

# ICA of fMRI Group Study Data

Markus Svensén,<sup>1</sup> Frithjof Kruggel, and Habib Benali\*

Max-Planck-Institute of Cognitive Neuroscience, Stephanstraße 1A, D-04317 Leipzig, Germany; and \*INSERM U494, Paris, France

Received January 30, 2001

---

**This paper proposes to extend independent component analysis (ICA) of functional magnetic resonance imaging (fMRI) data from single subjects to simultaneous analysis of data from a group of subjects. This results in a set of time courses which are common to the whole group, together with an individual spatial response pattern for each of the subjects in the group. The method is illustrated using data from two fMRI experiments. The results show that: (a) ICA is capable of extracting nontrivial task related components without any *a priori* information about the fMRI experiment; (b) in analysis of group data, ICA identifies components common to the whole group as well as components manifested in single subjects only.** © 2002 Elsevier Science (USA)

---

## INTRODUCTION

McKeown *et al.* (1998a,b) proposed using independent component analysis (ICA) as a method for analyzing functional brain imaging data from individual subjects, obtained using functional magnetic resonance imaging (fMRI). They suggested that ICA would separate out not only signals originating from the stimulation, which subjects receive during fMRI experiments, but also signals from other sources, such as “slowly varying” sources and subject movements. ICA requires no *a priori* information about the kind of signals to extract and thereby circumvents the need for specifying, e.g., the hemodynamic response convolving the stimulus related neuronal activation. ICA has also been used successfully in analysis of EEG (Makeig *et al.*, 1996) and MEG (Vigário *et al.*, 1998) data.

In this article, we consider extending this approach to simultaneously analyzing fMRI data from a group of subjects. This yields a set of temporal patterns (time courses) common across the group and, for each time course, a separate image for each of the subjects. It is common practice to repeat fMRI experiments across groups of subjects, but the subsequent analysis normally computes a per-voxel average over the group,

which will reduce the effective spatial resolution and hide individual differences. In contrast, ICA retains the resolution of the original data and preserves individual differences in the spatial response patterns.

The fact that ICA of group data is restricted to use a common set of time courses for all subjects in the group has a number of consequences. It means that weak sources with different temporal characteristics across subjects, which should primarily include noise sources, will tend to get suppressed. However, as the results below illustrate, ICA still appears to have the ability to separate out distinct stimulus related components confined to single subjects. Another effect is that regions with small differences in their time courses (e.g., arising from noise), which may be represented by separate components following an ICA of data from a single subject, are merged into a single component when the data from the subject are included in a group ICA. Finally, compared to single subject ICA, components with similar spatial as well as temporal characteristics across subjects will be easier to detect by manual inspection.

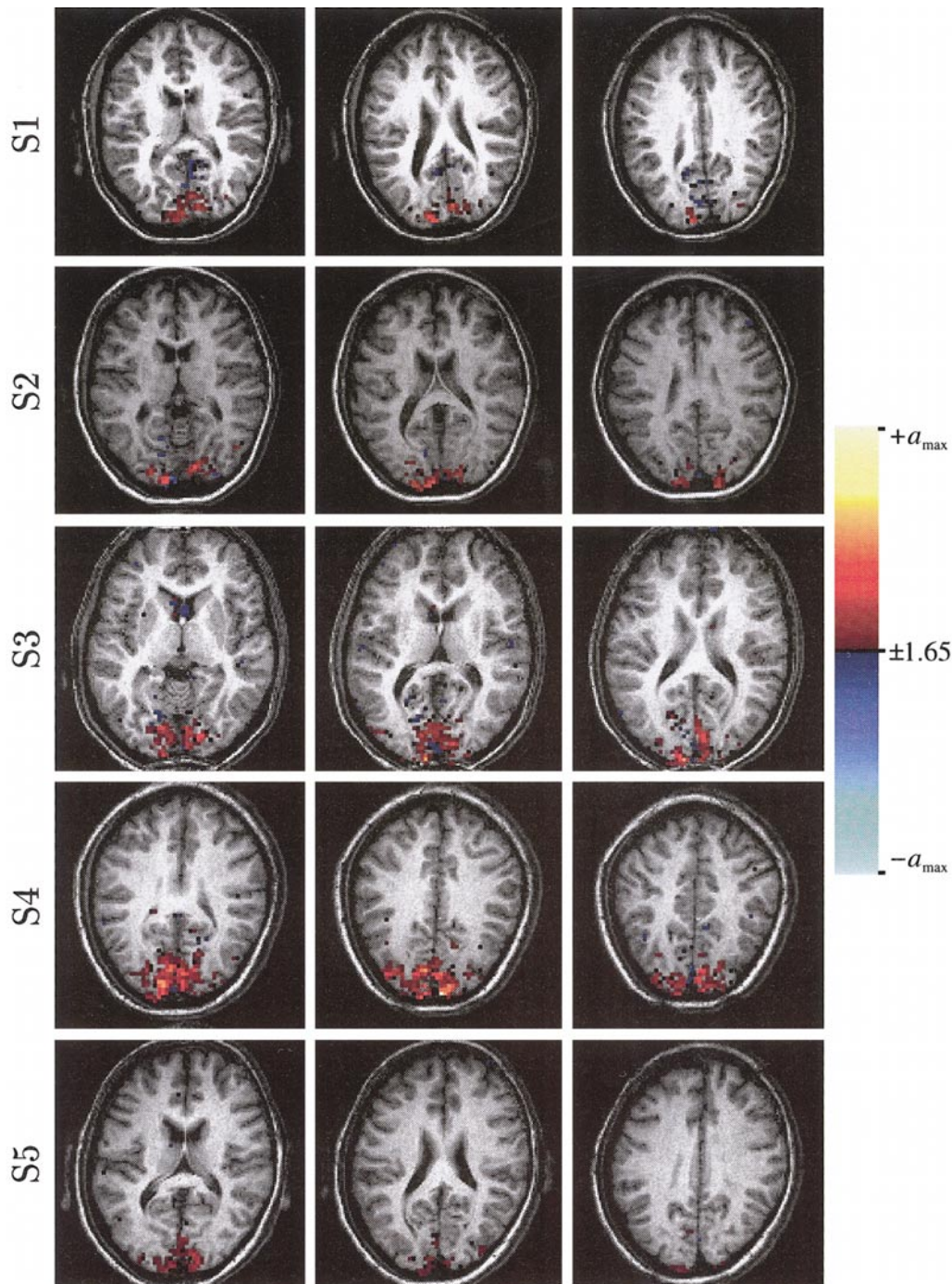
Next, we describe ICA, its application to fMRI single subject data and how it can be extended to deal with data from a group of subjects. We then give two examples of our new approach applied to data from two fMRI experiments: one with a block trial design and a visual stimulation, involving five subjects, the other with a single trial (event related) design and auditory stimulation, involving three subjects. We compare these results with those obtained from ICA of the data from the individual subjects.

