

The parcellation of cortical areas using replicator dynamics in fMRI

Jane Neumann,* D. Yves von Cramon, Birte U. Forstmann, Stefan Zysset, and Gabriele Lohmann

Max-Planck-Institute for Human Cognitive and Brain Sciences, Stephanstrasse 1a, D-04103 Leipzig, Germany

Received 2 February 2005; revised 24 February 2006; accepted 27 February 2006

Available online 27 April 2006

In this paper, we show that replicator dynamics can be used as an exploratory analysis tool to detect subregions of cortical areas on the basis of the similarity between fMRI time series. As similarity measure, we propose to use canonical correlation, a multivariate extension to the typically employed Pearson's correlation coefficient. We applied the replicator process to data obtained from two different experimental paradigms in the search for subregions within the left lateral frontal cortex (LFC). In both cases, the replicator process resulted in a parcellation that corresponds to a recently suggested subdivision of the LFC in anterior–posterior direction. Most notably, these results were very consistent when compared across different measurements of a single subject and across a group of subjects.

© 2006 Elsevier Inc. All rights reserved.

Introduction

One of the main goals in functional magnetic resonance imaging (fMRI) is the parcellation of cortical areas with respect to their different functionality. The usual approach is to design a functional experiment such that it consists of a number of experimental conditions which are assumed to involve different cognitive processes. Hemodynamic responses corresponding to two or more such conditions are then compared, and the statistical significance of the contrast is represented in a statistical parametric map. If a cortical area shows a significant contrast between two conditions, this area is assumed to pertain to a functionality that is necessary to process one condition but not the other. This model-based approach thus facilitates the dissociation of cortical areas by directly manipulating the underlying cognitive processes.

In recent years, exploratory analysis techniques such as Principal Component Analysis (Friston et al., 1999; Thirion and Fugeras, 2003; Viviani et al., 2005), Independent Component Analysis (McKeown et al., 1998; Calhoun et al., 2003), and clustering (Goutte et al., 1999; Baumgartner et al., 1998; Dodel et al., 2002) have provided a viable alternative to the model-

based statistical analysis of fMRI data. Such techniques explore the inherent structure of the data independently of the experimental design or any pre-defined model of the hemodynamic response. Clustering, for example, is based on the assumption that the pattern of cortical activation is structured such that activated voxels can be divided into few groups according to their similarity (Goutte et al., 1999). The most commonly used clustering methods like hard or fuzzy k -means or hierarchical clustering perform this grouping directly on the fMRI time series. Other techniques such as cross-correlation or spectral clustering operate on an interaction or similarity matrix which is derived from the correlation or covariance of fMRI time series (Dodel et al., 2002; Voultsidou et al., 2004; Worsley et al., 2005).

Recently, replicator dynamics was introduced as an exploratory data analysis technique to functional imaging (Lohmann and Bohn, 2002). The aim of the present study is to show that the iterative application of replicator dynamics can be employed to detect subregions in cortical areas which in model-based statistical analyses appear as homogeneous activated regions. The method can be viewed as a form of clustering in the sense that similar items are grouped together and separated from less similar ones. Specifically, it is related to cross-correlation or spectral clustering in that it exploits the structure of a similarity matrix derived from the correlation between fMRI time series. However, replicator dynamics differs from most clustering techniques in two aspects which, we would argue, make it a particularly suitable tool for this application.

Firstly, groups or clusters emerging from a replicator process exhibit different coherence properties than results from most other clustering techniques. While clustering typically identifies star-shaped topologies, where each cluster member is closely related to a single representative in the cluster center, replicator dynamics searches for network topologies where each network member is closely related to all other network members (Lohmann and Bohn, 2002). This complete pairwise similarity seems a plausible assumption on a functional cortical area, as such a region should only consist of voxels that all reflect the same underlying functionality. While in some applications this difference to other clustering techniques might be negligible, we will show that it can also lead to quite different solutions when compared, for example, with results from the closely related Principal Component Analysis.

* Corresponding author. Fax: +49 34 19940 221.

E-mail address: neumann@cbs.mpg.de (J. Neumann).

Available online on ScienceDirect (www.sciencedirect.com).

Secondly, in contrast to most other clustering techniques, replicator dynamics does not attempt to partition the entire feature space, i.e. to assign each voxel to one cluster. Rather, the replicator process selects from all input voxels groups of voxels exhibiting strong coherence properties among each other. This suits the task of detecting groups of strongly coherent voxels within a region of interest as voxels not belonging to the “core” of the region will not be artificially assigned to one of the clusters.

Replicator dynamics originated from theoretical biology (Fisher, 1930; Schuster and Sigmund, 1983) but today has a wide range of applications, including game theory, optimization, hierarchical pattern matching, and finding maximum cliques in graphs (Taylor and Jonker, 1978; Hofbauer and Sigmund, 1988a; Menon et al., 1995; Pelillo, 2001; Pelillo et al., 1999; Bomze et al., 2000). In the context of functional neuroimaging, it can be employed on various levels, reaching from the analysis of single subject data to meta-analytic studies (Lohmann and Bohn, 2002; Neumann et al., 2005).

We applied replicator dynamics to two fMRI experiments addressing cortical activation related to cognitive control processes: a color–word matching Stroop task (Stroop, 1935) and a cuing version of the task switching paradigm (Forstmann et al., 2005). Our results show a parcellation in the left lateral frontal cortex (LFC) in anterior–posterior direction. This is in accordance with recent findings suggesting a functional gradient within the left LFC (Forstmann et al., 2005).

The only prerequisite for the application of a replicator process is a non-negative, real-valued, symmetric similarity matrix, which in our case encodes the similarity between fMRI time series. This facilitates the use of different similarity measures such as the previously used Pearson’s correlation coefficient and Spearman’s rank correlation (Lohmann and Bohn, 2002). We propose to use canonical correlation as a measure for the similarity of fMRI time series. The rationale behind this choice is as follows.

One critical question in the analysis of functional neuroimaging data is whether the obtained results are generalizable. That is, we wish to be able to compare results between different imaging sessions and subjects in order to arrive at conclusions that can be generalized to the population from which the subjects are drawn. This requires analysis tools that are robust against the spatial variability of the raw data which is unavoidably introduced by measurement artifacts and anatomical differences between subjects. Bivariate measures such as Pearson’s correlation coefficient are strongly affected by such variability as they only represent the relationship between two single voxels. A slight difference in the location of one voxel, for example, the center of a cortical activation, can cause significantly different correlation coefficients. This will in turn affect any process that takes the similarity between voxels as input. Therefore, a similarity measure which compensates for the spatial variability between subjects is desirable. One such measure is the canonical correlation.

Canonical correlation is a multivariate extension to bivariate correlations, providing a measure of similarity between two sets of variables. Different applications of canonical correlation in the analysis of fMRI data have recently been proposed by Friman et al. (2003) and Worsley et al. (2004). In our application, canonical correlation is used to represent the similarity of fMRI time series between two groups of voxels. Our variability analysis shows that using canonical correlation as input to the replicator dynamics yields results that are very consistent across repeated measurements of single subjects as well as across different subjects.

The remainder of the paper is organized as follows. We first introduce replicator dynamics and canonical correlation as the two main building blocks of our analysis method. We then present the application of our method to the left LFC in the two experimental tasks and provide a detailed analysis of the within-subject and the between-subjects variability of the results.

Methods

Replicator dynamics

In the following, we will only present the basic principles of the replicator process that are necessary for the understanding of the paper. The mathematical framework of replicator dynamics in the context of fMRI data analysis is described in great detail in Lohmann and Bohn (2002). A modification of the method for the analysis of meta-analytic imaging data can be found in Neumann et al. (2005).

The use of replicator dynamics in fMRI is based on the idea that functional networks or groups of functionally coherent voxels can be detected by analyzing pairwise similarity measures between time series. Given a matrix $\mathbf{W} = (w_{ij})$ where w_{ij} represents a similarity measure between time series of voxel i and j , we wish to find a maximally coherent network, cluster, or group¹ of voxels. In this context, a coherent network is defined as a network for which each member is closely connected with every other member.

A group of voxels exhibiting this coherence property can be found by determining the vector $\mathbf{x} = (x_1, x_2, \dots, x_n)$ with $x_i \geq 0$ and $\sum x_i = 1$ that maximizes $\mathbf{x}^T \mathbf{W} \mathbf{x}$. Here, n is the number of voxels included in the analysis, and x_i represents the degree of membership of voxel i in the maximally coherent group. Membership in this group is a fuzzy concept since x_i can take on any value between 0 and 1. It can be “defuzzified” by defining a membership function that declares all voxels to be members of the group for which $x_i > 1/n$. In other words, after the vector \mathbf{x} has been found, all voxels i are regarded as member of the maximally coherent group for which $x_i > 1/n$.

