Deterministic and stochastic features of fMRI data: implications for analysis of event-related experiments

Martin J. McKeown [a,b,c,*], Vijay Varadarajan [d], Scott Huettel [a,f],
Gregory McCarthy [a,e,f,g]

[a] *Brain Imaging and Analysis Center (BIAC), Center for Cognitive Neuroscience, 254E Bell Research Building, Box 3918, Duke University Medical Center, Durham, NC 27710, USA*
[b] *Department of Medicine (Neurology), Duke University, Durham, NC, USA*
[c] *Department of Biomedical Engineering, Duke University, Durham, NC, USA*
[d] *Department of Electrical and Computer Engineering, Duke University, Durham, NC, USA*
[e] *Department of Radiology, Duke University, Durham, NC, USA*
[f] *Department of Psychology, Duke University, Durham, NC, USA*
[g] *Department of Veterans Affairs Medical Center, Durham, NC, USA*

Abstract

As the limits of stimuli presentation rates are explored in event-related fMRI design, there is a greater need to assess the implications of averaging raw fMRI data. Selective averaging assumes that the fMRI signal consists of task-dependent signal, random noise, and non-task dependent brain signal that can be modeled as random noise so that it tends to zero when averaged over a practical number of trials. We recorded a total of four fMRI data series from two normal subjects (subject 1, axially acquired; subject 2, coronally acquired) performing a simple visual event-related task and a water phantom with the same fMRI scanner imaging parameters. To determine which fraction of the fMRI data was deterministic as opposed to random, we created different data subsets by taking the odd or even time points of the full data sets. All data sets were first dimension-reduced with principal component analysis (PCA) and separated into 100 spatially independent components with independent component analysis (ICA). The mutual information between best-matching pairs of components selected from full data set–subset comparisons was plotted for each data set. Visual inspection suggested that 45–85 components were reproducible, and hence deterministic, accounting for 79–97% of the variance, respectively, in the raw data. The reproducible components exhibited much less trial-to-trial variability than the raw data from even the most activated voxel. Many (22–47) of reproducible components were significantly affected by stimulus presentation ($P < 0.001$). The most significantly-stimulus-correlated component was strongly time-locked to stimulus presentation and was directly stimulus correlated, corresponding to occipital brain regions. However, other spatially distinct task-related components demonstrated variable temporal relationships with the most significantly stimulus-correlated component. Our results suggest that the majority of the variance in fMRI data is in fact deterministic, and support the notion that the data consist of differing components with differing temporal relationships to visual stimulation. They further suggest roles for restricting interpretations of the spatial extent of activation from event-related designs to a specific region of interest (ROI) and/or first separating the data into spatially independent components. Averaging the time courses of spatially independent components time-locked to stimulus presentation may prevent possible biases in the estimates of the spatial and temporal extent of stimulus-correlated activation and of trial-to-trial variability. © 2002 Published by Elsevier Science B.V.

Keywords: Deterministic; Stochastic; Independent component analysis; Data decomposition; Event-related design

1. Introduction

1.1. Event-related designs

Increasingly, functional magnetic resonance imaging (fMRI) experimental designs incorporate event-related

---

* Corresponding author. Tel.: +1-919-684-0065; fax: +1-919-681-7033; http://www.duke.edu/∼mckeown/index.htm
E-mail address: martin.mckeown@duke.edu (M.J. McKeown).

paradigms (e.g. (Buckner et al., 1998; McCarthy et al., 1996)). With event-related designs, stimulus events are not clumped together within an extended block, but are presented singly and sufficiently separated in time so that the mean latency, rise, amplitude, duration, and recovery in response to the experimental stimulus can be estimated.

A major disadvantage of event-related fMRI designs is that the slow time course of the blood-oxygenation-level dependent (BOLD) response requires long inter-stimulus intervals on the order of 15 s to avoid overlapping responses from successive events (Bandettini and Cox, 2000; Howseman et al., 1998). Rather than accept these long intertrial intervals, many investigators have allowed overlapping responses in order to have more trials per experiment, relying on post-processing methods to estimate the mean effect of a single stimulus (Buckner et al., 1996, 1998; Rosen et al., 1998; Burock et al., 1998). In general, such post-processing methods have assumed a linear, time-invariant model of overlapping evoked responses (Friston et al., 1994; Boynton et al., 1996; Buckner et al., 1998), despite evidence for non-linearity under realistic trial conditions (e.g. Vazquez and Noll, 1998; Huettel and McCarthy, 2000; Huettel et al., 2001).

1.2. Statistical features in fMRI data

Post-processing models typically explicitly or implicitly assume that fMRI data consist of deterministic and random features. In this paper, we refer to a model as deterministic if the state of a dynamical system at selected time points or points in space can be modeled predictably from knowledge of the system at other time points or other points in space. In contrast, we define a signal as random if knowledge of the signal at some time points or points in space provides no information (i.e. is statistically independent from) the signal at other points in time or space. For example, consider the standard regression equation:

$$X = G\beta + \varepsilon,$$

where, X is the n by v matrix of fMRI data, with n being the number of time points and v is the number of voxels, G is an n by p design matrix, where p is the number of regressions, β is a p by v matrix of regression coefficients, and ε is a n by v matrix of errors, typically assumed to be independent and identically distributed (iid) in a Gaussian manner. The above equation contains a completely deterministic portion ($G\beta$) and a completely random portion (ε). Successful models allocate the deterministic and random aspects of the data to the appropriate terms of the model (McKeown, 2000).

Even the simple act of averaging data, a standard way to analyze data from traditional electrical event-related

potential (ERP) studies, incorporates these ideas. When ERP studies are analyzed, the EEG data are divided into epochs time-locked to stimulation presentation, which are then averaged to obtain a mean response to the stimuli (McCarthy, 1999). This method of analysis implicitly assumes that the data can be accurately modeled as a deterministic signal that is precisely time-locked to stimulus presentation and is corrupted with random noise that will tend to zero when averaged over many trials. Although ERP researchers appreciate that most artifact-free EEG signal is in fact representative of brain activity that is not time-locked to the stimulus, all brain signals not precisely time-locked to stimulus presentation are handled as noise with this model. However, if underlying components of the data are not completely random, with respect to stimulus presentation, they may tend to average to values other than zero, introducing biases in the estimates of stimulus-locked signals. For example, it has long been known that various features of ERP data may have distinct temporal relationships to differing aspects of the experiment (Kutas et al., 1977; McCarthy and Donchin, 1981). Some ERP components may be related to a cognitive or behavioral response, while others may be directly related to stimulus presentation. Since response time is distributed over a range, averaging the data time-locked to response time will tend to accentuate components of the data related to response, but this will result in temporally misaligned stimulus response components and vice-versa.

