

Partial least squares analysis of neuroimaging data: applications and advances

Anthony Randal McIntosh^{a,*} and Nancy J. Lobaugh^b

^aRotman Research Institute of Baycrest Centre, University of Toronto, Toronto, Ontario, Canada M6A 2E1

^bSunnybrook and Women's College Health Sciences Centre, University of Toronto, Toronto, Ontario, Canada M4N 3M5

Available online 11 September 2004

Partial least squares (PLS) analysis has been used to characterize distributed signals measured by neuroimaging methods like positron emission tomography (PET), functional magnetic resonance imaging (fMRI), event-related potentials (ERP) and magnetoencephalography (MEG). In the application to PET, it has been used to extract activity patterns differentiating cognitive tasks, patterns relating distributed activity to behavior, and to describe large-scale interregional interactions or functional connections. This paper reviews the more recent extension of PLS to the analysis of spatiotemporal patterns present in fMRI, ERP, and MEG data. We present a basic mathematical description of PLS and discuss the statistical assessment using permutation testing and bootstrap resampling. These two resampling methods provide complementary information of the statistical strength of the extracted activity patterns (permutation test) and the reliability of regional contributions to the patterns (bootstrap resampling). Simulated ERP data are used to guide the basic interpretation of spatiotemporal PLS results, and examples from empirical ERP and fMRI data sets are used for further illustration. We conclude with a discussion of some caveats in the use of PLS, including nonlinearities, nonorthogonality, and interpretation difficulties. We further discuss its role as an important tool in a pluralistic analytic approach to neuroimaging.

© 2004 Elsevier Inc. All rights reserved.

Keywords: Multivariate statistics; Functional MRI; Event-related potentials; Positron emission tomography; Magnetoencephalography; Nonparametric statistics

Introduction

As the use of neuroimaging to study behavioral and cognitive processes has grown, so too has the development of techniques optimized to extract signals from these data. Indeed, neuroimaging analysis has become a field of research in its own right, as evidenced by this special issue of *NeuroImage*.

* Corresponding author. Rotman Research Institute of Baycrest Centre, University of Toronto, 3560 Bathurst Street, Toronto, Ontario, Canada M6A 2E1. Fax: +1 416 785 2862.

E-mail address: mcintosh@psych.utoronto.ca (A.R. McIntosh).

Available online on ScienceDirect (www.sciencedirect.com).

1053-8119/\$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2004.07.020

One class of neuroimaging analysis methods has focused on identification of reliable signal changes at the level of individual image elements (i.e., voxels, or in the case of electromagnetic imaging, sensors or timepoints). These are referred to as univariate statistical analyses. A second family of methods has focused on the examination of distributed patterns. These methods, called multivariate statistical analyses, take advantage of the spatial and temporal dependencies among image elements. While the difference between univariate and multivariate methods can be linked to a philosophical difference in assumptions related to the underlying nature of brain organization, for the present purposes, they may be considered to support different levels of inferences about neural signals. Univariate methods are optimal for identification of signal changes at specific points in time or space, while multivariate methods enable inferences about differences across space and/or time by combining information across these dimensions.

Partial least squares (PLS), which was introduced to the neuroimaging community in 1996 (McIntosh et al., 1996), has proven to be a robust method for extracting distributed signal changes related to changing task demands (Task PLS). It has also been applied to measuring distributed patterns that impact on task performance (Behavior PLS) and finally to task-dependent changes in the relation between brain regions. This latter application is an assessment of functional connectivity or the correlation between neural elements (Seed PLS).

The present paper serves to briefly review the development of PLS for neuroimaging data and the recent extension of the method to simultaneous spatial and temporal analysis (Duzel et al., 2003; Hay et al., 2002; Itier et al., 2004; Lobaugh et al., 2001; Lobaugh et al., submitted for publication; McIntosh et al., in press). We present both the theoretical derivation of spatiotemporal PLS and its application to functional MRI and event-related potential (ERP) data.

Method and application

Partial least squares

The term “partial least squares” refers to the computation of the optimal *least-squares* fit to *part* of a correlation or covariance

matrix (Wold, 1982). The part is the “cross-block” correlation between some set of exogenous and dependent measures. PLS is similar to principal components analysis (PCA), but one important feature of PLS is that the solutions are constrained to the part of the covariance structure that is attributable to experimental manipulations or that relates to behavior. Moreover, PLS is ideal for data sets where the dependent measures within a block are highly correlated (e.g., neuroimaging data) because items within a block are not adjusted for these correlations (c.f., canonical correlation).

The present implementation of PLS is written entirely in Matlab (Mathworks Inc). For positron emission tomography (PET) and functional magnetic resonance imaging (fMRI), it is compatible with Analyze format images (Mayo Clinic). For ERP and magnetoencephalography (MEG) data, it will work with delimited text files from any system, as well as Neuroscan “avg” files. The user has the choice of three layout options for displaying the results for up to 150 sensors: BESA theta/phi coordinates for 10/20 and 10/10 ERP systems; EGI 128 channel systems; and CTF MEG systems. The code is available at <http://www.rotman-baycrest.on.ca:8080>.

Data organization. The recent extension of PLS to the temporal domain (spatiotemporal PLS, ST-PLS) allows for the same mathematical formulation to be applied to fMRI, ERP, and MEG data, where there is meaningful information in a measured time series. For the purposes of the explanation that follows, we will use the term “element” to refer to voxels, electrodes, and MEG sensors.

As with all multivariate approaches, PLS operates on the entire data structure at once, which requires that the data be in matrix form. Every row of the matrix contains data for one observation in one condition. An observation may be a single subject in the case of a group study or a trial within a condition in the case of single subject analysis. These rows are arranged such that observations are nested within condition blocks. With n observations and k conditions, there are $n * k$ rows in the matrix. The columns of the data matrix contain the signal measured for each element at each timepoint. The first column has intensity for the first element at the first timepoint, and the second column has the intensity for the first element at the second timepoint. With m elements and t timepoints, there are $m * t$ columns in the matrix.

Mathematical description of task ST-PLS analysis. As indicated above, the data matrix contains the m elements, measured across t timepoints, which are made into single vectors for each of n observations measured in each of k conditions. Thus, the data matrix \mathbf{M} has $n * k$ rows and $m * t$ columns.

Two versions of PLS have been used in the literature: one using orthonormal contrasts and the other using mean centering. These two approaches give virtually identical results when scaling differences (see below) are taken into account.

Contrast approach. A design matrix \mathbf{C} is constructed with orthonormal contrasts, coding for the $k-1$ degrees of freedom in the experimental design. The contrasts are made for each subject so \mathbf{C} has $n * k$ rows and $k-1$ columns. The operation:

$$\mathbf{Y} = \mathbf{C}^T * (\mathbf{M} - \mathbf{I} * (\mathbf{I}^T * \mathbf{M}) / (n * k)) / (n * k - 1) \quad (1)$$

yields a $k-1 \times m * t$ matrix \mathbf{Y} containing the covariance of each timepoint for each element with each contrast in \mathbf{C} (superscript T represents a matrix transpose, and $\mathbf{1}$ is a vector of ones $n * k$ in length).

\mathbf{Y} is then subjected to a singular value decomposition (SVD):

$$[\mathbf{USV}] = SVD[\mathbf{Y}^T] \quad (2)$$

where:

$$\mathbf{U} * \mathbf{S} * \mathbf{V}^T = [\mathbf{Y}^T] \quad (3)$$

From the decomposition, \mathbf{U} is an $m * t \times k-1$ orthonormal matrix containing the element saliences, \mathbf{V} is an $k-1 \times k-1$ orthonormal matrix of saliences for the contrasts, and \mathbf{S} is a diagonal matrix of the $k-1$ nonzero singular values.

The multiplication:

$$\mathbf{B} = \mathbf{M} * \mathbf{U} \quad (4)$$

produces \mathbf{B} , an $n * k \times k-1$ matrix of *Brain Scores* that indicate the variation of task effects across subjects and conditions. The dot product computed at each timepoint, rather than summing across time, provides a brain score for each timepoint, or a *Temporal Score*, for each subject and condition. This provides a rapid assessment of where, in a temporal sense, the major distinctions among tasks are expressed.

The multiplication:

$$\mathbf{D} = \mathbf{C} * \mathbf{V} \quad (5)$$

produces \mathbf{D} , an $n * t \times k-1$ matrix of *Design Scores* that are the set of contrasts that code the effects resulting from the SVD.