The maximization of $\mathbf{x}^T \mathbf{W} \mathbf{x}$ under the given constraints is known to be NP-hard, if \mathbf{W} has positive eigenvalues (Gibbons et al., 1997; Pardalos and Vavasis, 1991). This is the case in our application as \mathbf{W} is not guaranteed to be negative semi-definite. Note that the problem can also be interpreted as the search for the maximum-weighted clique in an undirected graph which, again, is known to be an NP-hard problem (Bomze et al., 2000). Consequently, exact algorithms are guaranteed to return the global solution only in a time that increases exponentially with the number of entries in \mathbf{W} and are therefore only applicable to very small problems. However, a local maximizer $\bar{\mathbf{x}}$ can be easily found using the following dynamical system:

$$\frac{d}{du} x_i(u) = x_i(u) \left[(\mathbf{W} \mathbf{x}(u))_i - \mathbf{x}(u)^T \mathbf{W} \mathbf{x}(u) \right], \quad i = 1 \dots n \quad (1)$$

This equation is known as *replicator equation* from theoretical biology, where it is used to model the evolution of self-replicating interacting entities over time (Schuster and Sigmund, 1983,

¹ In the following, we will use the terms “network”, “cluster”, and “group” interchangeably, accounting for the fact that our application of replicator dynamics addresses the coherence of voxels within the same cortical area, resulting in topologically connected clusters of voxels rather than in distributed networks.

Hofbauer and Sigmund, 1988b). Replicators are entities in the evolutionary process. In each replicator i , $i = 1..n$, which in our application are the voxels in the region of interest, is associated with a fitness, encoded in the matrix \mathbf{W} , and with a proportion x_i , indicating the percentage of the population that is identical to i . The term $\mathbf{x}(u)^T \mathbf{W} \mathbf{x}(u)$ denotes the mean fitness of the population at a given time point u . The discrete-time version of the replicator equation is given by

$$x_i(u+1) = x_i(u) \frac{(\mathbf{W} \mathbf{x}(u))_i}{\mathbf{x}(u)^T \mathbf{W} \mathbf{x}(u)}. \quad (2)$$

The dynamics of a replicator system are described by the fundamental theorem of natural selection (Fisher, 1930, Kimura, 1958). It states that, if \mathbf{W} is a non-negative, real-valued, symmetric matrix, then the function $\mathbf{x}(u)^T \mathbf{W} \mathbf{x}(u)$ is strictly increasing with increasing u along any non-stationary trajectory $\mathbf{x}(u)$ under both continuous-time (1) and discrete-time (2) replicator dynamics. Furthermore, any such trajectory converges toward a stationary point \mathbf{x} , and a vector $\mathbf{x} \in R^n$ with $x_i = 0$, $i = 1..n$ and $\sum x_i = 1$ is asymptotically stable if and only if it is a strict local maximizer of $\mathbf{x}^T \mathbf{W} \mathbf{x}$ (Hofbauer and Sigmund, 1988b). This means that, starting from an initial vector $\mathbf{x} = (x_1, x_2, \dots, x_n)$, we can find in an iterative process a vector \mathbf{x} that maximizes $\mathbf{x}^T \mathbf{W} \mathbf{x}$. If no assumptions can be made about differences of the individual replicators, the vector $\mathbf{x} = (x_1, x_2, \dots, x_n)$ should be initialized with $x_i = 1/n$ for all $i = 1..n$ in order to avoid additional bias in the input data (Lohmann and Bohn, 2002).

During the iterative maximization process, replicators with a particularly good fitness will be assigned a high proportion, while the proportion of replicators with poor fitness will decrease. After convergence and applying the membership function, we arrive at a maximally coherent group of voxels, often referred to as “dominant network”. This group consists of voxels whose individual fitness increased during the maximization process above the average fitness of all voxels. These can be interpreted as the ‘strongest’ or ‘most important’ members of the system, according to the coherence measure chosen.

In order to find subsequent groups which again show a coherent behavior among all voxels, the replicator process can be started again and reapplied to a fitness matrix from which the members of previous groups are deleted. Thus, it is possible to extract a series of coherent groups from the data. The degree of coherence among the voxels is thereby decreasing from application to application. If the true number of coherent groups in the data is not known in advance, a suitable stopping criteria for the repeated application of the replicator process needs to be defined. In our application, we wish to find groups of voxels whose coherence most likely reflects a common underlying functionality. Therefore, members of a coherent group should form a topologically connected cluster. Consequently, the application of the replicator process is repeated, and groups of voxels are extracted until they no longer form a topologically connected region. This seems intuitively plausible since a common underlying functionality would not be expected for voxels which are not topologically connected, even if they show a stronger coherence than the remaining voxels in the investigated region.

The coherence within a cluster and the separability of the extracted groups of voxels can be visualized using multidimensional scaling (MDS). MDS is a technique that transforms similarity values of high-dimensional data into Euclidean distance values in low-dimensional space as closely as possible. This allows the visualization of high-dimensional data like fMRI time series in

2-dimensional space such that relative distances between data points represent the similarity between them. Details of this methodology and its application, for example, in clustering, can be found in Davidson (1983) and Everitt et al. (2001).

In our application, data points in MDS are the time series of the voxels extracted as members of a coherent network by the replicator process. Therefore, MDS is applied to the similarity matrix \mathbf{W} that serves as input to the replicator process. Labeling the voxels with their network membership then shows how well the different extracted networks are separable.

Given the constraints on the fitness matrix \mathbf{W} described above, a number of different similarity measures for fMRI time series are conceivable. Previously, Pearson’s linear correlation coefficient and Spearman’s rank correlation were used (Lohmann and Bohn, 2002). However, these measures only describe the similarity of pairs of voxels and ignore the spatial dependencies between neighboring voxels in functional MR images. We will in the following introduce canonical correlation as similarity metric for fMRI time series, a multivariate correlation that allows to take into account the spatial dependencies of neighboring voxels.

Canonical correlation

Canonical Correlation Analysis (CCA) was developed by Hotelling (1936) and has become a standard tool in multivariate data analysis. A full discussion of the method and its relation to other multivariate analysis techniques can be found, for example, in Anderson (2003) and Rencher (2002).

Consider two sets of variables $\mathbf{x} = (x_1, x_2, \dots, x_p)$ and $\mathbf{y} = (y_1, y_2, \dots, y_q)$. The covariance structure associated with \mathbf{x} and \mathbf{y} is

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_{xx} & \mathbf{C}_{xy} \\ \mathbf{C}_{yx} & \mathbf{C}_{yy} \end{bmatrix}, \quad (3)$$

where \mathbf{C}_{xx} and \mathbf{C}_{yy} are the $p \times p$ and $q \times q$ within-set covariance matrices of \mathbf{x} and \mathbf{y} , respectively, and $\mathbf{C}_{xy} = \mathbf{C}_{yx}^T$ is the $p \times q$ between-sets covariance matrix. Let then $\rho_1^2, \rho_2^2, \dots, \rho_s^2$ be the eigenvalues of

$$\mathbf{C}_{xx}^{-1} \mathbf{C}_{xy} \mathbf{C}_{yy}^{-1} \mathbf{C}_{yx} \quad (4)$$

with $s = \min(p, q)$. The square roots of the eigenvalues $\rho_1, \rho_2, \dots, \rho_s$ are called *canonical correlation coefficients*. Note that the same canonical correlation coefficients can be found as eigenvalues of

$$\mathbf{C}_{yy}^{-1} \mathbf{C}_{yx} \mathbf{C}_{xx}^{-1} \mathbf{C}_{xy}, \quad (5)$$

however, the corresponding eigenvectors differ. It can be shown that the largest squared canonical correlation ρ_1^2 is the maximum squared correlation between the linear combinations

$$\mathbf{a}^T \mathbf{x} = a_1 x_1 + a_2 x_2 + \dots + a_p x_p, \quad (6)$$

$$\mathbf{b}^T \mathbf{y} = b_1 y_1 + b_2 y_2 + \dots + b_q y_q, \quad (7)$$

where \mathbf{a} and \mathbf{b} are the eigenvectors corresponding to ρ_1^2 obtained from Eqs. (4) and (5), respectively (Rencher, 2002).