Recently, Jung et al. have employed the technique of independent component analysis (ICA) to separate ERP data into temporally independent components (Jung et al., 2000a). With this method, they demonstrate that ERP components having no direct relationship to stimulus presentation, but a strong temporal relationship to behavioral response to the stimulus, can be robustly separated from ERP components that are directly stimulus-induced. As such, raw data that are averaged and time-locked to stimulation presentation contain the temporally smeared effects of response-dependent components, as these components do not tend to average to zero, even though the response-dependent components are essentially independent to direct stimulus presentation.

Part of the appeal of event-related fMRI designs is that they build on the prior experience gleaned from the analysis of ERP designs, which in turn facilitates the integration of information gleaned from fMRI studies and established ERP studies. As in ERP studies, analysis of fMRI data from event-related designs assumes that non-stimulus-correlated signal will tend to zero when averaged over many trials. However, the number of non-overlapping trials that can be employed in an fMRI study is usually less than an ERP study, in that it is practically limited to a few hundred across all experi-

mental conditions. Even if a feature of the fMRI data would eventually tend to zero when averaged over hundreds of trials, this may not necessarily be the case when averaged over the more common case of 30–40 trials per condition. A further problem is that, in contrast to EEG data, characterization of artifacts such as cardiac and respiratory pulsation and movement are less well known and not as readily identified.

Accurate isolation of non-stimulus-correlated but non-random aspects of fMRI data is a worthwhile goal, as adequately modeling their effects is a first step toward minimizing their influence on estimates of stimulus-correlated brain activity. For example, in a standard regression approach using the general linear model (Friston, 1996), the total variance of the data is assumed to be appropriately allocated into deterministic and random components. The deterministic fraction may contain, as nuisance regressions, non-brain activity such as cardiac pulsation. Under the limitations imposed by the statistical assumptions of the model, the contribution of these nuisance regressions can be stripped away from the raw data to more accurately reveal underlying hypothesized stimulus correlated activity.

It is clear that, whether employing a standard GLM regression approach or averaging fMRI data in an event-related design, researchers make implicit assumptions about the underlying nature of the stochastic and deterministic features of the data (McKeown, 2000). In fact, the above discussion touches on a fundamental question: what fraction of the variance of fMRI data is actually deterministic? Conversely, what fraction is truly random? What fraction of the data is really deterministic but can be modeled as random processes when examined at the time points of stimulation? These fundamental questions have not been widely discussed in the literature despite the explicit requirement of the modeler to address these issues.

Some clues suggest that fMRI data is heavily corrupted with Gaussian noise, although this is a standard of many models. McKeown and Sejnowski, by calculating the log-likelihood of observing the data under an assumed linear model (ICA), demonstrated that true fMRI data is unlike what would be expected by a relatively small number of deterministic components corrupted by zero-mean Gaussian distributed random noise (McKeown and Sejnowski, 1998). In a related study, neither training on alternate even and odd time points of an fMRI data set nor corrupting the data with pure Gaussian noise significantly affected the estimates of stimulus-correlated components of interest, suggesting that the stimulus-correlated component was deterministic and reproducible (McKeown et al., 1998a,b,c).

An indirect way to determine if a model of fMRI data has adequately allocated stochastic and deterministic features of the data is to assess overall statistical model performance. Assessing model performance is a complicated issue, and involves a balance between adequately fitting the data, not over-fitting the data (i.e. 'fitting the noise'), and the number of parameters in the model. A practical method to assess the adequacy of a model is to force the model to predict new data. Often, because obtaining data is relatively expensive, other statistical measures, such as predictive residual sum of squares (PRESS) (Myers, 1986; McKeown and Sejnowski, 1998) or bootstrapping methods (Press et al., 1992) are performed.

In this paper, we use ICA (McKeown et al., 1998a,b,c; Makeig et al., 1997; McKeown and Sejnowski, 1998; McKeown, 2000) to separate fMRI data from three sessions with human subjects and from a single session with a phantom into spatially independent components and their associated time courses. By calculating independent components on a subset of time points, we will determine which components are reproducible, and hence represent deterministic features of the data. We show that some reproducible components are evident in the phantom, suggesting that the scanner itself imposes deterministic structure on the data, and that many more reproducible components are the result of endogenous brain signals. We then show that the contribution of some components, presumably based on endogenous brain signals, do not sum to zero when averaged time-locked to stimulus presentation after a reasonable number of trials, and have the potential to introduce biases into the temporal estimates of stimulus-locked activity.

2. Methods

In order to explore the deterministic and stochastic portions of fMRI data, we performed a total of three series of a simple event-related fMRI visual experiment from two normal subjects. For each series, ninety-five visual stimuli, consisting of black and white radial checkerboards that subtended about 20 by 15° of visual angle, were presented singly for 500 ms. The interval between successive stimuli varied randomly between 14 and 18 s. For the first subject, two series were recorded, each consisting of ten contiguous 5 mm-thick axial slices that were acquired parallel to the line connecting the anterior and posterior commissures (axial imaging plane) on a 1.5 T machine. Functional gradient echo-planar images were acquired at a TR of 1 s (TE: 40 ms, Flip Angle: 81°, FOV: 24 cm, matrix: 64^2, in-plane resolution: 3.75 mm^2). For the second subject, one series was recorded, consisting of 12 slices that were acquired coronally with the same imaging parameters. These resulted in 1464 volume acquisitions for both series from the first subject, 1476 for the series from the second subject. For comparison, a phantom, consisting of a

water-filled ball, was imaged with the same parameters. A total of 512 volume acquisitions of the phantom were collected.

All the fMRI data were temporally aligned to compensate for interleaved slice acquisition. Since we were investigating responses of the sluggish BOLD signals to individual stimuli, the data were low-pass filtered with a Hanning window of length four to increase the signal-to-noise ratio (Press et al., 1992). The data were not spatially smoothed. A minimum-intensity threshold was used to identify voxels in the head and/or phantom.

