NEUROIMAGE 2, 157–165 (1995)

Characterizing Evoked Hemodynamics with fMRI

K. J. FRISTON, C. D. FRITH, R. TURNER, AND R. S. J. FRACKOWIAK

The Wellcome Department of Cognitive Neurology, Queen Square, WC1N 3BG United Kingdom; and the MRC Cyclotron Unit, Hammersmith Hospital W12 OHS United Kingdom

Received April 17, 1995

We describe an implementation of the general linear model that facilitates the characterization of evoked hemodynamic responses to sensorimotor or cognitive processing, when the exact form of these responses is not known. The importance of this approach is that one can test for differential responses among tasks that may elude more conventional analyses. In particular, we suppose that an evoked response has early and late components and that a differential response may involve (i) both components to the same degree, as in a conventional "activation" or (ii) differential expression of the early and late components in two tasks, as might be seen in differential adaptation, or differences associated with the tasks (e.g., requiring and not requiring sustained attention). Using this approach we were able to demonstrate that the anterior cingulate differentiates, in terms of its response, between two motor tasks that did and did not require sustained attention. This differential response was observed even though there was no classical "activation" (i.e., there was no difference in the mean activity associated with the two conditions). It is suggested that these demonstration results point to the possibility of making greater use of the temporal resolution afforded by fast fMRI techniques.

INTRODUCTION

This paper represents the integration of two previous approaches that we have described for the analysis of fast fMRI time-series. In the first paper (Friston *et al.,* 1995a) we introduced a simple extension of the general linear model that allowed for correlations between error terms due to physiological noise or correlations that ensue after temporal smoothing. This extension used the *effective degrees of freedom* associated with the error term (a simple function of the number of scans and the temporal autocorrelation function). The second paper (Friston *et al.,* 1995b, the companion paper to this one) used multivariate statistics (MANCOVA and canonical variates analysis) to characterize the form of hemodynamic transients that are evoked during a cognitive or sensorimotor task. This analysis assumed that physiological responses could show profound task-dependent adaptation and time-dependent changes over the seconds following the onset of a task. By modeling hemodynamic responses using appropriate temporal basis functions, we were able to estimate their form within the general linear model using MANCOVA and CVA. The results revealed some compelling and somewhat unexpected insights about transient but stereotyped responses to changes in cognitive or sensorimotor processing. The most remarkable observation was that these responses can be biphasic and show profound differences in their form depending on the extant task or condition.

The aim of this short paper is to describe a specific variant of the univariate techniques described in the first paper (Friston *et al.,* 1995a) that allows for region-specific and task-dependent differences in the form of an evoked hemodynamic response or evoked transient. This variant of the general linear model uses temporal basis functions as in the second paper (Friston, 1995b) but with the added constraint that the number of these basis functions is kept to a minimum. In the general linear model one tests hypotheses about specific effects using a compound or contrast of the effects one has modeled in the design matrix. The analysis is usually more powerful when the degrees of freedom of the error terms are higher. The degrees of freedom due to error represent the effective number of independent observations minus the number of effects modeled. Clearly if we were to use a large number of basis functions the degrees of freedom would be compromised and the choice of contrast would become very complicated. Therefore we have chosen to use two basis functions that model the family of likely responses as efficiently as possible.

The importance of modeling a whole family of response transients (as opposed to one fixed and simple form) is demonstrated using a single-subject study of attention in motor sequencing. It is shown that the "activation profiles" that ensue from the analysis can

1053-8119/95 $6.00

vary enormously depending on whether one tests for a conventional (sustained) difference in activity or whether one tests for biphasic differences that result from a differential response in terms of early and late components. Biphasic differences can be thought of as an interaction between early and late components of the hemodynamic response to a task and the tasks in question. This form of differential response was suggested by our previous multivariate analysis (Friston *et al.,* 1995b) and is intuitive in the sense that subjects are likely to habituate to some tasks more rapidly than to others (particularly those involving sustained attention or new learning). This differential habituation, "warm up," or indeed "carry over" from one task to the next (see Discussion) may be associated with differential adaptation over several seconds in terms of the measured hemodynamic sequelae. A simple consequence of this is that the one task may evidence a stronger earlier activation but, following adaptation, the second task may supervene at later stages. In other words the differences could be biphasic. In this paper we provide a demonstration study to show that this effect can be seen empirically and suggest that this is a sufficient reason for modeling these effects generally.

The paper is divided into two sections. The first theory section summarizes the use of the general linear model in the construction of statistical parametric maps, with specific emphasis on the temporal basis functions proposed for fast fMRI data analysis. The second section presents an analysis of a motor sequencing task involving sustained attention to visual cues. We show that only when more subtle differential responses among the conditions are examined, do we see systematic, task-specific differences in the anterior cingulate and SMA.

THEORY

In functional imaging the general linear model is used to make statistical inferences by performing univariate tests at each and every voxel. This is known as statistical parametric mapping (Friston *et al.,* 1991). In what follows we present a specific implementation of the general linear model designed to facilitate the comparison of stereotyped hemodynamic responses to cognitive or sensorimotor processing, where the exact form of these responses is unknown. These stereotyped responses will be referred to as [hemodynamic] transients. The general approach has been described in detail elsewhere (Friston *et al.,* 1995a) but will be summarized for completeness in this paper. The implementation considered in this paper centers on the way in which the transients are modeled, using an efficient basis set of two functions that can approximate most forms of response seen in fMRI.

