Journal of Cerebral Blood Flow and Metabolism
13:5–14 © 1993 The International Society of Cerebral Blood Flow and Metabolism
Published by Raven Press, Ltd., New York

Functional Connectivity: The Principal-Component Analysis of Large (PET) Data Sets

K. J. Friston, C. D. Frith, P. F. Liddle, and R. S. J. Frackowiak

MRC Cyclotron Unit, Hammersmith Hospital, London, U.K.

Summary: The distributed brain systems associated with performance of a verbal fluency task were identified in a nondirected correlational analysis of neurophysiological data obtained with positron tomography. This analysis used a recursive principal-component analysis developed specifically for large data sets. This analysis is interpreted in terms of functional connectivity, defined as the temporal correlation of a neurophysiological index measured in different brain areas. The results suggest that the variance in neurophysiological measurements, introduced experimentally, was accounted for by two independent principal components. The first, and considerably larger, highlighted an *intentional* brain system seen in previous studies of verbal fluency. The second identified a distributed brain system including the anterior cingulate and Wernicke's area that reflected monotonic time effects. We propose that this system has an *attentional* bias. Key Words: PET—Principal-component analysis—Functional connectivity—Effective connectivity—Verbal fluency—Neural networks.

Cooperative and connectionist approaches to understanding the integration of brain function are well established (Sherrington, 1941; Hebb, 1949; Edelman, 1978; McClelland, 1988). The nature and organizational principles of extrinsic cortical connections, particularly the long corticocortical afferents (e.g., Goldman Rakic, 1988) has provided a basis for mechanistic descriptions of brain function (e.g., Mesulam, 1990). These descriptions refer to parallel, massively distributed, and interconnected (sub)cortical areas. Anatomical connectivity is a necessary underpinning for these models and has been used to infer functional connectivity (e.g., Zeki, 1990). This article describes one way of measuring functional connectivity using positron emission tomographic (PET) measurements of neural activity.

PET is in a unique position to acquire data for this sort of analysis because it samples the entire brain state in a uniform fashion. This allows all possible functional connections to be assessed using serial measurements of the same subject in different brain states. The shortcomings of PET include its relatively poor spatiotemporal resolution and the exact nature of the dependency of measured regional CBF (rCBF) on neural discharge rates. However, PET can be used to address large scale *functional connectivity* as an important supplement to observations on the gross aspects of extrinsic *anatomical connectivity* and fine time scale *effective connectivity* defined by electrophysiology.

FUNCTIONAL AND EFFECTIVE CONNECTIVITY

In the past two decades, the concept of functional or effective connectivity has been most thoroughly elaborated in the analysis of multiunit recordings of separable neuronal spike trains, recorded simultaneously from different brain areas (Gerstein and Perkel, 1969; Gerstein et al., 1989). Temporal coherence among the activity of different neurones is commonly measured by cross-correlating their spike trains. The resulting correlograms are then interpreted as the signature of functional connectivity. In current approaches, effective connectivity is

Received October 28, 1991; final revision received May 8, 1992; accepted June 18, 1992.

Address correspondence and reprint requests to Dr. K. J. Friston at The Neurosciences Institute, 1230 York Ave., New York, NY 10021, U.S.A.

Abbreviations used: ANCOVA, analysis of covariance; BA, Broadmann's area; DLPFC, dorsolateral prefrontal cortex; J-PSTH, Joint Peri Stimulus Time Histogram; PCA, principal component analysis; PET, positron emission tomography; rCBF, regional CBF.

assessed using normalized Joint Peri Stimulus Time Histograms (J-PSTHs). In particular, the PST coincidence histogram reflects effective connectivity as the joint probability of two neurones firing together as a function of time in the interstimulus interval (Aersten and Preissl, 1991). Effective connectivity can also be measured in terms of *efficacy* and *contribution*. These terms are best understood at a synaptic level, where in the linear equality

$$x_j = \Sigma W_{ij} \cdot x_i \qquad (1)$$

x_j is the postsynaptic response to many presynaptic inputs (x_i). Here efficacy of the connection between k and j can be thought of as the synaptic efficacy W_{kj}, whereas the contribution reflects the effect of k on j relative to all presynaptic inputs, i.e., $W_{ij}/\Sigma W_{ij}$. These two aspects of effective connectivity can be estimated empirically, given certain assumptions (e.g., Gochin et al., 1991). There is a close relationship between effective connectivity and efficacy: "It is useful to describe the effective connectivity with a connectivity matrix of effective synaptic weights. Matrix elements would represent the effective influence by neurone i on neuron j" (Gerstein et al., 1989). It has also been proposed that "the notion of effective connectivity should be understood as the experiment and time-dependent, simplest possible circuit diagram that would replicate the observed timing relationships between the recorded neurons" (Aersten and Preissl, 1991). These definitions are essential and useful abstractions but lack operational significance. In this article, we reserve the term functional connectivity to mean the observed temporal correlation between two electro/neurophysiological measurements from different parts of the brain. Effective connectivity will refer to the underlying efficacy (W_{ij}), which may or may not be measurable.

With respect to PET neuroimaging, the measurement of functional connectivity therefore requires the repeated assessment of neurophysiology over time in the same subject(s). This sort of data is obtained from activation or longitudinal studies. We defer a discussion of the relationship between within-subject and between-subject correlations until the discussion.

An exposition of functional connectivity based on PET neurophysiological data reduces to an examination of its correlation structure. Correlation structure refers to the correlations observed over time in the same subject(s) (e.g., Friston et al. 1991*a*; Lagreze et al., 1991). Principal component analysis (PCA), as a first step, is most suited to this examination. PCA extracts the important features of the correlation matrix in terms of principal components or eigenvectors. These vectors are the linear combinations that account for independent or orthogonal amounts of variance in the observed data. Only a few principal components are usually required to explain the majority of observed variance. In terms of functional connectivity, a principal component represents a truly distributed brain system within which there are high intercorrelations. Furthermore, because any one component is orthogonal to the remaining, these systems are functionally unconnected from each other. However, any single area may be implicated in more than one system.

The PCA of PET data is not straightforward. PET data sets are usually large. Large here has a special meaning, namely a very high dimensionality but a low sample size, where sample size is the number of repeated observation in the same brain(s) at different times and dimensionality is the number of observations (voxels). Usually, a low sample size:dimensionality ratio is considered undesirable; however, in this special case of PCA, it can be used to advantage. Typically, the number of voxels can exceed 10^4. This requires the PCA of a correlation matrix with 10^8 elements. This is beyond the capacity of most workstations available. However, we have developed a recursive PCA technique that can handle these large data sets with a vast reduction in computational overhead.

This recursive PCA technique is described and applied to data obtained from a study of six subjects each scanned 12 times during two word generation tasks. The functional connectivity revealed and the orthogonal brain systems identified are presented as an example of this approach.