In order to explore the reproducibility of the data, data subsets were created by taking the odd-numbered time points. In addition, the even numbered time points from the first series were also used. This resulted in nine data sets for analysis: (subject 1-series 1, all time points; subject 1-series 1, even time points; subject 1-series 1, odd time points; subject 1-series 2, all time points; subject 1-series 2, odd time points; subject 2, all time points; subject 2, odd time points; Phantom, all time points; and Phantom, odd time points).

All data sets were reduced in dimension using principal component analysis (PCA) (Jackson, 1991). One hundred components, capturing > 93% of the data in all cases (see Section 3) were used.

ICA was applied to each dimension-reduced data set separately, in each case separating the data into 100 spatially-independent components (McKeown et al., 1998a,b,c; Jung et al., 2000a,b; Makeig et al., 2002; McKeown and Sejnowski, 1998; McKeown, 2000). See reference (McKeown et al., 1998a,b,c; Jung et al., 2000b) for a more detailed description of applying PCA as a preprocessing step to ICA. As the ICA algorithm does not assign any direct importance to the order of the components, the different sets of spatially independent components were compared for similarity. A pairwise comparison between separated components was made between subject 1-series 1 (all, odd), subject 1-series 1 (even, odd), subject 1-series 2 (all, odd), subject 2 (all, odd) and Phantom (all, odd).

Selecting a component from one group of separated components and selecting another component from the other group of separated components first created component pairs. When two component sets were compared, the mutual information between all possible pair combinations was computed to determine the components that matched the most closely. This was accomplished by separating the first component values into 20 equally spaced histogram bins exactly spanning the range of the data. The voxels in the second component that corresponded to the voxels from the first component contained in the first histogram bin were also separated into 20 histogram bins. This process was repeated until a 20-by-20 contingency table of the data was created. The computations were normalized so

that a component compared with itself would result in a value of 1, and two completely random components would result in a value of about 0 (see Press et al., eqn. 14.4.17; Press et al., 1992). The same component was not allowed to contribute to more than one component pair.

The mutual information between best-matching component pairs was plotted to determine which components were reproducible. A cutoff for the mutual information, based on the inflection point of the mutual information versus component number curve was estimated by visual inspection. Component pairs that matched above the cutoff were deemed reproducible and deterministic, and component pairs whose mutual information was below the cutoff were deemed unreproducible and assumed to refer to random noise. Specifically, the data were modeled as:

$$X = A \cdot s = [A_{\text{rep}}|A_{\text{noise}}]\begin{bmatrix} S_{\text{rep}} \\ S_{\text{noise}} \end{bmatrix} \tag{1}$$

$$A = [A_{\text{rep}}|A_{\text{noise}}] = v^T(W \cdot S)^{-1} \tag{2}$$

Here, X is the all time-point data set, of dimension n by p, where n is the number of time points and p is the number of brain voxels in all slices. The A and s matrices were partitioned to correspond to the reproducible and unreproducible components defined by the cutoff described above, v corresponds to the 100 largest eigen vectors of X and is of dimension n by 100, and W and S refer to the 100 by 100 weight and sphering matrices, respectively, obtained from the ICA analysis.

In order to assess the relative contribution to the entire data of the reproducible components, the outer product:

$$X_{\text{rep}} = [A_{\text{rep}}|0]\begin{bmatrix} S_{\text{rep}} \\ 0 \end{bmatrix} \tag{3}$$

was calculated. The relative variance in the data that was deterministic was then estimated by:

$$\frac{\sigma_{\text{rep}}^2}{\sigma_{\text{total}}^2} = \frac{\text{tr}(\text{cov}(X_{\text{rep}}^T))}{\text{tr}(\text{cov}(X^T))} \tag{4}$$

where tr() refers to the trace of a matrix, and cov() refers to the covariance matrix.

The full time point data set from the first series of the first subject and the series from the second subject was then considered for further analysis. The reproducible components were then examined to see if they had a relationship to the stimulus presentation. The time course of each reproducible component was divided into epochs by taking the five time points preceding the 13 time points following checkerboard onset. Each 19 time point epoch was linearly detrended. The mean of each component across epochs was calculated to create a mean vector, of 19 time points, M. The entire time course of the component was then correlated with the

convolution of **M** and a vector containing ones at the times of stimulus presentation. The degree of correlation was assumed to represent the amount that component was time-locked to stimulus presentation. To estimate the significance of the correlation (Press et al., 1992), under the null hypothesis that the two waveforms were not correlated, we calculated:

$$p = \mathrm{erfc}\left(\frac{|r|\sqrt{n}}{\sqrt{2}}\right) \tag{5}$$

where, n is the number of time points, and $|r|$ is the magnitude of the correlation coefficient. Components that were significant to $P < 0.001$ were selected as being affected by stimulus presentation.

The five most significantly stimulus-correlated components were overlaid on an anatomical image obtained at the same sitting. The spatial maps of the ICA components were spatially smoothed with a 6 mm full-width-half-maximum (fwhm) Gaussian filter and were threshold at $z > 1.25$.

3. Results

Each fMRI series created a data set consisting of 1464 (subject 1) or 1476 (subject 2) time points by approximately 10 000 brain voxels. Approximately 15 300 voxels were considered within the phantom. Retaining 100 principal components (PCs) captured $> 93\%$ of the variance for the full, odd, and even time point data sets from the first series of the first subject, $> 99.9\%$ of the variance for the data sets from the second series of the first subject and from the second subject, and 100% of the variance for both data sets recorded from the phantom.

The plots of mutual information between component pairs versus component number for the series 1 (even time point, odd time point) and the series 1 (all time point, odd time point) comparisons were qualitatively similar (Fig. 1). The series 2 (full, odd time point) curve was also qualitatively similar, but did not drop off as abruptly (Fig. 1). Best matching pairs had a mutual information of about 0.9 with least matching pairs having a mutual information of < 0.05, suggesting almost complete statistical independence. The phantom (full, odd time point) curve demonstrated much less mutual information between best-matching component pairs, and dropped off more rapidly.