The General Linear Model

The general linear model (Chatfield and Collins, 1980) for a time-series can be written in matrix notation as

$$\mathbf{X} = \mathbf{H}.\eta + \mathbf{D}.\gamma + \mathbf{e}, \tag{1}$$

where $\mathbf{X}$ is a column vector of response variables, in this case values from a single voxel in a fMRI time-series. The columns of $\mathbf{H}$ model the effects of interest, in this case the temporal basis functions modeling the transient response to each occurrence of a particular task or condition. The columns of $\mathbf{D}$ model effects of no interest that are considered confounds, for example, time or the global activity of the scan. $\mathbf{H}$ and $\mathbf{D}$ are partitions of the design matrix $\mathbf{G} = [\mathbf{H}\ \mathbf{D}]$. The design matrix has one row for every scan and one column for every effect (covariate) in the model. $\beta = [\eta^T\ \gamma^T]^T$ is a column vector of parameters for the "effects" modeled by each column of the design matrix. The errors $\mathbf{e}$ are assumed to be independent and identically normally distributed with covariance $\Sigma = \sigma^2.\mathbf{1}$, where $\mathbf{1}$ is the identity matrix. This model can be extended to include temporal smoothing. Let $\mathbf{K}$ be a convolution matrix using a Gaussian kernel with parameter s where, by Eq. (1)

$$\mathbf{K}.\mathbf{X} = \mathbf{G}^* \beta + \mathbf{K}.\mathbf{e} \tag{2}$$

where $\mathbf{K}.\mathbf{X}$ represents the temporally smoothed data. $\mathbf{G}^* = \mathbf{K}.\mathbf{G}$ is a similarly convolved design matrix $\mathbf{G}$ and now the error terms $\mathbf{K}.\mathbf{e}$ are identically distributed with covariance $\Sigma = \sigma^2.\mathbf{K}.\mathbf{K}^T$.

The Design Matrix, the Temporal Basis Functions, and Confounds

We have chosen two basis functions that define a family of curves typical of response profiles seen in fMRI. These functions are exponentially modulated sine functions of the form

$$f(t) = \sin(\pi.t/(n + 1)).\exp(-t/(n.k)), \tag{3}$$

where $k = 4$ (early response) or -1 (late response) and n is the length of each task or condition epoch in scans. The two modulated functions are shown in Fig. 1 (upper panel). The lower panel shows the variety of response forms that linear combinations of these two basis functions can model. These curves can be unimodal, bimodal, biphasic, positive, or negative, always beginning and ending with zero. In the examples shown the length of each task or condition is assumed to be 10 scans. The $\mathbf{H}$ partition of the design matrix contains two columns for each different task or condition. Each

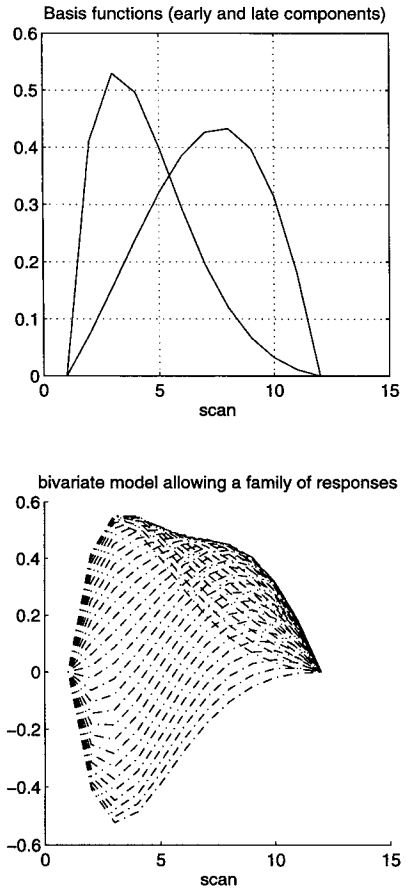


FIG. 1. The temporal basis functions modeling the early and late components of an evoked hemodynamic response. (Top) The two modulated basis functions spanning the duration of one epoch of a single condition or task (assumed here to be 10 scans). (Bottom) Some examples of the curves that obtain by linear combinations of these basis functions.

of these two columns contains repetitions of one of the two basis functions at the times during which the condition in question is present. The amount to which each of the two basis functions contributes to a particular task is given by the estimate of the respective coefficients in the vector of parameter estimates β. Least squares estimates of β, say **b,** are given by

$$\mathbf{b} = (\mathbf{G}^{*\mathrm{T}}\mathbf{G}^{*})^{-1}\mathbf{G}^{*\mathrm{T}}.\mathbf{K}.\mathbf{X}. \tag{4}$$

The null hypothesis that the effects embodied in **H** are not significant can be tested with the t statistic using linear compounds or contrasts of the parameter estimates **b.** A contrast (row vector) **c** is simply a set of weights that sum to zero. The principal advantage of using two basis functions is that these contrasts can be specified in a meaningful and intuitive fashion. One useful way of thinking about the variety of contrasts that could be specified is to consider the differences between any two conditions and the two basis functions as representing a factorial experimental design. In this framework we can ask three classes of questions. First, we can test for the main effect of the basis functions. This is equivalent to testing the hypothesis that there is a significant difference between the magnitude of the early component of the response (first basis function) and the late component (second basis function) that is prevalent in both conditions. For example, if pairs of columns in **G** modeled early and late components in conditions A, B, and C:

$$\mathbf{c} = [0\ 0\ 1\ -1\ 1\ -1\ .\ .\ .\ .\ .\ .]$$

would test for a main effect of the early component (relative to the late component) common to conditions B and C. The dots represent zeros for the confounds in the **D** partition of the design matrix. Clearly this contrast does not tell us anything about the differences between the responses to conditions B and C and will not be considered further here. The second sort of contrast tests for a main effect of condition, i.e., similar increases (or decreases) in both early and late components between two conditions. For example,

