NEUROIMAGE **3,** 167–174 (1996)
ARTICLE NO. 0018

A Multivariate Analysis of Evoked Responses in EEG and MEG Data

K. J. FRISTON,* K. M. STEPHAN,*,† J. D. HEATHER,* C. D. FRITH,* A. A. IOANNIDES,‡ L. C. LIU,‡ M. D. RUGG,§ J. VIETH,¶ H. KEBER,¶ K. HUNTER,‡ AND R. S. J. FRACKOWIAK*

**The Wellcome Department of Cognitive Neurology, National Hospital, Queen Square, London WC1N 3BG, United Kingdom; †Department of Neurology, University of Dusseldorf, Dusseldorf 40225, Germany; ‡Department of Physics, The Open University, Milton Keynes MK7 6AA, United Kingdom; §Wellcome Brain Research Group, School of Psychology, University of St Andrew's, St. Andrew's KY16 9JU, United Kingdom; and ¶Department Experimental Neuropsychiatry, University of Erlangen, Nuremberg, Germany*

Received August 21, 1995

This paper presents a multivariate analysis of evoked responses and their spatiotemporal dynamics as measured with electro- or magnetoencephalography. This analysis uses standard techniques (ManCova) to make possible statistical inference about differential responses, after the data have been transformed using singular value decomposition. The generality of this approach is limited only by the assumptions implicit in the general linear model and can range from simple analyses like Hotelling's T^2 test (in comparing evoked responses among different conditions) to complex analyses of a multivariate regression type (e.g., characterizing the response components associated with a behavioral or psychophysical parameter). To illustrate the technique we have characterized time-dependent changes (both within and between trials) in magnetic fields, evoked by self-paced movements. Our illustrative analysis showed that movement-evoked components were less prone to adaptation than premovement components, suggesting that functionally distinct (preparatory and early executive) biomagnetic signals show differential adaptation.

INTRODUCTION

This paper is about characterizing evoked electrical or biomagnetic brain responses and making statistical inferences about these (differential) responses. The technique employed uses multivariate analysis of covariance (ManCova) following dimension reduction using the singular vectors or spatiotemporal modes of the original data. The aim of this paper is to introduce the proposed analysis at a conceptual and procedural level. The importance of the approach lies in facilitating the application of standard multivariate techniques (of a powerful and general nature) to evoked responses in electroencephalographic (EEG) and magnetoencephalographic (MEG) data. To illustrate this application we have chosen a neurobiological question that would not be easy to address using conventional averaging approaches: Namely, "do the magnetic fields evoked by self-paced finger movements show significant adaptation with motor practice?" This paper is a methodology paper and focuses on the analysis of a single data set in order to provide operational details that could be more widely applied.

The Research Question

Our example pertains to time-dependent changes in the form of movement-related magnetic fields (MRMFs) during a self-paced, stereotyped motor act. The sequence of activation and the integration of cortical activity in the initiation and execution of voluntary movements are important issues in human motor physiology (Toro *et al.*, 1994). One key aspect of this integration relates to plasticity, adaptation, and other time-dependent neurophysiological changes during motor practice and learning. These changes can be seen over many time scales: In man, physiological adaptation during motor practice has been demonstrated in the cerebellum and SMA, over an hour or so, using PET (Friston *et al.*, 1992). Subsequently, functional imaging has been used to examine the neurophysiological correlates of motor learning within (Grafton *et al.*, 1992) and between (Karni *et al.*, 1995) scanning sessions. Here we examine time-dependent changes in evoked neural activity that are expressed over the first minute or so of motor practice using MEG data.

Self-paced voluntary movements are associated with the expression of discrete MRMFs that have been divided into a series of waveform components (Kristeva *et al.*, 1991). The readiness field corresponds to a slowly increasing magnetic field strength that precedes the movement. This component peaks at about 100 ms before movement onset and is called the motor field (MF). A further peak that occurs about 100 ms following movement onset has been labeled movement-evoked field 1 (MEF1). Additional movement-related peaks in the postmovement period have also been described. On

1053-8119/96 $18.00

the basis of their similar timing, the readiness field and MF have been identified with the Bereitshaftspotential and the peak of the negative slope seen in EEG recordings. These components are associated with motor intention and preparation (Kornhuber and Deecke, 1965; Libet *et al.*, 1982). Similarly, the MEF1 has been related to the frontal peak of the motor potential, a large amplitude peak in the EEG with frontocentral negativity and parietal positivity. For the purposes of this paper we divide the components of MRMFs into *preparatory* (i.e., the readiness field and MF) and *executive* (i.e., the movement-evoked MEF1 and subsequent components related to reafference and feedback). The hypothesis we consider is that the spatiotemporal dynamics of these two components can be dissociated in terms of their adaptation during continued performance: As any repetitive motor act becomes automatic, the preparatory and intentional elements could show time-dependent changes far in excess of the neural instantiation of the movement itself. For example, adaptation of evoked responses has been described by Walter *et al.* (1964) in their description of the contingent negative variation. The hypothesis then is that premovement components will show significant adaptation, whereas postmovement components will not.