In our context, \mathbf{x} and \mathbf{y} represent time series of groups of voxels, i.e. $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_p(t)]$ and $\mathbf{y}(t) = [y_1(t), y_2(t), \dots, y_q(t)]$ with $t = 1, \dots, T$, where T is the number of measured time steps. For the applications presented in the following section, we use single voxels and their respective 6-neighborhoods in 3D space

in order to form \mathbf{x} and \mathbf{y} , i.e. $p = q = 7$. Thus, the largest canonical correlation coefficient ρ_1 provides a measure of how well the time series of a voxel and its 6 immediate neighbors correspond to the time series of another voxel and its 6-neighborhood.

It should be noted that canonical correlation coefficients are in general larger than (or at least equal to) the Pearson correlation coefficients for the same voxels. This can be explained by the fact that the canonical correlation coefficient between $\mathbf{x} = (x_1, x_2, \dots, x_p)$ and $\mathbf{y} = (y_1, y_2, \dots, y_q)$ represents the maximum correlation between linear functions of \mathbf{x} and \mathbf{y} . Consequently, it exceeds the simple correlation between any x_i and y_j , $1 \leq i \leq p$ and $1 \leq j \leq q$ (Rencher, 2002). The absolute values of Pearson and canonical correlation coefficients are therefore not directly comparable. However, in our application, this is not problematic since the result of the replicator process is determined solely by the relative differences between canonical correlation coefficients of voxels.

The second, more interesting, difference between Pearson's and canonical correlation is illustrated in Fig. 1 where both coefficients are calculated between a seed voxel and voxels in the same cortical area. As can be seen, moving away from the seed voxel, the gradient of Pearson's correlation coefficient is much steeper than the gradient of the canonical correlation coefficient. While the canonical correlation coefficient is still relatively high in two voxels' distance from the seed voxel, Pearson's correlation coefficient drops very quickly as we move away from the seed voxel. This can be explained by the fact that voxels which show a relatively low similarity with the seed voxel in their fMRI time series, but are neighboring a voxel with a high similarity to the

seed voxel, benefit from this proximity when canonical correlation is applied. To be more precise, assume that the time courses of the seed voxel and its immediate neighbors are represented by \mathbf{x} and the time courses of another voxel of interest and its neighbors are stored in \mathbf{y} . If \mathbf{y} contains a time course that is very similar to the time courses in \mathbf{x} , it will be assigned a high coefficient in the linear combination in Eq. (7), resulting in a high eigenvalue corresponding to \mathbf{b} . This eigenvalue is then assigned as the canonical correlation coefficient to the voxel of interest, even if its own time course is less similar to those in \mathbf{x} . In other words, the similarity of a voxel's surrounding to the seed voxel is taken into account when determining the canonical correlation coefficient. In that sense, canonical correlation could be interpreted as a spatial filter on the correlation of time series. This property makes canonical correlation a particularly suitable measure for the similarity of fMRI time series if this similarity measure is to be compared across different imaging sessions and subjects.

For n voxels, the largest canonical correlation coefficients for any pair of two voxels i and j with $1 \leq i \leq n$ and $1 \leq j \leq n$ and their respective neighborhoods can be represented in an $n \times n$ similarity matrix. This matrix is symmetric, non-negative, and real-valued and thus meets the criteria for an input to the replicator process described above.

The algorithm

The various parts of our method for detecting clusters of coherent voxels within a region of interest can be summarized in the following steps:

1. After defining a region of interest (ROI), calculate for each voxel in this region the canonical correlation with all other voxels in the region.
2. Store the largest canonical correlation coefficient between any pair of voxels in the fitness matrix \mathbf{W} .
3. Initialize the elements of the proportion vector $\mathbf{x} = (x_1, x_2, \dots, x_n)$ with $1/n$, where n is the number of voxels in the ROI.
4. Apply the replicator process until convergence, thereby extracting the most coherent voxels.
5. Delete all members of the extracted group from the ROI and from \mathbf{W} .
6. Repeat the process from step 3 until the extracted group no longer forms a topologically connected cluster.

Relation to other methods

Replicator dynamics has close relations to other multivariate data analysis techniques. Like spectral clustering (e.g. Weiss, 1999) and cross-correlation clustering (Dodel et al., 2002), it explores the structural properties of some similarity matrix and can be reformulated as a graph partitioning problem for fully connected bidirectional weighted graphs. Cross-correlation clustering relies on graph-theoretic notions such as "cliques" and "connectivity components" in order to find subgraphs corresponding to clusters. Spectral clustering explores the eigenstructure of the Laplacian \mathbf{L} , which is related to the similarity matrix \mathbf{W} that serves as input to the replicator process. In the simplest form, \mathbf{L} is defined as $\mathbf{L} = \mathbf{D} - \mathbf{W}$, where \mathbf{W} is some $n \times n$ similarity matrix as defined above and \mathbf{D} denotes a diagonal matrix with entries $d_i = \sum_{j=1}^n w_{ij}$. The elements of the second eigenvector of \mathbf{L} then indicate a partitioning

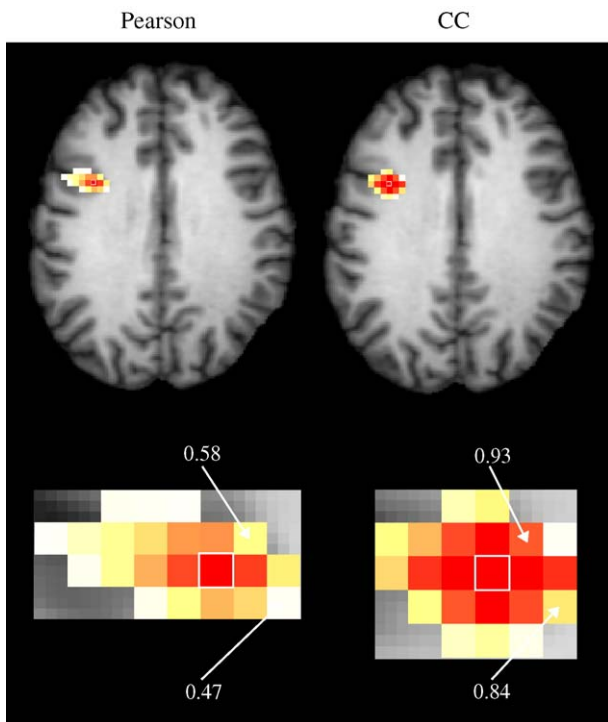


Fig. 1. A comparison of the Pearson correlation coefficient (left) and the canonical correlation (right). The seed voxel is marked with a white square. For the canonical correlation, sets were formed from a voxel and its 6 immediate neighbors in 3D space. The 21 voxels with the highest correlation to the seed voxel in this slice are enlarged in the bottom row. Data are taken from a single subject performing the color–word matching Stroop task described below.

of the input data into two disjunct groups. The inclusion of more eigenvectors provides a further sub-clustering.

Another related approach which, like replicator dynamics, operates directly on \mathbf{W} rather than on the Laplacian is Principal Component Analysis (PCA). Given some data \mathbf{z} , the first principal component of \mathbf{z} is the linear combination $\mathbf{y} = \mathbf{x}^T \mathbf{z}$ with maximum variance. This component is found by choosing the vector \mathbf{x} which maximizes the objective function $\mathbf{x}^T \mathbf{C}_z \mathbf{x}$ under the constraint $\sum x_i^2 = 1$. Here, \mathbf{C}_z is the covariance matrix of the data, and it turns out that the first eigenvector of \mathbf{C}_z is the solution for \mathbf{x} . In a similar manner, the replicator process finds a vector \mathbf{x} that maximizes the objective function $\mathbf{x}^T \mathbf{C} \mathbf{x}$. \mathbf{C} is now the matrix of canonical correlation coefficients for our observations. The only difference in the two methods lies in the different constraint $\sum x_i = 1$ with $x_i \geq 0, 1, \dots, n$ on the solution of the replicator process. Note that the latter is a priori met in our particular application of the replicator process since the matrix \mathbf{C} contains only positive entries.

As one of our reviewers pointed out, the difference between PCA and replicator dynamics might be negligible in some applications. However, the results of both methods can differ significantly as is shown in the following example. Consider the graph in Fig. 2 where nodes represent voxels within some region of interest and arcs between nodes represent strong similarities between voxels. Our aim is to find groups of voxels, where each voxel shows a strong similarity with all other voxels in the group. Consequently, in the example graph, nodes should be chosen that have connections to other strongly connected nodes and are also all connected to each other. The replicator process provides the following result vector: $(0.001 \ 0.001 \ 0.499 \ 0.499 \ 0.000 \ 0.000)^T$, where the value at position i indicates the importance of the i th node in the graph. Applying the membership function of our method selects nodes 3 and 4 as solution of this problem as the values for these two nodes increased above the initialization of $1/6$ during the replicator process. In contrast, applying PCA results in the first eigenvector $(0.468 \ 0.468 \ 0.437 \ 0.437 \ 0.298 \ 0.298)^T$. Here, nodes 1 and 2 have the highest values. While these two nodes are indeed connected to many other nodes in the graph, they do not have a direct connection and thus do not meet the requirements on our solution. Replicator dynamics thus seems to be the method of choice for this particular application.