$$\mathbf{c} = [0\ 0\ -1\ -1\ 1\ 1\ .\ .\ .\ .\ .\ .]$$

tests for an activation in C, relative to B that is expressed in both early and late components. This contrast would correspond most closely to the tests used in conventional fMRI time-series analysis (e.g., correlations with a box-car or sine waveform, or performing unpaired t tests on the data from conditions B and C). The third and final sort of contrast is effectively an interaction or difference in the relative contributions of the early and late components depending on condition. For example,

$$\mathbf{c} = [0\ 0\ 1\ -1\ -1\ 1\ .\ .\ .\ .\ .\ .]$$

would test the hypothesis that condition B evokes a phasic early component that adapts quickly, while condition C is associated with a slower and more sustained response with a more pronounced late component (relative to B). In other words a difference in the shape of the hemodynamic transients. It is this difference that the current framework is designed to accommodate, and we suggest that these differences can be as important as the simple differences tested for with conventional contrasts. The significance of a particular linear compound of effects is tested with

$$T = \mathbf{c}.\mathbf{b}/(\mathbf{c}.\epsilon^2.(\mathbf{G}^{*\mathrm{T}}\mathbf{G}^{*})^{-1}.\mathbf{c}^{\mathrm{T}})^{1/2}, \tag{5}$$

where

$$\epsilon^2 = \mathbf{r}^{\mathrm{T}}.\mathbf{r}/\nu \tag{6}$$

ν is the degrees of freedom associated with the residuals $\mathbf{r} = \mathbf{X} - \mathbf{G.b}$, and T has the Students t distribution with degrees of freedom ν. If the error terms (elements of $\mathbf{r}$) were independently distributed ν would simply be the number of scans minus the number of effects estimated, i.e., $N - \text{rank}(\mathbf{G}^*)$. However, the error terms are known to be correlated because of the temporal smoothing ($\Sigma = \sigma^2.\mathbf{K}.\mathbf{K}^T$) and $\nu = [N - \text{rank}(\mathbf{G}^*)]/\sqrt{(2\pi s^2)}$. See Friston *et al.* (1995a) for details. By using the above expressions, for every contrast or compound $\mathbf{c}$, we obtain a value for T at each and every voxel. These constitute a statistical parametric map or SMP{t}.

Statistical Inference in SPMs

To simplify subsequent analysis, the SPM{t} is transformed to a SPM{Z} using a probability integral transform or other standard device. The problem now is that an extremely large number of nonindependent univariate comparisons have been performed and the probability that any region of the SPM will exceed an uncorrected threshold by chance is high. Standard procedures have been developed in statistical parametric mapping that correct for the multiplicity of voxels and the spatial correlations among them. These corrections are based on either the height or the spatial extent of a local excursion of the SPM (i.e., cluster of voxels above a threshold) to give a corrected P value based on height $P(Z_{\max} > u)$ or spatial extent $P(n_{\max} \geq k)$. The distributional approximations required for these corrections derive from the theory of Gaussian random fields (see Friston *et al.,* 1991; Worsley *et al.,* 1992; and Friston *et al.,* 1994a, for the development of this theory in functional imaging). In this paper we present P values that are based on both the spatial extent and the peak height of a region.

A DEMONSTRATION STUDY

In these sections we describe the experimental design and data used to illustrate the above theory. The example chosen is typical of fMRI activation studies that use more than two conditions. There were three conditions: a rest condition, a motor sequencing condition where the subject moved his or her right and left hand in a fixed alternating order (in response to a visual cue), and a motor sequencing condition in which the subject moved either the right or the left hand as instructed visually in a random sequence. The only difference between the "fixed" and "random" conditions was that the subject had to attend to the instruction that specified the movement. Clearly in the random condition the subject could not anticipate or prepare the exact movement before seeing the cue. The movements involved raising the forefinger and the cues were presented at pseudorandom intervals of 2, 3, or 4 s. The details of this paradigm and a full discussion of the results will be presented elsewhere.

Data Acquisition

One hundred and twenty T_2* weighted volume images (128 × 64 × 10 voxels) were obtained from a single male subject using a GE/ANMR 1.5T system with EPI capabilities. The volumes consisted of 10 sequential transverse sections and were acquired every 3 s. Voxel size was 3 × 3 × 7-mm voxels with 0.5-mm slice separation. The three conditions were presented in blocks of 10 scans, in pseudorandom order. Each condition was therefore repeated four times, each time constituting a 30-s epoch for that condition.

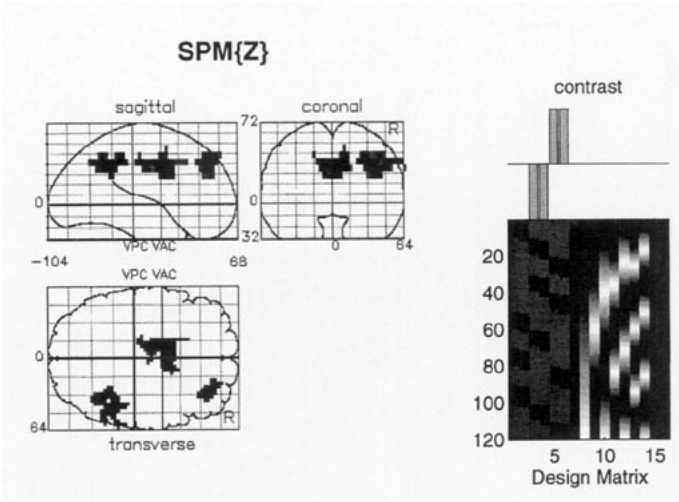
Data Prepossessing

The 120 volume images were realigned to the first as described elsewhere (Friston *et al.,* 1995c) and resampled to 3 × 3 × 6-mm voxels. The data were then smoothed with an isotropic Gaussian kernel with FWHM of 8 mm. Voxels that had values greater than 0.8 of the volume mean, in all the images, were selected to restrict the analysis to intracranial regions. The data were mean corrected to form the matrix $\mathbf{X}$ above with 14476 columns (one for each voxel) and 120 columns (one for each scan).