Mean-centering approach. Starting with the same data matrix \mathbf{M} , described above, columnwise averages are created within each task, yielding matrix \mathbf{T} , a k by $m * t$ matrix of task means.

$$\mathbf{T}_{dev} = \mathbf{T} - \mathbf{1} * (\mathbf{1}^T * \mathbf{T}) / k \quad (6)$$

Where matrix $\mathbf{1}$ is a column vector of ones of length k . \mathbf{T}_{dev} is a columnwise mean-centered matrix. The operation:

$$[\mathbf{U}_{dev} \mathbf{S}_{dev} \mathbf{V}_{dev}] = SVD[\mathbf{T}_{dev}^T] \quad (7)$$

where:

$$\mathbf{U}_{dev} * \mathbf{S}_{dev} * \mathbf{V}_{dev}^T = [\mathbf{T}_{dev}^T] \quad (8)$$

provides \mathbf{U}_{dev} , an $m * t \times k$ orthonormal matrix containing the element saliences, \mathbf{V}_{dev} an $k \times k$ orthonormal matrix of task saliences, and \mathbf{S}_{dev} a diagonal matrix of the k singular values. In this case, the final diagonal element of \mathbf{S}_{dev} is zero, representing the grand mean, which is eliminated through mean centering.

As stated before, the contrast and deviation methods produce identical analytic results save for scaling differences. For example, when the first $k-1$ columns of \mathbf{U}_{dev} are used:

$$\mathbf{U}^T * \mathbf{U}_{dev} = \mathbf{I} \quad (9)$$

where \mathbf{I} is an identity matrix.

The contrasts in \mathbf{D} (the design score matrix) are identical to the pattern in \mathbf{V}_{dev} . The scaling differences are found in the values of \mathbf{S} versus \mathbf{S}_{dev} , because of the difference in the values of \mathbf{Y} versus \mathbf{T}_{dev} . If the raw cross-product, rather than the covariance, is used with the design contrast method, the values in the two singular value matrices are the same.

Behavior ST-PLS. The analysis of brain-behavior correlations starts with the same matrix \mathbf{M} used for task ST-PLS, and a second matrix \mathbf{A} containing the behavioral measures of interest. These could be

reaction time or accuracy measures acquired during the experiment or any other measures of interest (e.g., physiological measures, demographics). For simplicity, we assume there are behavior measures for each of the k conditions in the experiment. Matrix \mathbf{A} may contain any number of columns, although we assume a single measure for this illustration (i.e., \mathbf{A} is $n * k \times I$).

Within each of k conditions, the columns of \mathbf{A} and \mathbf{M} are z score-transformed such that:

$$\mathbf{R}_k = \mathbf{A}_k^T * \mathbf{M}_k / (n_k - I) \quad (10)$$

where \mathbf{A}_k and \mathbf{M}_k are the behavior measures and element time series measures during condition k , respectively. Matrix \mathbf{R}_k is the correlation of the element time series across subjects or trials with the corresponding behavioral measures for condition k . These matrices are stacked into one large matrix \mathbf{R} and subjected to SVD:

$$[\mathbf{U}_R \mathbf{S}_R \mathbf{V}_R] = SVD[\mathbf{R}^T] \quad (11)$$

where \mathbf{U}_R is an $m * t \times k$ orthonormal matrix containing the element saliences, \mathbf{V}_R is a $k \times k$ orthonormal matrix of task saliences, and \mathbf{S}_R is a diagonal matrix of the k singular values. As with the deviation analysis in the task ST-PLS, the weights in \mathbf{V}_R are interpreted as contrasts, indicating task-dependent differences in brain-behavior correlations. In most cases, one column of \mathbf{V}_R will contain the same values for all conditions, thus identifying patterns of brain-behavior correlations that are similar across conditions.

Brain scores (matrix \mathbf{B}) are computed as in Eq. (4) above. Calculation of the correlation between the scores within condition k and the corresponding behavior measure for each column of \mathbf{B} yields values with a pattern similar to those in \mathbf{V}_R . Confidence intervals calculated around these latent variable correlation profiles can provide a more conventional assessment of the reliability of brain-behavior correlations (see below).

Assessment of significance. The arbitrary decisions regarding the number of LVs to retain (e.g., scree plots, percentage of variance accounted for) and which of the task and element weights to consider important are minimized by providing a statistical assessment of the LVs. This is done using permutation tests for the LVs and bootstrap estimation of standard errors for the element saliences. The permutation test assesses whether the effect represented in a given LV is sufficiently strong, in a statistical sense, to be different from random noise. The standard error estimates of the saliences from the bootstrap tests are used to assess the reliability of the nonzero saliences on significant LVs. The resampling across different levels may approximate a mixed model design, yielding optimal sensitivity and level of inference, without increased false-positive rates, although the exact mathematical description of this relation requires further clarification (Strother et al., 2002).

Statistical significance of each LV is assessed by means of a permutation test using 500 permutations (Edgington, 1980; Good, 2000; McIntosh et al., 1996; Nichols and Holmes, 2002). This is accomplished using sampling without replacement to reassign the order of conditions for each observation. ST-PLS is recalculated for each new sample, and the number of times the permuted singular values exceed the observed singular values is calculated. Exact probabilities are presented for all LVs, providing an objective means for determining the number of LVs to be retained.

An important requirement for the use of permutation testing is the exchangeability of data between rows. By folding time and space into the same dimension of the data matrix, the exchangeability

requirement is met. Had the matrix maintained a three-dimensional form and the permutations extended into the time domain, exchangeability would not have been met because the temporal correlation in the element signal would not be maintained.

To determine the stability of the maximal element saliences identified on the LVs, the standard errors of the saliences are estimated through 100 bootstrap samples (Efron and Tibshirani, 1986). Bootstrap samples are generated using sampling with replacement, keeping the assignment of experimental conditions fixed for all observations. ST-PLS is recalculated for each bootstrap sample. A salience whose value depends greatly on which observations are in the sample is less precise than one that remains stable regardless of the sample chosen (Sampson et al., 1989). The ratio of the salience to the bootstrap standard error is approximately equivalent to a z score if the bootstrap distribution is normal (Efron and Tibshirani, 1986). The primary purpose of the bootstrap is to determine the elements whose responses show reliable experimental effects, thus no corrections for multiple comparisons are necessary since no statistical test is performed. The statistical assessment is done through permutation tests, applied at the level of the full spatiotemporal pattern, as described above.

Bootstrap estimation also is used to derive confidence intervals for the latent variable correlations in the behavior ST-PLS. As the correlations are bounded (-1 to $+1$), the upper and lower percentiles of the bootstrap distribution are used to establish confidence limits (see fMRI example below).

To summarize, the *significance* of each LV is determined from permutation tests, and the *reliability* of the contribution of each nonzero element salience is then determined using bootstrap estimates of the salience standard errors.

During resampling with either permutations or bootstrapping, there is a possibility of axis rotation (a change in the order of extracted LVs) and reflection (a sign change in saliences) when the SVD is performed on the resampled matrix. These arbitrary reflections and rotations can be corrected using a Procrustes rotation of the resampled SVD outcome to the original SVD outcome (Milan and Whittaker, 1995). In the case of permuted data sets, the Procrustes rotation ensures that the assessment is done on the exact same effect space for each iteration. For bootstrapped data, the Procrustes rotation reduces the estimation bias because of reflection and rotation.

