NEUROIMAGE **3,** 143–157 (1996)
ARTICLE NO. 0016

Spatial Pattern Analysis of Functional Brain Images Using Partial Least Squares

A. R. MCINTOSH,* F. L. BOOKSTEIN,† J. V. HAXBY,‡ AND C. L. GRADY*,§

Rotman Research Institute of Baycrest Centre, 3560 Bathurst Street, University of Toronto, Toronto, Ontario M6A 2E1, Canada; †Institute of Gerontology, University of Michigan, Ann Arbor, Michigan 48109; ‡Functional Brain Imaging Section, Laboratory of Psychology & Psychopathology, National Institute of Mental Health, National Institutes of Health, Bethesda, Maryland 20892; and §Laboratory of Neurosciences, National Institute on Aging, Bethesda, Maryland 20892

Received December 13, 1995

This paper introduces a new tool for functional neuroimage analysis: partial least squares (PLS). It is unique as a multivariate method in its choice of emphasis for analysis, that being the covariance between brain images and exogenous blocks representing either the experiment design or some behavioral measure. What emerges are spatial patterns of brain activity that represent the optimal association between the images and either of the blocks. This process differs substantially from other multivariate methods in that rather than attempting to predict the individual values of the image pixels, PLS attempts to explain the relation between image pixels and task or behavior. Data from a face encoding and recognition PET rCBF study are used to illustrate two types of PLS analysis: an activation analysis of task with images and a brain–behavior analysis. The commonalities across the two analyses are suggestive of a general face memory network differentially engaged during encoding and recognition. PLS thus serves as an important extension by extracting new information from imaging data that is not accessible through other currently used univariate and multivariate image analysis tools.

1. INTRODUCTION

Functional brain images are extraordinarily rich data sets revealing participation of all parts of the brain in a wide variety of perceptual and cognitive operations. It is a substantial analytic challenge to integrate the temporal, spatial, and statistical signals making up these data. Every method currently available emphasizes certain aspects of these images while deemphasizing or averaging over others. In this paper we introduce a new emphasis that is embodied in a novel computational strategy, partial least squares. The method extracts spatial patterns of brain activity that are optimally associated with aspects of the experiment design or some behavioral measure. Its greatest strength is the flexible treatment of images in the context of simultaneous prediction of those images by their causes (e.g., aspects of the task design) and prediction by those images of their effects (e.g., measures of behavior). Partial least squares extracts certain features that are inaccessible by other methods, while overlooking some complexities for which other methods may be more suited.

Most of the contemporary techniques for analysis of functional neuroimaging data are variations of univariate analyses: either single image elements or contiguous regions of voxels are treated as computationally independent. Beyond these univariate methods, global approaches that treat the set of image pixels as a whole, rather than as a tessellation of regions, are potentially well suited to analysis of brain-imaging data. This latter class of multivariate analyses will typically produce weighted combinations of pixel contents, summed all across the image, that relate to some variable of interest, for example, experimental condition or behavior. In this context, interdependency among areas is used to differing degrees. For instance, methods introduced to neuroimaging by McIntosh and Gonzalez-Lima (1994) for network analysis and by Friston (1994) for measures of functional and effective connectivity (Aertsen *et al.*, 1989; Gerstein *et al.*, 1978) represent the explicit use of interdependency as indices of neural interactions.

This paper calls attention to a newer method originally developed for applications to one-dimensional images from spectrographs, as in chemometrics or remote sensing, or toxicology: partial least squares (PLS). PLS operates on the covariation among two or more "blocks" of variables and seeks to obtain a new set of variables that optimally relate the blocks using the fewest dimensions. "Optimal" as used here is in the sense of maximal covariance, not maximal correlation—a distinction quite important for the algebraic

1053-8119/96 $18.00

development. In the present applications, one block contains the functional images for all subjects in all experimental conditions, a second block is for vectors of contrasts coding the experimental design, and a third block collects the behavioral outcomes. Separately analyzing the second and third blocks against the first, we show how PLS can carry out image-wide PET activation analysis or extract task-specific features of brain–behavior covariances and correlations. Within the space of image descriptors, the findings of PLS are carried by *singular images,* computed multivariate optima somewhat related to the previously introduced subscale profiles (Moeller *et al.*, 1987) and *eigenimages* (Friston *et al.*, 1993). The singular images are spatial patterns of pixel contents representing pixel-by-pixel covariances with task or behavior. For a thorough review of PLS, including a tour of its scientific antecedents and a comparison with structural equations methods, see Streissguth *et al.* (1993) or Bookstein *et al.* (1996). To our knowledge, this is the first application of PLS to pixellated functional neuroimaging data.

The method underlying the applications of PLS to PET rCBF data is presented through four illustrations. The first is a conceptual illustration presented in graphical form. The second and third are more technically oriented descriptions that assume some knowledge of multivariate statistics by the reader. Finally, the Appendix presents a PLS analysis on a small simulated data set. Readers who do not have a strong statistical background may wish to read the conceptual description and the Appendix before studying the two real examples.

2. A FIRST CONCEPTUAL ILLUSTRATION OF PLS

Figure 1 presents an outline of the steps in a PLS analysis for activation data. For this illustration there are three scan conditions each with seven subjects. Step 1 (top of Fig. 1): the experimental data are represented as a design matrix X of contrasts and a matrix Y of normalized PET rCBF images for each subject in each condition. Assume that there are two hypotheses of interest: the contrast between condition 1 (a baseline task) and the average of task 2 and 3, and the contrast between 2 and 3. These can be expressed as the orthogonal vectors (2, −1, −1) and (0, 1, −1). A cross-correlation matrix S is computed for each contrast in X against each pixel of the normalized images. Step 2 (middle of Fig. 1): the singular value decomposition of matrix S produces a series of paired "singular vectors" or "latent variables" that generate successively more accurate approximations of the matrix S. Here we show only the first pair (A_1, B_1). A_1 is the linear combination of design contrasts that contributed optimally to the reconstruction of the correlations in S; the contents of the singular vector B_1 indicate which pixels are most sensitive to those contrasts. Remapped into image space, as shown in the figure, the vector B_1 becomes the *singular image* B_1, elaborated below. The singular value, d_1, is the covariance between A_1 and B_1. The square of this value, when divided by the sum of squared correlations in S, yields the "proportion of the sum of squared cross-block correlations explained" by A_1 and B_1. Step 3 (bottom of Fig. 1): *scores* express the original data in terms of components along the singular vectors. As suggested in the figure, these are computed as simple sums of products of corresponding elements. For functional neuroimaging experiments, it is useful to plot the scores on the singular images against experimental condition. For our second example, relating brain activity and behavior, the score on the singular images would be plotted instead against the score on the corresponding latent (singular) outcome.

3. THE ALGEBRA OF PLS AND THE EXPLANATIONS IT SUPPORTS

Methods that summarize large covariance structures with a smaller number of components, especially those based on the singular-value decomposition (SVD), are becoming familiar algorithmic tools in the postprocessing of functional brain images. Applied to matrices of pixel-specific, voxel-specific, or region-specific counts, variously normalized, SVD produces eigenimages and experiment-specific profiles combining the eigenimages. The eigenimages are eigenvectors of the element-by-element covariance or correlation matrix around the baseline or average: spatial patterns of variation accounting for successively maximal portions of that covariance pattern. The profiles characterize the contribution of each eigenimage to this finding or deviation of each condition in the data set. In this paper, we use the same SVD but apply it to a different matrix: a matrix of covariances, not raw image contents. The covariances relate the separate pixels of the normalized image to any of a variety of exogenous blocks representing separately measured causes or consequences of the functional images.