A relatively abrupt change in the slope of phantom, subject 1-series 1, and subject 2 curves was evident on visual inspection at $n = 45$, 60 and 65, respectively. A less obvious change in the slope of the subject 1-series 2 (full, odd time point) curve was seen at $n = 80$. These values were used as the arbitrary cutoffs to differentiate between reproducible or deterministic features of the
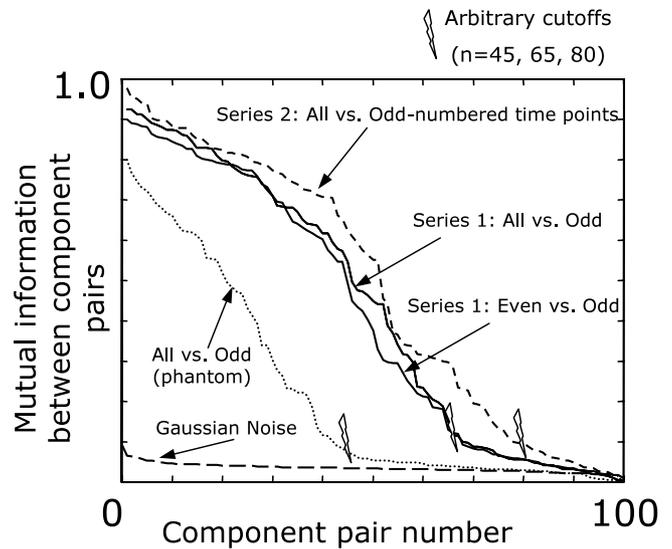

Fig. 1. Reproducibility of components. The mutual information between best matching component pairs (without duplication of any single component) is shown. The traces from the first series (solid lines), and second series (dashed line) were qualitatively similar. Note the substantially fewer number of reproducible components within the phantom data (dotted line). The data from the second subject (inflection point at 60) is not shown for clarity.

data, and non-reproducible or random features (cf. Eqs. (1) and (2)).

To assess whether our preprocessing steps of temporal smoothing and dimension reduction with PCA had a significant effect on the reproducibility of the data, we created an artificial data set with the same number of voxels and time points as subject 1-series 1, and a (full time point, odd time point) comparison was made. The data were selected from a zero mean, unit variance Gaussian distribution, preprocessed as before, and also separated into 100 maximally spatially independent components. The mutual information between all component pairs was about 0 (Fig. 1).

The ratio of deterministic variance to total variance (Eq. (4)) was 0.95 for subject 1-series 1, 0.79 for subject 1-series 2, 0.83 for subject 2, and 0.97 for the phantom data.

The 65 components from the subject 1-series 1 (all time point) analysis that corresponded to the best matching subject 1-series 1 (all time point, odd time point) comparison were used for further analysis. A total of 47 independent components were considered correlated with the stimulus presentation (Eq. (5)) to a significance of $P < 0.001$.

To ensure that the components deemed stimulus-correlated were not, in fact, independent low-frequency oscillations in the data that were by chance captured in phase with stimulus presentation, we created a pure sine wave the length of the fMRI series, with a half-period = 15 TRs (the mean ISI). Epochs of the sine wave were created by extracting out 19 time points, time-locked to

the same stimulus presentation times as before. The repeated average of the epochs was then correlated with the original sine wave (as described in the Section 2) resulting in a correlation $r = -0.0214$.

The five most significantly stimulus-correlated (all $P < 10^{-20}$) components are shown in Fig. 2. The most significantly stimulus-correlated ICA component demonstrated very little trial-to-trial variability and appeared directly task-related (Fig. 2a). The spatial distribution of this component (threshold at $z > 1.25$) was heavily weighted in occipital regions.

At least one significant component was presumably the result of task-dependent head motion (Fig. 2c), a phenomenon previously described (Rombouts et al., 1998; Bullmore et al., 1999; Thacker et al., 1999). Other significant components were weighted heavily in occipital regions (overlapping with the component in Fig. 2a), but with distinct, but still stimulus-dependent temporal profiles (Fig. 2d). Still other components involved temporal and frontal regions (Fig. 2b and e).

To determine if the significantly stimulus-correlated components had differing temporal latencies, we aligned each individual trial within each component by cross-correlating the individual trial with the mean response of that component across all trials. This epoch temporal alignment procedure (Fig. 2, column #2) had little effect on the most significant stimulus-correlated component (Fig. 2a), but had more significant effects on other components (see for example, Fig. 2d and e).

To determine the temporal relationship between the two stimulus-correlated components loading in the occipital cortex, we plotted a histogram of the differences in calculated alignment for the components depicted in Fig. 2a and d. This revealed a roughly Gaussian shape centered on zero (Fig. 3).

Since event-related data are usually analyzed by correlating averaged raw data with an empirically derived response, we estimated the effect of the stimulus-correlated components on this standard correlation procedure. The full data set was reconstructed with all significantly stimulus-correlated components (except the most significant one depicted in Fig. 2a) removed. This modified, reconstructed data set and the original full data set were then separately averaged, time-locked to stimulus presentation. The correlation at each voxel between an empirically-derived hemodynamic response function (HRF) and the two data sets was calculated. Since correlation values do not follow a Gaussian distribution, the correlation values were Fisher z-transformed (Press et al., 1992):

$$z = \frac{1}{2} \ln\left(\frac{1 + r}{1 - r}\right) \tag{6}$$

where, r is the correlation coefficient between the empirically derived HRF and the averaged data set, and z is the Fisher z-transformed result.

A scatter plot was created comparing the Fisher z-transformed correlation between the HRF and each of the two data sets at each voxel (Fig. 4). This demonstrated clear effects of the stimulus-correlated components on the correlation estimates.

To assess the spatial distribution of the changes in correlation estimates, we overlaid voxels that were considered significantly correlated with the HRF with the raw data ($P < 10^{-3}$, Bonferroni corrected) but were no longer considered significant in the modified data set. This revealed 401 voxels ($\sim 4\%$ of brain voxels, 16% of voxels previously considered activated) in a wide distribution throughout the brain, without any definitive pattern (Fig. 5).