The Procrustes rotation is computed as follows (for complete explanation and proof, see Milan and Whittaker, 1995): given \mathbf{V}_{orig} , the original matrix \mathbf{V} from the SVD step in a PLS analysis (Eqs. (7) and (11)), and $\mathbf{V}_{\text{resamp}}$, matrix \mathbf{V} from the SVD of a resampled data set, the goal is to define a $k \times k$ orthogonal matrix \mathbf{Q} that rotates the points of $\mathbf{V}_{\text{resamp}}$ to a position as close as possible to \mathbf{V}_{orig} in a least squares sense. Assuming both \mathbf{V}_{orig} and $\mathbf{V}_{\text{resamp}}$ are orthonormal, then \mathbf{Q} can be calculated by:

$$[\mathbf{NOP}] = SVD(\mathbf{V}_{\text{orig}}^T * \mathbf{V}_{\text{resamp}}) \quad (12)$$

$$\mathbf{Q} = \mathbf{N} * \mathbf{P}^T \quad (13)$$

The rotation is then applied to both $\mathbf{V}_{\text{resamp}}$ and $\mathbf{U}_{\text{resamp}}$:

$$\hat{\mathbf{V}}_{\text{resamp}} = \mathbf{V}_{\text{resamp}} * \mathbf{S}_{\text{resamp}} * \mathbf{Q} \quad (14)$$

and

$$\hat{\mathbf{U}}_{\text{resamp}} = \mathbf{U}_{\text{resamp}} * \mathbf{S}_{\text{resamp}} * \mathbf{Q} \quad (15)$$

Rescaling $\mathbf{V}_{\text{resamp}}$ and $\mathbf{U}_{\text{resamp}}$ by $\mathbf{S}_{\text{resamp}}$ before rotation allows the redistribution of the covariance accounted for by each LV to be

carried through the transformation. The singular values of the rotated matrix can be obtained by calculating the columnwise square root of the sums-of-squares.

In the statistical assessment of PLS, it may seem that using the bootstrap procedure on each element is somewhat redundant. At first blush, it would seem reasonable to perform a permutation test at both the level of the singular values and individual voxels. For reasons described below, we have favored the use of bootstrap to estimate the reliability of each voxel's contribution to the pattern captured by a given LV. Permutation tests indicate whether a signal can be differentiated from noise but do not index signal reliability. Although detection and reliability are strongly related, they are not mutually exclusive.

Fig. 1 serves to further illustrate this point. In all three panels, the estimated voxel salience from a single LV from a PLS analysis of PET data is plotted on the x -axis. The y -axis is the bootstrap estimated standard error for each voxel. Voxels in the top panel (Fig. 1A) are circled if the estimated salience for that voxel was statistically different from noise using permutation tests (two-tailed $P < 0.01$), while voxels in the middle panel (Fig. 1B) are circled if the ratio of salience to standard error exceeded a threshold of 2.57, which has an approximate two-tailed probability of 0.01 assuming a unit normal distribution. In the bottom panel (Fig. 1C), we conducted a repeated-measures ANOVA using the same contrast derived by the PLS analysis to assign a parametric P value to the voxels in the image. As with the top panel, voxels are circled if the estimated P value was less than 0.01. The two plots based on voxelwise P values (Figs. 1A, C) are virtually identical, so we will treat them as the same for the present discussion.

While the overlap across plots is obvious, there are also unique features in each. For the permutation test, there are clusters of voxels that are uniquely identified as significant (circled in green, Fig. 1A). These voxels are not circled in the bootstrap plot because of their relatively large standard error. On the other hand, in the bootstrap plot, two clusters of voxels are uniquely circled in green (Fig. 1B). These not only have relatively low standard error, but also low saliences. Because the permutation test and ANOVA are most sensitive to the “strength” of a signal, voxels with large saliences are favored, whereas the bootstrap ratio will reflect both large salience and reliability. By emphasizing reliability, the bootstrap adds an important complement to the overall assessment of significance done in PLS, in that the technique allows one to disregard signals that are not reliable. Note that the reliability issue would be equally applicable to parametric tests, as the permutation test and ANOVA identified the same significant voxels. A second feature of bootstrap estimation is the ability to make inferences about voxels having low but reliable signals. In terms of neural processes, such voxels will contribute to the shaping of a distributed response. Put another way, voxels with the strongest signals could determine the overall direction of a distributed response, while lower intensity signals may “fine tune” the response. Another interesting consequence of using reliability estimates is the ability to make inferences about “stable” zero effects. If one follows the line of voxels whose salience is estimated as zero, there is a distinct continuum of standard errors for such voxels. Voxels with a zero estimate and a lower standard error are essentially stable at zero, indicating noninvolvement in the distributed pattern captured by PLS. Conversely, voxels with a zero estimate and large standard error would be considered unreliable. The ability to identify stable zeros

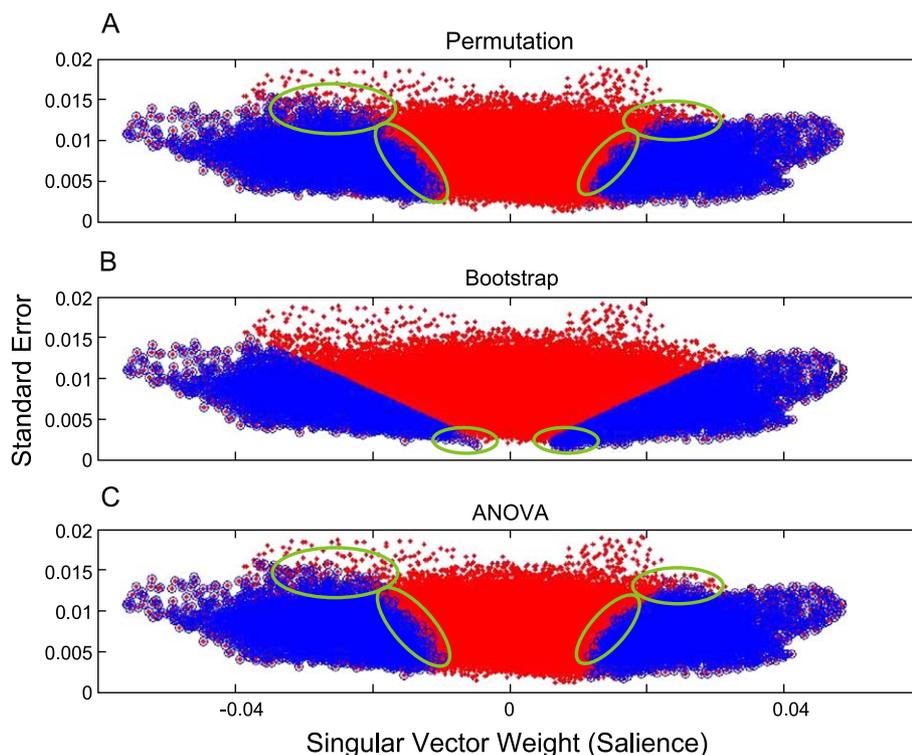


Fig. 1. Plot of voxelwise bootstrap estimated standard errors versus singular vector weights (salience) from a PLS analysis of PET data. (A) Voxels circled in blue have saliences showing a significant P value determined by permutation; (B) voxels circled blue have a ratio of salience to standard error greater than 2.57 (absolute value); (C) voxels circled in blue have a weight with a significant P value determined by repeated-measures ANOVA. In each panel, the green circle indicates the approximate region where each method of statistical assessment identifies unique voxels.

provides a level of inference not attainable by derivation of P values. This would be of particular utility when comparing different age or patient groups in being able to attribute group differences to noninvolvement of a particular region (indicated by a stable zero) versus increased variability within a group (indicated by a large standard error). The advantage of using reliability estimation, whether using standard errors or confidence intervals, enables a level of precision not accessible through more commonly reported probability assessments.

Interpreting ST-PLS results. As described above, the ST-PLS analysis provides pairs of patterns defining experimental differences or brain-behavior correlations in a spatiotemporal signal. Fig. 2 contains simplified results from a task ST-PLS analysis (modified from Lobaugh et al., 2001). The top panel shows grand average ERP waveforms from 1 of 32 electrodes from a two-condition simulation. In this example, Condition 1 latency was modeled to be longer than Condition 2 latency for the early peak (timepoints 10 to 25).