METHODS

Recursive PCA Analysis

The technique is modeled on "L" systems or string rewriting systems used in the construction of fractal and self similar patterns. L systems were introduced by Linenmayer in 1968 to model the growth of living organisms. In these systems, a pattern (axiom) is defined that is composed of line segments. According to (production) rules, each segment is replaced by the pattern primitive. This primitive is itself constructed from line segments that are recursively replaced with smaller scaled primitives. No "drawing" actually occurs until the scale reaches a specified lower limit [see Voss (1988) for a full discussion]. The charm of these systems is that the algorithm that replaces each line segment of the primitive with smaller versions calls itself recursively but only implements pattern drawing at the smallest scale. In a similar way, the recursive PCA used here recursively calls itself until the size of subpartitions of the original data matrix reach a lower limit. Let $\theta(M)$ denote the operation of the

PCA operator $\theta\{\cdot\}$ on a data matrix M where M can be bisected ($M = [M_1 M_2]$). The algorithm is defined by the following equivalence (where $\cdot$ denotes matrix multiplication):

$$\theta\{M\} = \begin{pmatrix} \theta\{M_1\} & 0 \\ 0 & \theta\{M_2\} \end{pmatrix} \cdot \theta\{(M_1 \cdot \theta\{M_1\} \quad M_2 \cdot \theta\{M_2\})\} \tag{2}$$

until the size of M reaches a lower limit (S); then,

$$\theta\{M\} = \epsilon\{C\{M\}\} = Q_k \tag{3}$$

where Q_k are the largest $S/2$ eigenvectors of the covariance matrix of M ($= C\{M\}$). The operator $\theta\{\cdot\}$ recursively calls itself until the multiply bisected subpartitions reach a stopping criterion in terms of size (S). The recursion relationship (2) essentially implies the splitting of a data matrix, the rotation of the observed scores of each half into principal component (PC) scores and the elimination of the half of these (redundant) scores before computing the eigenvectors of $C\{M\}$. The eigenvector solution is obtained by postmultiplying the original transformation matrix with the eigenvectors of the rotated and reduced data matrix. This elimination or reduction means that the largest matrix actually operated on by $\theta\{\cdot\}$ never exceeds size S. This holds for any size of M. The justification for eliminating half of the PC scores [implicit in Eq. (3)] relates to the sample size to dimensionality ratio. The subspace spanned by the data can only be an $(n - 1)$-dimensional subspace of the S-dimensional space defined by the PCs. In other words, the PC scores, although S in number, only describe variance in the $n - 1$ ($< S/2$) eigenvectors with nonzero eigenvalues. The rest, being zero, can be eliminated with no loss of information. The algorithm used is provided in PRO-MATLAB in Table 1 and a worked example is given in the Appendix. The saving in terms of computational overhead is illustrated in Fig. 1. As the size of M increases, the number of computational steps in a normal PCA increases geometrically. For the recursive approach, these increments are more arithmetical. This is seen in Fig. 2, which is a regression (in log space) of the number of floating point operations required on the size of the data matrix. In practice, the PCA takes about 14 s for a {6,2048} data matrix on a contemporary (SPARC) workstation.

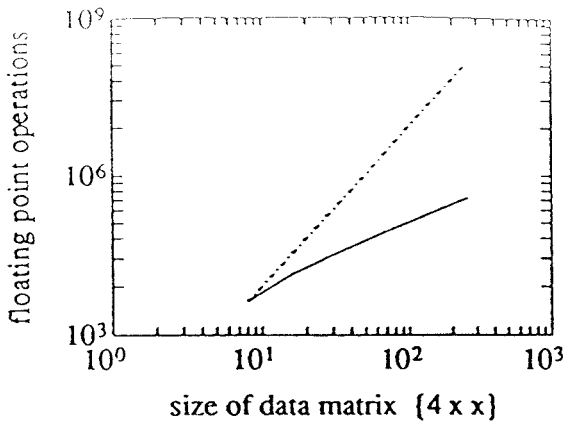

FIG. 1. Regression of number of floating point operations on size x of data matrix {4, x} for normal (nonrecursive) PCA (broken line) and recursive PCA (solid line). This regression is plotted in log space.

TABLE 1. *An example of the algorithm $\theta(\cdot)$ for a data matrix M of $O \times 2^n$ elements where $O < 16$*

```
function [e,v] = theta(M)
% recursive PCA analysis
% format [e,v] = theta(M)
% M = data matrix with 2 n columns and O rows where O < k = 16
% e—eigenvectors
% v—eigenvalues
[x,y] = size(M)
k = 16
% s—stopping criteria
s = 2*k
if y = = s
  [e,v] = eig( cov(M) ); [v i] = sort( diag(v) ); e(:,i) = e;
  e = e(:,[(1 + k):s]); v = v([(1 + k):s]);
  return
end
y = y/2;
x = [1:y]; y = [1:y] + y;
e1 = theta( M(:,x) ); e2 = theta( M(:,y) );
[e,v] = theta( [M(:,x)*e1 M(:,y)*e2 ] );
e = [ e1 zeros(e1) ; e2 zeros(e2) ]*e;
```

%, comment; *, matrix multiplication; eig, returns the matrix of eigenvectors and eigenvalues; cov, return the covariance matrix; sort, returns values in ascending order and the indices used; diag, returns a vector corresponding to the leading diagonal of a matrix; zeros, returns a matrix of a given size whose elements are zero.

We have presented an algorithm that uses bisection of M. This requires the number of columns to be a multiple of 2. More elaborate schemes are possible using n-way splits.

A recursive PCA analysis was applied to the mean [analysis of covariance (ANCOVA) adjusted] rCBF of six subjects time locked to the same stimuli or tasks. It is perfectly natural to apply the technique to individuals but we presented an analysis of mean data for two reasons: (i) the derived functional connectivity patterns have greater generic validity, being common to all subjects, and (ii) the elements of M are the mean of six independent observations and, by the central limit theorem, more normally distributed over the 12 observations.

Data acquisition

Six normal male volunteers with no neurological or psychiatric history were scanned 12 times in the same session whilst performing one of two verbal tasks. Permission to perform these studies was obtained from the local ethical committee and Advisory Committee for the Administration of Radioactive Substances (U.K.).

Scans were obtained with a CTI (model 953B CTI Knoxville, TN, U.S.A.) PET camera (as a fully three-dimensional acquisition). Reconstructed (Townsend et al., 1991) images had a resolution of 5.2 mm (Spinks et al., 1992). The volume images contained 128 × 128 × 31 voxels corresponding to 2 × 2 × 3.1 mm. ^{15}O was administered intravenously as radiolabeled water infused over 2 min. The total counts per pixel during the buildup phase of radioactivity served as an estimate of rCBF (Fox and Mintun, 1989).