In summary, the aim of this paper is to introduce a general multivariate approach to evoked magneto- or electrophysiological transients. To illustrate how this approach works we have used it to see if intentional and preparatory components of MRMFs show short-term time-dependent changes that are distinct from executive and movement-evoked components. The implications for averaging approaches (that assume all evoked transients are realizations of the same event) are obvious. This paper is divided into two parts. The first deals with the statistical background and operational equations required to implement the analysis and the second deals with an application to multichannel MEG data obtained from a normal subject during the repetitive performance of self-paced movements. The analysis is performed on the original data; we make no attempt to model discrete or distributed sources because the questions here relate to temporal dynamics. However, the endpoint of the analysis is a (canonical) spatiotemporal mode that can be considered as a short de-noised time series. This could be subject to source estimation techniques in the usual way [e.g., dipole models (Hamalainen *et al.*, 1993) or magnetic field tomography (Ioannides, 1995)]. To comment on the spatial topography of our results we use a V_3 transformation (Ioannides *et al.*, 1990) following decomposition of the spatiotemporal mode into a series of spatial modes. The statistical procedures described in this paper can be found in any standard introductory text on multivariate statistics. We have used Chatfield and Collins (1980).

STATISTICAL BACKGROUND

Dimension Reduction

The notation adopted here uses italics for indices and scalars and boldface for column vectors (lowercase) and matrices (uppercase). Consider N time series (e.g., MEG data from N channels) divided into I blocks, trials, or epochs (e.g., epochs that are time-locked to a sensory or behavioral event). Each epoch contains E time bins. A single spatiotemporal multivariate observation is taken to be all the $N \cdot E$ observations in one epoch. Note that there is no averaging of the data—each epoch corresponds to a single trial. The first step in multivariate analysis is to ensure that the dimensionality of the data, $N \cdot E$, is smaller than the number of observations, I. Clearly this is not the case, because there are generally more time bins than epochs; therefore the data have to be transformed. The dimension reduction proposed here is straightforward and uses a singular value decomposition (SVD) of the data to give a reduced set of components for each multivariate observation. Each component corresponds to a singular vector, eigenvector, principal component, or spatiotemporal mode (in the present context all these are equivalent). We use the term *spatiotemporal* mode to distinguish it from a *spatial* mode (see Friedrich *et al.*, 1991; Fuchs *et al.*, 1992; Friston *et al.*, 1993). A spatiotemporal mode spans not only space (channels) but also time (within an epoch). The rationale for this particular transformation is that it preserves the greatest amount of epoch to epoch variance. The singular vectors and their expression over epochs are given by

$$[\mathbf{VSU}] = \mathrm{SVD}[\mathbf{X}^*],$$

where

$$\mathbf{X}^* = \mathbf{V} \cdot \mathbf{S} \cdot \mathbf{U}^{\mathrm{T}} \tag{1}$$

and

$$\mathbf{X} = \mathbf{V} \cdot \mathbf{S}$$

Here $\mathbf{X}^*$ is an $I \times N \cdot E$ matrix of data with one mean-corrected column for each component (channel and time bin) and one row for each epoch. $\mathbf{S}$ is an $I \times I$ diagonal matrix of singular values and $\mathbf{V}$ is an $I \times I$ matrix of singular vectors (over epochs). Both $\mathbf{V}$ and $\mathbf{U}$ are orthogonal matrices. The spatiotemporal modes (over channels and time) are the columns of $\mathbf{U}$ and their expression over epochs are the columns of $\mathbf{X}$. The $I \times I$ matrix $\mathbf{X}$ has one column for every mode and one row for each epoch. The elements of $\mathbf{X}$ are x_{ij}—the activity of the jth mode in epoch i. Only the J columns of $\mathbf{X}$ and $\mathbf{U}$ associated with the J largest singular values are used in the ensuing analysis. We use a threshold of unity (after the singular values have been scaled such that their sum of squares = I). This is an arbitrary but commonly employed criterion.

The Statistical Model

In matrix notation the general linear model is

$$\mathbf{X} = \mathbf{G}\boldsymbol{\beta} + \boldsymbol{\rho}, \tag{2}$$

where $\mathbf{X}$ is designated the *response* variable. The general linear model assumes that the errors in each of the J columns of $\boldsymbol{\rho}$ have a multivariate normal distribution. The $I \times K$ matrix $\mathbf{G}$ is called the design matrix. The design matrix has one column for each of the K effects (factors or covariates) in the model. $\boldsymbol{\beta}$ is the parameter matrix with one column of parameters for each mode, containing K parameter estimates. The elements of $\mathbf{G}$ are *explanatory* variables relating to the conditions under which the observation (e.g., epoch) was made. These coefficients can be (i) covariates (e.g., reaction time, performance, interstimulus interval), in which case Eq. (2) is a familiar multivariate regression model, or (ii) indicator-type variables, taking integer values to indicate the level of a factor (e.g., experimental condition, drug) under which the response variable or event-related transient was measured. Some of these effects will be of interest and others will not. $\mathbf{G}$ can be partitioned into effects of interest $\mathbf{G}_h$ and confounds $\mathbf{G}_c$, i.e., $\mathbf{G} = [\mathbf{G}_h\ \mathbf{G}_c]$, and similarly $\boldsymbol{\beta} = [\boldsymbol{\beta}_h^T\ \boldsymbol{\beta}_c^T]^T$. We assume here that these two partitions are orthogonal, i.e., $\mathbf{G}_h^T \cdot \mathbf{G}_c = 0$, and, if they are not, they are made to be (i.e., by replacing the original $\mathbf{G}_c$ with $\mathbf{G}_c - \mathbf{G}_h \cdot (\mathbf{G}_h^T \cdot \mathbf{G}_h)^{-1} \mathbf{G}_h^T \cdot \mathbf{G}_c$). A confound is used to designate an effect that is not of interest and would otherwise "confound" the analysis of effects that are of interest. An example might be interstimulus interval. In this case $\mathbf{G}_c$ would be a mean-corrected column vector whose elements corresponded to the interstimulus interval for each epoch. The effects of interest specify the effects, differences, or changes under investigation. The simplest form for $\mathbf{G}_h$ would be a column of 1's, and this would model the effects that were common to all epochs (i.e., the averaged evoked response). Least-squares estimates of $\boldsymbol{\beta}$, say $\mathbf{B} = [\mathbf{B}_h^T\ \mathbf{B}_c^T]^T$, are given by

$$\mathbf{B} = (\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}^T\mathbf{X}. \tag{3}$$