Following a significant statistical assessment using permutation tests, the pattern expressed on the LV is examined. This starts with inspection of the task saliences, which indicates the contrast among the conditions (V_{dev} , Fig. 2, middle row, left panel). The pattern indicates that the effect distinguishes Condition 1 and Condition 2. Since there are only two conditions in this example, the weights are equal and opposite in sign. A first approximation of the temporal expression of those differences can be seen by examining the temporal brain scores (Fig. 2, middle row, right panel). The positive temporal scores, peaking at timepoint 20, are larger for Condition 1 than Condition 2. This indicates that Condition 1 amplitudes are generally higher than Condition 2 amplitudes during that interval. The negative temporal scores at earlier points indicate a period where Condition 2 amplitudes are higher than Condition 1 amplitudes. The spatiotemporal expression of this pattern of differences is shown in the ST-PLS electrode saliences for this sensor (U_{dev} ; Fig. 2, bottom row, left panel). As for the temporal scores, the sign of the saliences indicates the direction of the task differences. Saliences greater than zero indicate timepoints

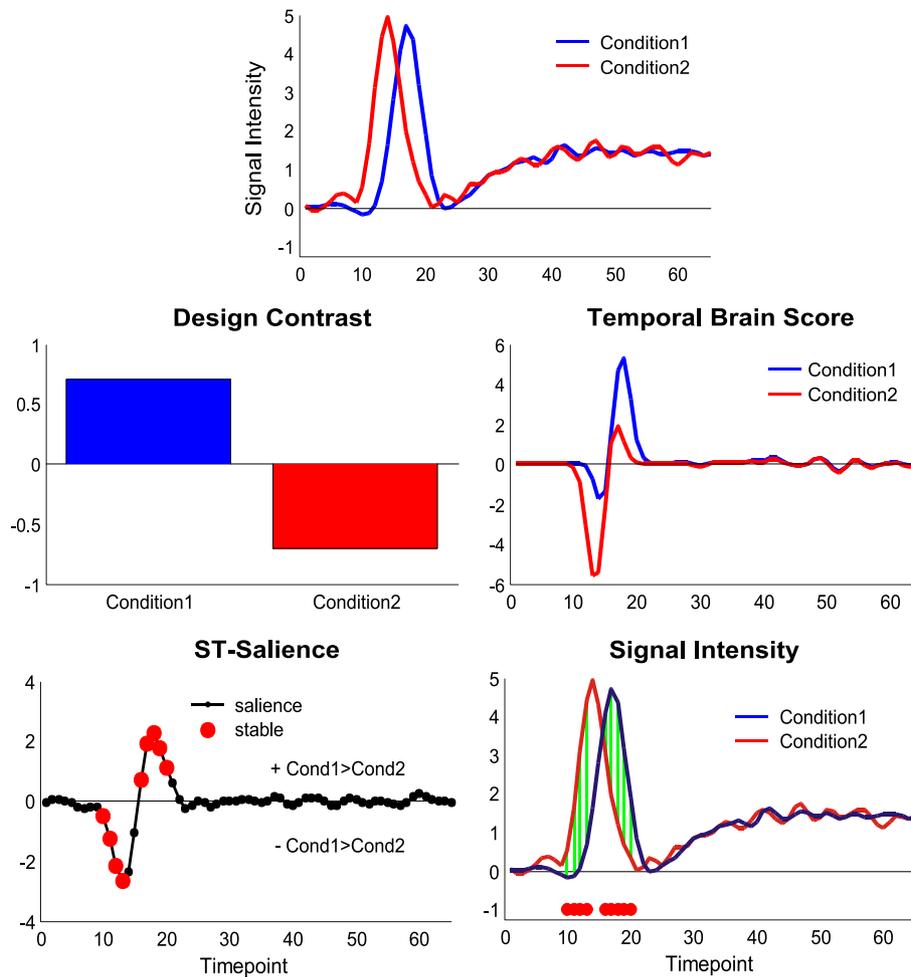


Fig. 2. Interpreting spatiotemporal PLS. (Top) One of 32 channels from a simulated two-condition spatiotemporal data set. Blue: Condition 1; Red: Condition 2. (Middle) Left: task saliences for the first latent variable; right: temporal brain scores for each condition. These are the dot products of the ERP or brain saliences with the ERP or fMRI data for each timepoint, thus are a summary of the task differences across all elements. (Bottom) Left: ERP saliences for the selected channel. Positive saliences indicate timepoints where the signal differences have a positive relation to the design saliences (Condition 1 > Condition 2). Negative saliences indicate timepoints where the signal differences have a negative relation to the design saliences (Condition 1 < Condition 2). Right: Stable differences between Condition 1 and Condition 2 are indicated by the red markers at the bottom of the plot and by the shading between the two curves. Note that the peak of Condition 2 was not reliably larger than the signal in Condition 1 at that same timepoint.

where the Condition 1 signal is higher than Condition 2. Saliences less than zero indicate that Condition 1 signal is lower than Condition 2. Zero and near-zero saliencies indicate periods where the two conditions did not differ. The bootstrap results indicated that the amplitude differences were reliable at nine timepoints (large markers overlaid on the electrode salience). The strongest saliencies are for the timepoints at the first peak, which maps onto the latency shift seen in the grand averages (Fig. 2., bottom row, right panel).

Application: ERP experiments

We demonstrate the application of ST-PLS for ERP data in a study examining the impact of configural manipulations on face processing (Itier and Taylor, 2002; Itier et al., 2004). This study measured the impact of two configural changes—contrast reversal and inversion—on the encoding and immediate and delayed recognition of faces.

Methods. Data from 32 young adults (16 females; 20 to 33 years, mean age 25.3 years) had normal or corrected-to-normal vision. Stimuli were 720 gray-scale pictures of unknown faces (50% female) and were either upright, inverted, or upright, but contrast-reversed. A single face type was presented every block, and within a block, one third of the faces repeated one time each. Half of the repeated faces repeated immediately and half repeated after one intervening face. Condition ‘Encode’ was defined as the first presentation of faces correctly recognized later. Similarly, the ‘Memory’ conditions consisted of only correctly recognized trials. Stimuli were presented centrally on a computer screen for 500 ms with a 1320-ms ISI. To minimize eye movements, subjects fixated on a centered cross appearing during the ISI. They pressed the spacebar of a keyboard to repeated stimuli.

Electrophysiology. EasyCaps containing 35 electrodes including three ocular sites were used to record ERPs. The EEG was recorded continuously via NeuroScan with a 500-Hz sampling rate and an amplification gain of 500 via SynAmps. Cz was the reference lead during acquisition; an average reference was calculated off-line. EEG was epoched into 1-s sweeps, including a 100-ms baseline. Off-line averaging was performed for each of the face types and memory conditions; averages were digitally filtered (0.8–60 Hz). Trials contaminated with ocular movements (≥ 100 μ V) between 0 and 700 ms were rejected. The ST-PLS analysis was conducted on this 700-ms interval (excluding the prestimulus baseline) for the 32 scalp electrodes; the EOG channels were not included.

Results. The ST-PLS analysis indicated that the primary impact of configural manipulations was similar across the three memory conditions. Three latent variables were significant, LV1 identified a difference between upright and inverted faces; LV2 was the difference between contrast-reversed faces and the other two conditions; only the third LV indicated an interaction of face-type and memory condition. To illustrate how the ST-PLS results are interpreted, we show partial data for the first LV; the full report can be found in Itier et al. (2004).

The task saliencies for LV1 ($P = 0.001$) are shown in Fig. 3 (top) and indicate that the largest difference was between upright and inverted faces, with contrast-reversed faces intermediate. The ERP saliencies for this effect are shown in the middle of Fig. 3. The blue lines indicate the spatiotemporal pattern across the 32 ERP

channels, and the markers at the top of each channel indicate the maximal saliencies stable by bootstrap estimation. One of the first features that can be seen when examining the saliencies is that the face inversion effect was seen at two time periods. Strong negative saliencies are seen posteriorly, peaking just before 200 ms, followed by an extended period of negative saliencies, maximal around 350 ms. The effects reverse polarity at frontal electrodes, but as the saliencies are smaller frontally, the conclusion here is that the differences over posterior, lateral channels dominate the inversion effect. The topographies of the peak saliencies are shown in the lower part of Fig. 3, highlighting the posterior lateral distribution for the inversion effect.

Once the peak saliencies are identified, a full interpretation, especially for ERP data, is facilitated by examining the grand-averaged ERP waveforms. To simplify the explanation, we have plotted three electrodes from the right hemisphere (P8, PO10, O2) in Fig. 4. Saliencies are shown in the left plots, and the grand-averaged ERPs for the three face types, collapsed across memory condition, are shown in the right plots. For both, timepoints of stable differences are indicated by markers at the top of each plot. The change in where the differences are expressed is clearly seen: the early saliencies are strongest at the more lateral channel (P8). This maps onto a difference after the P1 component, up to and including the N170 component, and the transition to the P2 component for the inversion effect, which dominates the differences at P8. The later saliencies map onto the much longer duration upright/inverted differences after the peak of the P2.