## MATERIALS AND METHODS

### ICA

ICA (Cardoso, 1998; Hyvärinen, 1999b; Lee *et al.*, 2000) can be regarded as a generalization of the more well-known principal components analysis (PCA) (Jolliffe, 1986). Just like PCA, ICA linearly separates an  $N$ -dimensional data set into  $N$  components. However, whereas PCA yields components which are uncorrelated, ICA tries to find components which are independent. These two conditions are identical when the com-

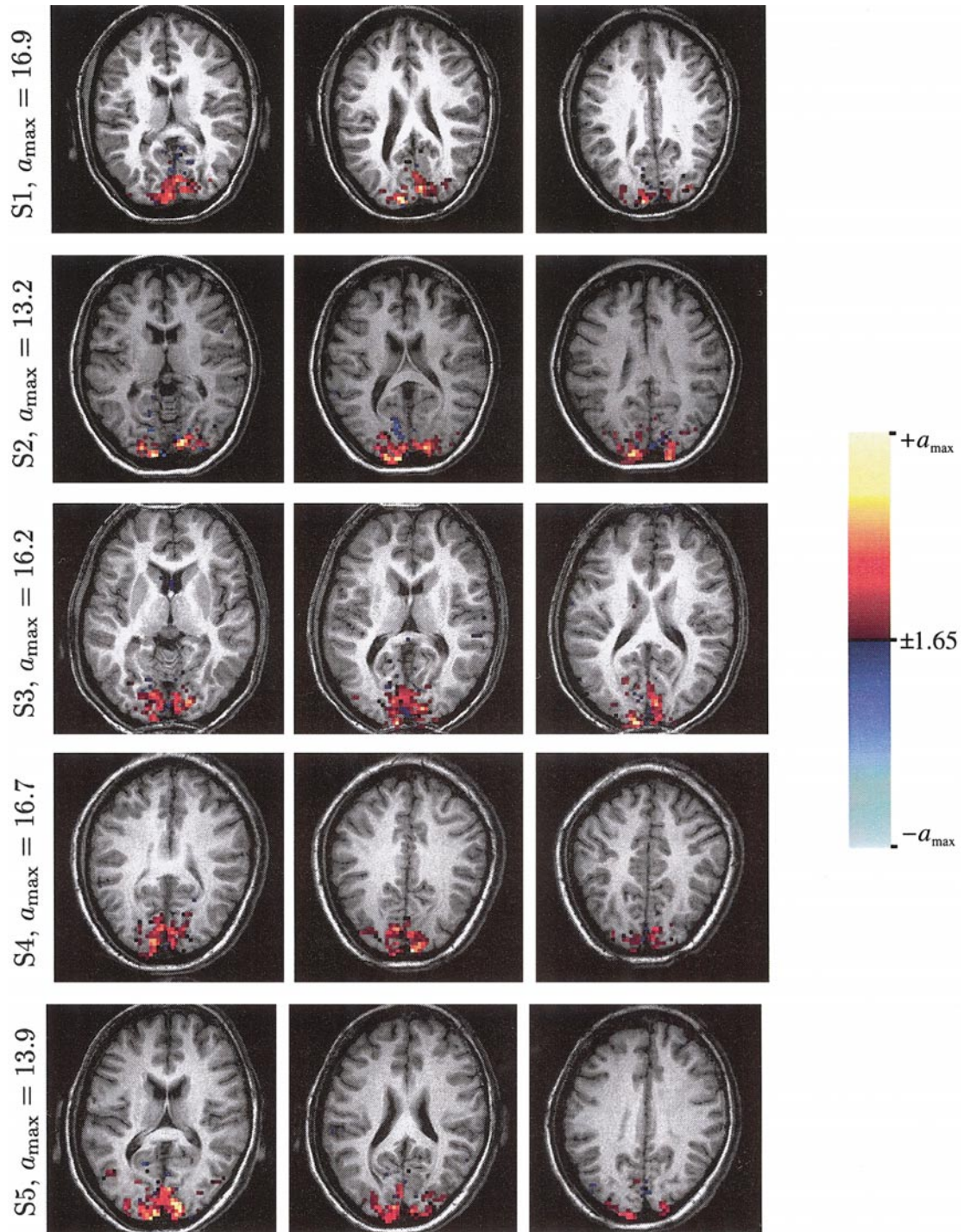
<sup>1</sup> To whom correspondence and reprint requests should be addressed.  
Fax: 49-0-341-994 0221. E-mail: {svensen, kruggel}@cns.mpg.de.



**FIG. 1.** The independent image, resulting from group ICA of the data from the visual experiment, corresponding to the time course showing the highest correlation with the performance of the task; this time course is plotted in the uppermost plot of Fig. 3. Each row corresponds to one subject and the columns correspond to brain slices, the with the lowest slice shown to the left. The independent image has been normalized to zero mean and unit variance, thresholded at  $\pm 1.65$ , and overlaid on the anatomical images of the five subjects. The blue-cyan voxels represent negative correlation with the time course, red-yellow voxels positive ditto;  $a_{\max} = 28.5$ .

ponents have Gaussian distributions, as is assumed by PCA (Tipping and Bishop, 1999), but ICA assumes that this is not the case. Given that the observed data were generated by linearly mixing independent sources with non-Gaussian distributions, ICA will separate the mix-

ture into the original sources. It achieves this by computing a linear transform that unmixes the mixed sources; if we think of the mixing transform as a matrix, ICA computes a matrix comparable to its inverse but for scaling and row order.



**FIG. 2.** The independent images, resulting from single subject ICA of the data from the visual experiment, corresponding to the time courses showing the highest correlation with the performance of the task for each individual; these the time courses are plotted in the second upper to lowest plots of Fig. 3, matching the order of the rows in this figure. The images have been visualized as the image shown in Fig. 1, except that the normalization was done individually for each subject.

To make this more concrete, let  $\mathbf{x}$  denote the  $N$ -dimensional random variable corresponding to the observed signals, and  $\mathbf{s}$  the  $N$ -dimensional random variable corresponding to the underlying sources. We can then write

$$\mathbf{x} = \mathbf{A}\mathbf{s}, \quad (1)$$

where  $\mathbf{A}$  is a  $N \times N$  full rank unknown mixing matrix. Note that we assume that the observed variables are

observed without noise, but ‘noise’ may still be represented in the data by one or more of the sources.

Most ICA estimation procedures take their starting point in the assumed independence of the original sources, which simply means that the probability distribution function of the random vector  $\mathbf{s}$  can be written as the product of marginal distributions of the vector elements,

$$p(\mathbf{s}) = \prod_n^N p(s_n). \quad (2)$$

Based on this assumption, the goal is to find an unmixing matrix,  $\mathbf{W}$ , such that the recovered sources,

$$\mathbf{y} = \mathbf{W}\mathbf{x} = \mathbf{W}\mathbf{A}\mathbf{s}, \quad (3)$$

are maximally independent. Several measures for independence have been proposed in the ICA literature and they can all be related to each other (Hyvärinen, 1999b). Here we restrict our attention to the mutual information, defined as

$$I(\mathbf{y}) = \sum_n^N H(y_n) - H(\mathbf{y}), \quad (4)$$

where

$$H(\mathbf{y}) = - \int p(\mathbf{y}) \ln p(\mathbf{y}) \, d\mathbf{y} \quad (5)$$

is the differential entropy for the random vector  $\mathbf{y}$  with a probability distribution  $p(\mathbf{y})$ . We see directly from (2), (4) and (5) that the mutual information is zero when the recovered sources,  $y_1, y_2, \dots, y_M$ , are independent and it can be shown to be strictly positive otherwise.