In a technique that, by accident of history, bears the quite uninformative name "partial least squares," the same singular-value decomposition now yields quantities that take on wholly different scientific interpretations. While the algebra and the computer implementation of the SVD are unchanged, the interpretation of the singular images and corresponding profiles is com-

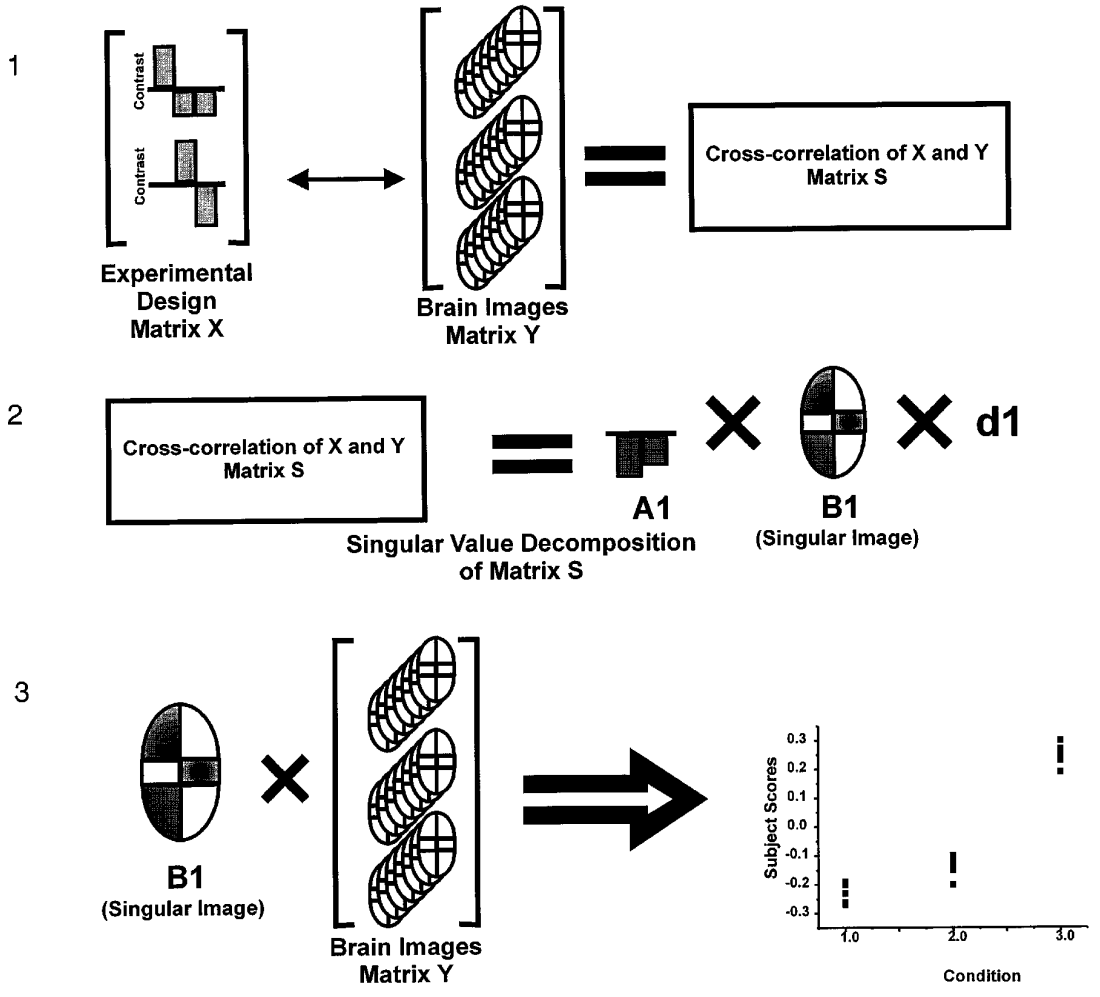

FIG. 1. Schematic illustration of the steps in the PLS analysis of activation effects. Step 1 represents the computation of the cross-correlation between vectors coding the experimental design and the normalized "brain images" for all subjects across all tasks (though represented as slices in the figure, the images used are actually multisliced volumes). Step 2: A singular value decomposition is performed on the cross-correlation matrix S. This gives several pairs of singular vectors, the first of which is shown, that approximate matrix S. Step 3: The raw brain images are multiplied through by the singular image B1, to obtain subject scores which are then plotted with respect to condition in the design.

pletely altered from the interpretation of eigenimages. We also introduce a third class of descriptors, the *subject scores*. These scores, not to be confused with either of the quantities of the raw data matrix decomposition, characterize both the image contents and the experimental tasks in a way quite different from the profiles that arise in the application of SVD to raw data matrices.

In the present setting, we apply PLS in two different ways. In the first setting, the exogenous block is the familiar design matrix of tasks for the original PET data collection. The matrix of "covariances" becomes a matrix of additive components of image covariation with the design. In the second setting, the exogenous block will be a set of quantities characterizing a behavioral measure—reaction time. The interpretations of the resulting quantities are quite different between the two contexts, even though the means by which they are produced are identical.

The scientist for whom PLS is designed is faced with two lists of ordinary statistical variables; call them the X's and the Y's. In the application here, the X's will sometimes be dummy variables for design features and sometimes behavioral measures, and the Y's will always be the contents of normalized pixels or voxels. Number the X's from X_1 to X_m, the Y's from Y_1 to Y_n. Write S for the covariance matrix of the X's by the Y's, m rows by n columns. (Actually, the data blocks X and Y need not be mean centered, so that the matrix S could be a raw cross product matrix instead of a covariance matrix; we will not need this extension here). Its element S_{ij} is the covariance of X_i with Y_j. This matrix

does not have variances down the diagonal; usually it is not even square.

Then a certain interesting computation can be phrased in any of several different ways. All involve the production of a column vector A of m coefficients A_i, one for each X, together with a column vector B of n coefficients B_j, one for each Y. The elements of either vector, A or B, are scaled so as to have squares that sum to 1. There is a pair of vectors A, B, the *first singular vector pair for the matrix* S, and a scalar d, the *first singular value* of S, such that the following four assertions (which are algebraically equivalent) are simultaneously true:

- The *outer product* AB^t (where t indicates transpose), the $m \times n$ matrix whose ijth entry is A_iB_j, is, after multiplication by d, the best (least-squares) fit of all such matrices to the covariance matrix S between the blocks. The goodness of this fit serves as a figure of merit for the overall PLS analysis. It is the ordinary explained sum-of-squares from the "model" $S_{ij} = dA_iB_j^t + error$.
- The *latent variables* $\mathrm{LV}_X = \Sigma_{i=1}^{m} A_iX_i$ and $\mathrm{LV}_Y = \Sigma_{j=1}^{n} B_jY_j$ have covariance d, and this is the greatest covariance of any pair of such linear combinations for which the coefficient vectors A and B both sum in square to 1. LV_X and LV_Y are also the *scores* defined in the previous illustration in Section 2. Conceptually, these latent variables differ somewhat from those of the structural-equation models (cf. McIntosh and Gonzalez-Lima, 1994) in that they are ordinary linear combinations of the observed data, not factors, and they have exact values. The quantity PLS is optimizing is the covariance between these paired LV's, *not* the variance explained in the course of predictions of the outcomes individually or collectively. Also, PLS is remarkably different from the similar-looking technique of canonical correlations analysis, which optimizes correlations between composites rather than, as here, covariances. These important differences are discussed in Bookstein (1994) and Sampson *et al.* (1989).
- The elements A_i of the vector A are proportional to the covariances of the corresponding X-block variable X_i with the latent variable LV_Y representing the Y's, the *other* block in the analysis; similarly, the elements B_j of the vector B are proportional to the covariances of the corresponding Y-block variable Y_j with the latent variable LV_X representing the A's. When it is known a priori that there is a causal relation between the X's and the Y's, as in most PET experiments, these coefficients may be called *saliences*. Each A_i is the salience of the variable X_i for the latent variable representing the Y-block, and each B_j is the salience of the variable Y_j for the latent variable representing the X-block.
- The vector A is the first, uncentered, unstandardized principal component of the columns of the matrix S considered as if it were raw data, and the vector B is similarly the first uncentered, unstandardized principal component of the rows.

Let us translate all this into the specific context of the data sets under discussion here. The block Y is an array of normalized pixel contents. (The normalization rules are crucial to the computations that follow; a careful consideration of rules for their imposition is the subject of another paper. For a very simple example, a single contrast between groups, see Bookstein, 1996.) Then the vector B will always be an image of unit amplitude (sum-of-squares) representing variation of normalized pixels around the grand mean image. Because these images are singular vectors of the SVD, they can be called singular images. When X is a design matrix of dummy variables, S is just a "matrix" of image contrasts: design effects around the grand mean. Then B is the single image that best captures the variation of the mean target images about their grand mean; A is a profile of the covariances of the elements of the design block X with this singular image; the scores LV_Y are the components of the singular image B in each of the experimental runs; and the scores LV_X are proportional to the components of this same image in each of the design conditions.

The eigenimages obtained by Friston (1994) can be thought of as a special case of this algorithm when the two blocks involved are identical—when the covariances of the matrix S are of the image elements among themselves. We will pursue differences between eigenimages and these singular images further under Discussion (Section 7 below).

When X is instead a vector of behavioral measurements, as in the second example, exactly the same computations result in a matrix S of covariances between image content and behavioral scores. This produces a set of three interlinked PLS descriptors as follows: B is still an image of unit amplitude; the score vector LV_Y still represents the component of the singular image B in the individual images of the data set. However, A is the profile of covariances of the X-block measures with the component of B in each image of the data set—their separate saliences for predicting the score LV_Y or being predicted by it—and the scores LV_X now characterize the individual experimental runs of the data set by their profiles on the particular pattern given by the saliences A.