Thus, the conclusion from this part of the ST-PLS analysis is that amplitude differences over two time periods, over the same parieto-occipital channels, are required to explain upright/inverted differences. When measured using traditional measures of peak amplitude and latency, the P1 appeared to be involved in face inversion (Itier and Taylor, 2002), but this peak was not consistently part of the pattern when the whole epoch and all electrodes were analyzed. In a separate analysis, ST-PLS was run across a shorter interval (0–250 ms, data not shown here), and the P1 was part of the face-inversion effect. One factor that contributes to whether effects are identical between traditional and ST-PLS analyses is whether an effect identified at a single peak is the best reflection of the global ERP differences involved in the experimental effect. As the distribution of the long-latency differences was similar to the short-latency differences, this provides support for the idea that processing inverted faces involves a reactivation of face-processing regions.

Application: fMRI experiment

Data were obtained from a perceptual memory study comparing auditory and visual processes. The full report of this experiment is in preparation. This paper considers data from four of the eight conditions in the study to illustrate the use of ST-PLS for fMRI (McIntosh et al., in press).

Methods. Data from eight participants (aged between 23 and 36, three males) were used in the analyses. All were right-handed and reported normal hearing, and normal or corrected-to-normal vision.

The experimental conditions involved making judgments on triplets of either auditory or visual stimuli in two fMRI runs. For the auditory condition, bandpass-filtered noise stimuli were played through air-conducting headphones. In the visual condition, bandpass-filtered noise stimuli were displayed as visual textures,

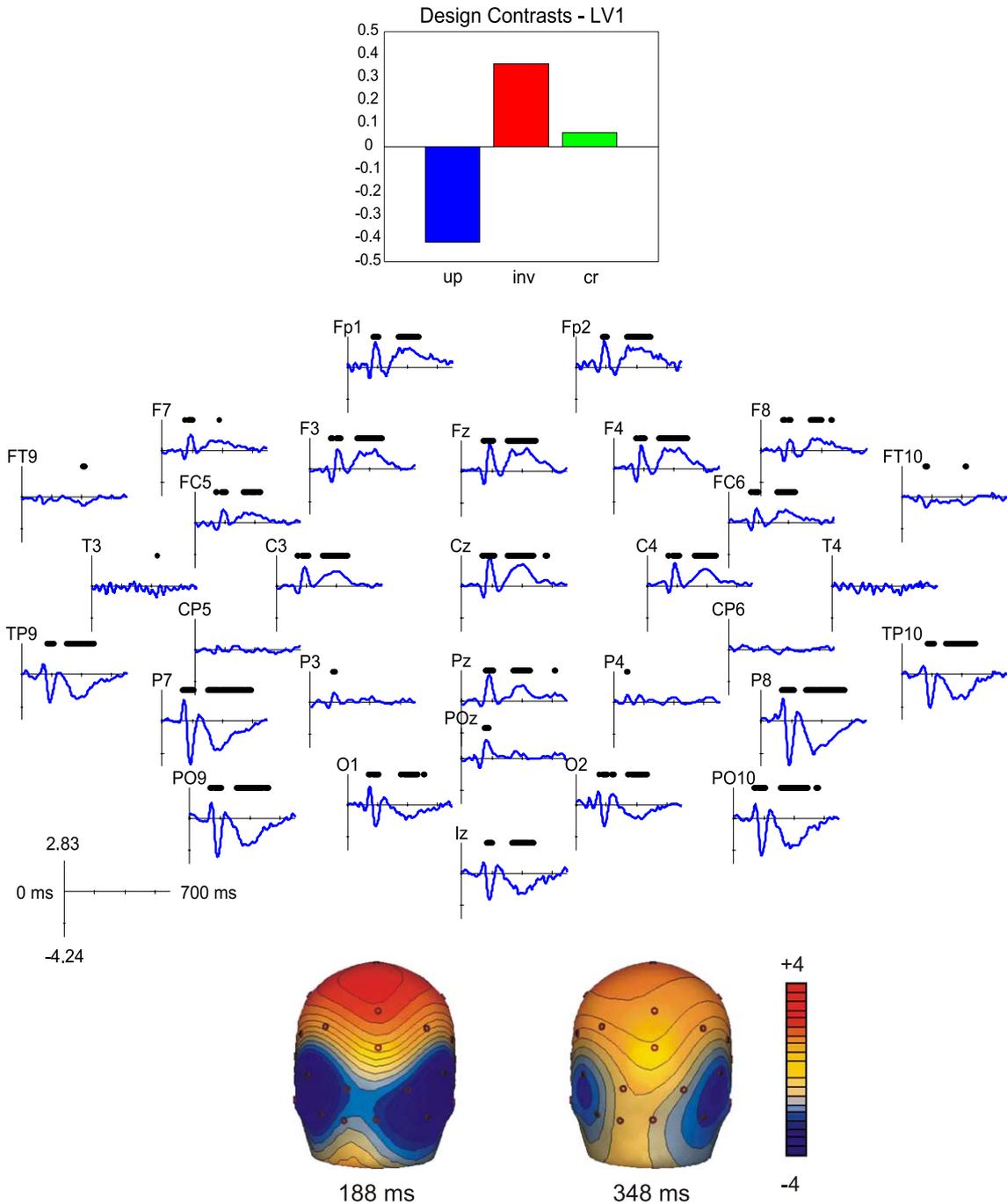


Fig. 3. Spatiotemporal PLS on ERP data. (Top) Task saliences for LV1 indicate that the primary distinction was between upright (UP, blue) and inverted faces (INV, red). For simplicity, only the design saliences for the encoding condition are shown; the values for the other two conditions were essentially the same. The near zero weighting for the contrast-reversed (CR, green) indicates that across the scalp, CR amplitudes were intermediate to UP and INV. (Middle) ERP saliences for all channels. Timepoints stable by bootstrap estimate (bootstrap ratio > 4) are indicated by markers on top of each channel. Y-axis reflects the range of ERP saliences for this LV. (Bottom) BESA head plots of the peak saliences, illustrating the posterior-lateral distribution of the differences.

projected onto a rear-projection screen and viewed by the participants through a mirror mounted in the MR head coil. On each trial, three 500-ms noise stimuli appeared successively, with a blank (silent/gray) interstimulus interval of 500 ms. The center frequency of the bandpass filter differed for each stimulus. After the offset of the third stimulus, there was a 1500-ms response window, during which observers pressed one of three keys. For the auditory temporal sequencing task, observers indicated whether the

tone with the highest pitch was presented first, second, or third. For the visual temporal sequencing task, observers indicated when the visual texture with the highest spatial frequencies appeared. The intertrial interval was chosen pseudorandomly and was 3, 5, 7, 9, or 11 s. The auditory baseline was the fMRI data from the 11-s intertrial intervals across all auditory tasks. The visual baseline was the fMRI data from the 11-s intertrial intervals across all the visual tasks.