Data Analysis

The design matrix $\mathbf{G}$ is seen in the upper right panel of Fig. 2 and models early and late response components (basis functions) for each of the three conditions. These effects are seen in the first six columns of the design matrix and correspond to the $\mathbf{H}$ partition. The remaining columns were confounds (i.e., $\mathbf{D}$ in the theory section). These confounds comprised low frequency sine and cosine functions of the scan number (up to a maximum of 3.5 cycles per 120 scans). These confounds effect a high-pass filter, removing low frequency artifacts due to aliased cardiorespiratory and other cyclical components. The remaining two columns were a column of ones and global or whole brain activity. The convolution matrix $\mathbf{K}$ used for temporal smoothing was a Toeplitz matrix with a Gaussian kernel with standard deviation $s = (\sqrt{8})/3$ scans or $\sqrt{8}$ s. This corresponds to a dispersion of about 8 s and rendered the effective degrees of freedom due to error 44.

SPM{Z} were constructed using Eq. (5) and two contrasts. The first contrast corresponds to a conventional analysis and tested for a main effect of the random condition relative to the fixed condition. This contrast is seen in graphical form displayed above the design matrix in Fig. 2 ($\mathbf{c}$ = [0 0 −1 −1 1 1 . . . . . .]). The elements of the contrast are shown above the corresponding columns or effects. The resulting SPM{Z} was thresholded at 0.01 (uncorrected) to show regions that



region	size {k}	P(n max > k)	Z	P(Z max > u)	(Uncorrected)	{x,y,z mm}
1	**92**	**0.014**	**5.29**	**0.002**	(0.000)	**36 −57 36**
			4.10	0.246	(0.000)	30 −51 30
			4.01	0.341	(0.000)	48 −45 36
2	**58**	**0.096**	**4.47**	**0.060**	(0.000)	**36 39 30**
			4.16	0.199	(0.000)	27 45 36
3	**136**	**0.002**	**4.18**	**0.189**	(0.000)	**9 3 30**
			4.07	0.280	(0.000)	9 0 42
			4.04	0.306	(0.000)	−3 3 24
4	**7**	**0.998**	**3.63**	**1.206**	(0.000)	**60 −39 36**

Threshold = 2.33; Volume [S] = 14476 voxels; df = 44

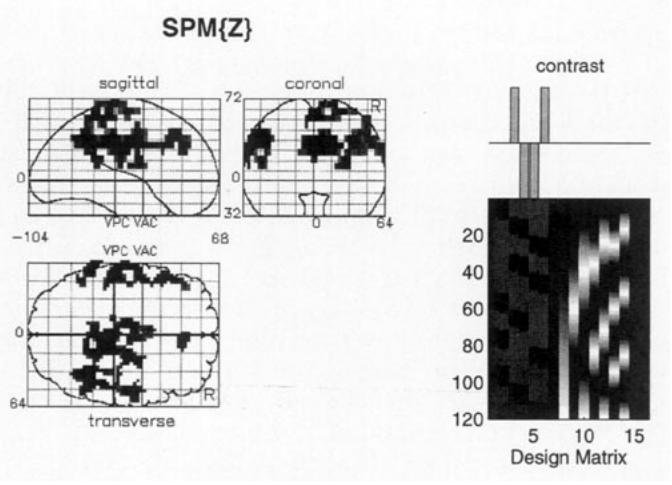
FWHM = [10.8 10.9 11.7] mm (i.e. 569 RESELS)

FIG. 2. Results of the test for activations in the random condition relative to the fixed condition. (Top right) Above: A contrast **c** = [0 0 −1 −1 1 1 . . . .] testing for the significance of the activation. Below: The design matrix **G** with 6 covariates on the left (early and late components for each of the three conditions) and 10 confounds on the right. These confounding covariates correspond to cyclical time effects, a constant and to global or whole volume activity. Because elements of this matrix can take negative values the gray scale is arbitrary and has been scaled to the minimum and maximum. The form of the design matrix is the same as in the text—**K.**[**H D**]. Note that the length of the contrast is the same as the number of columns, or effects, in the design matrix, which is the same as the number of parameters one is explicitly estimating. (Top left) SPM{Z}: This is a maximum intensity projection of the SPM{*t*} following transformation to the *Z* score. The display format is standard and provides three views of the brain from the front, below, and the left-hand side. Data are presented only for regions with $P < 0.1$ corrected. The gray scale is arbitrary and the space conforms to that described in the atlas of Talairach and Tournoux (1988). (Bottom panel) Tabular data are presented for "significant" regions ($P < 0.1$ corrected). The location of the maximal voxel in each region is given (positive *x* is left) with the size of the region (k) and up to three *Z* maxima. For each maxima the significance is assessed in terms of $P(Z_{max} > u)$ a corrected *P* value based on peak height and $P(n_{max} > k)$ a corrected *P* value based on the spatial extent of the region.

activate, in terms of the overall response (early and late components) when the motor movement is specified by the cue. Based on our understanding of the functional anatomy of attention and response selection we anticipated that the anterior cingulate and SMA would be involved in these differences. We were disappointed. The SPM{*Z*} is shown in the upper right panel of Fig. 2. Only regions reaching $P < 0.1$ (corrected, i.e., based on peak height or spatial extent) are shown. Tabular data on these regions is provided in the lower half of the figure. The three regions implicated included the left dorsolateral prefrontal cortex, the middle regions of cingulate cortex, and the left angular gyrus (in the parietal region). Although these areas are all very plausible the context of motor sequencing and attention they did not include the areas that we had hypothesized. However, the responses in these areas were significantly different, as we now demonstrate.