The results from the second subject were qualitatively similar. Twenty two components out of 60 were correlated with stimulus presentation ($P < 0.001$), with the most significantly task-correlated component (Fig. 6, top) also located in occipital regions. To determine if the different task-correlated components were merely phase-shifted versions of each other, we correlated the time courses the most significantly task-related component with four other significantly task-related components, using a moving window consisting of 200 TR intervals ($= 200$ s) (Fig. 6). The correlations varied throughout the trial, and occasionally switched sign from positive to negative correlation (Fig. 6, bottom).

4. Discussion

Our results suggest that the majority of the variance of fMRI data is based on deterministic features, so that knowing the value of the fMRI signal at some time points or voxels can theoretically give a good estimate of the values at other time points or voxels. For subject 1-series 1, 95% of the variance was captured by spatially independent components that were reproducible when only a subset of time points was used to calculate the components, 79% of the variance was captured by reproducible components in the subject 1-series 2, 83% of the data were captured by reproducible components from Subject 2, and 97% in the phantom data.

Fig. 2. Five independent components most significantly (all $P < 10^{-20}$) affected by task (column 1). The time courses of these components were divided into 19-time-point epochs that were time-locked to stimulus presentation, linearly detrended, and overlaid. The heavy white lines correspond to the mean response. (a, inset) The raw time course of a typical activated voxel ($P < 10^{-13}$, Bonferroni corrected) demonstrated much more trial-to-trial variability than the respective ICA component time courses (column 2). Component time courses were aligned by cross-correlating the independent component time course from an individual stimulus presentation with the mean from all stimulus presentations. Note the minimal effect the alignment procedure had on component (a), compared with the other components (b)–(e).
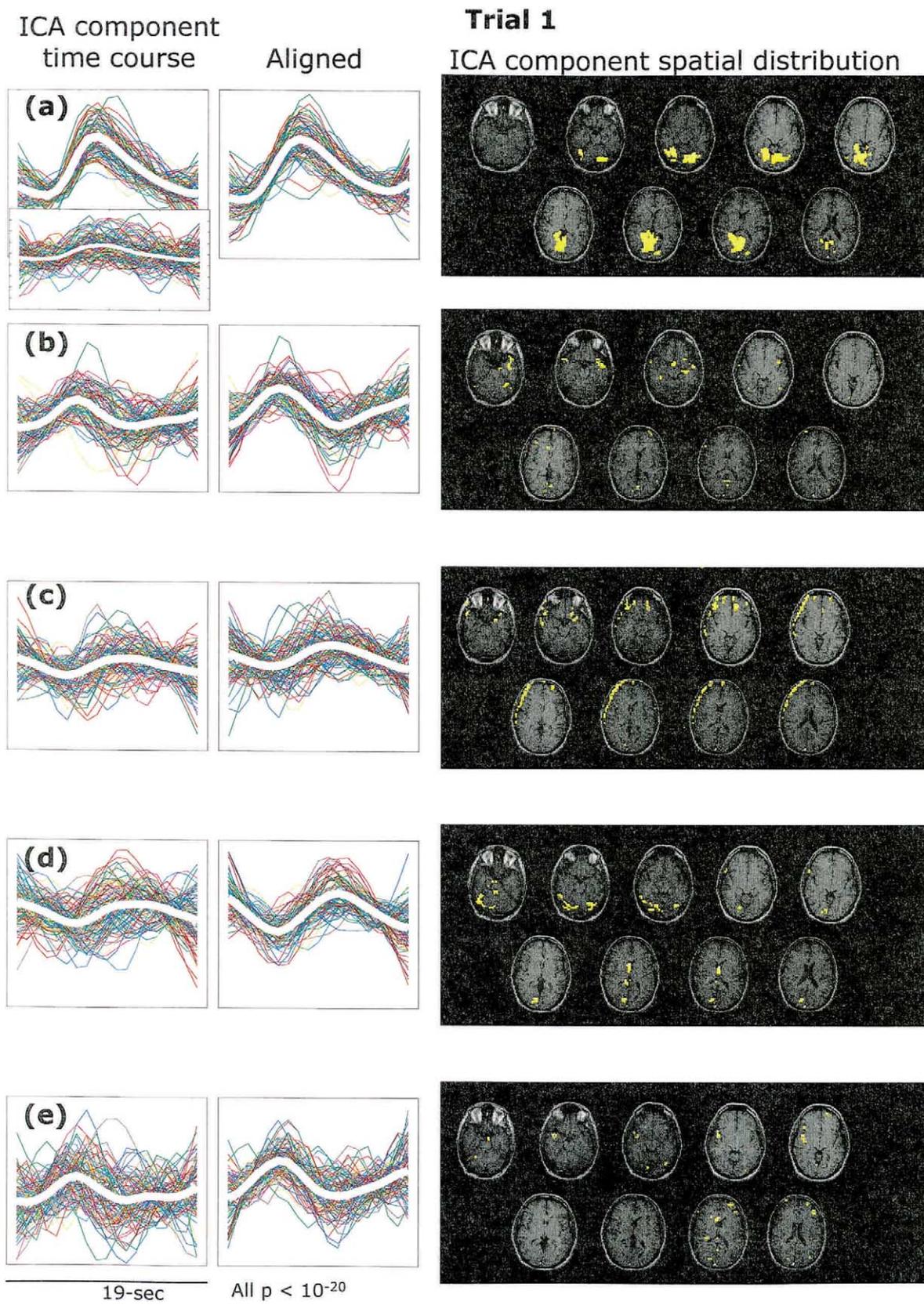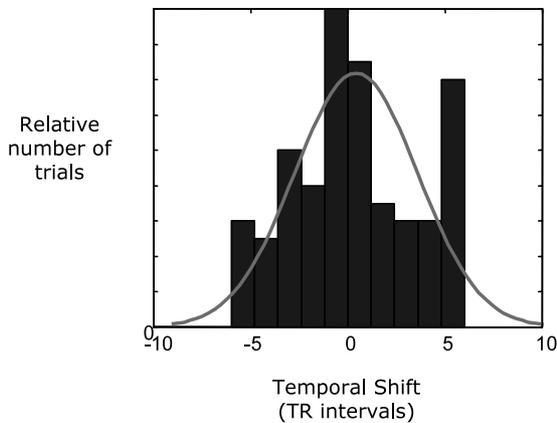
Fig. 2

Fig. 3. Relative temporal shift between two task-dependent components. The relative temporal shift of the component shown in Fig. 2(d) compared with the component in Fig. 2(a) is shown. Both these components were heavily weighted in occipital distributions (Fig. 2), and had distinct temporal profiles. Note the variability in the temporal alignment, demonstrating that these two components were not time-locked to one another. A Gaussian curve with the same variance is overlaid (solid line).