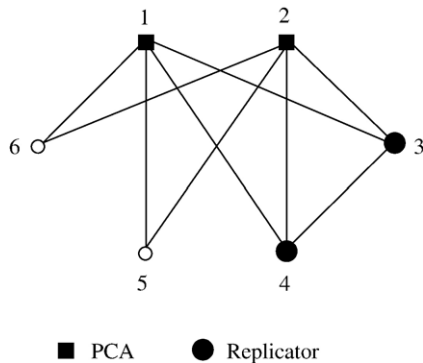


Fig. 2. Example graph showing different solutions for the application of PCA and replicator dynamics to the problem of finding highly interconnected subgraphs.

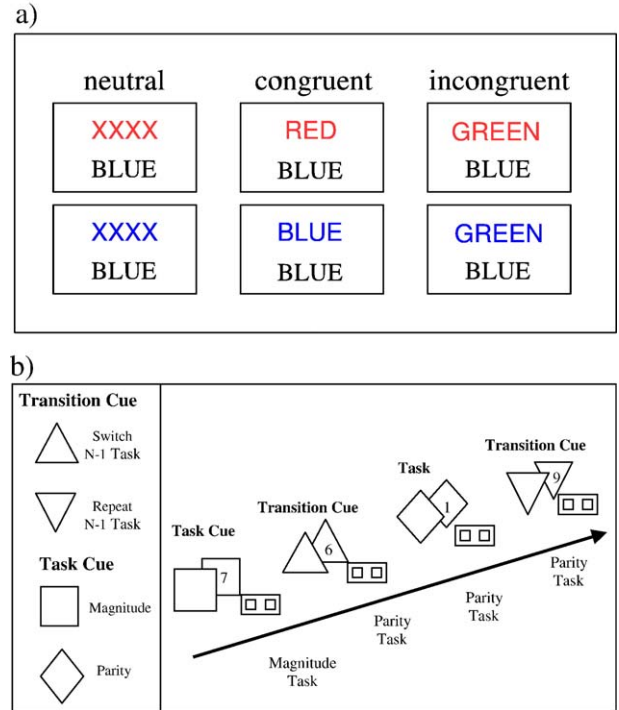


Fig. 3. The two experimental paradigms. (a) Examples of the three conditions in the color–word matching Stroop task. For the top row, the correct answer would be “NO”, for the bottom row “YES”. (b) The two different cue types (left) and an example of the cue and target presentation in the task switching paradigm.

Experimental data

Given a region of voxels that are activated within the experimental contrast of interest, replicator dynamics may be able to further subdivide this region into voxels belonging to different coherent clusters, thus indicating different functionalities of subregions or a functional gradient. Such a gradient was recently suggested for the lateral frontal cortex (LFC) in the anterior–posterior direction (see e.g. Brass and von Cramon, 2004 and Forstmann et al., 2005). More specifically, it should be possible to distinguish between an area lying in the vicinity of the inferior frontal sulcus and the precentral sulcus, the so-called inferior frontal junction (IFJ), and an area anteriorly located to the IFJ along the inferior frontal sulcus. Given this clear hypothesis, we used this cortical area as a test case. While its functional subdivision was impossible to detect on the basis of statistical significance of contrasts alone (Brass and von Cramon, 2004; Forstmann et al., 2005), we would expect to find a clear parcellation when the replicator process is applied to this region and clusters of coherent voxels are extracted.

We applied the replicator process to two data sets from previously performed fMRI experiments addressing cortical activation related to cognitive control processes. The first experiment employed an event-related single-trial version of the color–word matching Stroop task (Zysset et al., 2001, Neumann et al., 2003). Four subjects were examined each performing nine experimental sessions taking place within a time range of 9 weeks. There were three experimental conditions (neutral, congruent, and incongruent) which are exemplified in Fig. 3. In each session, trials were presented in a different randomized order.

In a second experiment, data were analyzed from 14 subjects performing a cuing version of the task switching paradigm (Forstmann et al., 2005). Subjects were presented with digits between 1 and 9 (except 5) on the screen and had to perform two different tasks, judging whether the presented digit was smaller or larger than 5 (magnitude task), or judging whether it was even or odd (parity task). A task cue or a transition cue instructed participants on which task they had to perform next. Examples of the cue and target presentation are shown in Fig. 3. Trials were presented in a different pseudo-randomized order to each subject, with equal numbers of task switch and repetition trials and equal transition probabilities between conditions and cue types.

Both experiments were carried out on a 3 T scanner (Medspec 30/100, Bruker, Ettlingen). Axial slices, parallel to the AC–PC plane and covering the whole brain, were acquired using a single shot, gradient recalled EPI sequence (TR 2000 ms, TE 30 ms, 90° flip angle). Prior to the functional runs, corresponding anatomical T1-weighted MDEFT (Ugurbil et al., 1993, Norris, 2000) slices and T1-weighted EPI slices were acquired.

Data were processed with the software LIPSIA (Lohmann et al., 2001). This software package contains tools for preprocessing, registration, statistical evaluation, and presentation of fMRI data. In LIPSIA, statistical evaluation is implemented as a two-stage random-effect analysis (Holmes and Friston, 1998) based on the least-squares parameter estimation of a GLM for serially autocorrelated observations (Friston, 1994, Worsley and Friston, 1995, Zarahn et al., 1997).

Functional data were corrected for motion using a matching metric based on linear correlation. To correct for the temporal offset between the slices acquired in one session, a sinc interpolation based on the Nyquist–Shannon Theorem was applied. A temporal high-pass filter with a cut-off frequency of 1/100 Hz and 1/84 Hz was used for baseline correction of the signal in the task switching and the Stroop experiment, respectively. Furthermore, a spatial Gaussian filter with 4.24 mm FWHM ($\sigma = 0.6$) was applied. To align the individual functional data slices onto the corresponding 3D stereotactic coordinate reference system, a rigid linear registration with six degrees of freedom (3 rotational, 3 translational) was performed. The rotational and translational parameters were acquired on the basis of the MDEFT and EPI-T1 slices to achieve an optimal match between these slices and the individual 3D reference data set.

More specific information on the experimental design and modeling in the two experiments can be found in Neumann et al. (2003) and Forstmann et al. (2005), respectively. For the application of the replicator process, a region of interest was determined for each experiment separately using the group average (one-sample Student's *t* test of contrasts maps across subjects and sessions) of the incongruent vs. neutral contrast in the Stroop paradigm and the contrast between the two different cue types in the task switching paradigm. These were the primary contrasts of interest for the investigated paradigms. In previous studies, the left IFJ was localized at Talairach coordinates *x* between -47 and -35 , *y* between -4 and 10 , and *z* between 27 and 40 (Brass et al., 2005, Derrfuss et al., 2004, 2005, Forstmann et al., 2005, Neumann et al., 2005). Therefore, ROIs were determined from the peak activation coordinate within these boundaries and all topologically connected voxels showing significant activation ($z > 3.09$, corresponding to $P < 0.001$). The resulting ROIs consisted of 100 and 210 voxels for the Stroop and the task switching paradigm, respectively. For three subjects performing the task switching paradigm, no functional

data could be obtained for some voxels included in the mask volume. These three subjects were excluded from the subsequent analysis. Note again that the ROIs consisted of voxels that all showed significant activation in the contrasts relevant for the experimental tasks. The ROIs could not be further subdivided based on experimental manipulations.

Experimental results

Subdivision of the ROIs

The replicator process was applied repeatedly to both data sets in order to find separable groups of coherent voxels. For each subject and measurement, the application was stopped after the voxels in the extracted groups did not form a single topologically connected cluster anymore. For the Stroop paradigm, this procedure resulted in three consecutive clusters for all nine imaging sessions of each of the four subjects. Subsequent groups consisted of very few voxels distributed over the entire mask volume. The extracted voxels covered on average 75% of the mask volume. A randomly chosen result is presented for a single imaging session of one subject in the top row of Fig. 4.