The second contrast tested for a biphasic or interaction between early and late components and the fixed and random conditions i.e., $c = [0\ 0\ 1\ -1\ -1\ 1\ \ldots.]$. This contrast and the resulting SPM{*Z*} are seen in Fig. 3. It is immediately obvious that the profile of differential responses is more extensive than that obtained with the previous contrast. Furthermore this profile now includes the anterior cingulate and a small portion of ventral SMA. In addition right superior frontal gyrus (premotor), right angular gyrus, and extensive somatosensory regions are implicated. It is pleasing to note the anterior cingulate is the second most significant area with a *Z* value of 4.96. The responses at this location are shown in Fig. 4. The top panel plots the observed activity against the linear compound of basis functions tested for by the contrast (i.e., $\mathbf{G.c}^{T}$). A strong positive correlation is observed. The lower panel plots the observed [adjusted] activity over time (dotted line) and that estimated in terms of the basis functions. Although the fit is not perfect there is a very reasonable agreement, with the sustained and late components of the random condition being clearly evident. The observed responses are shown in a different format in Fig. 5, by plotting each condition epoch separately for the same anterior cingulate voxel. This figure shows more clearly the difference in the early and late components of the response among the three conditions. The rest condition evoked a response with little systematic structure. The fixed conditions start off high and adapt within 5 to 10 s (the initial rise has been obscured by the temporal smoothing). In contrast the random condition is associated with a protracted and sustained increase in signal until about 20 to 25 s and thereafter it declines. It is interesting to note that the average signal in the fixed and random conditions is about the same. This is why this region was not identified using a more conventional analysis.

The relationship of these regions to the underlying anatomy is depicted in Fig. 6 which renders the acti-



region	size {k}	P(n max > k)	Z	P(Z max > u) (Uncorrected)	{x,y,z mm}
1	**97**	**0.011**	**5.56**	**0.000** (0.000)	**−3 −45 60**
			4.74	0.019 (0.000)	15 −30 66
			4.38	0.086 (0.000)	24 −15 60
2	**30**	**0.506**	**4.96**	**0.007** (0.000)	**6 33 24**
3	**302**	**0.000**	**4.83**	**0.014** (0.000)	**3 −9 30**
			4.76	0.018 (0.000)	−9 −15 30
			4.56	0.043 (0.000)	36 −48 30
4	**49**	**0.165**	**4.70**	**0.024** (0.000)	**54 −12 30**
			3.84	0.620 (0.000)	45 −9 42
			3.00	6.680 (0.001)	39 −24 30
5	**160**	**0.001**	**4.65**	**0.029** (0.000)	**−54 −27 36**
			4.26	0.139 (0.000)	−63 −18 24
			4.13	0.221 (0.000)	−51 24 36
6	**26**	**0.622**	**4.65**	**0.030** (0.000)	**27 −6 18**
			3.33	2.862 (0.000)	24 −15 12
7	**37**	**0.340**	**4.35**	**0.100** (0.000)	**−57 −48 36**
			4.28	0.129 (0.000)	−63 −42 18
			3.30	3.106 (0.000)	−63 −45 30

Threshold = 2.33; Volume [S] = 14476 voxels; df = 44
FWHM = [10.8 10.9 11.7] mm (i.e. 569 RESELS)

FIG. 3. The same as Fig. 2 but for a contrast that tests for a differential response in terms of early and late components.

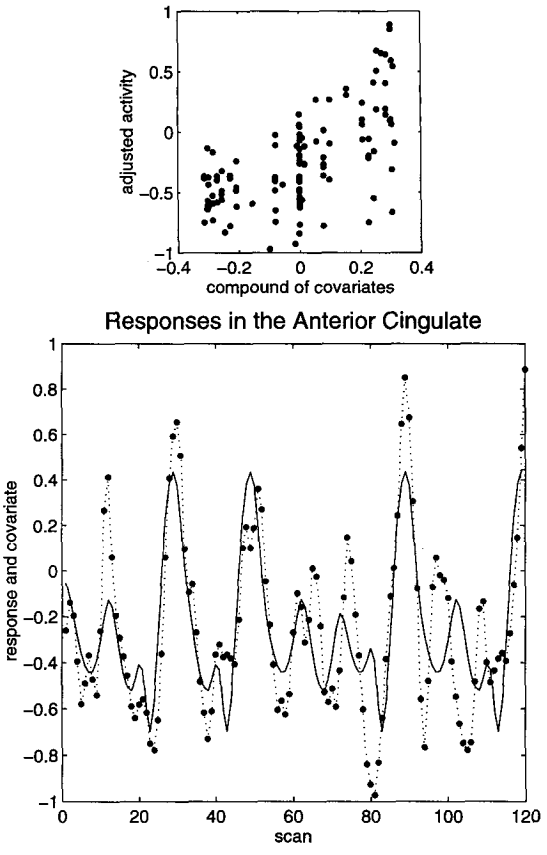


FIG. 4. Plots of the fMRI signal from a voxel in the cingulate gyrus. (Top) The observed [adjusted] activity as a function of the covariates (basis functions) in the design matrix weighted by the contrast in Fig. 3. (Bottom) The same data but plotted as functions of time. The solid line is the estimated activity according to the model used (i.e., in terms of the basis functions) and the broken line and dots represent the empirical adjusted data **X**.

vations ($P < 0.01$ uncorrected and $P < 0.1$ corrected) onto the original data. The lower panels show the anterior cingulate region, the left angular gyrus, and right superior frontal areas. In addition, we see differential responses in the middle and posterior cingulate gyrus and bilaterally around the central sulcus.