What has just been described is the first step in an algorithm that ends after the lesser of m or n steps. (When Y is a full image, this will be the count m of the exogenous block). The singular-value decomposition of S is actually an exhaustive reconstruction $S = \mathbf{ADB}^t$, where $\mathbf{A}$ is an orthogonal matrix of order m of which the vector A of the preceding discussion is the first column,

B is an orthogonal matrix of order n of which the vector B of the preceding discussion is the first column, and **D** is an $m \times n$ diagonal matrix of which the scalar d of the preceding discussion is the upper left entry (see Mardia *et al.*, 1979). This can be written more explicitly as the reconstruction $S = \mathbf{ADB}^t = \mathbf{d_1 A_1 D B_1^t} + \mathbf{d_2 A_2 B_2^t} + \ldots$. The interpretation in terms of explained sums-of-squares of the matrix S begins with the rank-1 approximant $\mathbf{A_1 d_1 B_1^t}$ and improves by adding in successive additional terms $\mathbf{A_2 d_2 B_2^t}$, etc. All singular vectors, not just the first, come in pairs $\mathbf{A_k}$, $\mathbf{B_k}$, and each pair is accompanied by a singular value $\mathbf{d_k}$. The total number of these is no greater than $\min(m,n)$. It is in *this* decomposition, not the eigenanalysis of the Y-block per se, that a threshold of significance is appropriately imposed for relevance of image contents to the experimental design or to behavioral measures (see Streissguth *et al.*, 1993, for a two-dimensional example). The number of terms that prove significant in this decomposition typically is small.

A PLS analysis thus combines the elements of two (or more) blocks of variables into systems of parameters, separately block by block, that sort contrasts/behavior and functional images (by saliences A_i or B_j), and individual tasks (by scores $\mathrm{LV}_X = \Sigma_{i=1}^{m} \mathrm{A_i X_i}$ and $\mathrm{LV}_Y = \Sigma_{j=1}^{n} B_j Y_j$). In the initial application below, the scores LV_X are invariant across design subcategories, but not the scores LV_Y, which vary from run to run within a design category. In more general contexts, both blocks, the X's as well as the Y's, contain values that are unique across the experiments of a data set. The resulting systematic summary of the pattern of cross-block correlation links characterization of images with characterization of their causes or effects, whether designed or behavioral.

A complete list of parameters for the principal hypothesis includes the saliences A of the design features, the saliences B of the singular images (do they regionalize? do they conform to models of neural circuitry?), and the values for scores on both latent variables over the experiments of a data set. This totals $i(m + n + 2 - i)$ parameters, where i is the number of pairs of singular vectors (singular images and the corresponding singular vectors for design and behavior) retained. Partial least squares computes all these parameters at once in a single matrix decomposition, and it treats them all with equal seriousness: there are many algebraic constraints relating these sets of descriptors, and there is only one statistical test (realized below as a permutation test on the explained variance of the task design) for the entire procedure. The scientific leverage afforded by such highly structured but arithmetically simple multivariate approaches is relatively unfamiliar to the applied statistician and deserves to be better known.

Statistical Significance

The statistical significance of a PLS analysis can be assessed using any of the quantities that characterized the computation—the saliences of either set of singular vectors or the scores that characterize the experimental tasks. In this application we use a permutation test (Edgington, 1980; Good, 1994) to assess the extent to which the scores characterize tasks. The test involves randomly reordering the rows of the design matrix (i.e., destroying the association of experimental conditions with images) and, for each new ordering, computing a new SVD and a new ANOVA on the analogous scores. At each permutation, the statistic is compared to the obtained statistic for the original data. The obtained statistic is assigned a probability value based on the number of times a statistic from the permuted data exceeds the obtained value. Using this method, the significance of a particular LV can be assessed without relying on the distributional assumptions underlying most conventional parametric statistical methods. A univariate version of permutation tests for functional mapping has recently been introduced (Holmes *et al.*, 1996).

The test used in this application was a multiple regression approach to ANOVA (Pedhazur, 1982). The collected condition contrast vectors were used as predictor variables and the scores from each singular image were used as dependent variables. The computation of the test statistic is straightforward in matrix notation. The regression weights (β) are computed by $\beta = R_x^{-1} r$, where R_x is a matrix of correlations among the predictor variables and r is the cross correlation of the predictor variables (the contrasts, which are orthogonal) and the particular set of scores. As an indication of the significance of the regression, R^2 (the proportion of variance explained in the LV scores by the predictor variables) can be computed as $R^2 = \beta^T r$.

The significance of our ANOVA was determined simply by calibrating the R^2 value actually obtained against the distribution of these R^2's arising from random permutations. Data were permuted 10,000 times to assign a probability for the obtained R^2. For the PLS analysis using behavioral measures, a multiple regression was performed on the scores and the measures (which in this case are not orthogonal) and the R^2 from this regression was assessed through permutation. The principal computation load of our procedure is incurred in the course of this loop. But each iteration is only linear in n, the number of pixels in the images. (The SVD itself is roughly cubic in the length of the smaller data set; in these and similar applications, that part will be trivial.) Hence a loop over sufficiently many hundreds of iterations is handled with dispatch by a modern workstation. Using an interpreted code, MATLAB,

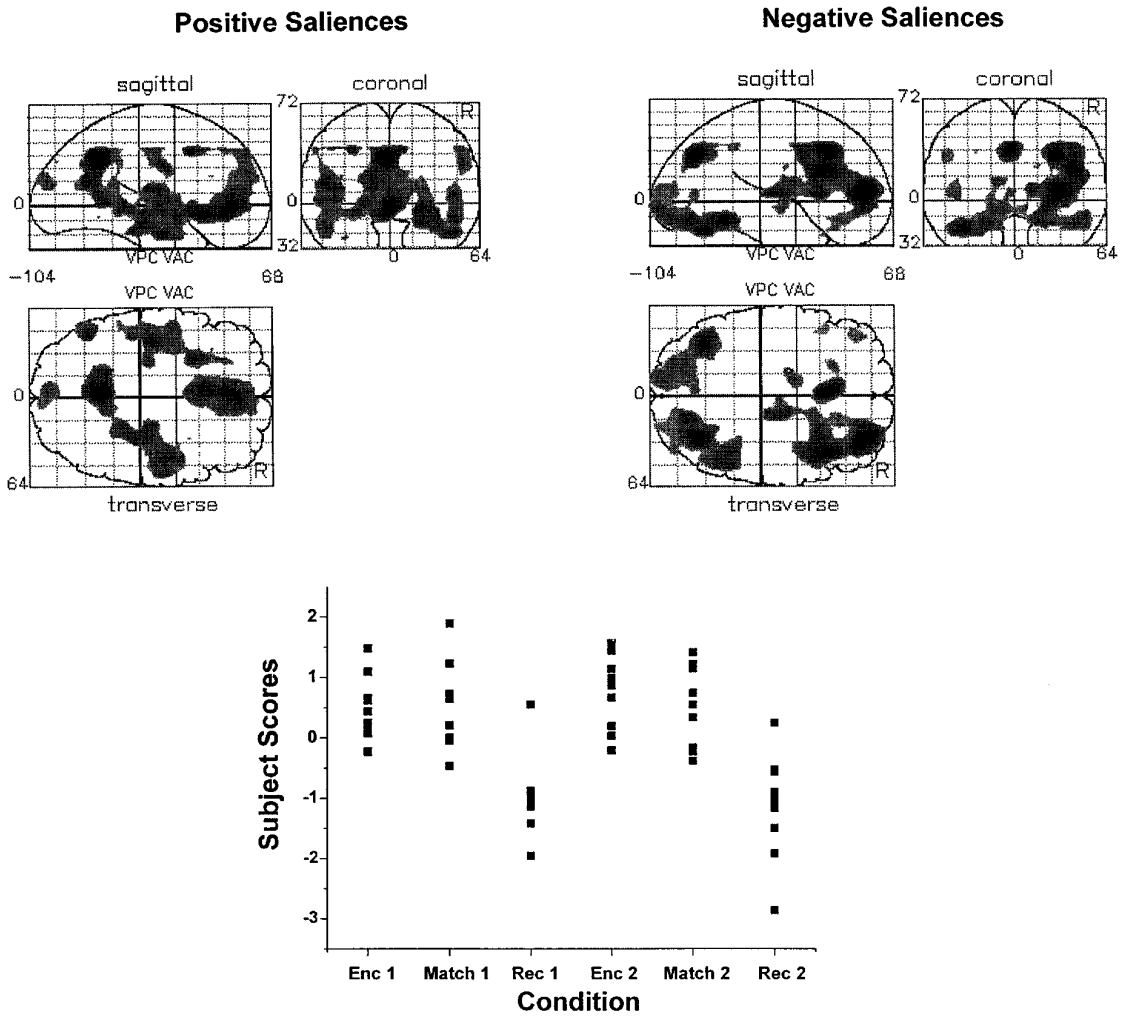

FIG. 2. Singular images and subject scores from dominant latent variable representing experimental variance. The singular images (top) present orthogonal projection plots of a "glass brain" showing positive and negative saliences separately. Saliences greater than an arbitrary threshold of 0.15 are shown. Scores for each subject are plotted by condition. The scores have been transformed to standard space.

permutation tests for all singular images were run simultaneously and completed over a weekend. It is anticipated that conversion to a compiled code, such as C, will greatly enhance the computation speed.