Assuming that  $\mathbf{W}$  in (3) is of full rank, (4) can be written

$$\begin{aligned} I(y_1, y_2, \dots, y_M) \\ = \sum_n^N H(y_n) - H(\mathbf{x}) - \ln(\text{abs}(\det(\mathbf{W}))). \end{aligned} \quad (6)$$

If we consider minimizing this expression with respect to  $\mathbf{W}$ , restricting its row vectors to be unit length, we note that the last term on the right hand side will be minimized when  $\mathbf{W}$  is orthogonal, while the first term is minimized when the marginal distributions  $p(y_m)$  are as non-Gaussian as possible.<sup>2</sup> However, these marginal

distributions are unknown, a complication which also affect all related measures of independence.

Several approximate schemes have been proposed, most of them primarily aimed at obtaining an estimate of the gradient of the objective function (e.g., (6) above) with respect to its parameters ( $\mathbf{W}$  in (6)); these estimates are computed by averaging over a data set representing a sample from the observed variables. Once the gradient has been obtained, the optimization can be carried out using stochastic gradient or Newton methods.

### ICA of fMRI Data

Applied to fMRI data, ICA tries to separate the sequence of recorded MR images into a set of independent source images.<sup>3</sup> Each recorded image is treated as finite sample from an observed random variable, and the voxels in the image are instances in that sample. McKeown *et al.* (1998b) motivate this with the argument that the spatial localization of brain areas activated by performing the task in an fMRI experiment should be independent of the spatial localization of signals arising from sources such as head movement or system noise.

Note that this is opposite to how ICA has been applied to EEG and MEG data, where the aim is to extract independent time courses, with accompanying spatial configurations. As pointed out by Friston (1998), this is usually not possible for fMRI data, since the number of images is (much) smaller than the number of recorded voxels. This means that the collected images will not span the space they live in, in which case we can find as many independent time courses as there are samples, but these will normally not tell us anything about eventual common sources.

Sticking with the spatial ICA of McKeown *et al.* (1998b), the mixing matrix ( $\mathbf{A}$  in Eq. (1)) will contain the corresponding set of time courses, specifying how the source images have been mixed to form the observed set of images. For example, we expect that the images that represent the cognitive task have time courses that somehow reflect this. We obtain the mixing matrix directly as the inverse of the unmixing matrix found by ICA.

### ICA of fMRI Group Study Data

The idea of using ICA to analyze fMRI group data is based on the following observations:

- All subjects in an fMRI experiment are carrying out the same task sequence. Thus, the individual source images corresponding to the performance of the task ought to have similar time courses.
- The union of two samples from  $N$  independent

<sup>2</sup> For a random variable with given mean and (co)variances, the Gaussian distribution is the distribution that maximizes the differential entropy.

<sup>3</sup> Throughout the rest of this article, we will use the terms sources, components, and images interchangeably, to refer to the independent source images (the rows of  $\mathbf{s}$  in Eq. (1)).

sources simply gives a larger sample where the  $N$  sources are still independent. This holds even when the two samples contains different sources.

With reference to the last point, we of course still expect that two samples in the form of fMR images from the same experiment will primarily contain sources with similar characteristics.

These two observations lead us to propose the following model (using the notation of Eq. (1)):

$$[\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_K] = \mathbf{A}[\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_K]. \quad (7)$$

Here,  $\mathbf{x}_k$  represents the data collected for subject  $k$ , in the form of a  $N \times L_k$  matrix;  $N$  is the number of images collected for each subject during the experiment and  $L_k$  is the number of voxels inside the brain mask for subject  $k$ . Accordingly,  $\mathbf{s}_k$  represents the matrix of independent source images of subject  $k$ .  $K$  is the number of subjects in the group and the  $[\cdot, \cdot, \cdot]$  operator denotes row-wise concatenation of matrices.  $\mathbf{A}$ , finally, denotes the mixing matrix, which is common to all subjects.

This means that we are extracting spatial components with common time courses across all subjects. Compared to the approach of McKeown *et al.*, we are also trying to obtain independent images, but the images now contain the voxels of all subjects. This approach does not require coregistration of the images from different subjects, since the spatial locations of the voxels are irrelevant to ICA. Thus, the resulting independent images have the same effective resolution as the original fMRI data.

Note that the ordering of the sources for an individual subject is arbitrary to ICA, and is determined only by the ordering of the columns in the common mixing matrix. Thus, we can readily assume that the task related source images have the same indices (row indices in the  $\mathbf{s}_k$  matrices) for all subjects. Note also that, there is nothing preventing physical sources, such the task induced activation, to manifest themselves in more than one source image.

### Data Sets

For the results presented in this paper we used data from two fMRI experiments. The first data set comes from a block trial experiment with a visual stimulation, whereas the second comes from an experiment with a single trial design and an auditory stimulation. Both data sets were recorded using a Bruker Medspec 30/100 3.0T MR system.

*Visual stimulation.* This experiment used an ON/OFF block trial design and a visual alternating checkerboard stimulus (Kruggel *et al.*, 2000). The experiment was performed with five different subjects, from each of which data were collected from three slices ( $64 \times 64$  matrix, 5-mm slice thickness with 2-mm gap,

TE 30 ms, TR 1333 ms, EPI). In addition to the functional data, a high resolution anatomical image was obtained for each slice and individual. We restrict our attention to a selected set of 228 functional images containing the first 9 blocks.

*Auditory stimulation.* In the second example below, we use data from an fMRI experiment with a single trial design intended for investigation of the neuronal correlates of sentence comprehension in the brain (Meyer *et al.*, 1998). Subjects had to decide whether an aurally presented sentence contained a syntactical violation or not. The experiment was performed with three different subjects, from each of which data were collected from four slices ( $128 \times 64$  matrix, 6 mm slice thickness with 2 mm gap, TE 40 ms, TR 2 s, EPI). Just like in the previous experiment, the functional data were accompanied by high resolution anatomical images obtained for each slice and individual.