It is perhaps not surprising that an inert water phantom would result in predictable (largely T2-weighted) fMRI signal. This predictability was unlikely to be due to our preprocessing techniques, because applying the same techniques to pseudorandom noise resulted in very little deterministic signal (Fig. 1). Spatial randomness is influenced by the spatial resolution of the
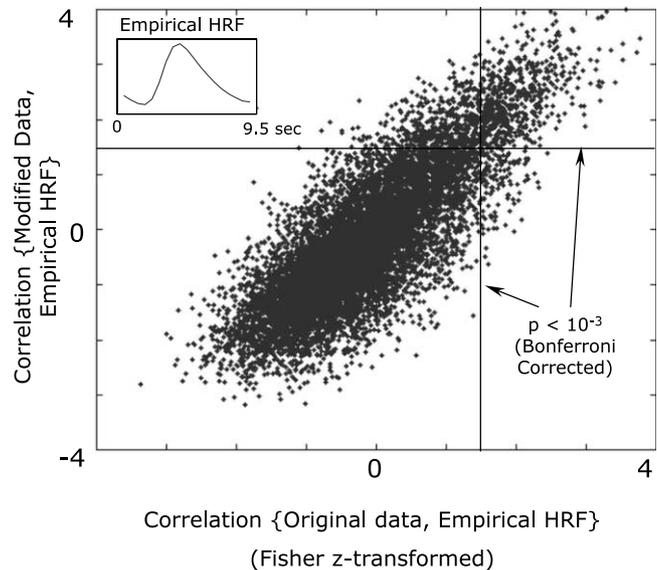

Fig. 4. Effects on correlation between averaged data and empirical response of stimulus-correlated components. The data were reconstructed with all stimulus-correlated components except the most significant (Fig. 2a) removed. After averaging time-locked to stimulus presentation, the correlation (Fisher z-transformed—see Eq. (6)) between the empirical response (inset) and the averaged reconstructed data is plotted against the correlation between the empirical response and the averaged original data for each voxel. Thresholds for significant correlation ($P < 10^{-3}$, Bonferroni corrected) (solid lines) demonstrate that some voxels became more significant (upper left quadrant), and others became less significant (lower right quadrant) after data reconstruction.
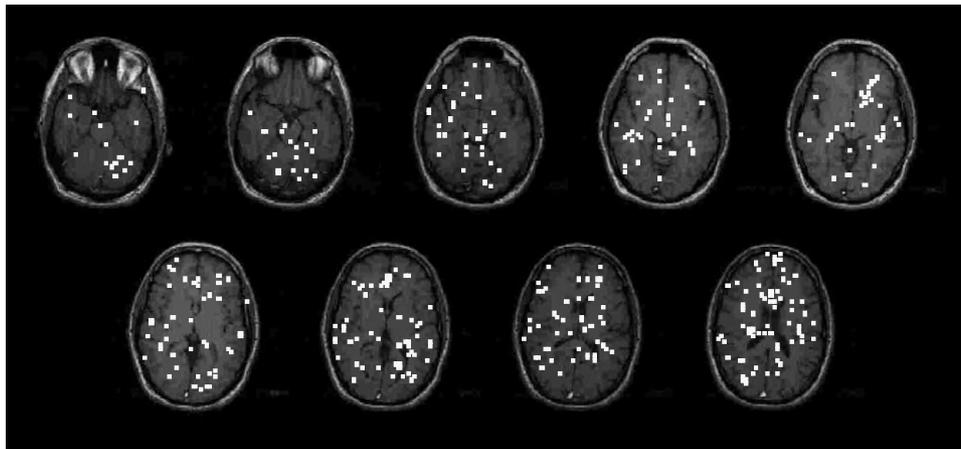
Voxels no longer significant (p < 10⁻³, Bonferroni corrected)

Fig. 5. Voxels no longer significantly active after removal of all but the most-stimulus-correlated component. After the data were reconstructed with some stimulus-correlated components removed (see Fig. 4), 401 voxels, constituting about 16% of the voxels initially considered activated, were no longer considered significantly activated ($P < 10^{-3}$, Bonferroni corrected). As shown, these newly non-significant voxels did not follow an obvious spatial pattern.