A similar resampling method can be employed for the saliences themselves. In this case, bootstrap estimates of standard errors provide a relatively simple method to determine the precision of the saliences within a given LV (Efron and Tibshirani, 1986; see also Sampson *et al.*, 1989 for an application in PLS). In this paper, we simply interpret the saliences as images of spatial pattern—the apparent concentrations in Figs. 2–4 are a guide to neurophysiological interpretation.

4. PET DATA

The data were obtained from a PET rCBF study that compared encoding and recognition of faces in 10 young subjects (Haxby *et al.*, 1996). The experiment had four conditions: (1) simultaneous match-to-sample of faces (face matching), (2) encoding of faces, (3) two-alternative recognition of faces, and (4) a sensorimotor control task. For the face-matching task, subjects were presented with three faces simultaneously. They indicated which of two choice faces at the bottom of the stimulus array corresponded to the sample face at the top of the array by pressing a button on the side corresponding to the correct choice with the thumb. During encoding, subjects were presented with a series of faces one at a time and instructed to memorize them, and during recognition they were presented with two faces and asked to indicate which one they had seen during the encoding phase by pressing the response button as in the face-matching task. The order of the tasks were as follows: control task, encoding 1, matching, recognition of faces from 1, encoding 2, matching, recognition 2, and control task. New faces were used in the first and second encoding tasks, and different faces were used in

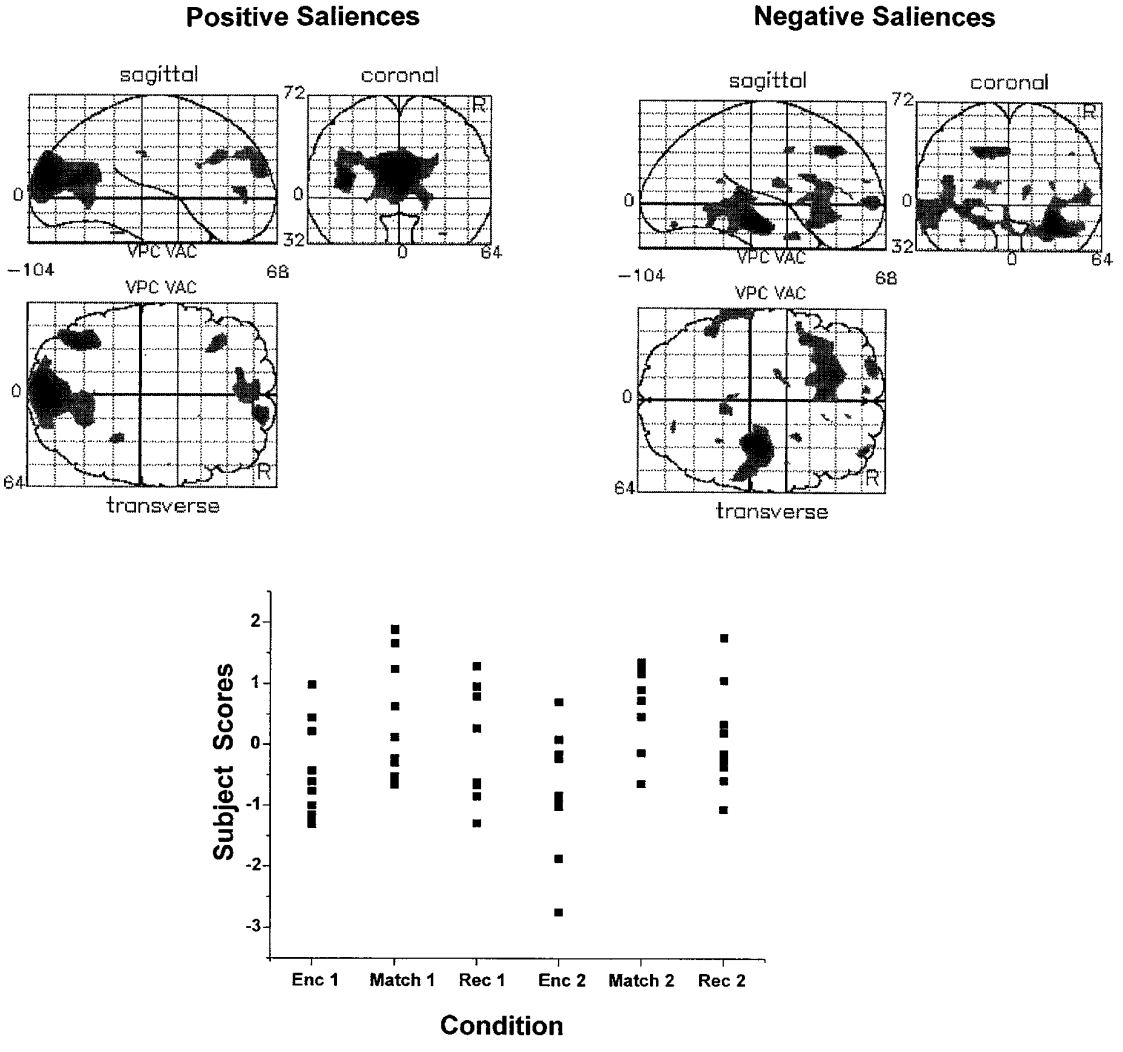

FIG. 3. Singular images and subject scores from second latent variable representing experimental variance. The singular images (top) present orthogonal projection plots of a "glass brain" showing positive and negative saliences separately. Scores for each subject are plotted by condition.

both matching tasks. This paper does not consider data from the control task.

The behavioral data acquired from all subjects were reaction time (RT) measures during face matching and recognition. Reaction time data were not obtained for 3 subjects owing to equipment failure; the analysis of rCBF-RT correlation is based on 7 subjects of 10.

5. ANALYSIS

PET data were analyzed using two methods: a conventional univariate voxel-by-voxel analysis using SPM (Version 4) and the PLS analysis. Because the data sets under study here had already been processed before the methods of anatomically driven image warping (e.g., Bookstein, 1996) became widely available, spatial normalization of the PET images was performed using algorithms that are part of the SPM package (Friston *et al.*, 1989, 1991). Changes in the method of registration would alter the findings here to some unknown extent. Both SPM analyses were run using a combination of C-code and PRO-MATLAB (Version 4, Mathworks) and the PLS analyses were run completely in MATLAB. Computations were performed on a Sun Sparc 20 workstation.

The SPM comparisons were performed to determine if there were brain regions specifically activated in encoding or recognition relative to face matching. The average of the two encoding tasks and the average of the two recognition tasks were each compared to the average of the two matching tasks.

Two PLS analyses were performed: one regarding the condition effects (activation analysis) and the other the correlation of RT with rCBF (brain–behavior analysis). The analyses used the same spatially transformed subject images that were used for SPM, but with a