The same procedure was applied to the data obtained for the task switching paradigm. Again, the application of the replicator process resulted in three topologically connected clusters of voxels for each subject. Subsequently, extracted voxels did not form topologically connected regions. The three extracted networks covered on average 40% of the mask volume. Note, however, that the mask volume was about twice as large as for the Stroop paradigm due to the higher anatomical variability between the 11 analyzed subjects. A randomly chosen result from a single subject can be seen in the bottom row of Fig. 4.

The separation between groups was then visualized using MDS. MDS maps for the examples in Fig. 4 are shown in Fig. 5a. In both data sets, the first and third group can clearly be separated from the

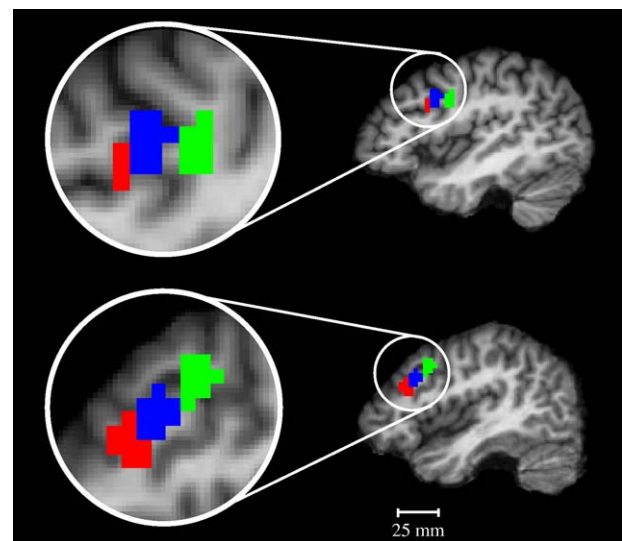


Fig. 4. Example results from applying the replicator process to the Stroop data (top) and the task switching data (bottom). Cluster 1 (blue), 2 (green), and 3 (red) are mapped onto anatomical slices ($x = -42$) of the individual subjects.

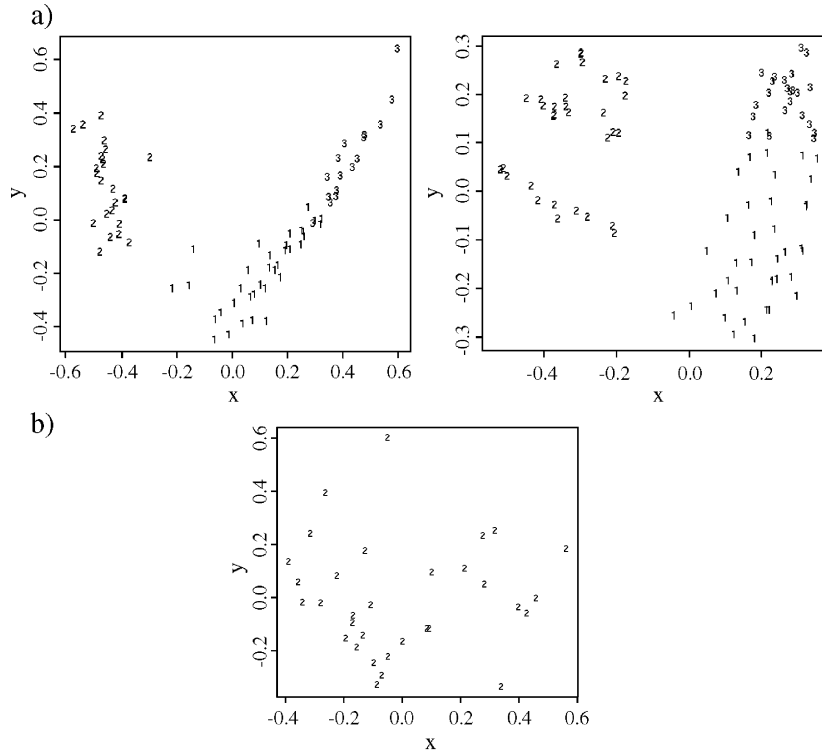


Fig. 5. Multidimensional scaling maps for the coherent groups presented in Fig. 4. (a) Left: Stroop data, right: task switching data. Note that relative distances as determined by MDS are represented by the x and y axes. However, absolute distance values are arbitrary and cannot be interpreted. (b) Individual MDS map of group 2 for the task switching data. The mapping shows that the apparent subdivision of group 2 in the task switching data (top right) is the result of simultaneously mapping all three groups rather than the reflection of an inherent sub-clustering.

second. Group 3 shows a relatively high similarity to group 1 but is still clearly separable. This similarity corresponds to the physical location of the groups, with the third group neighboring the first one in anterior direction. Note that the apparent subdivision of group 2 in the task switching data is a result of the MDS projection together with groups 1 and 3 rather than a representation of an inherent clustering. This can be seen in Fig. 5b where group 2 is mapped independently of the other two groups. This mapping shows an almost homogeneous cluster.

As stated in Methods, canonical correlation can be viewed as a spatial filter on the correlation between time series. Specifically, one could argue that it favors the coherence between spatially close voxels as they share some voxels in their respective neighborhoods. In order to verify that the parcellation of the ROIs is not exclusively driven by this effect but reflects true similarity of voxels within the same coherent group, we applied CCA and replicator dynamics to ROIs containing spatially randomized voxels. Spatially permuted ROIs were constructed

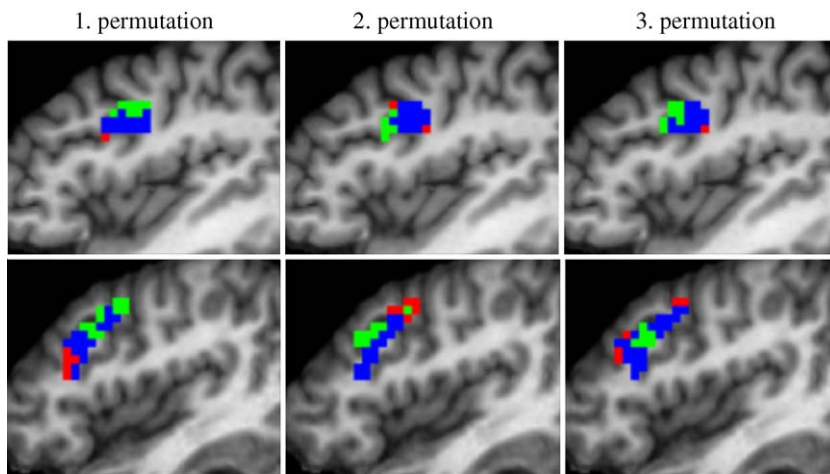


Fig. 6. Three example results from applying the replicator process to the spatially randomized ROI of the Stroop data (top) and the task switching data (bottom). The same subjects and anatomical slices as in Fig. 4 are shown. The extracted groups vary between different randomizations, and some groups no longer form topologically connected clusters.

by 1000 swaps of randomly selected pairs of voxels. This simple procedure ensures that the permuted ROIs are of the same size and physical location and contain exactly the same input time courses as the original data, only the positioning of the voxels within the ROI has changed. For comparability with the original data sets, a spatial Gaussian filter with 4.24 mm FWHM was applied after permutation.

Results of three permutation tests for the two subjects presented in Fig. 4 are shown in Fig. 6. For comparison, three groups of coherent voxels were extracted. If the parcellations obtained from the original data were exclusively driven by the use of canonical

correlation, one would expect the application to spatially permuted ROIs to result in patterns very similar to the original data. In particular, extracted groups should have the same size and the same number of extracted coherent groups should form topologically connected clusters. Visual inspection of the results in Fig. 6 is sufficient to see that this is not the case. Positioning and size of the extracted groups varied between different permutation tests, and voxels within the same group did not form topologically connected clusters. We would therefore argue that the results obtained from the original data truly reflect the intrinsic structure of the investigated ROIs.