Fig. 6. The temporal relationship between different task-related components (Subject 2). The most significantly task-related component overlapped closely with regions determined with averaging the raw data (top). Other significantly task-related components had distinct spatial profiles (middle panels). A moving correlation window ($n = 200$ TRs) reveals the changing temporal relation between the most significantly task-related component and the other components (bottom panel). Note that the correlation sign changes in a few places.
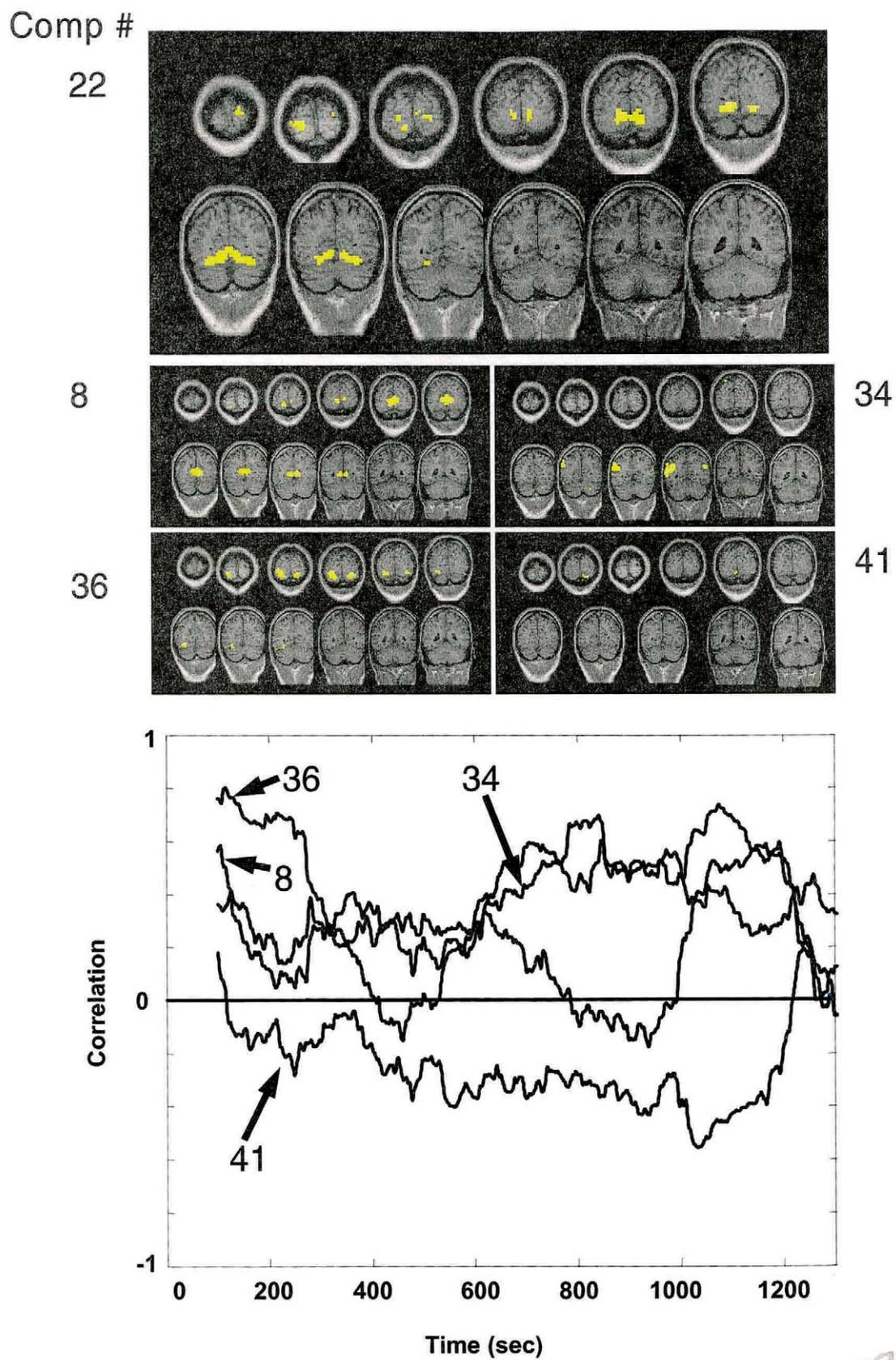
Fig. 6

scanner. Even an infinitely small point of fMRI signal will result in a finite volume of signal measured by the scanner, the so-called 'point spread function'. The distribution of an underlying signal across more than one voxel (or more than one TR interval) will tend to further reduce the inherent spatial and temporal randomness of the measured signal.

As expected, the human subject data resulted in a greater number of reproducible spatially independent components. (Fig. 1). The brain has a heterogeneous spatial structure, in addition to being a site of non-random events such as cardiac and respiratory pulsations.

Perhaps surprisingly, we also found spatially independent components whose mean time course remained significantly correlated with the empirical waveform when averaged over a reasonable number ($n = 95$) of trials (Fig. 2). We refer to these components as 'stimulus correlated' because their mean, when averaged time-locked to stimulus presentation appeared significantly correlated with the presentation of stimuli. The most significantly stimulus-correlated component appeared directly task related and was temporally and spatially similar to that expected from prior studies of averaging the raw data (Fig. 2a). However, other stimulus-correlated components appeared related to motion-induced artifact (Fig. 2c), overlapping occipital regions with differing temporal profiles (Fig. 2d), and frontal and temporal regions of unknown significance (Fig. 2b and e, Fig. 6). The presence of these stimulus-correlated components influenced the correlation with the raw data (Figs. 4 and 5), albeit without a definitive spatial structure (Fig. 5). This suggests that restricting interpretations of activation to predefined regions of interest (ROIs) may be necessary to prevent excessive false positive activation.

The different temporal alignment of these task-affected components (Fig. 2, second column and Figs. 3 and 6) is consistent with the notion that fMRI data consists of spatially and temporally overlapping components that may behave differently to the same stimulus. It also lends support to the concept, now becoming popular in the ERP literature, of aligning each ICA component individually (Jung et al., 2000a; Makeig et al., 2002) rather than all components en masse, which is in effect the result of averaging the raw data. However, aligning each component individually raises the possibility of erroneously representing as task-related any slowly undulating noise that may be present in the signal, but is completely independent from brain activity. In our post-hoc analysis, we used a pure sine wave whose half-period was the same as our mean interstimulus interval to demonstrate that our method of inferring task-related activity from an independent waveform was robust to this possibility. Examining the

baseline of a component's time course before the stimulus onset would also help in this regard.

Of the stimulus-correlated components, clearly not all were related to cortical processing, as at least one was related to stimulus correlated movement (Fig. 2c). However, isolation of stimulus-correlated components is still important, as their effects will tend to become more pronounced when data are averaged time-locked to stimulus presentation, as is typically done with event-related fMRI studies.

We note that the components themselves exhibited much less trial-to-trial variability than a typical stimulus-correlated voxel (Fig. 2a, inset). This suggests that the variability seen in an individual voxel is more the effect of spatially-overlapping deterministic processes than of truly random noise, and suggests a role for deterministic models (such as ICA) to isolate the stimulus correlated changes at a voxel. This may ultimately lead to a reduced number of total trials that may need to be averaged before a stable value is reached. Further work needs to be done to directly compare the number of trials required for stabilization of the spatial and temporal estimates of the individual stimulus-correlated components to the number of trials required when averaging the raw data.

In this study, we defined components as 'reproducible' if they were robust to the selection time points contained in the training set that was used. It is interesting to speculate how the results would differ if we trained on a subset of the voxels, as opposed to a subset of time points. In a prior study, we demonstrated that, at least for the consistently stimulus-correlated ICA components, the results were relatively robust to the exact subset of voxels used for training (McKeown and Sejnowski, 1998). We note that qualitatively similar results were obtained when different random matrices were used as the starting point of the ICA iterative training process (not shown).