Each scan lasted 2 min followed by an 8 min interscan interval. The tasks began 20 s prior to delivery of radiolabeled water. Subjects performed two tasks alternatively (order balanced across subjects). The first task served as

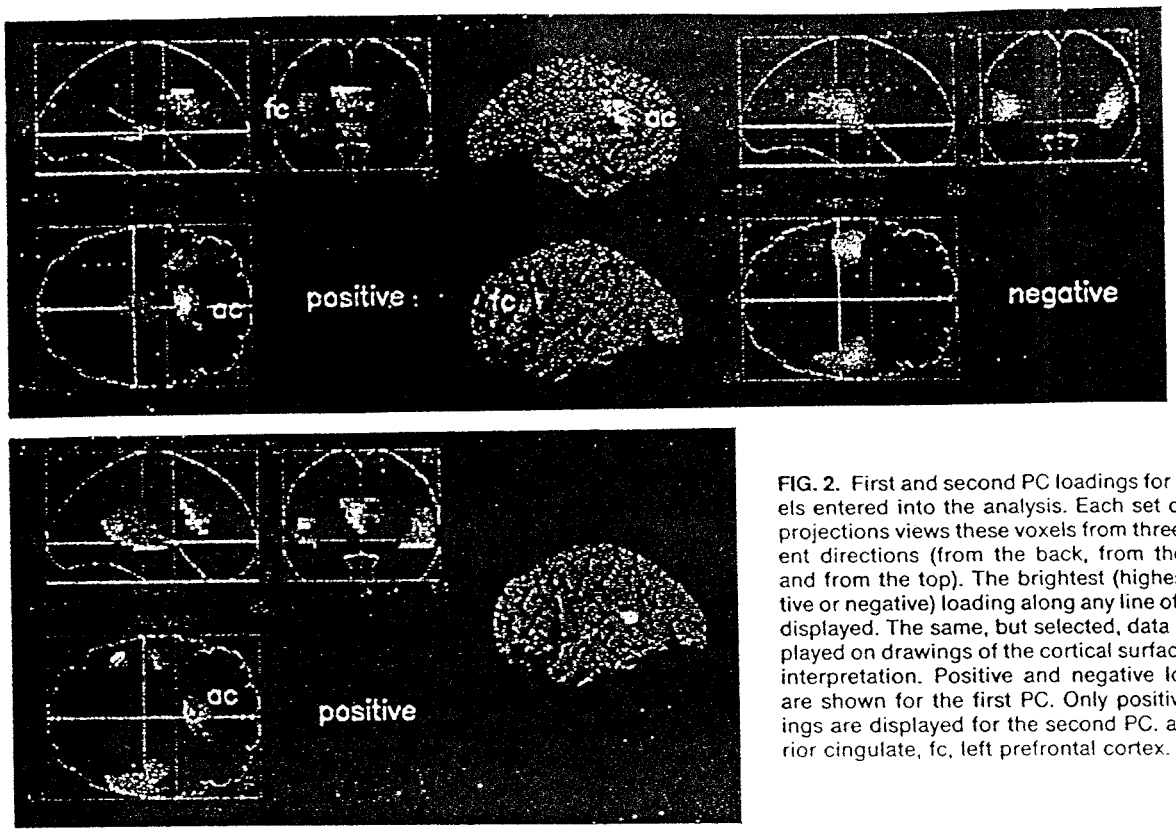

FIG. 2. First and second PC loadings for all voxels entered into the analysis. Each set of three projections views these voxels from three different directions (from the back, from the right, and from the top). The brightest (highest positive or negative) loading along any line of view is displayed. The same, but selected, data are displayed on drawings of the cortical surface to aid interpretation. Positive and negative loadings are shown for the first PC. Only positive loadings are displayed for the second PC. ac, anterior cingulate, fc, left prefrontal cortex.

baseline and involved repeating a heard letter. Letters were presented aurally at one per 2 s. The letter was changed four times during the task. The second task corresponded to a paced verbal fluency with the same stimuli (but different letters) that the subjects responded to by generating a word that began with that letter. Subjects had their eyes closed in all conditions.

Data analysis

Data were stereotactically normalized (Friston et al., 1989, 1991*b*) and mean rCBF equivalents derived for each condition in parallel for every voxel. These mean rCBFs were adjusted for the confounding effect of whole brain differences using ANCOVA (Friston et al., 1990). Only voxels in which significant differences between the 12 scans were detected (ANCOVA $F > 1.971$, df 11, 54, $p < 0.05$) were subjected to further analysis. The reason for using this subset was primarily that of computational expediency; however, it is easily justified by noting that voxels that do not contribute significantly to measured variance are unlikely to contribute to measured covariance. Stability of the PCA results was assessed by repeating the analysis using voxels with the highest F values with three different lower limits. Recursive PCA of these ANCOVA-adjusted mean rCBF data generated a series of PCs, the corresponding eigenvalues, and the PC scores for each condition. A PC score reflects how much a PC contributed to any given condition (Hope, 1968). The loadings on PCs for each voxel were displayed as a statistical parametric map as a volume image showing the brightest voxel along the line of view.

RESULTS

The number of voxels for which $F > 1.971$ (following an ANCOVA) was 8,277. The total number of voxels analyzed was 65,186. We would have expected 3,259 by chance. We observed over 2.5 times the expected number of voxels reaching criteria ($p < 0.001$). The 2^{13}, 2^{12}, and 2^{11} voxels with the highest F values were subject to recursive PCA (i.e., the largest numbers that were less than 8,277 that could be divided by 2). The results of these three analyses were stable and very similar (relative contributions of the PCs varying by only a few percent). The results of the middle (2^{12}) solution are presented.

The first two PCs accounted for almost all of the variance (86%). The first PC accounted for 71% and the second 15% of variance. The third accounted for only 4%. The corresponding profiles (loadings) are seen in Fig. 2. The PC scores are shown in Fig. 3 for each of the 12 alternating conditions (baseline–fluency–baseline–fluency . . .). The first PC had positive loadings in the anterior cingulate [Brodmann's area (BA) 24, 32], the left dorsolateral prefrontal cortex (DLPFC BA 46), Broca's area (BA 44), the thalamic nuclei, and the cerebellum. Negative loadings were seen bitemporally and in the posterior cingulate. This profile is a verbal fluency profile we have observed in two previous independent studies (Friston et al., 1991*a*; Frith et al., 1991). We have not observed subcortical activation to be so reliable in previous data. The PC scores (Fig. 3) testify to this interpretation with universally