Statistical Inference

In this section we address statistical inference about the effects of interest (modeled by $\mathbf{G}_h$). Significance is assessed by testing the null hypothesis that these effects do not significantly reduce the error variance (or alternatively the null hypothesis that $\boldsymbol{\beta}_h$ is 0). The null hypothesis can be tested in the following way. The sums of squares and products due to error $\mathbf{R}(\Omega)$ are obtained from the difference between the actual and the estimated values of $\mathbf{X}$:

$$\mathbf{R}(\Omega) = (\mathbf{X} - \mathbf{G} \cdot \mathbf{B})^T(\mathbf{X} - \mathbf{G} \cdot \mathbf{B}). \tag{4}$$

$\mathbf{R}(\Omega)$ represents an estimator of the dispersion matrix of the underlying error terms. The sums of squares and products due to effects of interest are given by

$$\mathbf{T} = (\mathbf{G}_h \cdot \mathbf{B}_h)^T \cdot (\mathbf{G}_h \cdot \mathbf{B}_h). \tag{5}$$

The error sums of squares and products under the null hypothesis $\mathbf{R}(\Omega_0)$ (i.e., if the effects of interest $\mathbf{G}_h$ do not exist) are simply

$$\mathbf{R}(\Omega_0) = (\mathbf{X} - \mathbf{G}_c \cdot \mathbf{B}_c)^T \cdot (\mathbf{X} - \mathbf{G}_c \cdot \mathbf{B}_c). \tag{6}$$

Significance can now be tested with

$$\Lambda = |\mathbf{R}(\Omega)|/|\mathbf{R}(\Omega_0)|, \tag{7}$$

where Λ is Wilk's statistic (known as Wilk's Lambda). $|\cdot|$ denotes the determinant. A special case of this test is Hotelling's T^2 test, which applies when $\mathbf{G}_h$ simply compares one condition with another (see Chatfield and Collins, 1980). Under the null hypothesis, and after transformation, Λ has a χ^2 distribution,

$$-(r - ((J - h + 1)/2)) \times \log(\Lambda) \sim \chi^2(J \cdot h), \tag{8}$$

where r is the degrees of freedom associated with the error terms and is the rank of the data minus the number of effects modeled, $I - \text{rank}(\mathbf{G})$. J is the number of modes in the J-variate response variable $\mathbf{X}$ and h is the degrees of freedom associated with the effects of interest, $\text{rank}(\mathbf{G}_h)$.

Characterizing the Effect

Once it has been established that the effects of interest are significant (e.g., the evoked response is not 0, or there are differences between activation and baseline epochs, or there is a significant effect of reaction time) the final step is to characterize these effects in terms of their spatiotemporal dynamics. This characterization uses canonical variates analysis. The objective is to find the linear combination (compound or contrast) of the components of $\mathbf{X}$, in this case the spatiotemporal modes, that best expresses the activation effects when compared to error effects. More exactly, we want to find $\mathbf{c}_1$ such that the variance ratio,

$$(\mathbf{c}_1{}^T \cdot \mathbf{T} \cdot \mathbf{c}_1)/(\mathbf{c}_1{}^T \cdot \mathbf{R}(\Omega) \cdot \mathbf{c}_1),$$

is maximized (Chatfield and Collins, 1980). Let $\mathbf{z}_1 = \mathbf{X} \cdot \mathbf{c}_1$, where $\mathbf{z}_1$ is the first canonical variate and $\mathbf{c}_1$ is a canonical vector (defined in the space of the spatiotemporal modes) that maximizes this ratio. $\mathbf{c}_2$ is the second canonical image that maximizes the ratio subject to the constraint $\text{Cov}\{\mathbf{c}_1\ \mathbf{c}_2\} = 0$, and so on. The matrix of canonical images $\mathbf{C} = [\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_J]$ is given by the

solution of the generalized eigenvalue problem,

$$\mathbf{T} \cdot \mathbf{C} = \mathbf{R} \cdot \mathbf{C} \cdot \Theta, \tag{9}$$

where Θ is a diagonal matrix of (scaled) canonical values. The corresponding canonical (spatiotemporal) modes $\mathbf{m}_j$ is obtained by rotating $\mathbf{c}_j$ back into "channel" space with the original modes over channels and time $\mathbf{U}$:

$$\mathbf{m}_j = \mathbf{U} \cdot \mathbf{c}_j. \tag{10}$$

The column vector $\mathbf{m}_1$ now contains the $N \cdot E$ values defining the spatiotemporal mode that best characterizes the effects one is interested in. In other words, $\mathbf{U} \cdot \mathbf{c}_1$ is the channel-space representation of $\mathbf{c}_1$ and $\mathbf{c}_1$ is the profile of spatiotemporal modes that maximizes the sum of squares due to interesting effects, relative to error. Subsequent canonical spatiotemporal modes $\mathbf{m}_2$, . . . , $\mathbf{m}_h$ are obtained using Eq. (10). The next step involves determining how many of the canonical modes are significant. This is done by testing for the dimensionality of the result using the canonical values in the leading diagonal of Θ (i.e., Θ_1, Θ_2, . . . , Θ_J). Under the null hypothesis the probability that the dimensionality is greater than t is tested with

$$(r - ((J - h + 1)/2 \cdot \log[\prod_{t+1}^{J} (1 + \Theta_j)] \sim \chi^2((J - t) \cdot (h - t)), \tag{11}$$

which is distributed according to the χ^2 distribution. Note that when $t = 0$ this is the same statistic as Wilk's Lambda. By computing a P value for each value of t we can infer how many of the canonical modes are significant.

The final step in this characterization is to describe the significant canonical modes. Because the canonical modes $\mathbf{m}_j$ span time, they are time series and each can be decomposed into a series of spatial modes using Eq. (1) (replacing $\mathbf{X}^*$ with the $N \times E$ matrix formed from the appropriate elements of the vector $\mathbf{m}_j$). It should be noted that this SVD is arbitrary in the sense it simply orthogonalizes the variance over time (within the epoch) that constitutes the canonical mode. Nevertheless spatial modes can be a very convenient way of characterizing and reporting complex spatiotemporal dynamics. We hope to illustrate this point below.

AN APPLICATION TO MEG DATA

The MEG Data

MEG data were obtained from a normal subject during self-paced unilateral movements of a joystick using a Siemens KRENIKUN 37-channel machine. The subject was trained to perform a joystick movement, with the right hand, every 2 s or so. The data were acquired every millisecond for 76 movements. ECG artifacts were removed using linear regression. The MEG data were high pass filtered by convolving with a 512-ms Gaussian kernel and subtracting the smoothed time series from the original. The data were then smoothed with a 16-ms Gaussian kernel. Each time series was divided into 76 epochs spanning 1000 ms before movement onset to 1000 ms after. Movement onset was defined as the onset of joystick movement as recorded by a transducer and registered in a trigger "channel." The data were decimated to 8-ms binwidths and corrected to a mean of 0 within each epoch. A large data matrix $\mathbf{X}^*$ was constructed. Each of the $I = 76$ rows of $\mathbf{X}^*$ contained the data over $E = 250$ time bins for each of the $N = 37$ channels.

MOVEMENT-RELATED MAGNETIC FIELDS

In the following sections we describe how the steps described above were applied to (i) test for evoked responses, (ii) test for time-dependent changes in these responses, and (iii) identify whether these changes represented pre- or postmovement adaptation. We also include a section that examines the validity of some of the assumptions that are implicit in the general linear model.

Dimension Reduction, the Design Matrix, and Statistical Inference

The data were reduced using SVD as described in Eq. (1). These were $J = 25$ singular values greater than unity and the epoch by epoch expression of the corresponding modes $\mathbf{X}$ were subject to ManCova as described above. In this example $\mathbf{G}_h$ contained three columns representing the effects of interest $\mathbf{G}_h = [\mathbf{G}_1\ \mathbf{G}_{\text{time}}]$. The first column $\mathbf{G}_1$ was a column of 1's modeling a time-invariant evoked response. The remaining effects were time-dependent changes in this response modelled in terms of two temporal basis functions constituting the two columns of $\mathbf{G}_{\text{time}}$. These functions comprised an exponential decay (decay parameter = 1 min) and its derivative. Both were mean corrected and Euclidean normalized. Linear combinations of the these functions can model a wide variety of time-dependent forms of adaptation. In this analysis there were no confounds and consequently $\mathbf{G}_c$ was empty. The resulting design matrix $\mathbf{G}$ is shown on the top left of Fig. 1 in image format. Note that this approach to detecting evoked responses could not be effected with conventional averaging because we have explicitly allowed for, and modeled, (unspecified) nonlinear changes in the evoked response with time. Wilk's Lambda suggested that the evoked responses were very significant ($P < 0.0001$, $\chi^2 = 215.7$, $df = 75$).