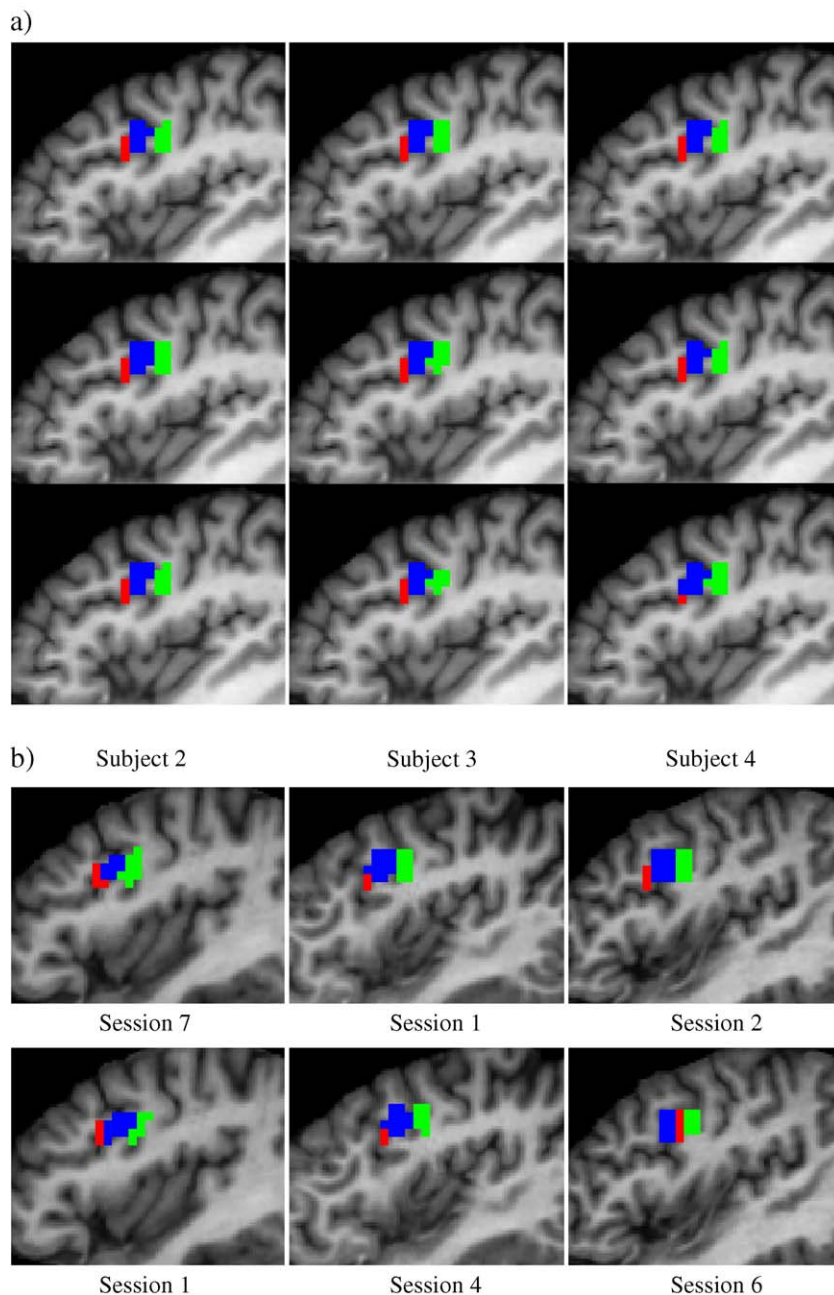


Fig. 7. (a) Within-subject variability of the extracted groups from subject 1 performing the Stroop task. A sagittal view ($x = -42$) of the three coherent groups as extracted in all nine experimental sessions is shown. (b) Example results ($x = -42$) for the remaining three subjects performing the Stroop paradigm. Out of the nine repeated measurements, two measurements were chosen at random for each subject.

Table 1
Within-subject variability of extracted coherent groups

Subject	Average no. of voxels			Overlap (9 sessions)			Overlap (6 or more sessions)		
	Group 1	Group 2	Group 3	Group 1	Group 2	Group 3	Group 1	Group 2	Group 3
1	29.3	28.0	18.8	25	18	14	27	26	19
2	31.7	27.7	18.6	17	14	13	29	26	18
3	32.4	20.7	22.0	14	12	14	31	20	22
4	30.3	23.8	17.1	11	7	–	24	19	10

Note. The table shows for all four subjects the average number of voxels forming a coherent group, the number of voxels that were members of the same group in all nine experimental sessions, and the number of voxels that were members of the same group in at least two thirds of the experimental sessions.

Within-subject variability

The reliability of the results was first investigated for repeated measurements of single subjects performing the Stroop interference task. For the first three subjects, the groups of coherent voxels were extracted in very similar anatomical locations and in the same order from all nine experimental sessions. This is exemplified for the first subject in Fig. 7a where results from all nine imaging sessions are shown.

From two sessions of subject 3, group 2 and group 3 were extracted in reversed order, that is, the most anterior group was found second, the most posterior one last. For subject 4, the first and second groups were comparable across all nine sessions. However, in three sessions, the third group was not located anterior but posterior to the first, neighboring the second group in dorsal direction. Two randomly chosen example results for subjects 2 to 4 are presented in Fig. 7b.

The results can be quantified by calculating the overlap of voxels in the extracted groups in all or the majority of imaging sessions. Table 1 shows for all subjects the average number of voxels forming a coherent group, the number of voxels that were members of the same group in all nine experimental sessions, and the number of voxels that were members of the same group in at least two thirds of the experimental sessions.

For the first three subjects, at least half of the extracted voxels were members of the same group in all experimental sessions, in some cases even more than two thirds. For subject 4, at least some voxels were found consistent members of groups 1 and 2 in all sessions. The size of the overlap in two thirds of the sessions almost matched the average size of the extracted groups for all subjects.

For comparison, the process was repeated using Pearson's correlation coefficient first without spatial smoothing then with spatial filters of different sizes. Gaussian filters of sizes between

4.24 mm and 8.48 mm FWHM ($\sigma = 0.6$ and $\sigma = 1.2$) were used. In Table 2, results are exemplified for subject 1. Similar results were obtained for the other three subjects. The analysis revealed again three separable groups of coherent voxels, showing that the strong coherence within the detected groups is not caused by spatial filtering alone. However, the size and location of the obtained groups were less stable across repeated measurements.

Between-subject variability

The between-subject variability was assessed in the same manner, now using the multi-subject data of the task switching experiment. The location of the three coherent groups was again very consistent across subjects. The order in which the groups were extracted was more variable than in the within-subject analysis. Most consistently, for 8 subjects, the most posterior group was extracted second, in the other three cases, it was extracted last. For 7 subjects, the first coherent group was located between group 2 and 3 in anterior–posterior direction, in the remaining 4 cases, the most anterior group was extracted first followed by the most posterior one.

Table 3 presents the average number of voxels in the coherent groups as well as the number of overlapping voxels for 6 and more subjects. As the table shows, the maximum number of subjects for which voxels were found as members of the same group is 8, 9 and 11 for the three groups, respectively. Moreover, all three coherent groups contained a reasonable number of voxels that were group members in at least 6, i.e. more than half of the subjects.

The overlap of groups across subjects is visualized in Fig. 8. The top row shows group borders when all voxels were included that were group members in at least 6 subjects. The bottom row shows the three groups separately in sagittal slices that contained the group centers, i.e. the voxels with the maximum overlap of subjects. The number of subjects where a voxel was a group

Table 2
Within-subject variability of extracted groups using Pearson correlation and Gaussian spatial filters

Filter (FWHM)	Average no. of voxels			Overlap (9 sessions)			Overlap (6 or more sessions)		
	Group 1	Group 2	Group 3	Group 1	Group 2	Group 3	Group 1	Group 2	Group 3
None	15.3	13.1	11.8	–	2	2	4	17	18
4.24 mm	15.3	13.1	11.7	–	2	2	4	17	16
7.06 mm	17.7	15.2	13.0	–	5	2	10	20	21
5.65 mm	16.7	14.1	11.9	–	8	3	11	22	18
8.48 mm	19.3	16.4	13.8	3	4	4	17	19	27

Note. The table shows for the first subject and different filter sizes the average number of voxels forming a coherent group, the number of voxels that were members of the same group in all nine experimental sessions, and the number of voxels that were members of the same group in at least two thirds of the experimental sessions. Coherent groups were obtained using the Pearson correlation coefficient instead of the canonical correlation coefficient and Gaussian spatial filtering.

Table 3
Between-subject variability of extracted groups

Cluster	Average no. of voxels	Overlap (6 or more subjects)					
		6	7	8	9	10	11
Anterior	26.8	18	12	7	–	–	–
“Middle”	29.0	20	14	9	3	–	–
Posterior	26.9	26	15	13	6	4	2

Note. The table shows the average number of voxels in the coherent groups and the number of overlapping voxels for 6 and more subjects.

member is color-coded from white (1) to red (11). As can be seen, the most consistent voxels, i.e. voxels within the borders in the top row and shown in red in the bottom row, form group centers that are clearly separated.