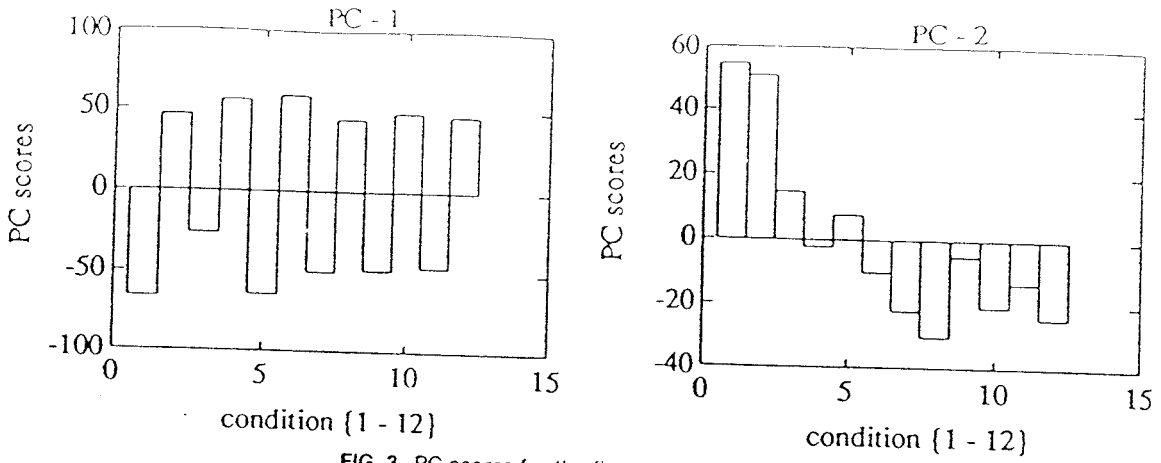

FIG. 3. PC scores for the first two components.

high loading on the verbal fluency tasks and low scores on the baseline. Furthermore, these scores are largely invariant over time. The second PC had its highest positive loading in the anterior cingulate and appeared to correspond to a monotonic time effect with greatest prominence in the first three conditions (Fig. 4). It is interesting that the second PC included bitemporal regions, but on the left there was a selective involvement of Wernicke's area in the posterior superior temporal region (BA 22).

DISCUSSION

Functional connectivity has been defined as the temporal correlation between neurophysiological (functional) measurements made in different brain areas. We have used a recursive PCA of such data obtained longitudinally from the same subjects with PET to demonstrate orthogonal (independent) functionally connected brain systems. The two systems evident in our data may represent an intentional system critical for the intrinsic generation of words and a second system whose physiology changes monotonically with time irrespective of the tasks the subject was engaged in. It is possible that this represents a more attentionally orientated system, which reflects the declining need for acquisition of perceptual set as the tasks become more familiar.

The system corresponding to the first PC accounted for 71% of the observable differences in adjusted mean rCBF from the 12 scans. This is a remarkable observation in that 71% of the variance in brain physiology was introduced by experimental design. This is a clear vindication of the PET technique in the investigation of functional anatomy and connectivity. Furthermore, the distributed system highlighted is in exact accord with that which has been predicted from anatomical connectivity. All of the components of this system (anterior cingulate, DLPFC, posterior cingulate, and superior temporal region) have dense and reciprocal connections (Goldman-Rakic, 1986, 1988).

The second system centered on the anterior cingulate seems to be involved in time-dependent changes; probably of attentional or perceptual set [see Wise (1989) for a brief discussion of set]. We infer this from previous PET studies on attention (Petersen et al., 1989; Pardo et al., 1990; Corbetta et al., 1991) and other ideas relating to the distribution of attentional systems (Posner et al., 1990). The anterior cingulate is an interesting example of an area that belongs, coincidentally, to two functionally unconnected systems (in this experiment). In other words, the activity of this area increases during the verbal fluency task and declines with time (possibly with acquisition of perceptual set). Yet these two effects are totally independent (see Fig. 4 for the effects on rCBF in the anterior cingulate and compare the time effects with rCBF in the DLPFC, where the absolute levels are relatively stable over trials).

These results highlight the simple point that the measured functional connectivity is implicitly dependent on the functional states of the brain at the time of measurement. The neurophysiological variance–covariance introduced experimentally forms the basis of functional connectivity and this depends on the tasks chosen.

Functional connectivity: electrophysiological and PET neurophysiological approaches

Separable spike trains can be interpreted in terms of single unit activity. PET rCBF reflects the activity of large neuronal populations. This distinction is not fundamental given the fact that multielectrode studies are striving to delineate population or "as-