Characterizing the Response

The nature of this significant response is itself multivariate and can be characterized by the canonical (spatiotemporal) modes $\mathbf{m}_1$, $\mathbf{m}_2$, . . . , $\mathbf{m}_h$. In this case $h = \text{rank}(\mathbf{G}_h) = 3$ and there are only three canonical modes. Their relative importance is reflected in their respective canonical values, which are shown in the bottom left of Fig. 1. The filled bars represent canonical modes that can be considered significant according to the test for the dimensionality of the response above. It is seen that only the first two modes are significant. The expressions over epochs $\mathbf{z}_1$, $\mathbf{z}_2$, and $\mathbf{z}_3$ are seen on the bottom right. The first spatiotemporal mode $\mathbf{m}_1$ is relatively constant in its expression over epochs. The second mode shows a profound time-dependent behavior. The first canonical mode $\mathbf{m}_1$ is shown in more detail in Fig. 2. The top is again $\mathbf{z}_1$—the expression over epochs plotted as a function of time. The bottom shows the spatial distribution at selected points in time, within an epoch. These data are displayed using a V_3 transformation (Ioannides *et al.*, 1990). The V_3 transformation is a simple linear operator that uses the partial derivatives of the signals, in space, to produce an image that gives a rough idea of the underlying current source distribution (it can be likened to edge enhancement in image restoration). It has been applied previously to the analysis of single-trial MEG data (Liu and Ioannides, 1995). Following V_3 transformation the lengths of the ensuing flux vectors were projected onto an ellipsoidal surface for display. It is immediately obvious that the dynamics of this spatiotemporal mode are complicated and not easy to describe as they stand. A more revealing characterization of this mode is provided in Fig. 3. Here the first canonical mode has been decomposed into a series of orthonormal spatial modes. The eigenvalue (singular values squared) spectrum (Fig. 3, top) shows that only three spatial modes had an eigenvalue greater than unity. The spatial distribution of these modes is seen in the middle and their time-dependent expressions are shown at the bottom. The first mode is most prominent over the central regions and peaks some 400 ms after movement onset (solid line). This may correspond to the late movement-related components. The second (broken line) spatial mode is most expressed in posterior regions and peaks first just after movement onset. This may correspond to the MEF1. The third (dotted line) mode covers more of the frontocentral regions and increases progressively in

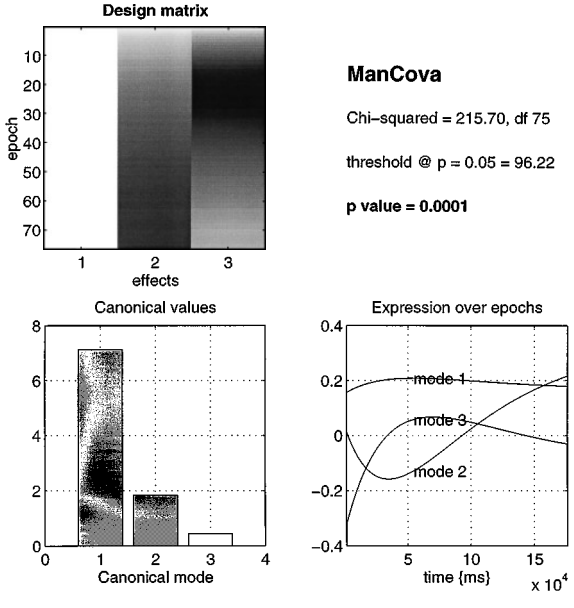

FIG. 1. Multivariate analysis of movement-related magnetic fields. (Top left) The design matrix used in the analysis of evoked responses. The first column is a column of 1's, $\mathbf{G}_1$, and models time-independent response components. The remaining two columns, $\mathbf{G}_{\text{time}}$, represent two temporal basis functions that model time-dependent effects. (Bottom left) Eigenvalues of $\mathbf{T}$ reflecting the relative contributions of the spatiotemporal modes that embody the effects of interest (evoked response). (Bottom right) The expression of the first three spatiotemporal modes over epochs as a function of time (i.e., $\mathbf{z}_1$, $\mathbf{z}_2$, and $\mathbf{z}_3$).

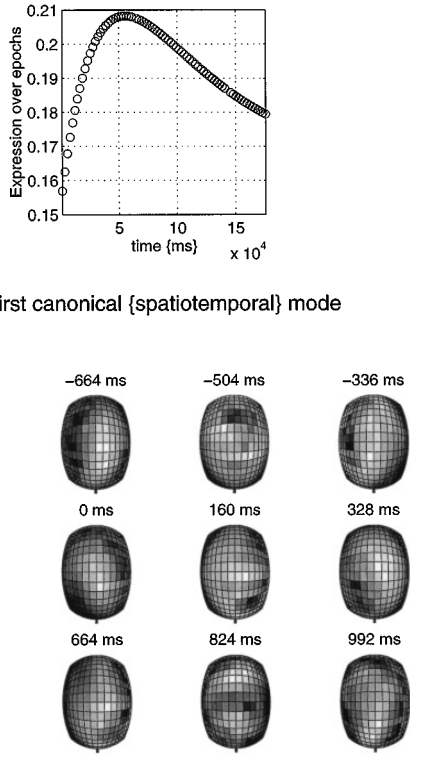

FIG. 2. The first spatiotemporal mode $\mathbf{m}_1$. (Top) Expression of the first spatiotemporal mode over epochs as a function of time (i.e., $\mathbf{z}_1$). Each circle corresponds to an epoch. (Bottom) Time-series representation at 12 equally spaced time bins during the epoch. At each time bin the activity over channels was used to construct this image representation of the spatial components of this spatiotemporal mode using a V_3 transformation. The images are viewed from the top of the head with the nose pointing downward. 0 ms represents movement onset. Each image has been scaled to its maximum.