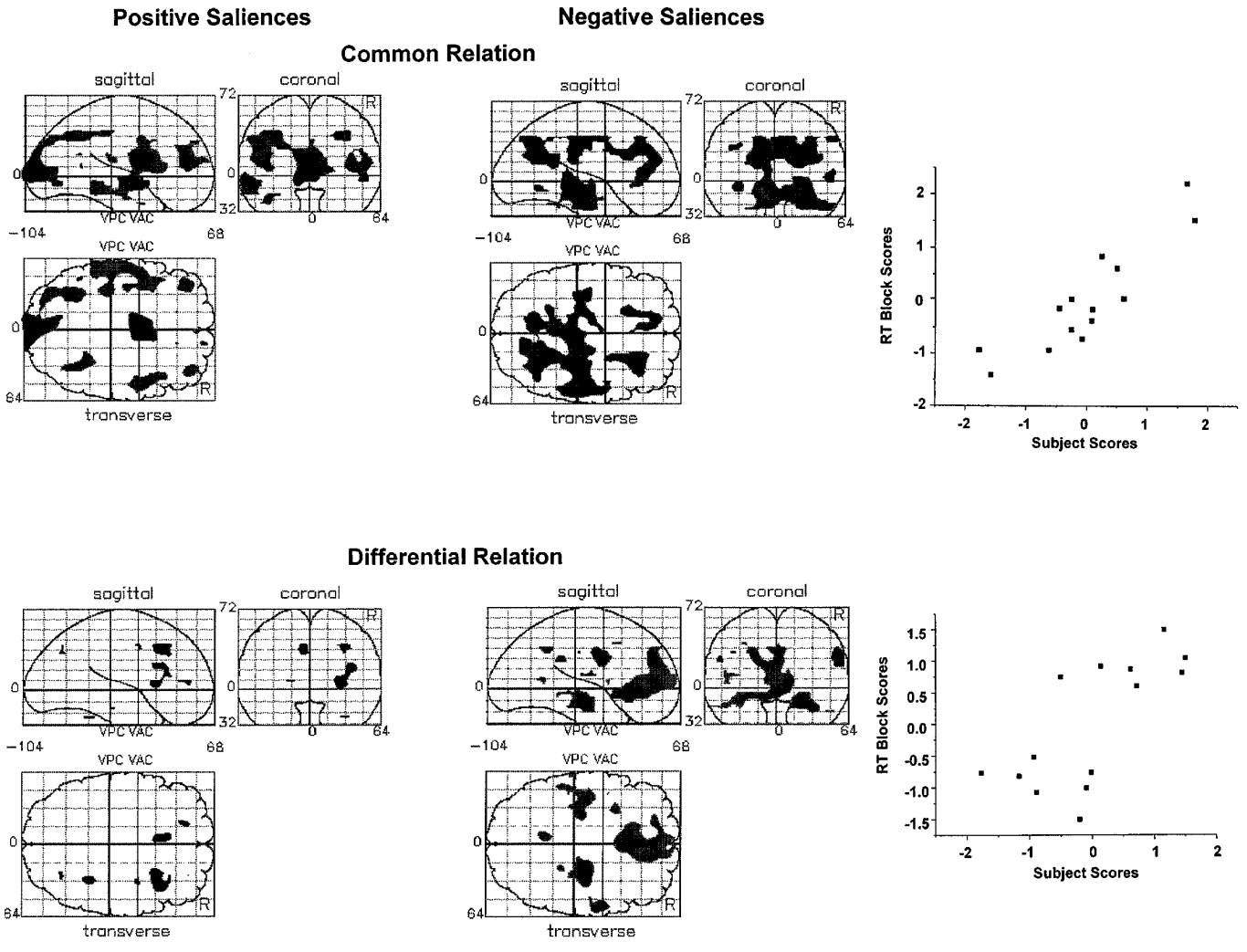

FIG. 4. Singular images from PLS analysis of brain–behavior relations. Top projections display regions showing a common correlation with behavior in both face matching and recognition conditions. The subject scores are plotted against the corresponding scores from the *X*-block containing the RT measures and the task by RT interaction. The relative continuity across the plot suggests the singular image depicts a common relation between RT and rCBF across task. The bottom projections are regions where the correlation with behavior differed between tasks, along with the plot of the subjects scores and the *X*-block. The clustering in the plot is suggestive of a differential relation across the two tasks.

smaller filter (15 mm instead of 20 mm) to reduce the possible confound of spatial autocorrelation. Slight differences in the stereotaxic location of some local maxima between the SPM and the PLS analyses will be partly a result of different smoothing filters (cf. Tables 1 and 2). Pixel intensities for individual images were adjusted for global CBF by dividing each image voxel by the within-task within-subject whole-image mean. The transformed images were then stacked together for all subjects and all conditions so the resulting data matrix had $n \times k$ rows (n, number of subjects; k, number of scans; the two scans for each of three tasks were not averaged so $k = 6$) and one column for each voxel in an image. The design matrix also had $n \times k$ rows and two columns for the orthogonal condition contrasts. In the order (encoding, matching, recognition), contrast 1 is $(-1, 2, -1)$—the average of encoding and recognition versus matching, and contrast 2 is $(1, 0, -1)$—encoding versus recognition.

The association of the images with reaction time was tested by a separate cross-correlation matrix of two columns representing an orthogonal rotation of the separate vectors of correlations of pixel by RT within the two tasks from which RT measures were obtained. The first column of this design matrix is the sum of the correlations of RT with rCBF in the two tasks separately, and the second column is the difference of those correlations. Because PLS is invariant against rotation of either block separately (Bookstein, 1991), this could as well have been represented as the unrotated matrix of the two task-specific correlations separately, followed by a rotation in the course of reporting the results.

6. RESULTS

SPM Results

For a full exposition of the SPM results, the reader is referred to papers by Grady *et al.* (1995) and Haxby *et al.* (1996). The essential findings from the activation results are that encoding of faces led to greater rCBF, compared to matching, in the left inferior and middle prefrontal cortex, right hippocampal gyrus, and left inferior temporal cortex. Decreased rCBF was observed in the dorsal medial occipital cortex. Recognition of faces, compared to matching, resulted in rCBF increases in right inferior and middle prefrontal, anterior cingulate, left and right parietal, and ventral posterior cortical and cerebellar regions and in decreases in dorsal medial occipital and superior temporal cortices. The local maxima and stereotaxic coordinates for these areas are presented in Table 1.

PLS Results—Activation Analysis

Projection plots of the two singular images (SI's) identified by the PLS analysis are presented in Figs. 2 and 3 and local maxima in Table 2. Figure 2 depicts a singular image with positive saliences extending from bilateral anterior temporal lobe to the right hippocampal gyrus and into posterior cingulate. Ventral aspects of the anterior cingulate have positive saliences on the first SI, as has a large part of the limbic system. The pattern follows the classic "C-shape" extending from anterior medial temporal lobes, posterodorsal hippocampus into retrosplenial, posterior, and ventral anterior cingulate. Regions showing negative saliences were the right prefrontal, dorsal anterior cingulate, inferior parietal, and ventral posterior regions and thalamus. Note that the appropriate confidence intervals around these regions, which do not concern us here, would be set by bootstrap analysis pixel by pixel rather than by manipulation of a "threshold." There is thus no analogue to the overall false positive and false negative rates of the computations in univariate methods.

The singular value for this was 27.51, accounting for 60.88% of the summed squared cross-block correlation. The saliences for the two task contrasts were 0.445 for the comparison of the average of encoding and recognition versus perceptual matching and 0.895 for the comparison of encoding versus recognition. Both contrasts are salient on this SI, but the contrast between encoding and recognition is dominant. The contrast that actually predicts this first SI can, in fact, be computed from the saliences: it is $0.445(-1, 2, -1) + 0.895(1, 0, -1) = (0.45, 0.891, -1.341)$, which is partly a contrast of encoding against recognition.

The plot of scores for the first SI (bottom of Fig. 2) shows the clustering of scores in encoding and matching apart from recognition. The multiple regression on

TABLE 1

Local Maxima in Areas of Significant Relative Activation and Deactivation in SPM Comparison of Encoding and Retrieval Conditions with Perceptual Matching Task ($P < 0.001$; Adapted from Haxby *et al.*, 1996)

Location	*X*	*Y*	*Z*	Comparison
L inf frontal	−33	27	−12	Enc vs match
L mid frontal	−26	42	16	Enc vs match
L mid frontal	−20	36	32	Enc vs match
R hippocampus	28	−10	−20	Enc vs match
R hippocampus	34	−24	−12	Enc vs match
L inf temporal	−56	−34	−12	Enc vs match
Medial occipital	−4	−86	12	Enc vs match[a]
R inf frontal	36	22	20	Rec vs match
R mid frontal	34	54	4	Rec vs match
Cingulate	−2	24	40	Rec vs match
L parietal	−38	−62	36	Rec vs match
R parietal	26	−64	28	Rec vs match
L ventral occipital	−36	−84	−12	Rec vs match
L lat cerebellum	−41	−63	−28	Rec vs match
R ventral occipital	24	−84	−8	Rec vs match
Cerebellum	14	−78	−20	Rec vs match
Medial occipital	−4	−86	12	Rec vs match[a]
L sup temporal	−46	−8	4	Rec vs match[a]

Note. Coordinates are in reference to the atlas of Talairach and Tourneaux (1988).

[a]Significant relative deactivation.