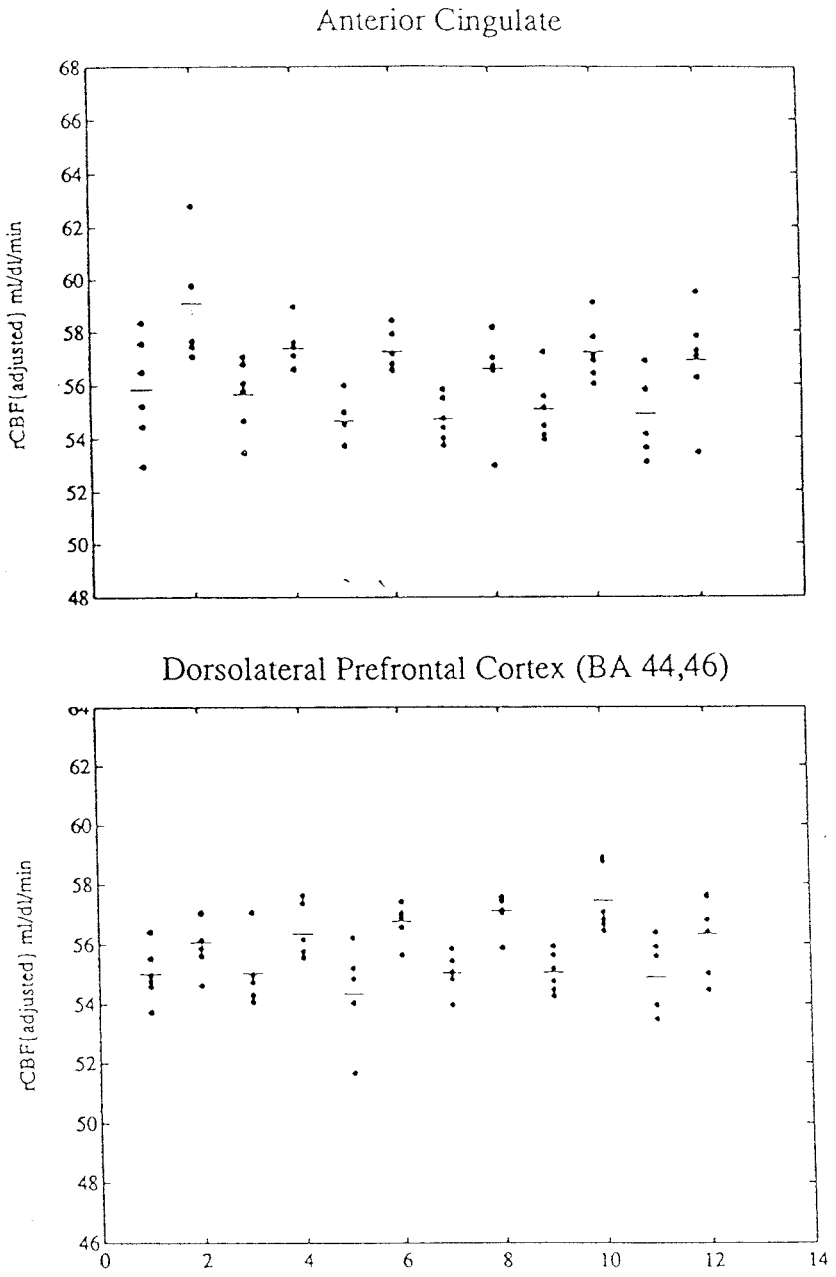

FIG. 4. Adjusted rCBF equivalents for each subject in each condition {baseline–fluency–baseline–fluency . . .} for two brain locations that both activate with verbal fluency but only one (anterior cingulate) evidences a time effect over trials. **Above:** The anterior cingulate {BA 24; −2 18 24 mm *x*, *y*, *z*} (Talairach and Tournoux, 1988) demonstrates a fall in rCBF levels over the first four trials irrespective of their nature. **Below:** Equivalent data for the DLPFC {BA 44/46; −36 12, 24 mm} are shown.

sembly" (Gerstein et al., 1989) dynamics. Indeed, population thinking is central to some interpretations of neuronal correlations (e.g., Sporns et al., 1989).

A key difference between electrophysiological and PET indices of physiology is time scaling. This is important given the possible functional significance of high-frequency stimulus-specific oscillatory events in extended regions of visual cortex (Eckhorn et al., 1988; Gray and Singer, 1989). Functional connectivity using the two techniques, however, can be linked at two levels: (i) A unifying concept is provided by coherence [$\sigma(w)$]. Coherence is a measure of the correlation at a particular frequency (w) (Cox and Miller, 1980). Consequently, coherence and functional connectivity at a frequency w [$fc(w)$] are directly related:

$$fc_{ij}(w) = \sigma_{ij}(w) = |g_{ij}(w)|^2 / g_{ii}(w) \cdot g_{jj}(w) \quad (4)$$

where $g_{ij}(w)$ is the cross-spectral density and $g_{ii}(w)$ and $g_{jj}(w)$ are the autospectral densities of the neurophysiological processes in question. Equation (4) explicitly relates functional connectivity to EEG coherence. Multielectrode recording and EEG measures deal with coherence at frequencies with periods of milliseconds whereas PET covers the low-frequency component of the coherence profile on a

time scale of minutes. There is no requirement or expectation that coherence is invariant over frequencies. Indeed, even on very small time scales, the typically nonuniform characteristics of the PST coincidence histogram suggest that "near-coincidence firing is strongly modulated as time proceeds through the stimulus cycle with . . . switching, in a stimulus-locked fashion from a condition of incoherent firing to coherent firing" (Aersten and Preissl, 1991).

If functional connectivity is distributed over temporal frequencies, then electrophysiological and PET neurophysiological assessments can be complementary, each sensitized to different frequency domains of coherence.

We know of no empirical evidence to suggest that short-term (millisecond) coherence and long-term (minutes) coherence are dependent; however, simulations have shed some light on this relationship: to explore the notion that modulation of short-term functional connectivity might be an "emergent property of dynamic cooperativity," Aersten and Preissl (1991) investigated the behavior of artificial networks, analytically and through simulation. They concluded that the short-term effective connectivity varies strongly with, or is modulated by, pool activity. Pool activity is the product of the number of neurones in the pool and the mean firing rate. The mechanism is simple; the efficacy of subthreshold excitatory postsynaptic potentials in establishing dynamic interactions is a function of postsynaptic depolarization, which in turn depends on the tonic background of pool activity.

This sort of analysis suggests the second and more fundamental link between short- and long-term coherence, namely (ii) that slow (co)variation in mean activity may be a necessary condition for the emergence of fast dynamic correlations.

Anatomical, effective, and functional connectivity

These three connectivities are at the same time interdependent and yet different concepts. Neuronal interactions can receive contributions from direct connections (possibly mono- or polysynaptic) or contributions from shared input (Gerstein et al., 1989) originating in a stimulus-related or modulating source. Although direct or indirect anatomical connectivity is necessary for effective connectivity, there is no simple one-to-one mapping. A similar necessary, but not sufficient, relationship exists between effective connectivity and functional connectivity, in that context-dependent and dynamic modulation of functional connectivity has been demonstrated (by simulation) in the context of constant efficacy W_{ij} (Aersten and Preissl, 1991).

If functional connectivity was modulated by diffuse and divergent anatomical connectivity, e.g., the thalamocortical system in "the generation of internal functional modes" (Llinas and Pare, 1991), then the relationship between functional and anatomical connectivity would be tenuous by virtue of its nonspecificity. However, empirical evidence from studies of EEG coherence is consistent with the notion that measured coherence is mediated by long corticocortical association fibers (Thatcher et al., 1986). Whether this is the case for PET functional connectivity remains to be seen; however, observations in our laboratory suggest a correspondence between (PET) functional connectivity and major anatomical pathways delineated in nonhuman primates.

Hitherto, we have dealt with the correlational analysis of longitudinal data collected over time from a single subject (or the mean from several subjects time locked to the same stimuli). This is in contradistinction to correlations between observations on different subjects in the same state. Correlational analyses of these cross-sectional data have been an important theme in PET data analysis for many years [e.g., Horwitz et al., 1984, interregional correlations; Metter et al., 1984; Moeller et al., (1987), Scaled Subprofile Model; Friston et al., (1992*a*), canonical correlational analysis]. Usually, the objective of these analyses is to uncover the characteristic profile or pattern of cerebral physiology that underlies a neuropathophysiology or particular brain state. These distributed profiles certainly depend on the concept of functional connectivity for interpretation but are not direct measures of *functional connectivity*. We mean this in the sense that temporal correlations cannot be measured without a time series of observations. There have been promising attempts to abstract a measure of *effective connectivity* from cross-sectional data (e.g., Horwitz et al., 1991). The interpretation of these correlations relies on a number of assumptions and is best understood with reference to specific models (e.g., Horwitz, 1990).

It is hoped that the method presented in this article will allow, if not a paradigm shift, a balance between categorical "cognitive subtraction" techniques and correlational approaches to cognitive dimensions using graded tasks. See Friston and Frackowiak (1991) for a discussion of the implications for activation study design. The data used to illustrate the recursive PCA algorithm represent "state-of-the-art" in that they were three-dimensional reconstructions and acquired with a system sensitive enough to permit 12 repeated measurements. We are now routinely applying the al-

gorithm to more conventional data (with six observations per subject) and getting valid results comparable to those presented here.