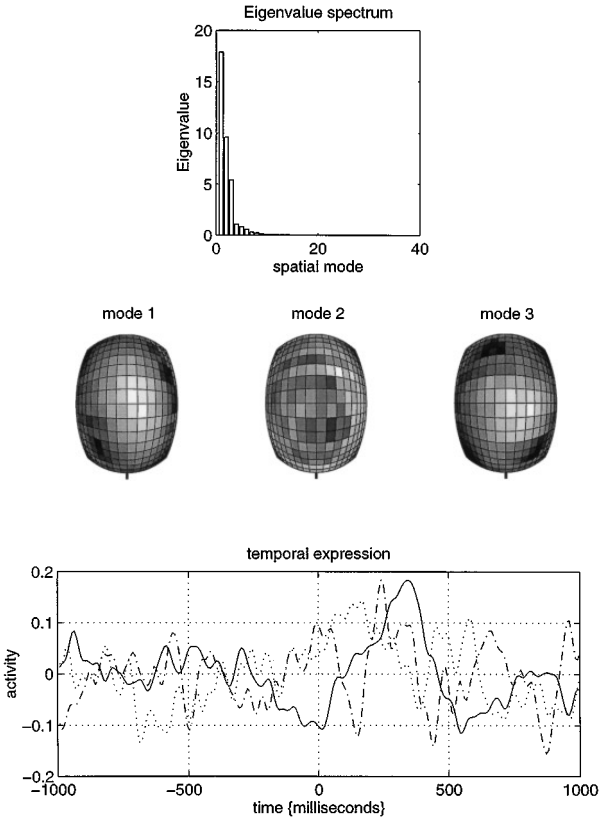

FIG. 3. Spatial modes of the first spatiotemporal mode. (Top) Eigenvalue spectrum showing that the first spatiotemporal mode is largely accounted for by three orthogonal spatial modes. (Middle) The spatial distribution of these spatial modes. The spatial modes were computed in the space of the original 37-channel data and subject to a V_3 transformation. (Bottom) Temporal expression of the first three spatial modes—solid (1st), broken (2nd), and dotted (3rd) lines.

the 500-ms premovement period. Although this mode peaks after movement onset it may include the readiness field and MF. All three modes show a degree of lateralization toward the left, as would be expected with right-hand movements.

Were the Parametric Assumptions Valid?

The general linear model assumes that the error terms are independent and normally distributed. In EEG and MEG data analysis it is sometimes necessary to ensure that this is the case. To demonstrate that the error terms in this analysis do conform to parametric assumptions we have taken the distribution of error terms in the columns of $\boldsymbol{\rho}$ above and scaled then to unit standard deviation. The pooled distribution of these terms is seen in Fig. 4 (solid line) and compares well with the required Gaussian distribution (dotted line).

Adaptation

In the previous section we demonstrated that evoked responses (with or without some time-dependent adaptation) were significant. In this section we ask "was the time-dependent component significant in itself and if so was this adaptation pre- or postmovement (or both)." This allows us to demonstrate how to use ManCova to ask if particular components of a response are significant in isolation. In this case do the effects modeled by $\mathbf{G}_{\text{time}}$ contribute significantly to the overall responses? This is answered by making $\mathbf{G}_h = \mathbf{G}_{\text{time}}$ and relegating the time-invariant component modeled by $\mathbf{G}_1$ to the confound partition of the design matrix, i.e., $\mathbf{G}_c = \mathbf{G}_1$. By repeating the above analysis with this new design matrix $\mathbf{G} = [\mathbf{G}_h\ \mathbf{G}_c]$ we were able to demonstrate that the adaptation was itself significant ($P < 0.0011$, $\chi^2 = 88.95$, $df = 52$).

Differential Adaptation in Pre- and Postmovement Components

In this the final section we address the hypothesis that the intentional and preparatory (i.e., premovement) response components would show more adaptation than executive postmovement components. To answer this question we repeated the analysis of the previous section but using only (i) time bins from the 500 ms preceding, and (ii) the 500 ms following, movement onset. Only the premovement analysis was, as predicted, significant ($P < 0.017$, $\chi^2 = 58.85$, $df = 38$). The corresponding result for the postmovement period was $P = 0.49$ ($\chi^2 = 35.58$, $df = 36$). This suggests that, whatever the functional significance of the time-dependent changes may be, they are differentially expressed in the preparatory and executive components of the MRMFs. The first canonical (spatiotempo-

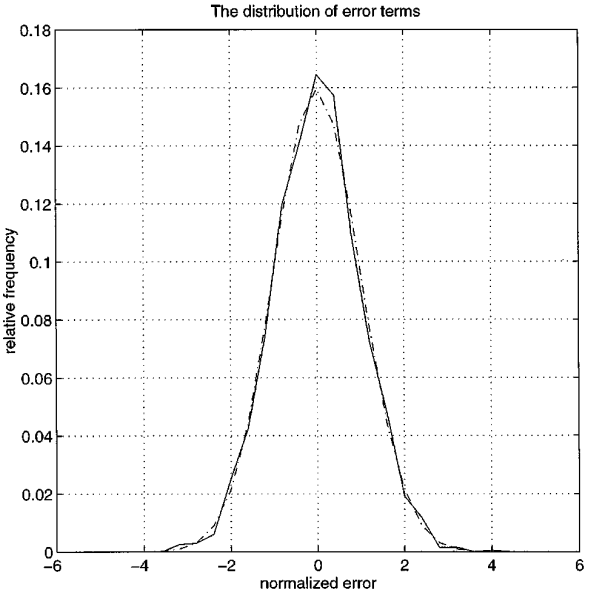

FIG. 4. Distribution of error terms. (Solid line) Empirically determined distribution of the error terms after normalization to unit standard deviation. (Dotted line) Unit Gaussian distribution.

ral) mode for premovement adaptation is shown in Fig. 5. It can be seen that this mode has a complicated time course with pronounced attenuation over the 1st min, followed by a more protracted rebound. The first three spatial modes of this canonical mode show that the adaptation is much more bilateral than the MRMFs in Fig. 3 and involve mostly frontal and central regions (Fig. 6).