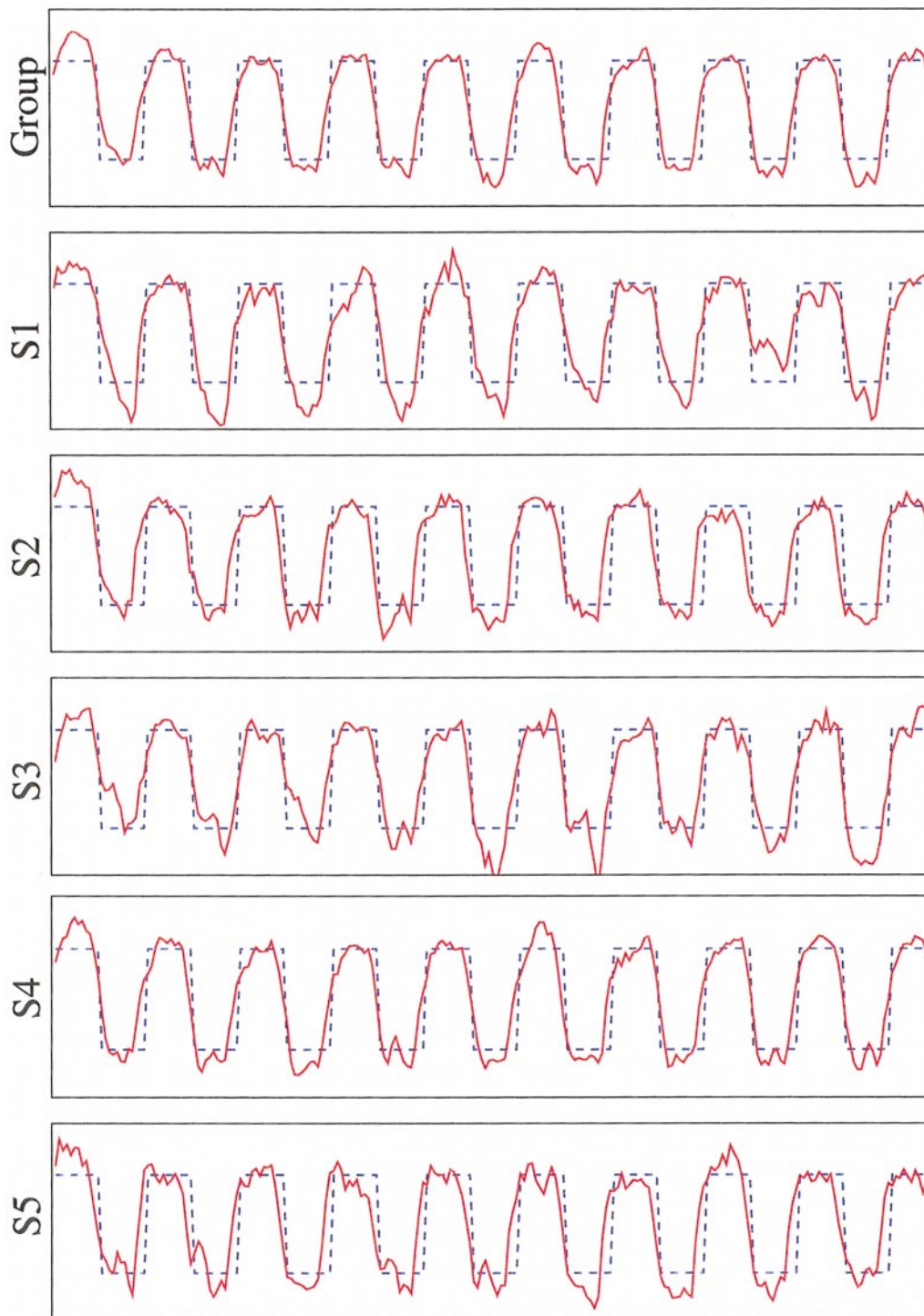
*Data processing.* The images from each subject of the respective experiments were arranged into matrices, as described in Eqs. (1) and (7), which were then passed to the ICA algorithm. For the examples presented below, we have used the extended infomax algorithm (Lee *et al.*, 1999), but we have also obtained similar results with the FastICA algorithm (Hyvärinen, 1999a). The ICA algorithm was run with a learning rate of 0.0001 performing up to 500 sweeps through the training set for the single subject data sets and up to 400 sweeps for the group data sets. Each sweep consisted of one pass through the entire data set divided into blocks of 500 data points (voxels), with the unmixing matrix being updated after each block. The algorithm terminated if the sum of the squared differences between the elements of the unmixing matrices from two subsequent sweeps was less than  $10^{-6}$ .

Most ICA algorithms, including the ones just mentioned, whiten the data (a transformation also known as sphering) using PCA (see, e.g., Hyvärinen, 1999b, sec. 5.2), as a first step in the extraction of independent components. Thus, there is the possibility to reduce the dimensionality of the data and thereby the number of independent components, by only retaining the  $M$  most principal components in the whitening. This technique was used to reduce the dimensionality of the single trial design auditory data from 864 to 200. This dimensionality reduction retained 99.76% of the variance of the original data set.

## RESULTS AND DISCUSSION

### Results

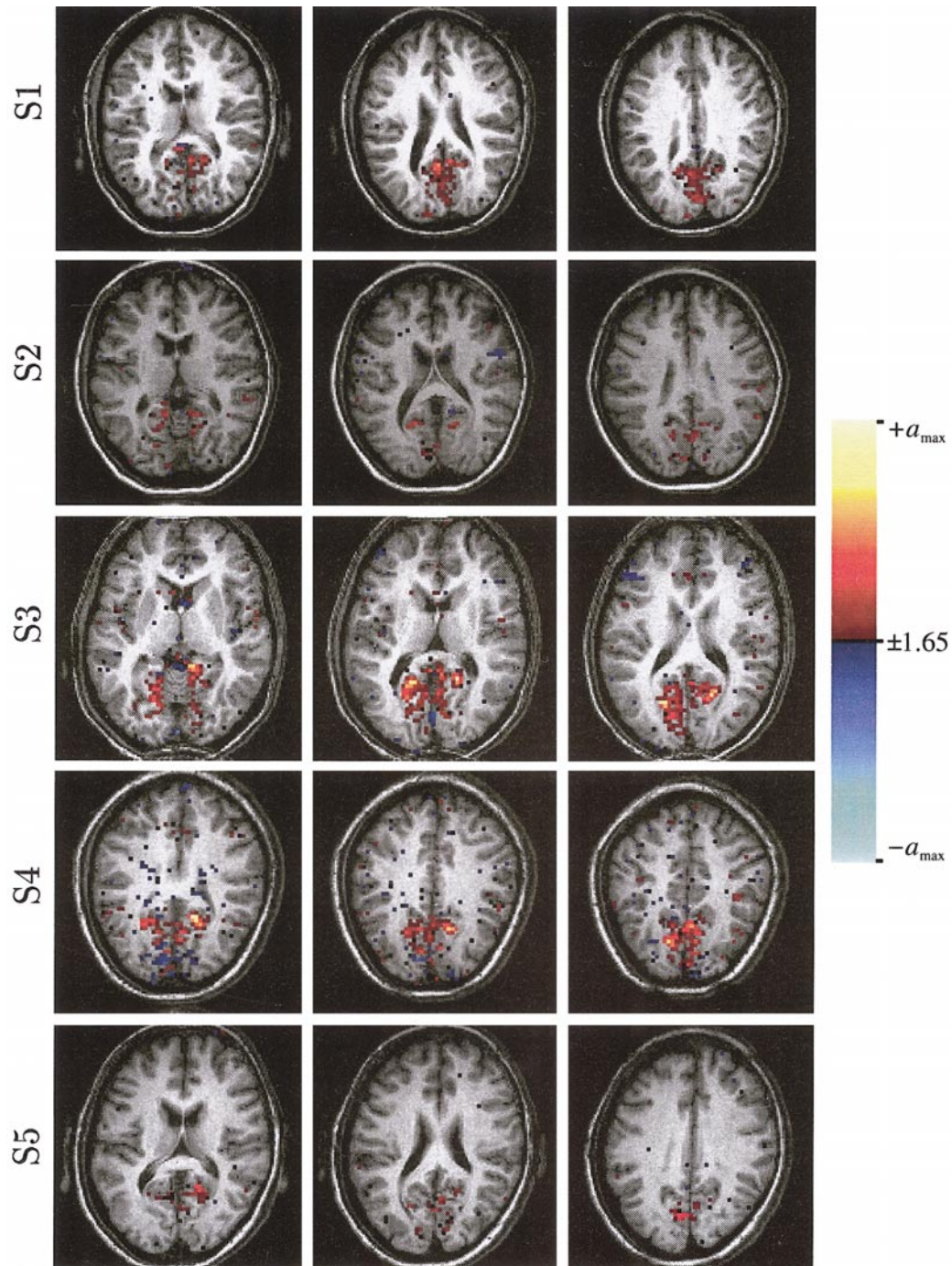
Following ICA, task correlation scores—correlations between the time courses in the resulting mixing matrix and a function representing the performance of the task after having taken account of the delay in the



**FIG. 3.** Time courses (solid red) corresponding to the independent images shown in Fig. 1 (top) and 2 (second to bottom, matching the rows of Fig. 2), plotted together with a time-shifted “box-car” function (dashed blue) representing the performance of the task after hemodynamic correction. The time courses and the function representing the task have been normalized to zero mean and unit variance prior to plotting. The scale on the horizontal axes runs from 0 to 304 (s), on the vertical axes from  $-2.0$  to  $2.0$ .

hemodynamic response—were computed, to identify task related components (McKeown *et al.*, 1998b). Furthermore, the results were inspected manually, to detect other interesting components.