Discussion

For two experimental paradigms, we were able to assemble voxels located in the left LFC into different coherent groups based on purely exploratory data analysis techniques. This subdivision was predicted from previous studies (e.g., Forstmann et al., 2005). For the task switching paradigm, the present analysis revealed a distinction of the IFJ from two clusters more anteriorly located along the inferior frontal sulcus. These two clusters showed a relatively high similarity in their fMRI time series revealed by their close proximity in the MDS maps. However, the IFJ could be clearly separated from both of them. For the Stroop paradigm, three distinct groups were found, again arranged in anterior–posterior direction. Interestingly, Müller et al. (2003) reported very similar results from applying an entirely different methodology to the same region and experimental paradigm. Using a combination of wavelet-based statistics and the general linear model for the evaluation of contrasts, they found a chain of clearly dissociable

activations reaching from the anterior tip of the inferior frontal sulcus to the IFJ. Note, however, that our automatically determined mask volume extended beyond the IFJ in posterior direction. One cluster was therefore found in posterior direction at the border of the inferior part of the precentral sulcus.

An important aspect of the presented work is the reproducibility of the results and their comparability across multiple imaging sessions and subjects. Correlation-based analysis methods typically employ bivariate measures which capture the similarity between pairs of voxels. Using such bivariate measures as input to a replicator process has been shown to yield good results in single subject analyses (Lohmann and Bohn, 2002). However, they are strongly affected by the spatial variability of the input data when results from repeated measurements or different subjects are compared. As the results in Table 2 show, hardly any overlap of clusters can be observed for multiple imaging sessions, even after applying large spatial filters to the input data. Such overlap would have been expected, however, given that the data were obtained from the same subject.

Our analysis revealed that canonical correlation as similarity measure between fMRI time series much alleviates this problem. It compensates for the variability of the input data, resulting in a largely increased overlap of clusters both between repeated measurements and across subjects. Canonical correlation accounts for the fact that fMRI time series always exists in a spatial context and captures the additional information provided by the relationship between neighboring voxels. We wish to point out, however, that care should be taken when using canonical correlation in the analysis of single subject data, in particular, if the precise anatomical localization of correlated cortical areas is to be determined. As canonical correlation functions like a spatial filter on the input data, boundaries between areas of correlated and uncorrelated areas can be blurred, and areas of highly correlated voxels are likely to appear larger than in analyses based on bivariate Pearson or rank correlation.

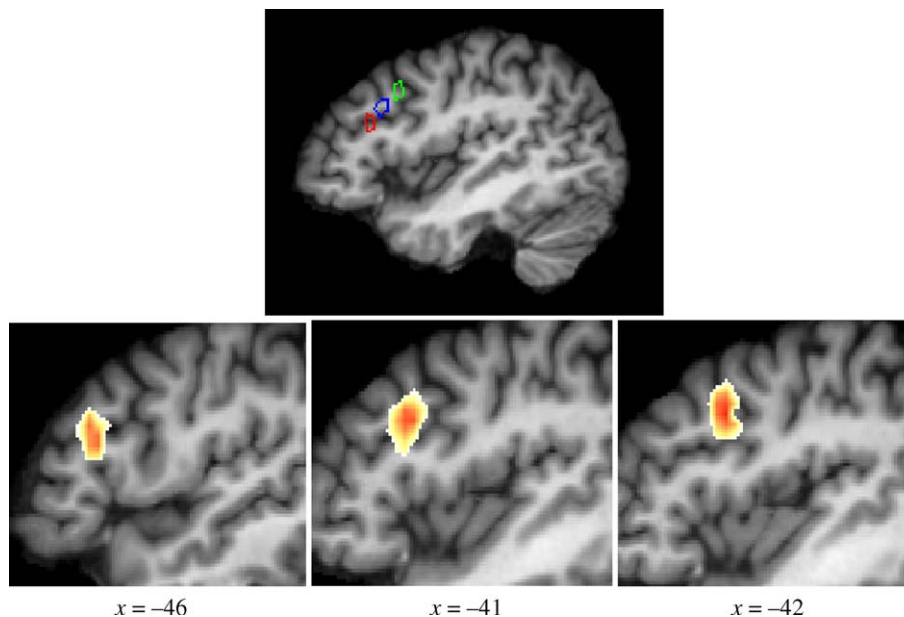


Fig. 8. Between-subject variability of the extracted groups from all subjects performing the task switching experiment. Top: sagittal view ($x = -42$) of the coherent groups including voxels that were group members in at least 6 subjects. Bottom: group centers shown individually in anterior–posterior direction from left to right. The number of subjects for which a voxel was member of the coherent group is color-coded from white (1) to red (11).

It is important to note that the replicator process is a selection process, that is, it always extracts a network of voxels with maximal coherence from the input data, whether or not this reflects the natural structuring inherent in the data. The degree of coherence in the network, i.e. the similarity between the network members, and the dissimilarity to the remaining voxels cannot be deduced directly from the result. This can be problematic if there exists no inherent structure in the input data at all. The extracted groups would then be meaningless and the underlying cortical area should really be treated as a homogeneous region. This problem is not exclusive to replicator dynamics. Most traditional clustering techniques would also result in some artificial structuring if no natural structure is inherent in the data, in particular, if the number of clusters has to be specified in advance. However, a post hoc assessment of the compactness and separability of the obtained clusters or the use of re-randomization techniques can help to distinguish between an inherent and an artificial clustering of the data. We have used spatial randomization of the time courses in the mask volume in order to verify that the results of the replicator process reflect some inherent rather than artificial structure of the data. When three coherent groups of voxels were extracted from spatially randomized time courses, these groups differed in size from the original results and no longer formed topologically connected clusters. This indicates that the inherent structure of the data was destroyed by the randomization process and, most importantly, the obtained groups are not simply a result of the spatial filtering, the particular similarity measure or the parcellation procedure used.

An important parameter in our method is the number of different networks or groups of voxels that should be extracted from the data, i.e. the number of subsequent applications of the replicator process once a network has been detected. Again, finding the optimal number of groups in some data is not specific to the multiple application of replicator dynamics but is in common with clustering techniques such as *k*-means. In general, clustering algorithms have to be provided with the expected number of clusters or, in case of hierarchical clustering, need some stopping criteria. Unfortunately, a universal solution to this problem does not exist as the appropriate number of clusters not only depends on the structure of the input data but also on the objective of the clustering process (Halkidi et al., 2001). In our application, we propose to repeatedly apply the replicator process until the extracted network does not form a topologically connected cluster of voxels anymore. This is in line with the assumption that voxels which are not topologically connected would not be expected to reflect a common functionality.

Finally, we wish to point out again the two important differences between replicator dynamics and traditional clustering techniques such as *k*-means as they reflect the advantages of our method over other clustering techniques when groups of functionally closely related voxels are sought. Firstly, clusters derived from traditional clustering techniques and networks detected by a replicator process exhibit different coherence properties. While the former contain elements which are all similar to some representative cluster center, the latter contain only elements which are closely connected to *all other* network members. In other words, traditionally clustering pertains to finding star-shaped topologies in the input data. In contrast, replicator dynamics results in network topologies. This can lead to quite different results in some cases, as was seen in the example of Fig. 2.

Secondly, in classical clustering procedures, each input element is assigned to one cluster. For example, if we wish to partition the input space into three clusters by *k*-means, each element in the input space will be a member of one of the three clusters. In contrast, the repeated application of a replicator process selects from a set of input elements different groups of elements which show a coherent behavior. Consequently, the extracted groups do not necessarily contain all elements. This way, outliers and voxels showing no strong correlation with other voxels are not forced to be a member of a cluster and thus do not affect the choice of cluster boundaries.

Conclusion

The present study was designed to investigate whether replicator dynamics facilitates the parcellation of cortical areas into subregions which were revealed by prior model-based statistical analysis as single homogeneous regions. We found that replicator dynamics indeed suits this task. However, we propose to use canonical correlation instead of the typically applied Pearson's correlation coefficient as a measure for the similarity of the fMRI time series. Combining replicator dynamics and canonical correlation yielded results that were very stable when compared across measurements of the same subject and, most notably, in the analysis of different subjects.