TABLE 2

Local Maxima in Areas Identified from PLS Analysis of Activation Data

Location	*X*	*Y*	*Z*	SI
R mid frontal	28	52	4	SI1+
R inf frontal	40	20	16	SI1+
Cingulate	2	22	32	SI1+
R ventral occipital	38	−84	4	SI1+
R cerebellum	18	−76	−20	SI1+
R thalamus	14	−14	4	SI1+
R parietal	30	−78	28	SI1+
L inf temporal	−44	−28	−24	SI1−
R inf temporal	44	−6	−24	SI1−
R hippocampus	30	−20	−12	SI1−
R parahipp gyrus	22	−18	−20	SI1−
Retrosplenial	28	−24	−2	SI1−
Cingulate	−4	−50	28	SI1−
Cingulate	2	40	−4	SI1−
R parahipp gyrus	32	−14	−20	SI2+
L inf temporal	−62	−34	−8	SI2+
L med frontal	−18	34	32	SI2+
L inf frontal	−42	16	16	SI2+
L inf frontal	−30	22	−8	SI2+
Med occipital	−4	−90	16	SI2−

Note. Singular image number (SI) is indicated in the right most column with +/− indicating whether the voxel showed a positive or negative salience.

the scores suggested that the first SI was extracting a significant task-related spatial pattern ($R^2 = 0.575$, permutation test probability $P \approx 0.0005$).

The second SI shows positive saliences mainly in dorsal medial occipital cortex (Fig. 3). Negative saliences are present in the right parahippocampal gyrus, more ventral and anterior than the location depicted in the first SI. Left inferior temporal and left prefrontal cortices are also represented strongly in this pattern. Saliences for the contrasts were 0.895 for contrast 1 and −0.445 for contrast 2. (The exact repetition of the absolute values for the two contrast saliences is because there are only two columns in the design matrix; this need not always be the case.) Applying these saliences to the original contrasts gives $0.895(-1, 2, -1) + -0.445(1, 0, -1) = (-1.341, 1.791, -0.45)$, which is mainly a contrast of encoding and matching. The singular value for the second SI was 22.05, accounting for the remaining 39.12% of the summed squared cross-block correlation.

The distribution of scores on the second SI shows matching separated from both encoding and recognition. Regression of the scores on the second SI on the contrast vectors was likewise significant ($R^2 = 0.289$, $P \approx 0.001$). (A reviewer has kindly pointed out to us that this second test is not conditioned to the "true" value of the first singular pair. It is thus somewhat biased toward acceptance of the null hypothesis. As that hypothesis has been rejected here, we need not pursue such a refinement.)

In summary, the PLS activation analysis shows that the dominant (first) pattern distinguished recognition of faces from encoding and face matching. The singular image incorporates positive saliences for posterior and ventral anterior cingulate cortices, anterior temporal cortices, and right hippocampus, and negative saliences for right prefrontal and dorsal anterior cingulate, ventral occipital and cerebellum, and thalamus. The scores are equal for encoding and face matching, suggesting that the areas identified in the first SI do not differentiate encoding and matching. The second SI distinguishes encoding from matching with positive saliences for dorsal occipital cortex and negative for ventral–anterior right parahippocampal gyrus and left prefrontal cortex. Scores in the recognition condition were most similar to those from encoding. In view of the similarity in scores for both memory conditions on the second SI, it is possible that these regions represent general memory operations. There have been suggestions that recognition of previously presented information requires reactivation of some of the same regions engaged in the initial encoding episode (Tulving and Thompson, 1973; Nyberg *et al.*, 1995). The PLS results are consistent with this possibility.

TABLE 3

Local Maxima in Areas Identified from PLS Analysis of Behavior/Brain Relations

Location	*X*	*Y*	*Z*	SI
Thalamus	−2	−2	12	SI1+
Medial occipital	−6	−98	4	SI1+
L mid frontal	−40	44	8	SI1+
R mid frontal	38	44	12	SI1+
L inf parietal	−28	−56	32	SI1+
R inf parietal	34	−58	32	SI1+
R fusiform	46	−12	−28	SI1−
R parahipp gyrus	26	−26	−20	SI1−
Cuneus/precuneus	−10	−60	28	SI1−
Cuneus	14	−74	28	SI1−
R caudate	26	16	0	SI2+
R parahipp gyrus	34	−14	−20	SI2−
R hippocampus	20	−14	−14	SI2−
L hippocampus	−32	−20	−16	SI2−
Cingulate	0	28	−8	SI2−
Cingulate	−4	50	16	SI2−
Cingulate	−16	48	20	SI2−

Note. Singular image number (SI) is indicated in the rightmost column with +/− indicating whether the voxel showed a positive or negative salience.

Brain–Behavior Analysis

The relation of RT to rCBF in the matching and recognition task complemented the preceding PLS analysis of experimental conditions. The maxima from this analysis are presented in Table 3. The dominant SI reflected common correlations in both conditions and is presented at the top of Fig. 4 ($R^2 = 0.75$, $P \approx 0.0005$). The singular value for this SI was 89.49, accounting for 72% of the summed squared cross-block correlation. Positive saliences represent activations associated with slower RT and negative saliences with faster RT. For the singular image, the strongest positive saliences are for the dorsomedial portion of thalamus and dorsomedial occipital cortex. The middle prefrontal gyrus is also represented bilaterally for the first SI, as is the inferior parietal. But there is an asymmetry of saliences for the superior temporal gyrus, with left greater than right. Negative saliences for the first SI are predominantly in the right hemisphere, with the strongest in the anterior portion of the fusiform, which extends more medially to include the right parahippocampal gyrus. The dorsal extent of the relation is mainly in the right medial parietal region.

The second SI (Fig. 4, bottom) represents the interaction of the RT–rCBF relation and task ($R^2 = 0.50$, $P = 0.006$; singular value = 55.76, 28% of cross-block correlation). Few areas showed positive saliences (activity associated with slower RT), most notably the right caudate nucleus. The greatest task differences appeared in negative saliences with the image encompassing the right and left hippocampal and parahippocam-

pal gyri and a large extent of anterior cingulate cortex. The relation of this singular image to the scores in matching and recognition suggested the areas identified in the negative pattern showed stronger negative correlations between rCBF and RT in recognition. For example, the correlation of the most salient voxel from the second SI located in the right hippocampus and RT in recognition was -0.91; in matching, -0.67. A concrete interpretation of this would be that subjects who activated these areas more were able to respond more rapidly. Some of the regions having strong saliences in this pattern are similar to those identified in the second SI from the analysis of the task effects. This is not unexpected since the interaction vector is a combination of task differences and RT. If the second SI from the activation analysis reflects general memory operations, then the RT–rCBF correlation suggests that subjects who respond more rapidly are able to more fully activate the regions engaged in these memory operations.

7. DISCUSSION

This paper has demonstrated how partial least squares methods of singular value decomposition can be used to describe the relation between brain activity and experimental design or behavior measures. The analysis involves three steps: computation of the cross-correlation matrix S between images and the design or behavior block, singular value decomposition of S, and derivation of scores.

While the results from the PLS analysis of brain activations were not completely dissimilar from those obtained through SPM (a typical univariate voxelwise method), limbic and thalamic regions of salience were identified by PLS and not by SPM. Liberalizing the criteria for significance in SPM would not restore that equivalence of the analyses. In PLS, associations of the image with its causes or effects are treated as a whole, consistent with our scientific stance of regarding brain operations as outcomes of cooperative interactions among areas (McIntosh and Gonzalez-Lima, 1994). At the same time, the permutation tests underlying significance here avoid problems of multiple comparisons that bedevil other approaches to the image as a whole such as the eigenimage analysis. To evade these problems with eigenimages seems to require that the feature space of image variability be restricted prior to consideration of interesting correlations or comparisons.

A second important difference between the typical univariate analysis and PLS is in the analytic approach. The most common practice in univariate image analysis, as epitomized by the cognitive subtraction paradigm, is to examine differences between tasks. PLS analysis, by being able to use all conditions in an experiment at once, provides indices of similarity as well as indices of differences. In the first application here, the first singular image extracted from the activation PLS analysis suggested that encoding and matching conditions were most similar on a singular image that included a large extent of limbic cortex, while the second singular image identified a different limbic region that distinguished encoding and matching. Such results have some important implications for data interpretation as it provides another dimension of theoretical examination—namely the commonalities in neural systems across different paradigms.

PLS is not designed to supplant either the univariate image analysis methods, such as SPM, or the extant multivariate methods, such as eigenimage analysis. Rather, these three classes of analysis are complementary and address somewhat different explanatory tasks. The scientist choosing among them must balance the goal of optimal localization *within* the image with that of optimal extraction of information *from* the image for subsequent analytic steps. The univariate methods emphasize the careful weighing of contrasts involving specific regions, contrasts that distinguish different cognitive processes. The multivariate methods, on the other hand, identify distributed systems of pixels that, *as a whole,* concentrate covariation with the causes or effects of these activations. Of these multivariate methods, PLS is unusual in its choice of a figure of merit for this assessment (covariance, not correlation); but, in exchange, it offers an analogue of the univariate "detection of regions" in the pixel-by-pixel coefficients of salience, which can be tested for localization even though they have not been computed independently.