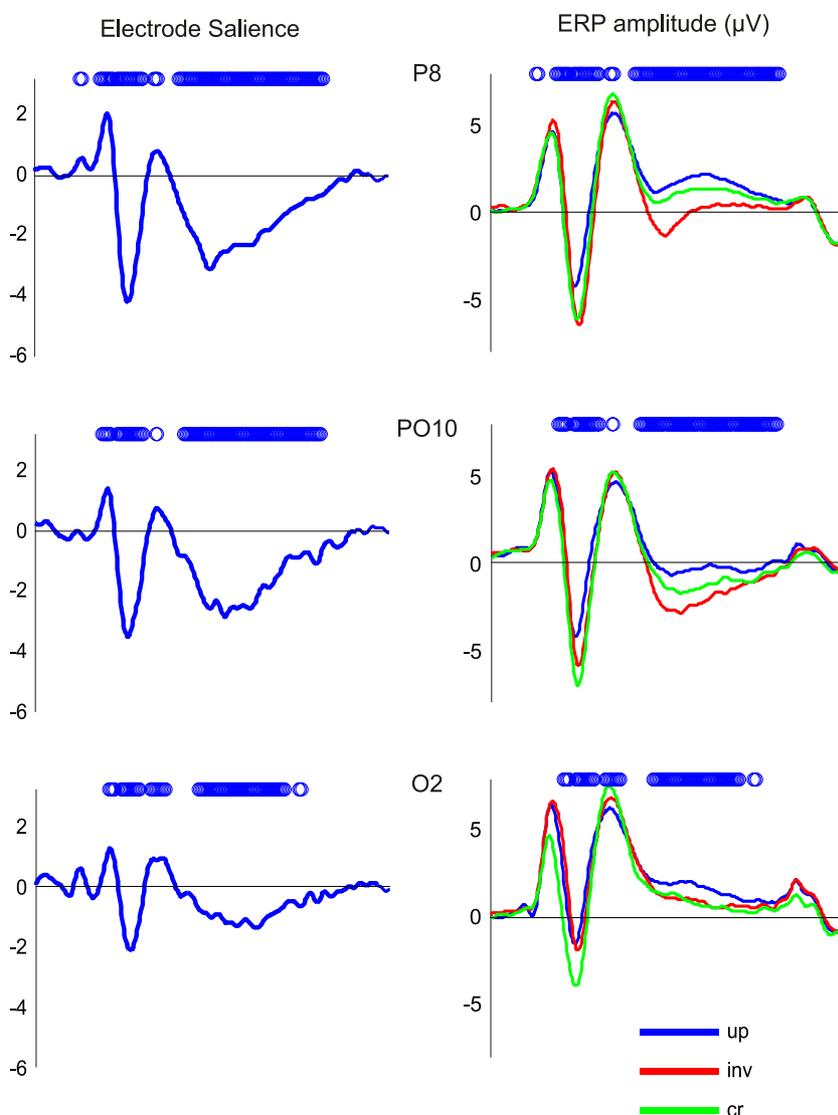


Fig. 4. Spatiotemporal PLS on ERP data. (Left) Saliences from Fig. 3 are plotted for three channels in the right hemisphere. (Right) Timepoints of stable upright/inverted differences are shown on the ERPs for upright (UP, blue), inverted (INV, red), and contrast-reversed (CR, green) faces. Note that as the CR amplitudes become more distinct from UP and INV (e.g., O2), the saliencies defining the upright/inverted distinction become smaller.

Performance was equated across subjects and conditions by establishing a psychophysical accuracy threshold for each subject. Thresholds were established the day before testing and were set such that a subject had roughly 80% accuracy across all tasks.

fMRI procedure. Regional activity was measured using a 1.5 T Signa MR scanner with a standard head coil (CVI hardware, LX8.3 software: General Electric Medical Systems, Waukesha, WI). Eighteen axial slices were acquired (64×64 , voxel size = 3.25 mm), each with a thickness of 7 mm. Functional scans were obtained using a single shot T2*-weighted pulse sequence with spiral readout, offline gridding, and reconstruction (TR = 2000 ms, TE = 40 ms, flip angle 80°, 90×90 effective acquisition matrix).

Data preprocessing was performed using Analysis of Functional NeuroImaging (AFNI) software (<http://afni.nimh.nih.gov/>, Cox, 1996). Time series data for each subject were spatially coregistered within a scanning run to correct for head motion by using a 3D Fourier transform interpolation and detrended to a constant reference scan by using a fifth-order polynomial. Motion-

corrected images were then spatially transformed to an fMRI spiral scan template which was registered to the MNI atlas used in SPM99 (<http://www.fil.ion.ucl.ac.uk/spm/>, Friston et al., 1995). The transformation of each subject to the spiral template was achieved in SPM99 using a 12-parameter affine transform with sinc interpolation.

ST-PLS analysis. As for ERP, in event-related fMRI, the time series length is a user-defined parameter, which can be modified depending on the particular design used, that is, rapid event-related versus extended ITI. The hemodynamic response (HR) for any voxel typically takes 6–20 s to rise and fall. A “temporal lag window” is defined to accommodate the major response changes. The absolute value of the MR signal is susceptible to low frequency drift from physiological or environmental noise. In constructing the data matrix, the MR signals are normalized within a trial with respect to the signal at the onset of the trial. The normalization maintains the amplitude information of the BOLD signals and is similar to baseline correction in ERP analyses. This

normalization step also allows averaging BOLD signals of the same condition across scan sessions to increase the signal-to-noise ratio.

Task and behavior ST-PLS analyses were performed on these data. For both, a 12-s temporal window was specified for each event (i.e., six TRs). For the experimental conditions, stimulus onset was defined at the onset of the first of three stimuli in the triplet. The onset of baseline was taken 4 s after the end of the trials preceding the 11-s ITIs.

Reaction time measures from each subject were used for the behavior ST-PLS. As there were no behavioral measures in the baseline conditions, only the auditory and visual temporal sequencing tasks were analyzed. Reaction time measures were expressed as z scores from each subject's mean and standard deviation across all conditions.

Results

Task ST-PLS. Two significant LVs were identified from the task analysis. The first LV (Fig. 5) was a main effect of Task versus Baseline ($P < 0.001$). The second LV depicted an interaction of task by stimulus modality. For the present paper, we focus only on the first of these LVs. The spatiotemporal activity pattern for LV1 is shown in Fig. 5 (top). From this pattern, the peak differentiation appears to occur at 6 to 8 s after stimulus onset (three to four TRs), which is corroborated by the temporal scores (Fig. 5 (bottom)). The response for individual voxels, selected from the maxima in the singular image, is shown in Fig. 6 and illustrates that most of the regions in the activity pattern show a response function similar to the temporal scores.

Behavior ST-PLS. The first LV from the analysis of brain with reaction time (RT) was significant by permutation test ($P = 0.010$).

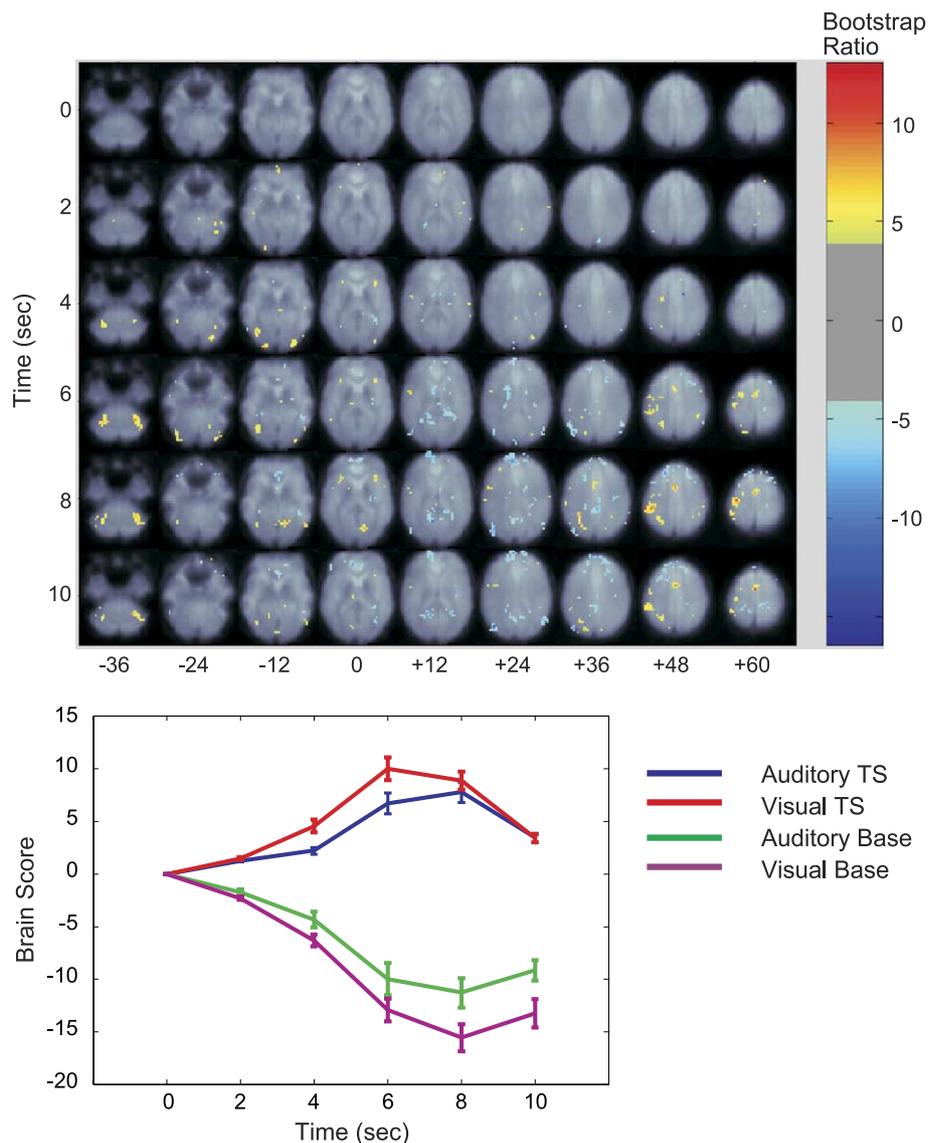


Fig. 5. Spatiotemporal PLS on fMRI data. Singular image and temporal brain scores for the first latent variable (LV1). Time from stimulus onset is indicated on the Y-axis of the singular image and is expressed in seconds from stimulus onset. The X-axis gives the approximate location of the axial slice in reference to the MNI atlas space. Voxels in the image are highlighted according to the magnitude of their stability (bootstrap ratio > 4). Temporal brain scores in the bottom panel indicate the expression of task effects across time and are computed for each task. Consistent with the more extensive activation differences at $T = 6$ to 10, the temporal scores show maximal differentiation 6 to 10 s after stimulus onset. Abbreviations: TS = temporal sequencing, Base = baseline.