These results confirm our initial hypothesis and suggest that (nonmonotonic) adaptation is prevalent in the 1st min or so of motor practice with the possibility that the movement-evoked (executive) responses are relatively exempt from these time-dependent changes relative to preparatory neuromagnetic signals.

CONCLUSION

In this paper we have presented a general multivariate analysis of evoked responses and their spatiotemporal dynamics as measured with MEG. The key contribution of this paper is a method of applying the general linear model, and the statistical inference that ensues, to evoked EEG and MEG data. This application uses ManCova following dimension reduction of the original data. To illustrate the analysis we have looked at time-dependent changes in MRMFs during self-paced movements. Using these techniques we were able to show that, in a single subject, MRMFs show significant adaptation over the 1st min or so of practice. These changes appeared to differentially affect pre- and post-movement responses where adaptation was most marked in preparatory, premovement components.

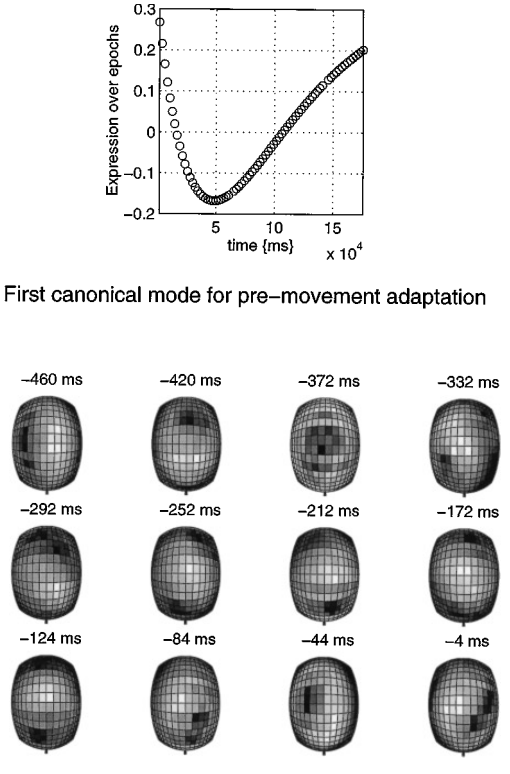

FIG. 5. The first spatiotemporal mode of premovement adaptation. As for Fig. 2, but in this case only data from the 500 ms preceding movement onset have been used and only the time-dependent changes were considered in the canonical variates analysis (the time-invariant component was modeled in the design matrix, but as a confound).

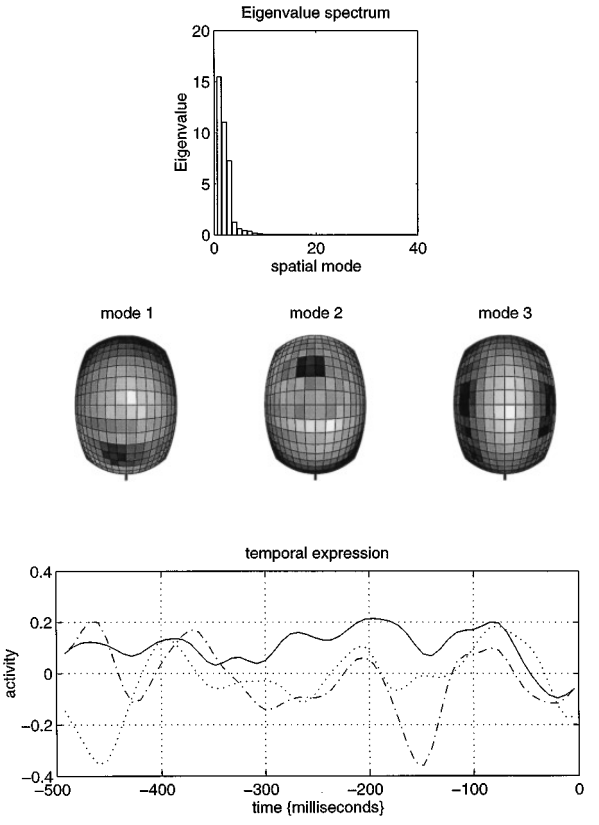

FIG. 6. Spatial modes of first premovement spatiotemporal mode. As for Fig. 3 but using the first premovement spatiotemporal mode shown in Fig. 5.

ManCova and Spatiotemporal Modes

The analysis presented above represents a novel application of standard multivariate statistics. The importance of this application relates to the latitude provided in experimental design and the inferences that are sought. In the present example we were able to ask questions about time-dependent changes in evoked responses and make statistical inferences that would not be possible using conventional evoked response averaging. The stages of the analysis involve (i) dimension reduction using the singular vectors of the original data, where the response variable includes all observations from an epoch over all channels (this can be thought of as a de-noising device, required for the subsequent multivariate analysis); (ii) ManCova, using

the general linear model to make statistical inferences; and (iii) characterization of the interesting effects (i.e., averages, differences, or changes in evoked responses) using spatiotemporal modes and subsequent decomposition into spatial modes.

Implication and Applications

Although the "average" evoked response is a convenient summary of the data, its use can entail a considerable loss of information (Ioannides, 1994). One implication of the results presented here, for conventional averaging approaches to evoked response data, is that the sensitivity to detecting evoked responses could be enhanced if short-term changes in the response were appropriately modeled. There are many parameters other than time that may account for differences among evoked responses, for example, the interstimulus interval, performance parameters, behavioral contingencies, and so on. All these attributes of motor performance can be brought under experimental control (or measured) and, with the techniques described here, used to dissociate evoked response components, in order to make inferences about their functional significance.