We end on a different note by suggesting that the same mechanics of a self-similar embedding of a local PCA transformation may provide an interesting metaphor for neural connections themselves. In some respects, the problem faced by the analyzer of PET data resembles those faced by the brain itself: the abstraction of PCs and reduction to a small number of perceptual dimensions of sensory input sequences distributed over a vast number of afferents. Connectionist models of elemental PCA-like transformations have already been described (e.g., Oja, 1982, 1989; Linsker, 1988; Foldiak, 1989). In the case of brain-like function, redundant dimensions can result not from the limited sample size, but from significant intercorrelations in the input sequence that underlie functional segregation and specialization. The constraint on the number of dimensions (columns of M) that can be operated on is imposed by the limited spatial extent of intrinsic connectivity and finally the elimination of redundant dimensions may be reflected in the inequality between the number of inputs (afferent extrinsic axons) to a small cortical region and the number of outputs (initial segments of projection neurons). See Friston et al. (1992*b*) for further discussion. In this context, the title of this article takes on a completely different, but not inappropriate, meaning.

Acknowledgment: K.J.F. was funded by the Wellcome Trust. We thank our colleagues at the MRC Cyclotron Unit for making this study possible.

APPENDIX

Worked example

To illustrate how the algorithm $\theta\{\cdot\}$ works, the calculations implicit in the recursive scheme are presented for a simple data matrix M. Let us suppose that we are required to find the eigenvector solution (PCs) of the covariance matrix of a data matrix M with size $\{x,y\}$, where $x \ll y$. We have at our disposal two operators $\epsilon\{\cdot\}$ and $C\{\cdot\}$ that return the eigenvectors (and eigenvalues) and covariance matrix of their respective operands. In the normal course of things, our solution is simply $Q = \epsilon\{C\{M\}\}$. However, the solution must now be obtained subject to the constraint that the number of columns (or rows) of any operand of $\epsilon\{\cdot\}$ or $C\{\cdot\}$ must not exceed S, where $S \ll y$. This is effectively the memory constraint for large y. The example below starts uses a $\{2,16\}$ data matrix M representing 16 measurements made twice (e.g., 16 voxels from two consecutive scans). S must be greater than or equal to $2 \cdot x$. Let S be 4. Because $x = 2$, there will be only one eigenvector [describing the line passing through the x $(=2)$ points in a 16-dimensional space]. This is the simplest example illustrating the recursive nature of the algorithm.

To help the reader follow the arithmetic, variable names are subscripted according to the appropriate level.^{T} denotes transposition

$$M = \begin{matrix} 0.29 & -0.10 & -0.29 & -1.39 & 0.15 & -1.65 & 0.21 & -0.65 \\ -0.24 & -1.16 & -0.50 & -0.02 & -0.01 & 2.27 & -1.05 & 1.02 \\ -1.83 & 1.90 & -1.78 & 0.24 & -0.13 & 1.68 & -1.30 & 0.30 \\ -1.79 & 0.63 & -1.66 & 0.82 & -0.98 & -0.03 & 1.28 & -0.15 \end{matrix}$$

the desired solution $Q = \epsilon\{C(M)\} =$

$$\begin{matrix} 0.0904 \\ 0.1782 \\ 0.0362 \\ -0.2288 \\ 0.0281 \\ -0.6556 \\ 0.2130 \\ -0.2811 \\ -0.0067 \\ 0.2136 \\ -0.0202 \\ -0.0965 \\ 0.1423 \\ 0.2887 \\ -0.4319 \\ 0.0771 \end{matrix}$$

with an eigenvalue of 17.9077. The algorithm proceeds as follows:

$$Q = \theta\{M\}$$

As the number of columns $= 16 > S = 4$, M is split into two subpartitions M_1 and M_2. Both are resubmitted to $O\{\cdot\}$:

$$Q_1 = \theta\{M_1\}$$

$$M_1 = \begin{matrix} 0.29 & -0.10 & -0.29 & -1.39 & 0.15 & -1.65 & 0.21 & -0.65 \\ -0.24 & -1.16 & -0.50 & -0.02 & -0.01 & 2.27 & -1.05 & 1.02 \end{matrix}$$

Again the number of columns $= 8 > S$ and M_1 is split into M_{11} and M_{12}. Both are resubmitted to $\theta\{\cdot\}$:

$$Q_{11} = \theta\{M_{11}\}$$

$$M_{11} = \begin{matrix} 0.29 & -0.10 & -0.29 & -1.39 \\ -0.24 & -1.16 & -0.50 & -0.02 \end{matrix}$$

Now the stopping criterion is satisfied and the number of columns $= S$. The covariance matrix is derived:

$$C\{M_{11}\} = \begin{matrix} 0.14 & 0.28 & 0.05 & -0.37 \\ 0.28 & 0.56 & 0.11 & -0.73 \\ 0.05 & 0.11 & 0.02 & -0.14 \\ -0.37 & -0.73 & -0.14 & 0.93 \end{matrix}$$

and the $S/2 = 2$ eigenvectors with the largest eigenvalues are returned. The remaining eigenvalues are zero, so no variance/covariance is lost:

$$Q_{11} = \epsilon\{C\{M_{11}\}\} = \begin{matrix} -0.45 & -0.29 \\ -0.33 & -0.58 \\ -0.62 & -0.11 \\ -0.54 & 0.74 \end{matrix}$$

return to previous level

Now M_{12} is submitted:

$$Q_{12} = \theta\{M_{12}\}$$

Again the number of columns = $S = 4$. The $x\ (=2)$ eigenvectors are given by $Q_{12} = \epsilon\{C\{M_{12}\}\}$; return to previous level.

A reduced form of M_1 is now obtained by taking the component scores over $S/2 = 2$ PCs for both M_{11} and M_{12} to make a new, smaller matrix M_1':

$$M_1' = [M_{11} \cdot Q_{11}\ M_{12} \cdot Q_{12}] = \begin{matrix} 0.84 & -1.03 & 0.28 & -1.77 \\ 0.84 & 0.79 & 0.28 & 2.68 \end{matrix}$$

M_1' is an orthogonal transformation of M_1 where $M_1' = M_1 \cdot L_1$ and L_1 has Q_{11} and Q_{12} in the upper left and lower right quadrants, respectively (and zeros elsewhere). The $S/2$ eigenvector solutions of $C\{M_1\}$ are the columns of Q_1 where $Q_1^T \cdot C\{M_1\} \cdot Q_1 = V_1$ and V_1 has $S/2$ eigenvalues on its leading diagonal. Now $C\{M_1'\} = L_1^T \cdot C\{M_1\} \cdot L_1$ and the $S/2$ eigenvector solutions are Q_1', where $= Q_1'^T \cdot L_1^T \cdot C\{M_1\} \cdot L_1 \cdot Q_1' = V_1' = V_1 = Q_1^T \cdot C\{M_1\} \cdot Q_1$. The equivalence between V_1 and V_1' is assured because the eliminated columns of M_1' were zero and account for no variance.