To demonstrate the consistency of this differential response pattern we repeated an identical analysis using an independent data set obtained from a female subject doing exactly the same paradigm. As before, a conventional analysis failed to show the anterior cingulate but this region was strongly implicated when we tested explicitly for a differential response. In this instance the biphasic differences in the evoked transients are even more pronounced. Figure 7 shows the observed activity in the anterior cingulate of the second subject. Here the phasic response and rapid adaptation during the fixed condition are very evident. The response to the random condition is more sustained but peaks earlier than in the first subject. The anatomic topography of the SPM{Z} is less extensive than the first subject's (see Fig. 8) but is composed of areas previously identified, namely, the anterior cingulate, the right superior prefrontal gyrus, and left angular gyrus.

DISCUSSION

We have described here a simple implementation of the general linear model that facilitates the characterization of evoked hemodynamic responses to sensorimotor or cognitive processing, when the exact form of these responses is not known. The importance of this approach is that one can test for differential responses among tasks that may elude more conventional analyses. In particular we suppose that an evoked response has early and late components and that a differential response may involve (i) both components to the same degree, as in a conventional "activation" or (ii) differential expression of the early and late components in two tasks, as might be seen in differential adaptation, or differences associated with the tasks (e.g., requiring

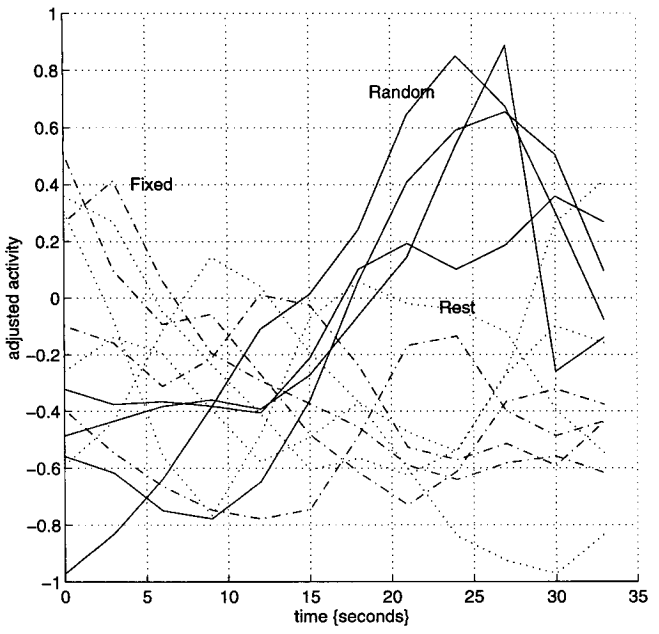


FIG. 5. The same data as in Fig. 4 but now plotted on an epoch by epoch basis. The dotted lines correspond to the rest epoches, the broken lines to the fixed epoches, and the solid lines to the random epoches. Each epoch lasts for 10 scans or 30 s.

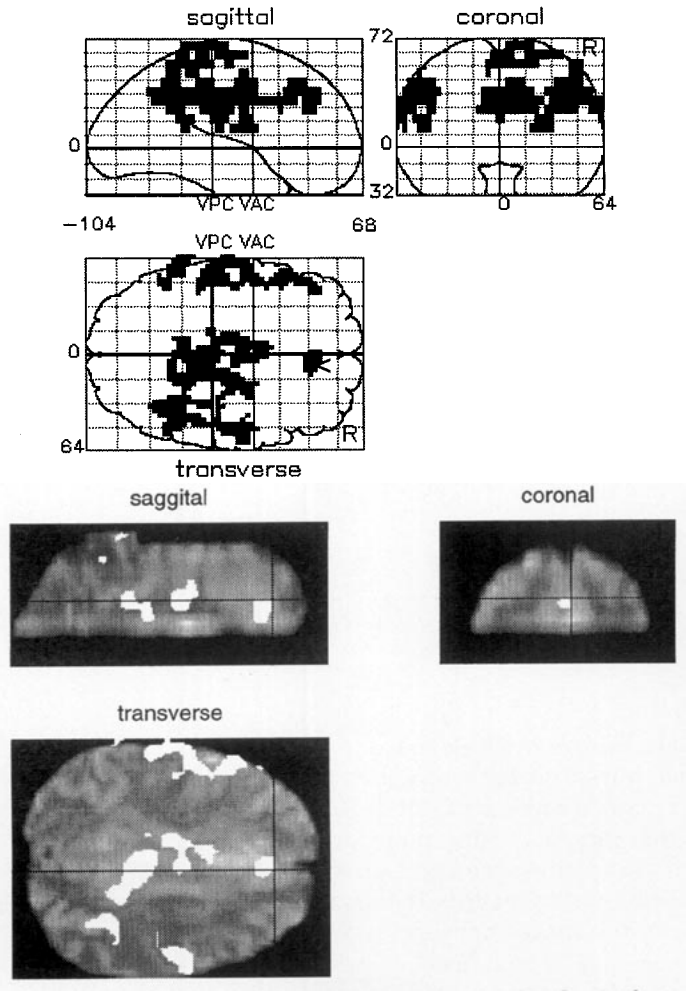


FIG. 6. The same SPM{Z} as in Fig. 3 (top) rendered onto three orthogonal sections of the original MRI data, through a point in the anterior cingulate (lower). The white regions correspond to areas of significant differential response and are described in the main text.

and not requiring sustained attention). Using this approach we were able to demonstrate that the anterior cingulate differentiates, in terms of its response, between two motor tasks that were and were not specified by a visual cue. This differential response was observed even though there was no classical "activation" (i.e., there was no difference in the mean activity associated with the two conditions).