Imprecision in defining the onset of an epoch means that ensuing errors in time locking will obscure any high frequency or transient evoked components. For example, to make statistical inferences about event-related desynchronization (Pfurtscheller and Aranibar, 1979) or systematic modulation of spectral density at a particular frequency, one would require both precise time locking (an experimental requirement) and precise phase locking (a physiological requirement) at the frequency in question. One alternative and obvious approach to characterizing event-related desynchronization is to replace the elements of the original MEG data matrix $\mathbf{X}^*$ with equivalent "instantaneous" estimates of power spectral density at the frequency of interest. In this application it is important to ensure that appropriate transformations are applied to the data to render the distribution of error terms Gaussian.

ACKNOWLEDGMENTS

K.J.F., K.M.S., M.D.R., and R.S.J.F. were funded by the Wellcome Trust. We also thank our colleagues, especially Dr. Cathy Price, for help and support in developing these ideas.

REFERENCES

Chatfield, C., and Collins, A. J. 1980. *Introduction to Multivariate Analysis,* pp. 189–210. Chapman & Hall, London.

Friedrich, R., Fuchs, A., and Haken, H. 1991. Modelling of spatio-temporal EEG patterns. In *Mathematical Approaches to Brain Functioning Diagnostics.* (I. Dvorak and A. V. Holden, Eds.), pp. 45–62. Manchester Univ. Press, New York.

Friston, K. J., Frith, C., Passingham, R. E., Liddle, P., and Frackowiak, R. S. J. 1992. Motor practice and neurophysiological adaptation in the cerebellum: A positron tomography study. *Proc. R. Soc. London Ser. B* **248:**223–228.

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

Fuchs, A., Kelso, J. A. S., and Haken, H. 1992. Phase transitions in the human brain: Spatial mode dynamics. *Int. J. Bifurcat. Chaos* **2:**917–939.

Grafton, S., Mazziotta, J., Presty, S., Friston, K. J., Frackowiak, R. S. J., and Phelps, M. 1992. Functional anatomy of human procedural learning determined with regional cerebral blood flow and PET. *J. Neurosci.* **12:**2542–2548.

Hamalainen, M., Hari, R., Ilmoniemi, R. J., Knuutila, J., and Lounasmaa, O. V. 1993. Magnetoencaphalography—theory instrumentation and applications to the noninvasive studies of the working human brain. *Rev. Mod. Phys.* **65:**413–497.

Ioannides, A. A. 1994. Estimates of brain activity using magnetic field tomography and large scale communication within the brain. In *Electrodynamics and Biocommunication* (F. A. Popp and U. Warnke, Eds.), pp. 319–353. World Scientific, Singapore.

Ioannides, A. A. 1995. Estimates of 3D brain activity millisecond by millisecond from biomagnetic signals: Method (MFT), results and their significance. In *Quantitative and Topological EEG and MEG Analysis* (E. Eiselt, U. Zwiener, and H. Witte, Eds.). Universitatsverlag Druchhaus-Maayer GmbH, Jena, Germany.

Ioannides, A. A., Hasson, R., and Miseldine, G. J. 1990. Model-dependent noise elimination and distributed source solutions for the biomagnetic inverse problem. In *Digital Image Synthesis and Inverse Optics,* Proc. SPIE, Vol. 1351, p. 471.

Karni, A., Meyer, G., Jezzard, P., Adams, M., Turner, R., and Ungerleider, L. G. 1995. Functional MRI evidence for adult motor cortex plasticity associated with skill learning. *Nature* **377:**155–158.

Kornhuber, H. H., and Deecke, L. 1965. Himpotentialanberungen bei Willkurbewegungen und passiven Bewegungen des Menschen: Bereitshaftspotential und reafferente Potentiale. *Pflugers Arch.* **286:**1–17.

Kristeva, R., Cheyne, D., and Deecke, L. 1991. Neuromagnetic fields accompanying unilateral and bilateral voluntary movements: Topography and source analysis. *Electroencephalogr. Clin. Neurophysiol.* **81:**284–298.

Libet, B., Wright, E. W., and Gleason, C. A. 1982. Readiness potential preceding unrestricted "spontaneous" vs. planed voluntary acts. *Electroencephalogr. Clin. Neurophysiol.* **52:**322–335.

Liu, L. C., and Ioannides, A. A. 1995. Single epoch analysis of MEG signals. In *BIOMAG'93.* Elsevier, Amsterdam, in press.

Pfurtscheller, G., and Aranibar, A. 1979. Evaluation of event-related desychronisation (ERD) preceding and following voluntary self-paced movement. *Electroencephalogr. Clin. Neurophysiol.* **46:**138–146.

Toro, C., Wang, B., Zeffiro, T., Thatcher, R. W., and Hallet, M. 1994. Movement-related cortical potentials: Source analysis and PET/MRI correlation. In *Functional Neuroimaging: Technical Foundations* (R. W. Thatcher, M. Hallet, T. Zeffiro, E. R. John, and M. Huerta, Eds.), pp. 259–267. Academic Press, San Diego.

Walter, M. G., Cooper, R., Aldridge, V., McCullem, W. C., and Winter, A. L. 1964. Contingent negative variation: An electrical sign of sensory–motor association and expectancy in the human brain. *Nature* **203:**380–384.