The fact that the majority of fMRI data are deterministic suggests that with careful modeling, greater effective signal-to-noise ratios or larger effect sizes are possible. With increasing field strength, the baseline variability of non-stimulus-correlated signals increases along with the amplitude of stimulus-correlated signals, placing a premium on the ability of a model to strip away deterministic signals from stimulus-correlated activity of interest.

References

Bandettini PA, Cox RW. Event-related fMRI contrast when using constant interstimulus interval: theory and experiment. Magn Reson Med 2000;43(4):540–8.

Boynton GM, Engel SA, Glover GH, Heeger DJ. Linear systems analysis of functional magnetic resonance imaging in human V1. J Neurosci 1996;16(13):4207–21.

Buckner RL, Koutstaal W, Schacter DL, Dale AM, Rotte M, Rosen BR. Functional-anatomic study of episodic retrieval. II. Selective averaging of event-related fMRI trials to test the retrieval success hypothesis. NeuroImage 1998;7(3):163–75.

Buckner RL, Koutstaal W, Schacter DL, Dale AM, Rotte M, Rosen BR. Functional-anatomic study of episodic retrieval. II. Selective averaging of event-related fMRI trials to test the retrieval success hypothesis. NeuroImage 1998;7(3):163–75.

Bullmore ET, Brammer MJ, Rabe-Hesketh S, Curtis VA, Morris RG, Williams SC, Sharma T, et al. Methods for diagnosis and treatment of stimulus-correlated motion in generic brain activation studies using fMRI. Hum Brain Map 1999;7(1):38–48.

Burock MA, Buckner RL, Woldorff MG, Rosen BR, Dale AM. Randomized event-related experimental designs allow for extremely rapid presentation rates using functional MRI. NeuroReport 1998;9(16):3735–9.

Friston KJ. In: Toga AW, Mazziotta JC, editors. Statistical Parametric Mapping and other Analyses of Functional Imaging Data. Brain Mapping, The Methods. San Diego: Academic Press, 1996:363–96.

Friston KJ, Lezzard P, Turner R. Analysis of functional MRI time-series. Hum Brain Map 1994;1:153–71.

Howseman AM, Porter DA, Hutton C, Josephs O, Turner R. Blood oxygenation level dependent signal time courses during prolonged visual stimulation. Magn Reson Imaging 1998;16(1):1–11.

Huettel SA, McCarthy G. Evidence for a refractory period in the hemodynamic response to visual stimuli as measured by MRI. NeuroImage 2000;11(5):547–53.

Huettel SA, Singerman JD, McCarthy G. The effects of aging upon the hemodynamic response measured by functional MRI. NeuroImage 2001;13(1):161–75.

Jackson JE. A User's Guide to Principal Components. New York: Wiley, 1991.

Jung T-P, Makeig S, Lee T-W, McKeown MJ, Brown G, Bell AJ, Sejnowski TJ. Independent Component Analysis of Biomedical Signals. The Second International Workshop on Independent Component Analysis and Signal Separation 2000a.

Jung TP, Makeig S, Humphries C, Lee TW, McKeown MJ, Iragui V, Sejnowski TJ. Removing electroencephalographic artifacts by blind source separation. Psychophysiology 2000b;37(2):163–78.

Kutas M, McCarthy G, Donchin E. Augmenting mental chronometry: the P300 as a measure of stimulus evaluation time. Science 1977;197(4305):792–5.

McCarthy G. Event-related potentials and functional MRI: a comparison of localization in sensory, perceptual and cognitive tasks. Electroencephalogr Clin Neurophysiol Suppl 1999;49:3–12.

McCarthy G, Donchin E. A metric for thought: a comparison of P300 latency and reaction time. Science 1981;211(4477):77–80.

McCarthy G, Luby M, Gore JC, Goldman-Rakic P. Functional magnetic resonance imaging in a visual oddball task. NeuroImage 1996;3:S548.

McKeown MJ. Detection of consistently stimulus correlated activations in fMRI data with HYBrid independent component analysis (HYBICA). NeuroImage 2000;11:24–35.

McKeown MJ, Sejnowski TJ. Independent component analysis of fMRI data: examining the assumptions. Hum Brain Map 1998;6:368–72.

McKeown MJ, Jung TP, Makeig S, Brown G, Kindermann SS, Lee TW, Sejnowski TJ. Spatially independent activity patterns in functional MRI data during the stroop color-naming task. Proc Natl Acad Sci USA 1998;95(3):803–10.

McKeown MJ, Makeig S, Brown G, Kindermann S, Sejnowski TJ. Modes or models: a critique on independent component analysis for fMRI-reply to commentary by Professor Friston. Trends Cogn Sci 1998;2:375.

McKeown MJ, Makeig S, Brown GG, Jung TP, Kindermann SS, Bell AJ, Sejnowski TJ. Analysis of fMRI data by blind separation into independent spatial components. Hum Brain Map 1998;6(3):160–88.

Makeig S, Jung TP, Bell AJ, Ghahremani D, Sejnowski TJ. Blind separation of auditory event-related brain responses into independent components. Proc Natl Acad Sci USA 1997;94(20):10979–84.

Makeig S, Westerfield M, Jung TP, Enghoff S, Townsend J, Courchesne E, Sejnowski TJ. Dynamic brain sources of visual evoked responses. Science 2002;295(5555):690–4.

Myers RH. Classical and Modern Regression with Applications. Boston, MA: PWS Publishers, 1986.

Press WH, Teukolsky SA, Vetterling WT, Flannery BP. Numerical Recipes in C: The Art of Scientific Computing. Cambridge: Cambridge University Press, 1992.

Rombouts SA, Barkhof F, Hoogenraad FG, Sprenger M, Scheltens P. Within-subject reproducibility of visual activation patterns with functional magnetic resonance imaging using multislice echo planar imaging. Magn Reson Imaging 1998;16(2):105–13.

Rosen BR, Buckner RL, Dale AM. Event-related functional MRI: past, present, and future. Proc Natl Acad Sci USA 1998;95(3):773–80.

Thacker NA, Burton E, Lacey AJ, Jackson A. The effects of motion on parametric fMRI analysis techniques. Physiol Meas 1999;20(3):251–63.

Vazquez AL, Noll DC. Nonlinear aspects of the BOLD response in functional MRI. NeuroImage 1998;7(2):108–18.