Therefore, $Q_1 = L_1 \cdot Q_1'$. L_1 is known and Q_1' is obtained by submitting M_1' to $\theta\{\cdot\}$:

$$Q_1' = \theta\{M_1'\}$$

The number of columns of $M_1' = S = 4$ so $Q_1' = \epsilon\{C\{M_1'\}\}$; return to previous level:
Compute

$$Q_1 = L_1 \cdot Q_1'$$

$$\begin{matrix} 0.27 & -0.11 \\ 0.53 & -0.22 \\ 0.10 & -0.04 \\ -0.69 & 0.28 \\ -0.01 & -0.03 \\ 0.33 & 0.81 \\ -0.10 & -0.26 \\ 0.14 & 0.34 \end{matrix} = \begin{matrix} -0.45 & -0.29 & 0 & 0 \\ -0.33 & -0.58 & 0 & 0 \\ -0.62 & -0.11 & 0 & 0 \\ -0.54 & 0.74 & 0 & 0 \\ 0 & 0 & -0.03 & -0.03 \\ 0 & 0 & -0.21 & 0.88 \\ 0 & 0 & -0.95 & -0.28 \\ 0 & 0 & -0.22 & 0.37 \end{matrix} \times \begin{matrix} 0 & 0 \\ -0.92 & 0.37 \\ 0 & 0 \\ 0.37 & 0.92 \end{matrix}$$

return to previous level

The above is now repeated for M_2:

$$Q_2 = \theta\{M_2\}$$

The number of columns $> S$ therefore split M_2 into M_{21} and M_{22} and invoke $\theta\{\cdot\}$:

$$Q_{21} = \theta\{M_{21}\}$$

number of columns = S calculate Q_{21}
return to previous level

Now compute Q_{22}:

$$Q_{22} = \theta\{M_{22}\}$$

return to previous level

Derive the reduced form of $M_2 = M_2' = [M_{21} \cdot Q_{21}\ M_{22} \cdot Q_{22}]$ and compute Q_2':

$$Q_2' = O\{M_2'\}$$

return to previous level

Compute $Q_2 = L_2 \cdot Q_2'$

return to previous level

Derive the reduced form of M; $M' = [M_1 \cdot Q_1\ M_2 \cdot Q_2]$ and compute Q':

$$Q' = \theta\{M'\}$$

return to previous level

Compute

$$Q = L \cdot Q' =$$

$$\begin{matrix} 0 & -0.09 \\ 0 & -0.17 \\ 0 & -0.03 \\ 0 & 0.22 \\ 0 & -0.02 \\ 0 & 0.65 \\ 0 & -0.21 \\ 0 & 0.28 \\ -0.04 & 0.00 \\ 0.36 & -0.21 \\ -0.14 & 0.02 \\ 0.83 & 0.09 \\ 0.01 & -0.14 \\ 0.03 & -0.28 \\ 0.09 & 0.43 \\ 0.36 & -0.07 \end{matrix} = \begin{matrix} 0.27 & -0.11 & 0 & 0 \\ 0.53 & -0.22 & 0 & 0 \\ 0.10 & -0.04 & 0 & 0 \\ -0.69 & 0.28 & 0 & 0 \\ -0.01 & -0.03 & 0 & 0 \\ 0.33 & 0.81 & 0 & 0 \\ -0.10 & -0.26 & 0 & 0 \\ 0.14 & 0.34 & 0 & 0 \\ 0 & 0 & 0.04 & -0.01 \\ 0 & 0 & -0.36 & 0.36 \\ 0 & 0 & 0.14 & -0.03 \\ 0 & 0 & -0.83 & -0.16 \\ 0 & 0 & -0.01 & 0.24 \\ 0 & 0 & -0.03 & 0.48 \\ 0 & 0 & -0.09 & -0.72 \\ 0 & 0 & -0.36 & 0.13 \end{matrix} \times \begin{matrix} 0 & -0 \\ 0 & 0.80 \\ -1 & 0 \\ 0 & -0.59 \end{matrix}$$

The right hand column of Q has a real positive eigenvector and is the required solution. At no point did either $\epsilon\{\cdot\}$ or $C\{\cdot\}$ have an operand with more than four columns.

REFERENCES

Aersten A, Preissl H (1991) Dynamics of activity and connectivity in physiological neuronal networks. In: *Non Linear Dynamics and Neuronal Networks* (Schuster HG, ed), New York, VCH Publishers Inc., pp 281–302

Corbetta M, Miezin FM, Dobmeyer S, Shulman GL, Petersen SE (1991) Selective and divided attention during visual discrimination of shape, color and speed. Functional anatomy by positron emission tomography. *J Neurosci* 11:2383–2402

Cox DR, Miller HD (1980) *The Theory of Stochastic Processes*. New York, Chapman & Hall, pp 272–336

Eckhorn R, Bauer R, Jordan W, Brosch M, Kruse W, Munk M, Reitboeck HJ (1988) Coherent oscillations: A mechanism of feature linking in the visual cortex? Multiple electrode and correlation analysis in the cat. *Biol Cybernet* 60:121–130

Edelman GM (1978) Group selection and phasic re-entrant signalling: A theory of higher brain function. In: *The Mindful Brain*, Cambridge, MA, MIT Press, pp 55–100

Foldiak P (1989) Adaptive network for optimal linear feature extraction (1989) *Proc IEEE* 1:401–405

Fox PT, Mintun MA (1989) Non-invasive functional brain mapping by change distribution analysis of averaged PET images of $H^{15}O_2$ tissue activity. *J Nucl Med* 30:141–149

Friston KJ, Frackowiak RSJ (1991) Imaging functional anatomy. In: *Brain Work and Mental Activity (Alfred Benson Symposium 31)* (Lassen NA, Ingvar DH, Raichle ME, Friberg L, eds), Copenhagen, Munksgaard, pp 267–277

Friston KJ, Passingham RE, Nutt JG, Heather JD, Sawle GV, Frackowiak RSJ (1989) Localization in PET images: Direct fitting of the intercommissural (AC–PC) line. *J Cereb Blood Flow Metab* 9:690–695

Friston KJ, Frith CD, Liddle PF, Dolan RJ, Lammertsma AA, Frackowiak RSJ (1990) The relationship between global and local changes in PET scans. *J Cereb Blood Flow Metab* 10:458–466

Friston KJ, Frith CD, Liddle PF, Frackowiak RSJ (1991*a*) Investigating a network model of word generation with positron emission tomography. *Proc R Soc Lond Ser B* 244:101–106

