbil, K. (1997) 'The nature of spatiotemporal changes in cerebral hemodynamics as manifested in functional magnetic resonance imaging', *Magnetic Resonance Medicine*, 37: 511–518.

Ogawa, S., Lee, T.M., Nayak, A.S. and Glynn, P. (1990) 'Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high magnetic fields', *Magnetic Resonance Medicine*, 14: 68–78.

Stanberry, L., Nandy, R. and Cordes, D. (2003) 'Cluster analysis of fMRI data using dendrogram sharpening', *Human Brain Mapping,* 20: 201–219.

Sun, F.T., Miller, L.M. and D'Esposito, M. (2004) 'Measuring interregional functional connectivity using coherence and partial coherence analyses of fMRI data', *Neuroimage*, 21: 647–658.

Turner, G.H. and Twieg, D.B. (2005) 'Study of temporal stationarity and spatial consistency of fMRI noise using independent component analysis', *IEEE Transactions on Medical Imaging*, 24: 712–718.

Virta, A., Barnett, A. and Pierpaoli, C. (1999) 'Visualizing and characterizing white matter fiber structure and architecture in the human pyramidal tract using diffusion tensor MRI', *Magnetic Resonance Medicine*, 17: 1121–1133.

Visscher, K.M., Miezin, F.M., Kelly, J.E., Buckner, R.L., Donaldson, D.I., McAvoy, M.P., Bhalodia, V.M., and Petersen, S.E. (2003) 'Mixed blocked/event-related designs separate transient and sustained activity in fMRI', *Neuroimage*, 19: 1694–1708.

Welchew, D.E., Honey, G.D., Sharma, T., Robbins, T.W. and Bullmore, E.T. (2002) 'Multidimensional scaling of integrated neurocognitive function and schizophrenia as a disconnexion disorder', *Neuroimage*, 17: 1227–1239.

Worsley, K.J., Chen, J.I., Lerch, J. and Evans, A.C. (2005) 'Comparing functional connectivity via thresholding correlations and singular value decomposition', *Philosophical Transactions of the Royal Society of London, B. Biological Sciences*, 360: 913–920.

Worsley, K.J., Evans, A.C., Marrett, S. and Neelin, P. (1992) 'A three-dimensional statistical analysis for CBF activation studies in human brain', *Journal of Cerebral Blood Flow and Metabolism*, 12: 900–918.

Yacoub, E., Shmuel, A., Pfeuffer, J., Van De Moortele, P.F., Adriany, G., Ugurbil, K., and Hu, X. (2001) 'Investigation of the initial dip in fMRI at 7 Tesla', *NMR Biomedicine*, 14: 408–412.

APPENDIX 1

Characterization of the hemodynamic response function

Great effort has been placed on characterizing the hemodynamic response function (HRF). Several approaches have been developed to achieve this goal. The most general approach is the finite impulse response (FIR) model, in which it is possible to capture any shape of response up to specified time scale from a linear combination of specific basis functions. In this model, the activation of a particular voxel at time t is defined as the weighted sum of the stimuli (s_i) at the preceding (n) points (Goutte, Nielsen, and Hansen, 2000). Formally:

$$y_t = \sum_{i=1}^{n} a_i s_{t-(i-1)} + a_o + \varepsilon \qquad \text{(A1.1)}$$

where y_t corresponds to the intensity value at time t, ε follows a Gaussian noise distribution, and a_0 is a constant parameter. Generally, a least squares approach is used to estimate those parameters (a_i) that minimize the

estimated function with the observed data. However, the signal to noise ratio provided by the hemodynamic response is quite low, and many parameters have to be modeled. Indeed, the most frequent choice is the canonical HRF. This approach achieves a reasonable good fit with the impulse response function but reduces the degrees of freedom in the model and consequently allows powerful statistical tests. The canonical HRF (or *double gamma function*) is defined from the difference of two gamma probability density functions:

$$H(t) = f(t; 6, 1) - \frac{1}{6} f(t; 16, 1) \qquad \text{(A1.2)}$$

where:

$$f(t; \alpha, \beta) = \frac{t^{\alpha-1} e^{-t/\beta}}{\beta^{\alpha} \Gamma(\alpha)} \qquad \text{(A1.3)}$$