*Visual stimulation.* Figures 1 and 2 show the independent images from the group and single subject ICA, respectively, associated with the time courses showing the highest correlation with the performance of the



**FIG. 4.** An independent image resulting from group ICA of the data from the visual experiment. The corresponding time course is plotted in Fig. 5. The image has been visualized in the same fashion as in Fig. 1;  $a_{\max} = 18.2$ .

task, overlaid on the anatomical images of the five subjects. The corresponding time courses are shown in Fig. 3, plotted together with the function representing the performance of the task.

In the response pattern in the group image, most of the response is located to the subject S4, but the response appears in the corresponding anatomical location also in the other subjects. In the images resulting

from ICA of the individual subjects, the spatial responses are more similar across the group. This difference between group and single subject ICA results is largely due to the varying strength (or scale) of the response between subjects in the group, combined with the normalization across all subjects; in the case of the single subject ICA results, the normalization is done for each subject individually. If the group image is

TABLE 1

Voxel Count Statistics from the Independent Images Resulting from Group and Single Subject ICA of the Data from the Visual Experiment

		> 1.65	> 1.96	> 2.58	> 3.27
Subject 1 4703	Group, grp norm.	153 (3.3%)	100 (2.1%)	64 (1.4%)	37 (0.8%)
	Group, s.s. norm.	240 (5.1%)	183 (3.9%)	102 (2.2%)	69 (1.5%)
	Single subject	232 (4.9%)	187 (4.0%)	127 (2.7%)	96 (2.0%)
Subject 2 5223	Group, grp norm.	135 (2.6%)	107 (2.0%)	65 (1.2%)	46 (0.9%)
	Group, s.s. norm.	207 (4.0%)	152 (2.9%)	111 (2.1%)	74 (1.4%)
	Single subject	306 (5.9%)	206 (4.0%)	113 (2.2%)	73 (1.4%)
Subject 3 6306	Group, grp norm.	309 (4.9%)	233 (3.7%)	163 (2.6%)	119 (1.9%)
	Group, s.s. norm.	306 (4.9%)	229 (3.6%)	162 (2.6%)	118 (1.9%)
	Single subject	250 (4.0%)	183 (2.9%)	127 (2.0%)	91 (1.4%)
Subject 4 4479	Group, grp norm.	352 (7.9%)	305 (6.8%)	227 (5.1%)	183 (4.1%)
	Group, s.s. norm.	202 (4.5%)	174 (3.9%)	121 (2.7%)	81 (1.8%)
	Single subject	194 (4.3%)	141 (3.1%)	88 (2.0%)	65 (1.5%)
Subject 5 5727	Group, grp norm.	113 (2.0%)	83 (1.4%)	38 (0.7%)	21 (0.4%)
	Group, s.s. norm.	283 (4.9%)	203 (3.5%)	130 (2.3%)	94 (1.6%)
	Single subject	298 (5.2%)	211 (3.7%)	114 (2.0%)	66 (1.2%)

*Note.* The first column simply labels the subjects and gives the total number of voxels inside the brain mask for that subject. The second column subdivides the rows according to group and single subject ICA and, for group ICA, between group and single subject normalization. Columns 3–6 give the number of voxels in the (part of the) independent image for the subject, whose absolute values exceed the limits given in the column headers; the percentages given parentheses express these numbers as fractions of the total number of voxels inside the brain mask for the subject.

normalized for each subject individually, the result is very similar to those obtained from the single subject data. These images are also in good agreement with the  $Z$  maps obtained by conventional analysis of this data set (Kruggel *et al.*, 2000). However, as pointed out by McKeown *et al.* (1998b), the spatial response pattern resulting from ICA of fMRI data are not directly comparable with the activation maps resulting from conventional methods, even though they often look similar. A more quantitative comparison between the results obtained from group and single subject ICA is given by the voxel count statistics in Table 1. For the group image these statistics are given for normalization across the group as well as for per-subject normalization.

The time courses from individual and group ICA, shown in Fig. 3, all look fairly similar. The time course from the group (top) is the most regular one, which also shows the highest task correlation (0.91). This is expected, since this time course needs to match all the subjects the group and thus becomes a kind of average. The task correlation scores computed for the individual subjects were, second to bottom, 0.87, 0.82, 0.89, 0.90, and 0.80. From the time courses, it is clear that these differences are partly due to individual differences in the hemodynamic response.

Figure 4 shows another independent image from the group ICA, whose time course is shown in Fig. 5. This time course is correlated with the absolute derivative of the function representing the task, i.e., it is correlated with the block transitions between resting and performing the task. The differences between the sub-

jects are also in this example primarily reflecting a difference in the scale of the response of the subjects. In comparison to the highlighted regions in the image shown in Fig. 1, which are situated in the visual cortex (V1), the region of the visual field which is activated during the ON-phase, the highlighted regions in this image correspond to the peripheral parts of the visual field. Thus, this component might reflect an attentional mechanism forcing the focus to the central visual field during the ON phase, while suppressing activity in the periphery (Smith *et al.*, 2000). Another possible explanation would be to relate this activity to a default mode of brain function, as suggested by Raichle *et al.* (2001).

This component is also found when doing ICA of the individual subjects. However, whereas group ICA gathers this component from individual subjects into a single group component, highlighting the similarities in the spatial response patterns, detection of these spatiotemporal similarities in the results from single subject ICA relies on manual inspection and comparison of the spatial components and associated time courses found in the different subjects. Unless some characteristics of these similarities are known *a priori*, e.g., a common time course or a common spatial response region, this is a tedious and potentially error prone task.

As a final example from this data set, Fig. 6 shows the spatial response pattern of a single subject found in one group component (bottom) and two components found doing ICA of the data from this subject alone (middle and top); the corresponding time courses are shown in Fig. 7. Both the spatial response patterns and



the time courses suggest that the group component is summarizing the two separate components found in the single subject ICA. In the bottom plot, the 'group time course' (red) has been plotted together with a time course (dotted black) constructed through a least squares regression on the two 'single subject time courses'. This linear combination lies notably closer to the group time course than any of the two single subject time courses alone (RMS difference 0.39, compared to 1.23 (top) and 0.53 (middle)). The time courses of all three components suggest that these are, at least partially, related to the performance of the task. The peaks in the spatial response patterns are located to secondary visual areas.