In our view, the principal contribution of PLS, apart from its computational simplicity, is to point out the possibility of a third form of explanation, just as cogent as the styles of explanation underlying univariate or eigenimage analysis. In doing so, it also emphasizes how answers to questions about functional images depend rather sensitively on the way in which the questions are phrased. PLS, for instance, splits the question about the "significance" of an image into two pieces: the significance of the contrast as a whole (the number reported here as R^2) and the significance of any possible localization of that contrast. Most univariate analysis methods have no equivalent of this first significance test (but see Worsley *et al.,* 1995), while somewhat improving the power by which the second is tested. The eigenimage methods, conversely, emphasize those image regions that correlate with other regions likewise salient among themselves, without separating out the part that is salient for a particular cross-block prediction task. PLS bridges this divide, combining both explanatory purposes. It thus circumvents the troublesome tensions in neuroscience between localized and holistic modes of explanation. PLS is neutral in this dispute, proffering methods by which

one can weigh either extreme postulate against the other. Indeed a recent study by Nyberg *et al.* (1996) demonstrated how a combination of univariate correlation analysis and PLS identifies a central role of the medial temporal lobe in a distributed system underlying episodic memory retrieval.

PLS analyses of the relation of brain activity to behavioral measures and task-related changes in the relation provide an important complement to conventional image analysis. In this manner, PLS can be used to examine relations across a number of blocks. Covariances may pertain to any set of behavioral measures. PLS can be extended to sets of demographic or psychological measures or to measures of neuroanatomy (e.g., structural MRI); in other versions of PLS, there can be multiple blocks of X's all competing to account for patterns of covariance with the same image data. Across this whole class of analyses, scientific understanding arises precisely when the saliences and the scores fall into patterns consistent with the causal nexus presumed to underlie the data that were measured.

The analysis here treats the experimental design as consisting only of two dimensions of "contrast." This omits two other aspects that are often included here: a set of subject effects (contrasts across subjects) and a set of run effects (contrasts across repetitions of the task). In principle, each of these constitutes an additional design block that could be wielded to produce singular images comparable to those here. There is no reason to believe that those images would be the same as those here, and thus it may be unwise to combine the task contrasts and the repeated-measures contrasts into a single enlarged design matrix a priori. Interindividual differences would be of importance if the study sample consisted of two groups (e.g., patients versus controls); run effects for tasks that show a strong component of learning. If one or the other of these facets is a determinant of image contents commensurate with the experimental design, it can be accounted for using PLS. The columns in the design matrix may be appended with columns for contrasts among subgroups of subjects, for run effects, and, if necessary, combinations of task, subjects, and run, just as in the setup of general crossed MANOVAs. The singular images from this PLS analysis are compared to those from a simpler analysis where only task effects are considered. If the singular images do not differ, then the task effects are robust and the added information provided using subject and run effects may serve to enhance the cross-block prediction of task effects. If they differ, then the data may be adjusted for the effects from subject or run, either by preprocessing steps (cf. Moeller *et al.*, 1987; Alexander and Moeller, 1994) or by adjusting out singular images representing such effects.

Three other image analysis methods have been proposed that attempt to extract images related to either experimental manipulations or experimental groups. Scaled subprofile modeling (SSM; Moeller *et al.*, 1987; Moeller and Strother, 1991; Alexander and Moeller, 1994) performs a principal components analysis on a covariance matrix obtained between brain regions across a sample of subjects containing both normal controls and patients. Before the covariance matrix is computed, the raw data are normalized by log transform and subtraction of regional and subject means. The extracted components represent eigenimages that show the most variation after the normalizations. Component scores are computed for all subjects to show how patients and controls relate to the particular eigenimage. Although they are not formally optimal for any discriminatory purpose, the eigenimages can be powerful aids in description of group differences.

The eigenimage analysis of Friston *et al.* (1993; also Friston, 1994) differs from that of SSM in that subject variance is completely removed by averaging and the analysis is performed across tasks. The goal of the analysis here is to identify those brain regions most related to the experimental condition. This is not too dissimilar from the PLS analysis of activations except that PLS makes use of subject variance as well as experimental variance. PLS identifies those regions most related to the experimental manipulation and derives scores within each condition to give an indication of how each subject relates to the manipulation (i.e., how "typical" that subject's brain activity was in comparison with the rest of the sample). In this way, PLS can be thought of as a combination of SSM and eigenimage analysis, with the important improvement that all the data can be examined simultaneously without exclusion of variance sources or additional transforms of the data.

The PLS analysis is quite different in principle from the analysis of eigenimages recently suggested by Friston *et al.* (1995). These are restricted to the subspace of the brain image block Y spanned by its largest principal components (eigenimages of the experiment irrespective of design). Vectors of this subspace are regressed on the elements of the design matrix by multiple regression (realized as a canonical variates analysis). The results do not agree with those of the PLS analysis here, even in perfectly balanced crossed designs, owing to the reduction of the image block to eigenimages before the multiple regression procedure. Such a filter would be appropriate if the confrontation of explained and observed images is by the MANOVA used by Friston *et al.*, which implicitly incorporates a multiple regression step on both sides; but in partial least squares there is no multiple regression at any time—only decomposition of one single covariance matrix. What is optimally accounted for in PLS are the elements of this matrix, not the values of the individual

pixels at any point of the image. The purpose of PLS is to optimize information between the blocks for scientific understanding, not to predict either set of values by the other.

The interest in identifying spatial patterns across the brain that relate to the experimental task or behavior is in part a reflection of the recent emphasis on interregional interactions and how these interactions change across tasks of an experiment. In this context, PLS can play an important role in identifying the key regions that form the nodes of a functional network differentially engaged across the experiment. The capacity for the examination of brain–behavior relations in PLS helps to identify portions of the functional networks that are most involved in organization of neural processes in response selection. Enriching the toolkit of image analysis by a least-squares multivariate technique such as PLS will help researchers extract more information from any functional neuroimaging data set.

APPENDIX—WORKING EXAMPLE

For the simulation, assume that measures from a four-pixel image (pixels Y_1 to Y_4) were obtained from three conditions. The data set was created using the random-number functions in MATLAB with the pixel values sampled from a population of normally distributed random numbers (mean = 10, SD = 1.5). Values were grouped into three "conditions" of five observations each. A constant, representing an effect, was added to certain pixels within a condition, plus uncorrelated error variance. The data are as follows:

	Data			
Condition	Y_1	Y_2	Y_3	Y_4
1	9.49	7.75	11.93	17.31
1	9.40	10.55	9.54	17.06
1	7.71	10.06	10.78	19.84
1	7.46	11.25	12.05	19.00
1	8.53	10.69	11.41	19.91
2	11.14	10.45	11.80	18.80
2	10.56	10.54	14.36	20.89
2	12.08	9.95	14.86	22.20
2	10.39	11.52	14.97	22.72
2	9.15	9.61	15.26	21.25
3	9.96	10.41	8.41	17.13
3	10.18	12.08	7.93	14.13
3	10.74	9.94	8.19	16.32
3	9.66	10.56	6.90	15.73
3	10.66	9.88	7.61	15.21
Mean 1	8.52	10.06	11.14	18.62
Mean 2	10.66	10.41	14.25	21.17
Mean 3	10.24	10.57	7.81	15.70

For the PLS analysis, assemble orthogonal contrast vectors to form a design matrix X:

	Design matrix (X)														
Contrast 1	2	2	2	2	2	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1
Contrast 2	0	0	0	0	0	1	1	1	1	1	−1	−1	−1	−1	−1

The cross correlations between design and data matrices are:

	Cross correlations (S)	
	Contrast 1	Contrast 2
Y_1	−0.7555	0.1432
Y_2	−0.217	−0.0692
Y_3	0.019	0.9403
Y_4	0.035	0.8796

The cross-correlation matrix is then decomposed using singular value decomposition ($S = \mathbf{ADB}^t$). The resulting solution yields two pairs of singular vectors (singular images, SI), because the smallest dimension in the cross-correlation matrix is 2.

The singular values, representing the covariance between SI pairs, for the solution are:

Singular values (D)	
SI1	SI2
1.2981	0.7858

The sum of the squared singular values is equal to the sum of squared cross correlations (2.30). The ratio of a squared singular value to this sum, is "the proportion of summed squared cross-block correlation accounted for" by the SI. For SI1, this is 0.73 (1.68/2.30), and for SI2, is 0.27 (0.62/2.30).