The exact model proposed is a simple variant of the general linear model that uses two temporal basis functions to model the response of interest. These basis functions are exponentially modulated sine waves and model the early and late components of a response respectively. Framing this analysis in terms of the general linear model allows for the production of statistical parametric maps of the t statistic and statistical inferences about them using established techniques. The nature of the differences in response between any two conditions is specified by a linear compound or contrast of the coefficients of the early and late components in each condition. A variety of differential responses can be examined; we have focussed on conventional differences due to an activation (involving early and late components to a similar degree) and more complicated interactions between the two conditions and early and late responses (involving a significant difference in the early and late components). Clearly there are other forms of difference one could test for (e.g., the differences between early components, while ignoring the late components) depending on the nature of the hypothesized effect.

The Nature of Differential Responses

A major difference between PET and fMRI is that with fMRI observations of brain physiology can be repeated immediately and typical experimental designs observe the subject switching from one condition to another. With sensory stimuli the consequences of this switching may be small; however, this may not be the case for more complex tasks that involve some "top down" (volitional) processing. At least two, closely related, effects are likely to occur which we can call "warm up" and "carry over." Both of these phenomena have been extensively examined in behavioral studies.

Warm up is a well established phenomenon in motor skill performance. Peak performance cannot be achieved immediately from rest, but only after a period of activity. The term comes from the not necessarily erroneous belief that the muscles function better if they are warm. Studies of warm up are reviewed in the

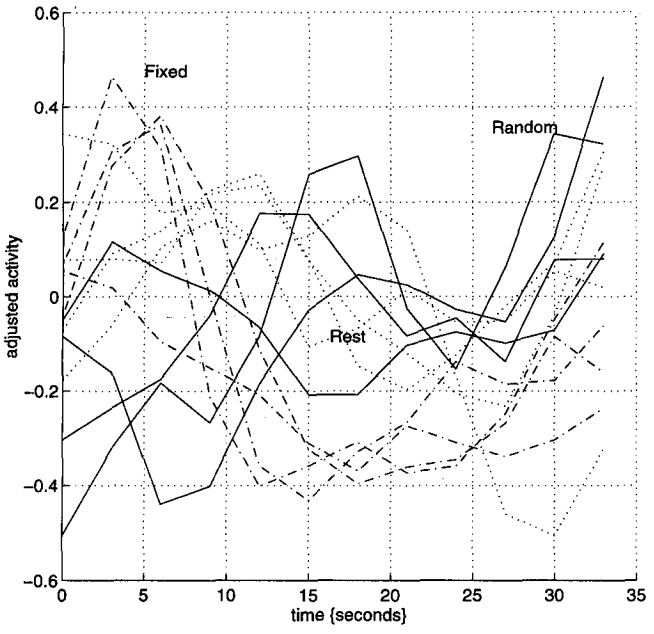


FIG. 7. The same format as Fig. 5 but showing the results of an independent study on a second subject, using the same analysis.

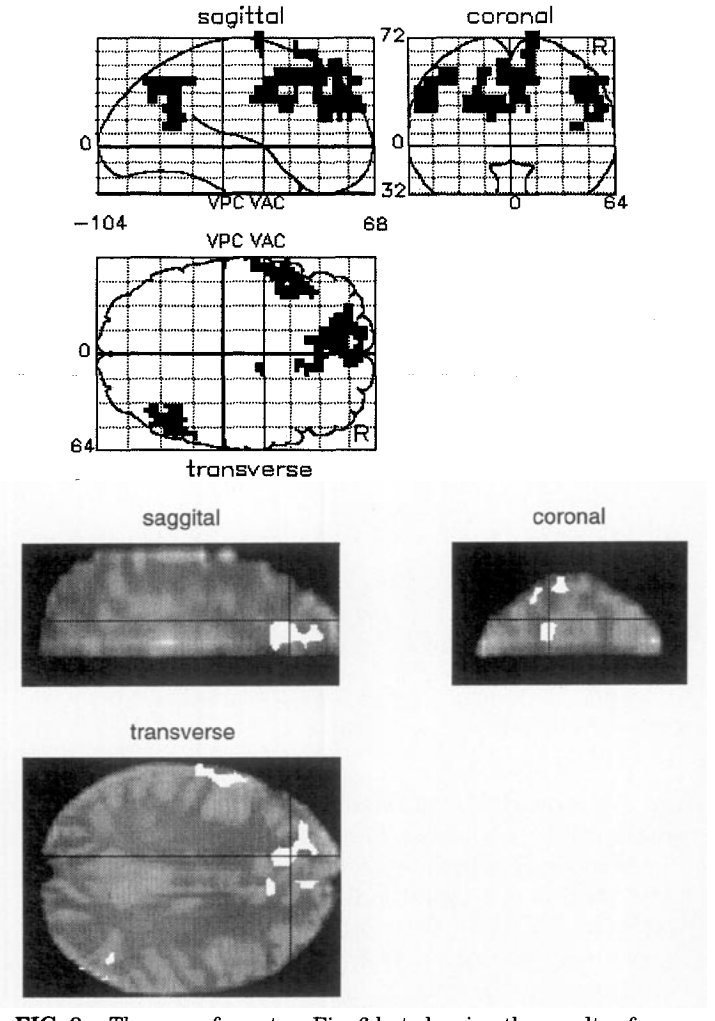


FIG. 8. The same format as Fig. 6 but showing the results of an independent study on a second subject, using the same analysis.

classic work of Eysenck and Frith (1976). Performance on pursuit rotor (a simple tracking task) shows a sharp increase during the first 10 to 20 s after a rest, with this increase being more marked the longer the time elapsed since the last period of practice. The simplest explanation for this is that time is needed to instantiate the most appropriate "set" for the performance of the task. Thus, for the first 10 to 20 s after a rest, performance may well engage difference or extra systems.