Also worth noting is that more than nine out of ten voxels (91%) in the group image with an absolute value above 3.27 were located to the subject shown in Fig. 6. This should be contrasted with the fact that this subject contributes with less than one fifth (17%) of the total number of voxels in the independent image from the group.<sup>4</sup> The few voxels located to other subjects did not appear in the same spatial regions or show any common spatial pattern. Thus, this component seems to primarily reflect a source from this subject only. We will return to this effect when discussing the results obtained with the data from the auditory experiment next.

*Auditory stimulation.* The results from ICA of the data from the auditory experiment were very similar to those obtained from the visual experiment. For the independent images associated with the time courses having the highest correlation with the performance of the task, the differences in the results from group and single subject ICA were primarily due to differences in the response strength between subjects in the group. As was the case with the data from the visual experiment, these independent images were in good agreement with those resulting from conventional analysis. As expected, the peaks in the spatial response patterns were located to the primary auditory cortex (Heschl's gyrus). Voxel count statistics of the kind given in Table 1 for the visual experiment data did not reveal any systematic differences between group and single subject ICA. Also here, the time courses from the group and single subject were very similar, where again the time course corresponding to the group image was the smoothest and the most regular. However, the highest correlation between the extracted time courses and the function representing the performance of the task was observed for a single subject component (0.82), closely followed by the group component (0.80). We have chosen not to illustrate these results here, because of their

similarity with the corresponding results obtained from the visual experiment data.

Instead we turn our attention to the following images, as ranked by the correlation between the corresponding time courses and the time-shifted stimulus function. These have spatial response patterns, where the regions containing the signal peaks are largely confined to a single subject. An example of this effect was seen in the results obtained with the data from the visual experiment above (see last paragraph preceding this section). For the auditory data, these images are shown in Fig. 8, for each image only including the part of the image (the subimage) corresponding to the subject where these signal peaks are localized. The top subimage contains more than 98% of all the voxels in the group image with an absolute value above 3.27, even though the number of voxels in this subimage only represents 30% of the total number of voxels in the group image. Similarly, the bottom subimage contains more than 99% of all the voxels in the group image with an absolute value above 3.27, even though the number of voxels in this sub-image only represents 35% of the total number of voxels in the group image. For both group images, the few voxels located to other subjects did not appear in the same spatial regions or show any common spatial pattern.

There is a notable similarity between the images in Fig. 8, even though they are localized in two different subjects and slice planes, and since the corresponding time courses are both correlated with the stimulus, these images may represent the same cognitive process in both subjects. We identify the region where peaks of the spatial response patterns appear to be Broca's area, believed to be associated with syntactical processing. A question is why this process is split over two components. The perhaps most probable reason is that the signal arising from the cognitive process is being mixed with signals corresponding to other sources such as noise and baseline drifts, which are different in the different subjects. Inspection of the time courses<sup>5</sup> suggests that this could be the case.

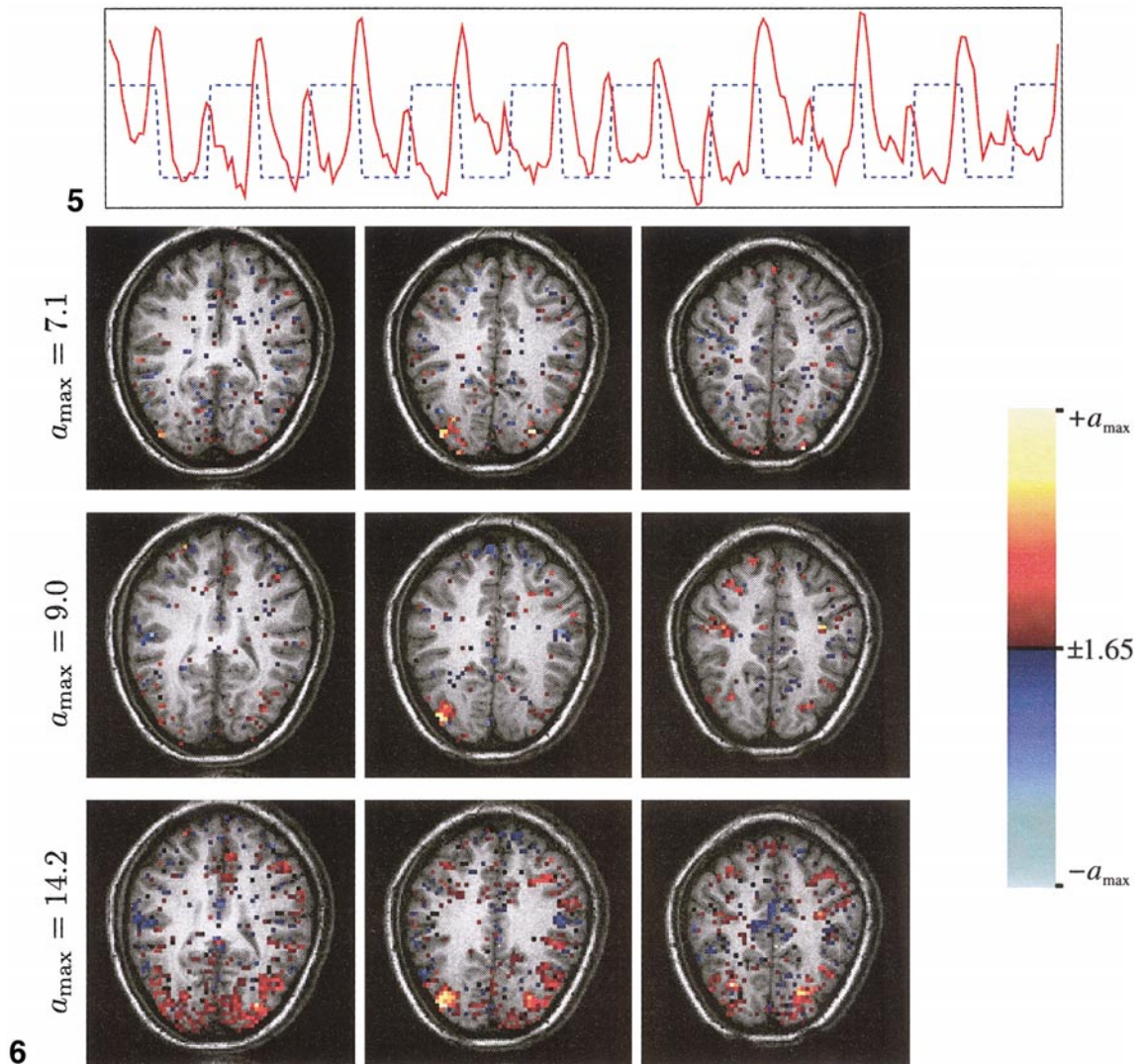
These results show that ICA of group data can extract components common to the group as well as components primarily localized to single subjects.

## Discussion

*ICA of fMRI data.* ICA offers a complementary approach to model-driven methods for the analysis of fMRI data. As shown above and in examples from the literature, ICA is capable of separating out the signal induced by the performance of the task as well as signal changes arising from other changes of the ex-

<sup>4</sup> We only count the voxels within the brain masks of the respective subjects.