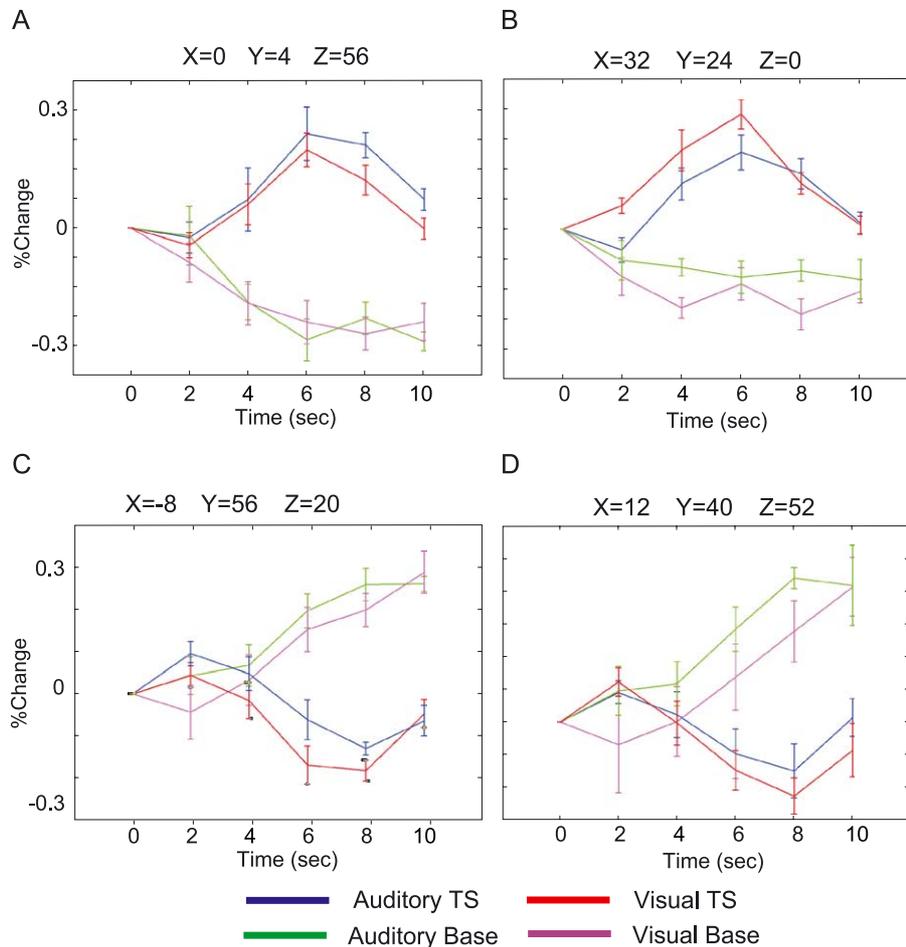


Fig. 6. Spatiotemporal PLS on fMRI data. Hemodynamic response functions from selected regions identified in LV1 (MNI voxel coordinates below each panel). Responses are expressed as percent change from stimulus onset ($T = 0$) and are averaged across subjects (\pm SE).

and reflected a common correlation of RT with brain activation in both temporal sequencing tasks. The second LV was remarkably nonsignificant ($P = 0.95$).

The spatiotemporal pattern for LV1 shows distinct temporal fluctuations in correlation patterns (Fig. 7). Positive saliences (yellow) are related to slower RT, and negative saliences (blue) are related to faster RT. Also shown are scatterplots of RT with the brain scores. While the correlation is strong for each condition, there is a potential outlier (circled symbol). Although the bootstrapping procedure reduces the impact of outliers (in the sense that the correlation would be less reliable), it is advisable in such cases to reanalyze the data without the potential outlier to ensure the reliability of the pattern. In the present case, reanalysis excluding the outlier did not change the pattern of ST-PLS results.

Fig. 8 shows RT-HRF correlations for four dominant voxels in the LV. The first, a voxel in occipital cortex (MNI coordinate: $X = 0$, $Y = -80$, $Z = -4$), shows a positive correlation with RT for both tasks, which was maximal 6 s after stimulus onset. A second voxel, situated much more dorsally near supplementary motor cortex ($X = -4$, $Y = 0$, $Z = 64$) also showed a positive correlation, peaking 2 s later than the occipital voxel. Regions showing a negative correlation with RT included left inferior prefrontal ($X = -40$, $Y = 36$, $Z = 8$) and medial prefrontal cortex ($X = -4$, $Y = 56$, $Z = 0$).

The correlation pattern in medial prefrontal cortex is of potential theoretical interest given its activity differences identified in the task ST-PLS.

Caveats in PLS

Computation

Computational overhead may be a concern when conducting PLS analysis, especially in the case of fMRI or ERP/MEG data when high density spatial or temporal sampling is used. For an average-sized event-related fMRI study, an analysis including 500 permutations and 100 bootstraps takes 30 to 60 min when running on a Linux-based Intel workstation with 4GB of RAM, depending on the workstation load. In systems with parallel architectures, it is possible to distribute the resampling operations across CPUs, reducing total computation time.

Linearity and orthogonality

The PLS model we use assumes the measured effects relate linearly to either the task manipulations or to behavior (i.e., the exogenous measures). While this is likely a reasonable assumption in most cases, the possibility of nonlinear relations is a consideration, especially in ERP and MEG data. Versions of PLS are used in other disciplines, where it is possible to quantify