References

- Anderson, T.W., 2003. An Introduction to Multivariate Statistical Analysis, 3rd ed. John Wiley and Sons, New York.
- Baumgartner, R., Windischberger, C., Moser, E., 1998. Quantification in functional magnetic resonance imaging: fuzzy clustering vs. correlation analysis. *Magn. Reson. Imaging* 16 (2), 115–125.
- Bomze, I.M., Pelillo, M., Stix, V., 2000. Approximating the maximum weight clique using replicator dynamics. *IEEE Trans. Neural Netw.* 11 (6), 1228–1241.
- Brass, M., von Cramon, D.Y., 2004. Decomposing components of task preparation with functional magnetic resonance imaging. *J. Cogn. Neurosci.* 16 (4), 609–620.
- Brass, M., Derrfuss, J., Forstmann, B., von Cramon, D.Y., 2005. The role of the inferior frontal junction area in cognitive control. *Trends Cogn. Sci.* 9 (7), 314–316.
- Calhoun, V.D., Adali, T., Hansen, L.K., Larsen, J., Pekar, J., 2003. ICA of functional MRI data: an overview. Fourth International Symposium on Independent Component Analysis and Blind Source Separation, Nara, Japan, pp. 281–288.
- Davidson, M.L., 1983. Multidimensional Scaling. John Wiley and Sons, New York.
- Derrfuss, J., Brass, M., von Cramon, D.Y., 2004. Cognitive control in the posterior frontolateral cortex: evidence from common activations in task coordination, interference control, and working memory. *NeuroImage* 23 (2), 604–612.
- Derrfuss, J., Brass, M., Neumann, J., von Cramon, D.Y., 2005. Involvement of the inferior frontal junction in cognitive control: meta-analyses of switching and Stroop studies. *Hum. Brain Mapp.* 25 (1), 22–34.
- Dodel, S., Herrmann, J.M., Geisel, T., 2002. Functional connectivity by cross-correlation clustering. *Neurocomputing* 44–46, 1065–1070.
- Everitt, B.S., Landau, S., Leese, M., 2001. Cluster Analysis, 4th ed. Oxford Univ. Press.
- Fisher, R.A., 1930. The Genetical Theory of Natural Selection. Oxford Univ. Press, London, UK.

- Forstmann, B.U., Brass, M., Koch, I., von Cramon, D.Y., 2005. Internally generated and directly cued task sets: an investigation with fMRI. *Neuropsychologia* 43 (6), 943–952.
- Friman, O., Borga, M., Lundberg, P., Knutsson, H., 2003. Adaptive analysis of fMRI data. *NeuroImage* 19 (3), 837–845.
- Friston, K.J., 1994. Statistical parametric maps in functional imaging: a general linear approach. *Hum. Brain Mapp.* 2, 189–210.
- Friston, K., Phillips, J., Chawla, D., Buechel, C., 1999. Revealing interactions among brain systems with nonlinear PCA. *Hum. Brain Mapp.* 8, 92–97.
- Gibbons, L.E., Hearn, D.W., Pardalos, P.M., Ramana, M.V., 1997. Continuous characterizations of the maximum clique problem. *Math. Oper. Res.* 22 (3), 754–768.
- Goutte, C., Toft, P., Rostrup, E., Nielsen, F.A., Hansen, L.K., 1999. On clustering fMRI time series. *NeuroImage* 9 (3), 298–310.
- Halkidi, M., Batistakis, Y., Vazirgiannis, M., 2001. On clustering validation techniques. *J. Intell. Inf. Syst.* 17 (2–3), 107–145.
- Hofbauer, J., Sigmund, K., 1988a. *Evolutionary Games and Population Dynamics*. Cambridge Univ. Press, Cambridge.
- Hofbauer, J., Sigmund, K., 1988b. *The Theory of Evolution and Dynamical Systems*. Cambridge University Press, Cambridge.
- Holmes, A.P., Friston, K.J., 1998. Generalisability, random effects and population inference. *NeuroImage* 7, S754.
- Hotelling, H., 1936. Relation Between Two Sets of Variates. *Biometrika* 28, 321–377.
- Kimura, M., 1958. On the change of population fitness by natural selection. *Heredity* 12, 145–167.
- Lohmann, G., Bohn, S., 2002. Using replicator dynamics for analyzing fMRI data of the human brain. *IEEE Trans. Med. Imag.* 21 (5), 485–492.
- Lohmann, G., Müller, K., Bosch, V., Mentzel, H., Hessler, S., Chen, L., Zysset, S., von Cramon, D.Y., 2002. LIPSIA—A new software system for the evaluation of functional magnetic resonance images of the human brain. *Comput. Med. Imag. Graph.* 25 (6), 4498–5457.
- McKeown, M.J., Makeig, S., Brown, G.G., Jung, T.-P., Kindermann, S.S., Sejnowski, T.J., 1998. Analysis of fMRI by blind separation into independent spatial components. *Hum. Brain Mapp.* 6 (3), 160–188.
- Menon, A., Mehrotra, K., Mohan, C.K., Rank, S., 1995. Optimization using replicators. In: Eshelman, L. (Ed.), *Proceedings of the Sixth International Conference on Genetic Algorithm*. Morgan Kaufmann, San Francisco, CA, pp. 209–216.
- Müller, K., Lohmann, G., Zysset, S., von Cramon, D.Y., 2003. Wavelet statistics of functional MRI data and the general linear model. *J. Magn. Reson. Imaging* 17 (1), 20–30.
- Neumann, J., Lohmann, G., Zysset, S., von Cramon, D.Y., 2003. Within-subject variability of BOLD response dynamics. *NeuroImage* 19 (3), 784–796.
- Neumann, J., Lohmann, G., Derrfuss, J., von Cramon, D.Y., 2005. The meta-analysis of functional imaging data using replicator dynamics. *Hum. Brain Mapp.* 25 (1), 165–173.
- Norris, D.G., 2000. Reduced power multi-slice MDEFT imaging. *J. Magn. Reson. Imaging* 11, 445–451.
- Pardalos, P.M., Vavasis, S.A., 1991. Quadratic programming with one negative eigenvalue is NP-hard. *J. Glob. Optim.* 1 (1), 15–22.
- Pelillo, M., 2001. Replicator dynamics in combinatorial optimization. In: Floudas, C.A., Pardalos, P.M. (Eds.), *Encyclopedia of Optimization*, vol. 5. Kluwer Academic Publishers, Boston, pp. 23–35.
- Pelillo, M., Siddiqi, K., Zucker, S.W., 1999. Matching hierarchical structures using association graphs. *IEEE Trans. Pattern Anal. Mach. Intell.* (11), 1105–1120.
- Rencher, A.C., 2002. *Methods of Multivariate Analysis*, 2nd ed. John Wiley and Sons, New York.
- Schuster, P., Sigmund, K., 1983. Replicator dynamics. *J. Theor. Biol.* 100, 533–538.
- Stroop, J.R., 1935. Studies of inference in serial verbal reactions. *J. Exp. Psychol.* 18, 643–662.
- Taylor, P., Jonker, L., 1978. Evolutionary Stable Strategies and Game Dynamics. *Math. Biosci.* 40, 145–156.
- Thirion, B., Fugeras, O., 2003. Dynamical components analysis of fMRI data through kernel PCA. *NeuroImage* 20 (1), 34–49.
- Ugurbil, K., Garwood, M., Hendrich, K., Hinke, R., Hu, X., Menon, R.S., Merkle, H., Ogawa, S., Salmi, R., 1993. Imaging at high magnetic fields: initial experiences at 4 Tesla. *Magn. Reson. Q.* 9, 259–277.
- Viviani, R., Groen, G., Spitzer, M., 2005. Functional principal component analysis of fMRI data. *Hum. Brain Mapp.* 24 (2), 109–129.
- Voultzidou, M., Dodel, S., Herrmann, J.M., 2004. Emergence of functional connectivity clusters via Hopfield Neural Networks. 10th Annual Meeting of the Organization for Human Brain Mapping (OHBM), Budapest, June.
- Weiss, Y., 1999. Segmentation using eigenvectors: a unifying view. *Proceedings IEEE International Conference on Computer Vision*, pp. 975–982.
- Worsley, K.J., Friston, K.J., 1995. Analysis of fMRI time-series revisited—again. *NeuroImage* 2, 173–181.
- Worsley, K.J., Taylor, J.E., Tomaiuolo, F., Lerch, J., 2004. Unified univariate and multivariate random field theory. *NeuroImage* 23 (Suppl. 1), S189–S195.
- Worsley, K.J., Chen, J.-I., Lerch, J., Evans, A.C., 2005. Comparing connectivity via thresholding correlations and SVD. *Philos. Trans. R. Soc.* (360), 913–920.
- Zarahn, E., Aguirre, G.K., D’Esposito, M., 1997. Empirical analyses of BOLD fMRI statistics. *NeuroImage* 5, 179–197.
- Zysset, S., Müller, K., Lohmann, G., von Cramon, D.Y., 2001. Color–word matching Stroop task: separating interference and response conflict. *NeuroImage* 13, 29–36.