<sup>5</sup> We have chosen not to include plots of these time courses, since plotting them at an adequate resolution—each time course contains 864 time steps—would require unreasonable amounts of space.



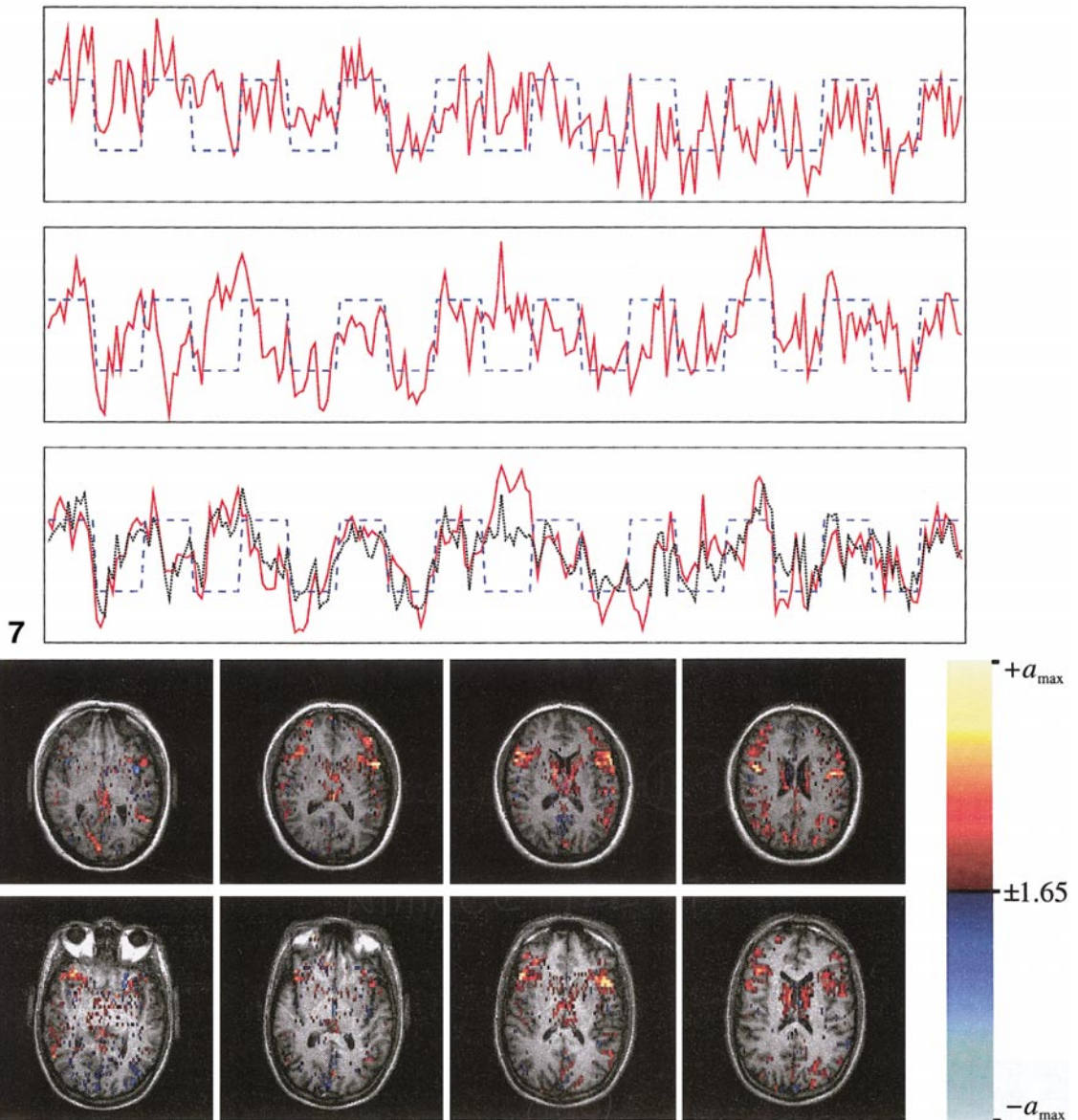
**FIG. 5.** Time course (solid red) corresponding to the independent image shown in Fig. 4, plotted together with a time-shifted “box-car” function (dashed blue) representing the performance of the task after hemodynamic correction. The time course and the function representing the task have been normalized to zero mean and unit variance prior to plotting. The scale on the horizontal axis runs from 0 to 304 (s), on the vertical axis from  $-1.7$  to  $2.6$ .

**FIG. 6.** The top two rows show two independent images resulting from ICA of the data from a single subject (S4). The bottom row shows the part corresponding to the same subject of an independent image resulting from ICA of the whole group of subjects. The images have been visualized just like the image shown in Figs. 1 and 2. The corresponding time courses are shown in Fig. 7.

perimental conditions (Biswal and Ulmer, 1999). Maybe most interesting is the possibility of detecting signals that would not necessarily emerge out of conventional analysis. One example of this was shown in Figs. 4 and 5; another one, involving components related to the stimulus in a nonlinear fashion, was presented by Berns *et al.* (1999). In contrast to conventional regression-based techniques, ICA requires no *a priori* description of what to look for in the raw data. It belongs to the group of so called unsupervised techniques, to which also PCA and clustering methods are counted. However, compared to PCA, ICA seems to be better suited for the separation of stimulus correlated

components from the observed raw data (McKeown *et al.*, 1998b; Biswal and Ulmer, 1999).

At the same time, whereas the statistical maps—e.g.,  $Z$  or  $F$  maps—resulting from conventional methods can be given a direct interpretation, indicating the confidence in the hypothesis that the agreement between observations and an *a priori* specified model is not coincidental, no such interpretation exists for the independent images. This is a direct consequence of the weak assumptions made in ICA, which neither include any *a priori* specified model for the time courses, nor any explicitly specified noise model that would account for discrepancies between the model and the observa-



**FIG. 7.** Time courses (solid red) corresponding to the independent images shown in Fig. 6, plotted together with a time-shifted “box-car” function (dashed blue) representing the performance of the task after hemodynamic correction. The additional time course plotted in black in the bottom plot is a linear combination of the time courses in the first and second row. The time courses and the function representing the task have been normalized to zero mean and unit variance prior to plotting. The scale on the horizontal axes runs from 0 to 304 (s), on the vertical axes from  $-3.0$  to  $3.0$ .

**FIG. 8.** Subimages from two different group ICA images containing the voxels of two different subjects from the auditory experiment. The images have been visualized as in Fig. 1.

tions.<sup>6</sup> The lack of common underlying assumptions make the formal justification for quantitative comparisons of independent images and statistical maps difficult. McKeown (2000) tries to bridge this gap between ICA and traditional analysis by proposing a hybrid model, where time courses extracted by ICA are subsequently used in traditional linear regression models.

<sup>6</sup> In ICA, any ‘noise’ is being accounted for by one or more independent components, but no assumptions are made about the nature of this ‘noise’.

*ICA of fMRI group study data.* Applied to group data, ICA has the same advantages and disadvantages as when applied to single subject data, but also some new aspects emerge. The most important distinction, compared to performing ICA of the data from the separate subjects in the group individually, is the extraction of a single set of time courses common to the whole group. This is accompanied by a set of individual spatial response patterns for the subjects in the group. For any particular component, the significant peaks of this

spatial response pattern may be isolated to a few or a single subject(s). Thus, ICA may extract components specific to a subset of the group or even to single subjects, as was illustrated in Fig. 8. In contrast to conventional methods for analyzing group data, our ICA approach avoids the need for any spatial normalization or *a priori* specified temporal behavior.