Friston KJ, Frith CD, Liddle PF, Frackowiak RSJ (1991*b*) Plastic transformation of PET images. *J Comput Assist Tomogr* 15:634–639

Friston KJ, Liddle PF, Frith CD, Hirsch SR, Frackowiak RSJ (1992*a*) The left medial temporal lobe and schizophrenia: A PET study. *Brain* 115:367–382

Friston KJ, Frith CD, Passingham RE, Dolan RJ, Liddle PF, Frackowiak RSJ (1992*b*) Entropy and cortical activity: Information theory and PET findings. *Cerebral Cortex* 2:259–267

Frith CD, Friston KJ, Liddle PF, Frackowiak RSJ (1991) Willed action and the prefrontal cortex in man. *Proc R Soc Lond Ser B* 244:241–246

Gerstein GL, Bedenbaugh P, Aersten AMHJ (1989) Neuronal assemblies. *IEEE Trans Biomed Eng* 36:4–14

Gerstein GL, Perkel DH (1969) Simultaneously recorded trains of action potentials: Analysis and functional interpretation. *Science* 164:828–830

Gochin PM, Miller EK, Gross CG, Gerstein GL (1991) Functional interactions among neurons in inferior temporal cortex of the awake macaque. *Exp Brain Res* 84:505–516

Goldman-Rakic PS (1986) Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. In: *Handbook of Physiology*—V (Mountcastle VB, ed), Baltimore, Williams & Wilkins, pp 373–417

Goldman-Rakic PS (1988) Topography of cognition: Parallel distributed networks in primate association cortex. *Annu Rev Neurosci* 11:137–156

Gray CM, Singer W (1989) Stimulus specific neuronal oscillations in orientation columns of cat visual cortex. *Proc Natl Acad Sci USA* 86:1698–1702

Hebb DO (1949) *The Organization of Behavior. A Neuropsychological Theory*. New York, Wiley

Hope K (1968) *Methods of Multivariate Analysis*. London, University of London Press Ltd.,

Horwitz B (1990) Simulating functional interactions in the brain: A model for examining correlations between regional cerebral metabolic rates. *Int J Biomed Comput* 26:149–170

Horwitz B, Duara R, Rappoport SI (1984) Intercorrelations of glucose rates between brain regions: Application to healthy males in a reduced state of sensory input. *J Cereb Blood Flow Metab* 4:484–499

Horwitz B, Grady C, Haxby J, Schapiro M, Carson R, Herscovitch P, Ungerleider L, Mishkin M, Rapoport S (1991) Object and spatial visual processing: Intercorrelations of regional cerebral blood flow among posterior brain regions [Abstract]. *J Cereb Blood Flow Metab* 11(suppl 2):S380

Lagreze HL, Hartmann A, Shaub A (1991) A factor imaging of cortical blood flow during behavioral activation: Interaction of neuronal networks in cognition [Abstract]. *J Cereb Blood Flow Metab* 11(suppl 2):S369

Llinas RR, Pare D (1991) Of dreaming and wakefulness. *Neuroscience* 44:521–535

Linsker R (1988) Self organization in a perceptual network. *Computer* March:105–117

McClelland JL, Rumelhart DE (1988) *Explorations in Parallel Distributed Processing*. Cambridge, MA, The MIT Press.

Mesulam MM (1990) Large scale neurocognitive networks and distributed processing for attention language and memory. *Ann Neurol* 28:597–613

Metter EJ, Riege WH, Kuhl DE, Phelps ME (1984) Cerebral metabolic relationships for selected brain regions in healthy adults. *J Cereb Blood Flow Metab* 4:1–7

Moeller JR, Struther SC, Sidtis JJ, Rottenberg DA (1987) Scaled subprofile model: A statistical approach to the analysis of functional patterns emission tomographic data. *J Cereb Blood Flow Metab* 7:649–658

Moler C, Little J, Bangert S (1987) *PRO-MATLAB User's Guide*. Sherborn, MA, The Math Works Inc.

Oja E (1982) A simplified neuron model as a principal component analyzer. *J Math Biol* 15:267–273

Oja E (1989) Neural networks, principal components, and subspaces. *Int J Neural Syst* 1:61–68

Pardo JV, Pardo PJ, Janer KW, Raichle ME (1990) The anterior cingular cortex mediates processing selection in the Stroop attentional conflict paradigm. *Proc Natl Acad Sci USA* 87: 256–259

Petersen SE, Fox PT, Posner MI, Mintun M, Raichle ME (1989) Positron emission tomographic studies of the processing of single words. *J Cogn Neurosci* 1:153–170

Posner ML, Sandson J, Dhawan M, Shulman GL (1990) Is word recognition automatic? A cognitive–anatomical approach. *J Cogn Neurosci* 1:50–60

Sherrington C (1941) *Man of His Nature, The Gifford Lectures. Edinburgh 1937–1938*. Cambridge, U.K., Cambridge University Press

Smith BT, Boyle JM, Dongarra JJ, Garbow BS, Ikebe Y, Klema VC, Moler CB (1976) *Matrix Eigensystem Routines—EISPACK Guide, Lecture Notes in Computer Science*, Vol. 6, second edition. New York, Springer-Verlag

Spinks TJ, Jones T, Bailey DL, Townsend DW, Grootnook S, Bloomfield PM, Gilardi MC, Casey ME, Sipe B, Reed J (1992) Physical performance of a positron tomograph for brain imaging with retractable septa. *Phys Med Biol* 37:1637–1655

Sporns O, Gally JA, Reeke GN, Edelman GM (1989) Reentrant signalling among simulated neuronal groups leads to coherence in their oscillatory activity. *Proc Natl Acad Sci USA* 86:7265–7269

Talairach J, Tournoux P (1988) *A Stereotactic Coplanar Atlas of the Human Brain*. Stuttgart, Thieme.

Thatcher RW, Krause PJ, Hrybyk M (1986) Cortico-cortical associations and EEG coherence: A two compartmental model. *Electroencephalogr Clin Neurophysiol* 64:123–143

Townsend DW, Geissbuhler A, Defrise M, Hoffman EJ, Spinks TJ, Bailey D, Gilardi MC, Jones T (1992) Fully three-dimensional reconstruction for a PET camera with retractable septa. *IEEE Trans Med Imag* MI-10:505–512

Voss RF (1988) Fractals in nature. From characterization to simulation. In: *The Science of Fractal Images* (Peitgen H, Saupe D, eds), New York, Springer Verlag, pp 21–69

Wise SP (1989) Frontal cortical activity and motor set. In: *Neural Programming* (Ito M, ed), Tokyo, Japanese Scientific Societies Press, p 26

Zeki S (1990) The motion pathways of the visual cortex. In: *Vision: Coding and Efficiency* (Blakemore C, ed), Cambridge, U.K., Cambridge University Press, pp 321–345