The first term models the peak of the HRF, whereas the second is responsible for the post-stimulus undershoot. Nevertheless, differences between hemodynamic response patterns exist from one brain region to another and across subjects. These differences are mostly due to variation in the vasculature, although non-linearities at the neuronal level (e.g., the adaptive behavior of neural activity) can induce hemodynamic differences as well (Logothetis, 2003). This variability can be accommodated by expanding the HRF in terms of temporal basis functions. The use of multiple basis functions provides more variability in the shape of the hemodynamic response, such as differences between the latency of the peak or in the peak delay, than the canonical HRF. However, it is important to comment that the BOLD response depends directly on the hemodynamic properties of the surrounding vasculature. Therefore, alterations in vascular dynamics might influence our ability to attribute BOLD signal changes to alterations in neural activity. In particular, since changes in vasculature are related to aging, it is important to be careful when interpreting BOLD changes in elderly populations or when vasculature alterations could be possible (D'Esposito, Deouell, and Gazzaley, 2003).

APPENDIX 2

Mathematical bases of the independent component analysis

The initial point of ICA involves supposing that signal *s*(t) is a mixture of statistically independent signals *x*(t) with the relation:

$$\mathbf{s} = \mathbf{A}\mathbf{x} \qquad \text{(A2.1)}$$

The goal of ICA is to find the inverse relation:

$$\mathbf{x} = \mathbf{W}\mathbf{s} \qquad \text{(A2.2)}$$

where W is the so-called unmixing matrix. We can suppose that our signal is composed of several elements: increase/decrease of cerebral activity related to behavioral processes of the experiment; noise that can be random, due to non-cerebral signals as movements, muscular, or cardiac activity, or due to electrical or magnetic activity; slow processes related to the experiment. All of these elements will present a temporal evolution x_j(t) and behave independently one each other, i.e.:

$$f(x_1 \ldots x_m) = f_1(x_1) \ldots f_m(x_m) \qquad \text{(A2.3)}$$

where $f()$ is the joint density of signals and $f_j()$ is the marginal density of x_j. In addition, all these processes can change in time (i.e., evoked cerebral activity by different stimuli can decrease with time due to habituation), given the above mentioned non-stationary quality of the signal. Even when the electrical activity changes on time, however, we suppose that its sources will not change. Hence, certain activity x_j(t) will be associated with a fixed map (e.g., a scalp potential distribution in EEG or 3D activity map in fMRI) that will be represented in matrix $\mathbf{W}.,j$.

It is important to note that the concept of 'statistical independence' is much more restrictive than other concepts used

in high-order statistics. For example, two variables x_1 and x_2 are not correlated if:

$$E\{x_1x_2\} = E\{x_1\}E\{x_2\} \quad \text{(A2.4)}$$

Non-correlation is a weak form of independence given that two sources statistically independent are also non-correlated. The opposite argument is only true when both distributions are Gaussian.

The application of ICA to study EEG or fMRI is performed under the following assumptions:

1. Component maps (activity maps for fMRI or projection of sources in the scalp in EEG) must be constant in time but not in their temporal evolution
2. Temporal activation of sources must be statistically independent
3. The statistical distribution of activations is not Gaussian. This condition has to be applied because ICA uses higher order measures than other methods traditionally used in the analysis of cerebral signals (e.g., Principal Component Analysis, PCA, and Factor Analysis, FA) that are based on second order measures (e.g., searching maximum variance in PCA),. The former methods need $x(t)$ distributions to be Gaussian, whereas $x(t)$ in ICA need to be sub- or super-Gaussian.

In addition, we have to assume a fourth condition in the EEG case:

4. The signal conduction times are equal (instantaneous in practice), and the sum of the sources in the electrodes is linear. This condition is followed in the EEG situation, since it is reasonable to apply the quasistatic approach to frequencies involved in brain electrical activity ($<$1 kHz).

The application of ICA to EEG data has been demonstrated to be very useful in denoising data and studying the components involved in the generation of an ERP (Marco-Pallares, Grau, and Ruffini, 2005). On the other hand, the use of ICA in the study of fMRI data is more recent than their application to cerebral electrical activity. However, some studies have demonstrated that ICA can be useful in denoising as well as studying BOLD-related activity to the events presented to the subject. In this sense, McKeown et al., (1998) demonstrated that the ICA applied to fMRI data revealed several statistical independent components, such as activations related to the events or blocks presented, slow activations, and activity related to fast and slow movements (Figure 29.6).