An unavoidable consequence of restricting ICA to a common set of time courses for the whole group is that weak sources with different temporal characteristics across subjects will get suppressed. This does not necessarily have to be negative, since noise sources in the different subjects will almost certainly have different temporal characteristics and hence be among those sources that get suppressed. Another possible consequence is that some more subtle differences between subjects may get lost in an ICA of a group data set. Figure 3, for example, illustrates what appears to be small differences in the hemodynamic response between the different subjects in the group, differences which cannot be seen from the group ICA result.

When choosing between doing group and single subject ICA for groups of subjects, one is probably best guided by the purpose of the analysis. With group analysis components common to the group can be captured and identified relatively easily, whereas analysis of single subjects will be better suited to reveal individual features. Our results do not suggest that there would be any substantial differences between group and single subject ICA in the spatial response patterns representing the same cognitive process, although a process that might be split into two or more components by single subject ICA may be represented as a single component by group ICA, as illustrated in Figs. 6 and 7.

From a computational point of view, working with data from a group of subjects, will obviously require more computation than if we are working with data from a single subject from that group. However, the computational effort does not have to grow linearly. Our approach effectively means that we are increasing the number of samples from the observed variables, and hence the computation may converge with fewer iterations through the (larger) data set.

In this paper, we have concentrated on the task related independent components. If we consider other components, it has been suggested that ICA finds components corresponding to e.g. head movements and baseline drifts. Whether this is also the case when doing group ICA remains to be investigated. However, since the corresponding sources are likely to have different temporal characteristics in different subject, they should either appear in components with the peaks of the spatial response pattern localized to a single subject or not at all, i.e. get suppressed.

Another interesting question is whether ICA could also be applied to group study data from EEG or MEG

experiments. As already pointed out, ICA of EEG/MEG data is carried out in the temporal rather than in the spatial domain, as is the case with fMRI data. Thus, ICA of group study EEG/MEG data would result in a set of spatial response patterns common to the whole group, combined with individual temporal responses for each subject.

During the final stages of preparation of this manuscript, we became aware that Calhoun *et al.* (2001) have recently proposed an alternative scheme for ICA of fMRI data from a group of subjects. However, their method differs from ours in several ways. The perhaps most important difference is that their method yield a mixing matrix which "will be approximately separable across subjects." In other words, their method yield separate time courses for the different subjects in the group. In contrast, our method result in a mixing matrix which is common to all the subjects in the group, although certain components may be primarily located to single subjects.

## CONCLUSION

ICA offers a new approach for analyzing fMRI group study data, to be added to the collection of already existing tools. Like any other method, ICA has its pros and cons, that have to be taken into account whenever its merits for analyzing a particular data set are being considered. We believe that its most useful role will be as an exploratory tool, for identifying components which are related to, although perhaps not directly correlated with the performance of the task, as well as for differentiating between components common to the whole group and components that only manifest themselves in a smaller set of subjects.

## ACKNOWLEDGMENT

The authors thank the anonymous reviewers for useful comments and suggestions.

## REFERENCES

- Berns, G. S., Song, A. W., and Mao, H. 1999. Continuous functional magnetic resonance imaging reveals dynamic nonlinearities of "dose-response" functions for finger opposition. *J. Neurosci.* **19**: RC17.
- Biswal, B. B., and Ulmer, J. L. 1999. Blind source separation of multiple signal sources of fMRI data sets using independent component analysis. *J. Comput. Aided Tomogr.* **23**(2): 265–271.
- Cardoso, J.-F. 1998. Blind signal separation: Statistical principles. *Proc. IEEE* **9**(10): 2009–2025.
- Friston, K. J. 1998. Modes or models: A critique on independent component analysis for fMRI. *Trends Cogn. Sci.* **2**: 373–375.
- Hyvärinen, A. 1999a. Fast and robust fixed-point algorithms for independent component analysis. *IEEE Trans. Neural Networks* **10**(3): 626–634.

- Hyvärinen, A. 1999b. Survey on independent component analysis. *Neural Comput. Surveys* **2**: 94–128. <http://www.icsi.berkeley.edu/~jagota/NCS>.
- Jolliffe, I. T. 1986. *Principal Component Analysis*. Springer-Verlag, New York.
- Kruggel, F., Wiggins, C. S., Herrmann, C. S., and von Cramon, D. Y. 2000. Recording of event-related potentials during functional MRI at 3.0 tesla field strength. *Magn. Reson. Med.* **44**(2): 277–282.
- Lee, T.-W., Girolami, M., Bell, A. J., and Sejnowski, T. J. 2000. A unifying information-theoretic framework for independent component analysis. *Comput. Math. Appl.* **31**(11): 1–21.
- Lee, T.-W., Girolami, M., and Sejnowski, T. J. 1999. Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. *Neural Comput.* **11**(2): 417–441.
- Makeig, S., Bell, A. J., Jung, T.-Z., and Sejnowski, T. J. 1996. Independent component analysis of electroencephalographic data. In *Advances in Neural Information Processing Systems* (D. S. Touretzky, M. C. Mozer, and M. E. Hasselmo, Eds.), Vol. 8, pp. 145–151. MIT Press.
- McKeown, M. J. 2000. Detection of consistently task-related activations in fMRI data with hybrid independent component analysis. *NeuroImage* **11**: 24–35.
- McKeown, M. J., Jung, T.-Z., Makeig, S., Brown, G. G., Kindermann, S. S., Lee, T.-W., and Sejnowski, T. J. 1998a. Spatially independent activity pattern in functional MRI data during the Stroop color naming task. *Proc. Natl. Acad. Sci. USA* **95**: 803–810.
- McKeown, M. J., Makeig, S., Brown, G. G., Jung, T.-Z., Kindermann, S. S., Lee, T.-W., Bell, A. J., and Sejnowski, T. J. 1998b. Analysis of fMRI data by blind separation into independent spatial components. *Hum. Brain Mapp.* **6**: 160–188.
- Meyer, M., Friederici, A. D., von Cramon, D. Y., Kruggel, F., and Wiggins, C. J. 1998. Auditory sentence comprehension: Different BOLD patterns modulated by task demands as revealed by a 'single-trial' fMRI-study. *NeuroImage* **7**(4): S181.
- Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., and Schulman, G. L. 2001. A default mode of brain function. *Proc. Natl. Acad. Sci. USA* **98**(2): 676–682.
- Smith, A. S., Singh, K. D., and Greenlee, M. W. 2000. Attentional suppression of activity in the human visual cortex. *NeuroReport* **11**(2): 271–276.
- Tipping, M. E., and Bishop, C. M. 1999. Probabilistic principal component analysis. *Stat. Method. J. R. Stat. Soc. Series B* **61**(3): 611–622.
- Vigário, R., Jousmäki, V., Hämäläinen, M., Hari, R., and Oja, E. 1998. Independent component analysis for identification of artifacts in magnetoencephalographic recordings. In *Advances in Neural Information Processing Systems* (M. I. Jordan, M. J. Kearns, and S. A. Solla, Eds.) Vol. 10, pp. 229–235. MIT Press.