Switching from one task to another is likely to require an even greater change of "set" than switch from rest to a task. Allport (among others) has studied this switching (Allport *et al.*, 1994) and has measured what he terms "switch costs." For example, he has studied switching in the Stroop test (color words written in incongruous ink, e.g., GREEN written in red). In condition 1 the subject names the ink color, while in condition 2 he reads the word. Immediately after the switch from one condition to the other, and for a few trials thereafter, the response time is greatly increased. Furthermore this carry over effect can be markedly asymmetric. Of particular interest is the observation that getting up to speed again after a switch depends not on the time since the switch, but on the number of responses. These carry over effects are potentially important for fMRI because (i) they demonstrate that even at a behavioral level there are profound changes in response during the seconds following a change in condition and (ii) they suggest that the effects of condition A preceded by condition B may not be the same as the effects of condition A preceded by C.

Alternative Approaches

In this section we consider some alternative ways of addressing the issue about differential responses from a statistical perspective. One simple and intuitive approach would be to use an extended set of basis functions (e.g., a Fourier set) and use the F ratio as the statistic to test for a systematic temporal response pattern. The disadvantage of this approach is that one cannot make statistical inferences about specific differences between tasks as would normally be effected with contrasts and a SPM{t}. An alternative would be to use a more elaborate set of basis functions and contrasts; however, the diversity of contrasts that one would have to explore may be prohibitively large and many contrasts would have no direct intuitive interpretation. We have previously suggested using a multivariate analysis in conjunction with canonical variates analysis to overcome this problem and let the data decide what the best "contrasts" are. However, this ap-

proach denies the possibility of testing specific hypotheses directly, and in our experience the anatomic or spatial information in the canonical images is difficult to interpret. In conclusion we consider the use of univariate tests to produce SPMs using just two basis function to be the most effective way of making statistical inferences about differential hemodynamic responses in a regionally specific fashion.

Conclusion

The analysis in this paper is not meant to devalue conventional approaches to fMRI data. The detection of "classical" activations (as opposed to significant differential responses that are not easily assigned to the class of activations) is probably the most important example of a differential response. In the sensory cortices it may be, biologically, the most prevalent sort of response difference. Indeed our previous work on the nature of the hemodynamic response function to underlying neural activity (Friston *et al.,* 1994b) was predicated on a box-car reference waveform and, implicitly, a sustained and fixed activation in [extra]striate cortex due to photic stimulation. What we are suggesting is that for some paradigms that involve learning or cognitive components that may show some short term (seconds) adaptation a more careful characterization of the hemodynamic responses may be appropriate. Although this may complicate the interpretation of the results, it does provide for a richer characterization and for the possibility of activation studies that explicitly use habituation and adaptation to probe functional organization in the brain. We have not presented a comprehensive analysis of these responses but have simply shown that they can exist. We consider this the first step in maximizing the potential of fast fMRI techniques.

ACKNOWLEDGMENTS

The authors are supported by the Wellcome Trust. We thank our colleagues for help and support during the development of these ideas. We are very grateful to colleagues at the Institute of Psychiatry, Denmark Hill, UK, for essential help in doing these experiments.

Note added in proof. It has pointed out recently that the estimators for the variance of the parameter estimates [Eq. (5) and Eq. (6)] are incorrect and biased (Worsley-personal communication). The correct expressions will be found in Worsley and Friston (Worsley, K. J. and Friston, K. J. Analysis of fMRI time-series revisited-Again. submitted for publication) and generally give a more conservative test. The latter communication also contains a more rigorous treatment of the effective degrees of freedom.

REFERENCES

Allport, D. A., Styles, E. A., and Hseih, S. 1994. *Shifting Intentional Set: Exploring the Dynamic Control of Tasks.* To appear in C Umilta and M Moscovitch (Eds.). *Attention and Performance XV: Conscious and Nonconscious Information Processing.* MIT Press.

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

Eysenck, H. J., and Frith, C. D. 1977. *Reminiscence, Motivation and Personality,* pp. 165–199. Plenum, New York.

Friston, K. J., Frith, C. D., Liddle, P. F., and Frackowiak, R. S. J. 1991. Comparing functional (PET) images: The assessment of significant change. *J. Cereb. Blood Flow Metab.* **11:** 690–699.

Friston, K. J., Worsley, K. J., Frackowiak, R. S. J., Mazziotta, J. C., and Evans, A. C. 1994a. Assessing the significance of focal activations using their spatial extent. *Human Brain Mapping* **1:** 214–220.

Friston, K. J., Jezzard, P., and Turner, R. 1994b. Analysis of functional MRI time-series. *Human Brain Mapping* **2:** 69–78.

Friston, K. J., Holmes, A. P., Poline, J-B., Grasby, P. J., Williams, S. C. R., Frackowiak, R. S. J., and Turner, R. 1995a. Analysis of fMRI time-series revisited. *NeuroImage* **2:** 45–53.

Friston, K. J., Frith, C. D., Frackowiak, R. S. J., and Turner, R. 1995b. Characterizing dynamic brain responses with fMRI: A multivariate approach. *NeuroImage* **2:** 166–172.

Friston, K. J., Williams, S., Howard, R., Frackowiak, R. S. J., and Turner, R. 1995c. *Movement-Related Effects in fMRI Time-Series.* Submitted for publication.

Talairach, J., and Tournoux, P. 1988. *A Co-planar Stereotaxic Atlas of a Human Brain.* Thieme, Stuttgart.

Worsley, K. J., Evans, A. C., Marrett, S., and Neelin, P. 1992. A three-dimensional statistical analysis for rCBF activation studies in human brain. *J. Cereb. Blood Flow Metab.* **12:** 900–918.