Two matrices are also obtained from the solution:

Singular vectors for design matrix (A)		
Contrast	SI1	SI2
1	0.0417	0.9991
2	−0.9991	0.0417

This states that the second contrast (condition 2 versus 3) is most related to SI1, and the first contrast (condition 1 versus 2 and 3) is more related to SI2.

The saliences for the data matrix indicate which variables are most related to each SI.

Saliences for data matrix (B)

Variable	SI1	SI2
Y_1	−0.1345	−0.9529
Y_2	0.0463	−0.2795
Y_3	−0.7231	0.0741
Y_4	−0.6759	0.0912

The saliences for SI1 show that Y_3 and Y_4 are most related to SI1. Since this SI also is related to the contrast of conditions 2 and 3, the interpretation is that these are the pixels that differ most between those conditions. SI2 has high weights for Y_1, this is the pixel most different between condition 1 and conditions 2 and 3.

The scores give an indication of the conditions that are being distinguished for a given SI.

	Scores	
Condition	SI1	SI2
1	−21.2465	−8.7456
1	−19.206	−9.6381
1	−21.7726	−7.5481
1	−22.0385	−7.6252
1	−22.3599	−8.4531
2	−22.2494	−10.9465
2	−25.4334	−10.043
2	−26.9216	−11.1651
2	−27.0417	−9.9361
2	−26.1812	−8.3368
3	−18.5139	−10.2132
3	−16.0927	−11.2029
3	−17.9365	−10.9182
3	−16.4309	−10.2113
3	−16.7595	−10.9665
Mean 1	−21.3247	−8.40202
Mean 2	−25.5655	−10.0855
Mean 3	−17.1467	−10.7024

The scores on the first SI (with high saliences for Y_3 and Y_4) show an even gradient across conditions, consistent with the raw means in the data matrix table. The scores on the second SI will be substantially identical to the values for pixel Y_1, which shows the greatest distinction between condition 1 and conditions 2 and 3.

ACKNOWLEDGMENTS

The authors thank Dr. N. J. Lobaugh for her valuable input in the development of the manuscript, and Drs. E. Tulving, L. Nyberg, M. Furey, and J. D. Van Horn for their comments on earlier drafts. This work was partially supported by the Natural Sciences and Engineering Research Council of Canada to A. R. McIntosh and by USPHS Grant DA-09009 to F. L. Bookstein. The latter grant is jointly sponsored by the National Institute on Drug Abuse, the National Institute of Mental Health, and the National Institute on Aging as part of the Human Brain Project.

REFERENCES

Aertsen, A. M. H., Gerstein, G. L., Habib, M. K., and Palm, G. 1989. Dynamics of neuronal firing correlation: Modulation of "effective connectivity." *J. Neurophysiol.* **61:**900–917.

Alexander, G. E., and Moeller, J. R. 1994. Application of the scaled subprofile model to functional imaging in neuropsychiatric disorders: A principal component approach to modeling brain function in disease. *Hum. Brain Map.* **2**(1, 2):79–94.

Bookstein, F. L. 1991. *Morphometric Tools for Landmark Data: Geometry and Biology.* Cambridge Univ. Press, Cambridge, UK.

Bookstein, F. L. 1994. Partial least squares: A dose–response model for measurement in the behavioral and brain sciences. [Revised] *Psycoloquy* [an electronic journal] **5**(23), 4/6/1994.

Bookstein, F. L. 1996. Biometrics and brain maps: The promise of the morphometric synthesis. To appear in *Neuroinformatics: An Overview of the Human Brain Project.* (S. Koslow and M. Heurta, Eds.), Progress in Neuroinformatics, Vol 1. L. Erlbaum, Hillsdale, NJ.

Bookstein, F. L., Sampson, P. D., Steissguth, A. P., and Barr, H. M. 1990. Measuring 'dose' and 'response' with multivariate data using partial least squares techniques. *Commun. Stat. Theor. Methods* **19:**765–804.

Bookstein, F. L., Streissguth, A., Sampson, P., and Barr, H. 1996. Exploiting redundant measurement of dose and behavioural outcome: New methods from the teratology of alcohol. *Dev. Psychol.* **32:**404–415.

Edgington, E. S. 1980. *Randomization Tests.* Dekker, New York.

Efron, B., and Tibshirani, R. 1986. Bootstrap methods for standard errors, confidence intervals and other measures of statistical accuracy. *Stat. Sci.* **1:**54–77.

Friston, K. J. 1994. Functional and effective connectivity: A synthesis. *Hum. Brain Map.* **2**(1, 2):56–78.

Friston, K. J., Passingham, R. E., Nutt, J. G., Heather, J. D., Sawle, G. V., and Fracowiak, R. S. J. 1989. Localization in PET images: Direct fitting of the intercommisural (AC-PC) line. *J. Cereb. Blood Flow Metab.* **9:**690–695.

Friston, K. J., Frith, C. D., Liddle, P. F., and Fracowiak, R. S. J. 1991. Plastic transform of PET images. *J. Comput. Assisted Tomogr.* **15:**634–639.

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.

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

Gerstein, G. L., Perkel, D. H., and Subramanian, K. N. 1978. Identification of functionally related neural assemblies. *Brain Res.* **140:**43–62.

Good, P. 1994. *Permutation Tests.* Springer-Verlag, New York.

Grady, C. L., McIntosh, A. R., Horwitz, B., Maisog, J. M., Ungerleider, L. G., Mentis, M. J., Pietrini, P., Schapiro, M. B., and Haxby, J. V. 1995. Age-related reductions in human recognition memory involve altered cortical activation during encoding. *Science* **269:**218–220.

Haxby, J. V., Ungerleider, L. G., Horwitz, B., Maisog, J. M., Rapoport, S. I., and Grady, C. L. 1996. Face encoding and recognition in the human brain. *Proc. Natl. Acad. Sci. USA* **93:**922–927.

Holmes, A. P., Blair, R. C., Watson, J. D. G., and Ford, I. 1996.

Nonparametric analysis of statistical images from functional mapping experiments. *J. Cereb. Blood Flow Metab.* **167:**7–22.

Mardia, K. V., Kent, J. T., and Bibby, J. M. 1979. *Multivariate Analysis*. Academic Press, New York.

McIntosh, A. R., and Gonzalez-Lima, F. 1994. Structural equation modeling and its application to network analysis of functional brain imaging. *Hum. Brain Map.* **2**(1, 2):2–22.

McIntosh, A. R., Grady, C. L., Ungerleider, L. G., Haxby, J. V., Rapoport, S. I., and Horwitz, B. 1994. Network analysis of cortical visual pathways mapped with PET. *J. Neurosci.* **14**(2):655–666.

Moeller, J. R., and Strother, S. C. 1991. A regional covariance approach to the analysis of functional patterns in positron emission tomographic data. *J. Cereb. Blood Flow Metab.* **11:**A121–A135.

Moeller, J. R., Strother, S. C., Sidtis, J. J., and Rottenberg, D. A. 1987. Scaled subprofile model: A statistical approach to the analysis of functional patterns in emission tomographic data. *J. Cereb. Blood Flow Metab.* **7:**649–658.

Nyberg, L., Tulving, E., Habib, R., Nilsson, L.-G., Kapur, S., Houle, S., Cabeza, R., and McIntosh, A. R. 1995. Functional brain maps of retrieval mode and recovery of episodic information. *Neuroreport* **7:**249–252.

Nyberg, L., McIntosh, A. R., Houle, S., Nilsson, L.-G., and Tulving, E. 1996. Relation of hippocampal and medial temporal activity with episodic memory retrieval in individual subjects. *Nature* **380:**715–717.

Pedhazur, E. J. 1982. *Multiple Regression in Behavioural Research: Explanation and Prediction*, 2nd ed. Holt, Reinhart & Winston, New York.

Sampson, P. D., Streissguth, P., Barr, H. M., and Bookstein, F. L. 1989. Neurobehavioural effects of prenatal alcohol. Part II. Partial least squares analysis. *Neurotoxicol. Teratol.* **11:**477–491.

Streissguth, P., Bookstein, F. L., Sampson, P. D., and Barr, H. M. 1993. *The Enduring Effects of Prenatal Alcohol Exposure on Child Development: Birth through 7 Years, a Partial Least Squares Solution*. Univ. Michigan Press, Ann Arbor.

Talairach, J., and Tourneaux, P. 1988. *Co-planar Stereotaxic Atlas of the Human Brain*. Thieme, New York.

Tulving, E., and Thomson, D. M. 1973. Encoding specificity and retrieval processes in episodic memory. *Psychol. Rev.* **80**(5):359–380.

Worsley, K. J., Poline, J.-B., Vandal, A. C., and Friston, K. J. 1995. Tests for distributed, nonfocal brain activations. *NeuroImage* **2:**183–194.