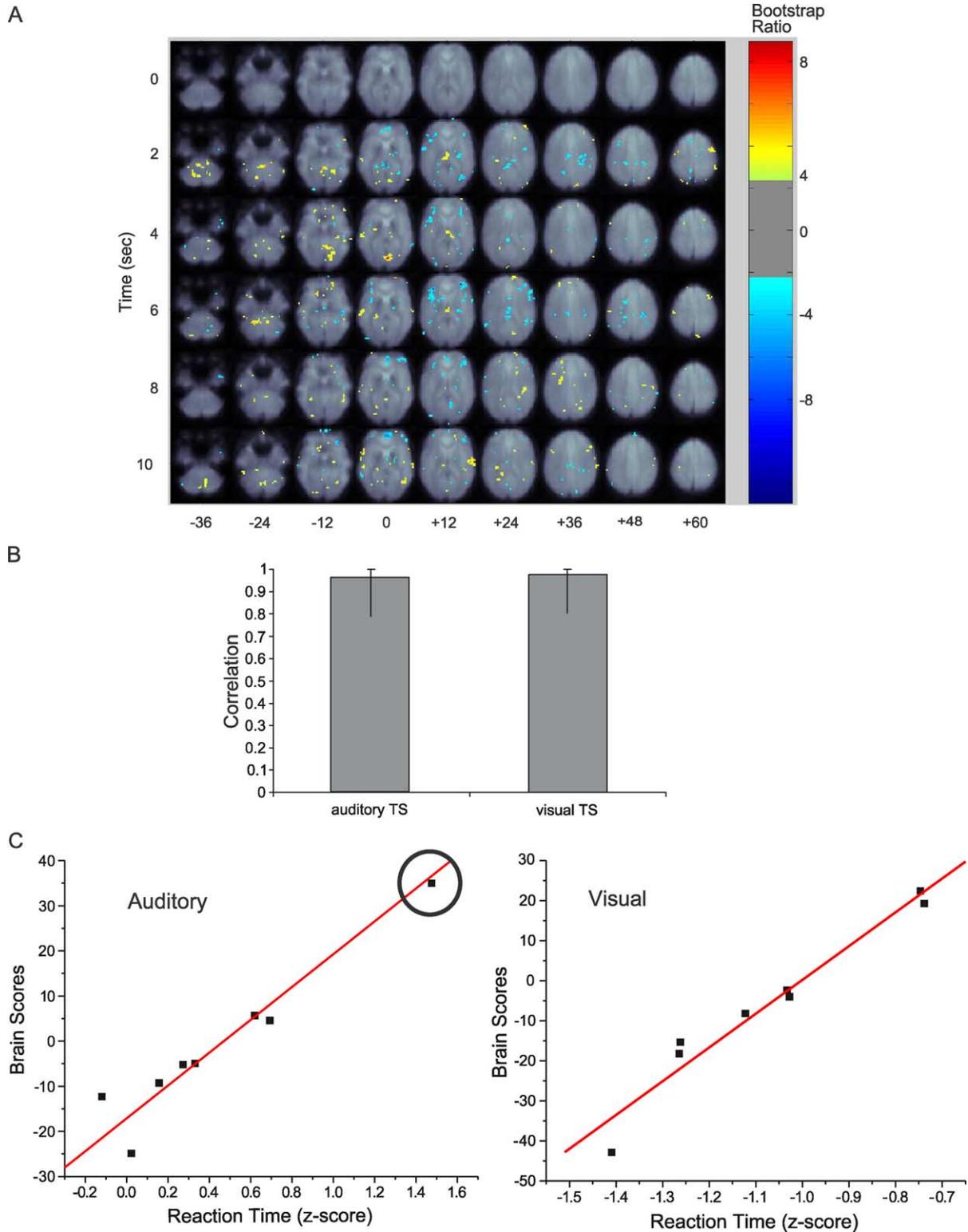


Fig. 7. Spatiotemporal PLS on fMRI data. Singular image (A) and correlations of brain scores and RT for the behavior ST-PLS analysis. Correlations are plotted both as bar graphs with bootstrap estimated 95% confidence intervals (B) and as scatterplots (C).

nonlinearities (Hasegawa et al., 1996; Martin et al., 1995), and we are exploring the use of these algorithms for neuroimaging data.

A related issue is that our version of PLS uses singular value decomposition to reexpress the relation between the neuroimaging

data and the exogenous measure. The extraction of mutually orthogonal patterns is mathematically elegant, but this may not adequately capture the “true” dependence between the patterns. For this reason, many researchers have advocated the use of Independent Components Analysis (ICA), where the algorithms

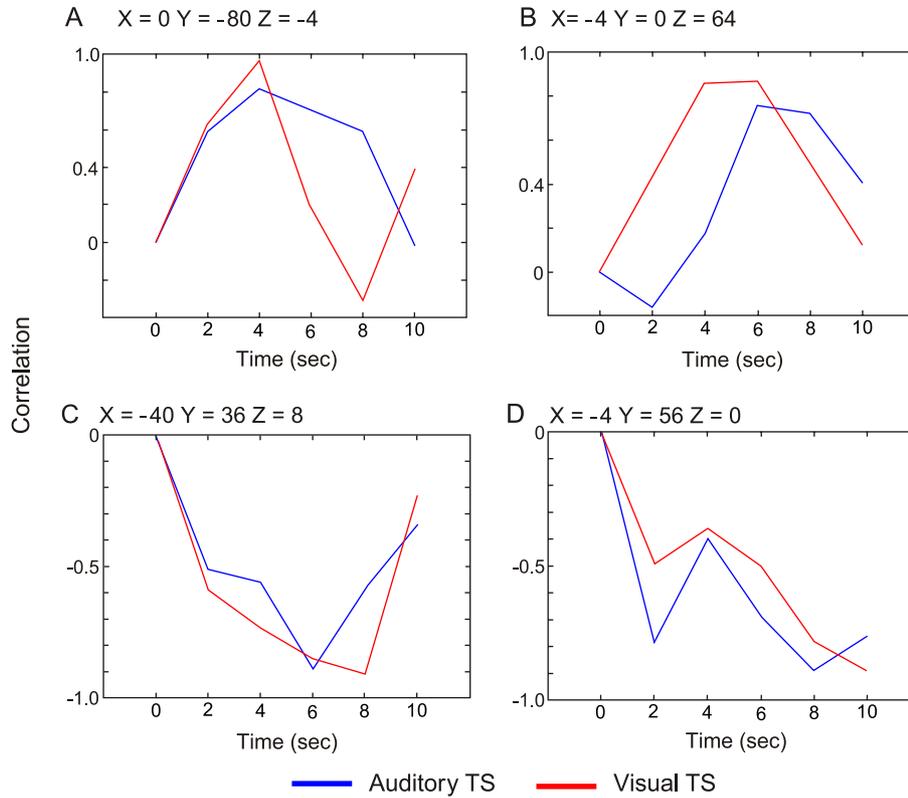


Fig. 8. Spatiotemporal PLS on fMRI data. Correlation profiles for selected voxels identified in the behavior ST-PLS analysis. Correlations are plotted for the two tasks across the time window from stimulus onset ($T = 0$). Coordinates below each panel are referenced to the MNI atlas space.

work only to maximize the independence between patterns rather than forcing orthogonality (McKeown et al., 1998).

While ICA is theoretically attractive because of this problem, we have found little difference in using ICA versus SVD as part of PLS and actually noted that SVD performed better than ICA in extracting reliable signals (Lin et al., 2003). Thus, the theoretical attraction of ICA seems to be problematic when considered from the perspective of statistical sensitivity and robustness.

Finally, it is important for researchers to make the distinction between mathematical constraints and underlying biology. For example, we have demonstrated that orthogonal patterns identified by PLS can be considered as interacting neural systems when the dependence among the constituents of these patterns is explicitly assessed (McIntosh et al., 2003).

Interpretation

Interpretational difficulties in PLS often arise when the dimensions identified do not map easily to the expectation of the researcher. While PLS identifies the maximal effects in the data set, it is possible that these effects do not correspond to the *a priori* expectations of the researcher. We have utilized a post hoc contrast procedure to aid in the interpretation of PLS solutions. For example, in a parametric delay-match to sample task, we projected linear and quadratic contrasts to the design scores to further characterize the dominant trends in a given effect (Grady et al., 1998).

More recently, we have developed a “nonrotated” version of PLS, wherein *a priori* contrasts are used to restrict the patterns derived from PLS. This is possible because of the relation of the singular values from the SVD to the total sums-of-squares from the data matrix. Specifically:

If \mathbf{S} is a vector of k singular values derived from SVD on matrix \mathbf{M} , then

$$\sum \mathbf{S}_k^2 = \sum \mathbf{M}_{ij}^2 \tag{16}$$

Each of the squared singular values divided by the sum of the squared singular values indicates the proportion of the total sums-of-squares accounted for from \mathbf{M} . It is also true, given Eqs. (7) and (8), that if one takes matrix \mathbf{V} , which contains the design saliences, from the SVD on matrix \mathbf{M} ,

$$\mathbf{V}^T * \mathbf{M} = \mathbf{U} * \mathbf{S} \tag{17}$$

substituting \mathbf{D} for $\mathbf{U} * \mathbf{S}$, then sums-of-squares for each vector in \mathbf{D} yields the squared singular values, which is the figure of merit for the permutation test in PLS.

If we introduce a new matrix \mathbf{X} , which is the same size as \mathbf{V} , and contains a different set of design saliences or contrasts, and write:

$$\mathbf{X}^T * \mathbf{M} = \mathbf{E} \tag{18}$$

then, as with \mathbf{D} , the sums-of-squares for each vector in \mathbf{E} are essentially the singular values and can be assessed using the same permutation algorithm used for the regular PLS.

The unrotated approach was applied to the fMRI example described earlier, using the contrasts: $[-1 -1 1 1]$; $[-1 1 -1 1]$ and $[1 -1 -1 1]$, which code the modality effect, the task effect, and the interaction. We observed that only the modality effect and interaction terms were significant by permutation. This is also what the regular PLS found.

The unrotated version of PLS has the advantage of allowing a direct assessment of hypothesized experimental effects, although the interpretation may be difficult if nonorthogonal contrasts are used. We tend to favor the rotated version since the solution produced is optimal for the detection of the maximal experimental effects, and the effects are mutually orthogonal. In most cases, this gives PLS a distinct advantage as an exploratory data analysis tool and provides an assessment of whether the intuitions one has about an experiment are consistent with the data. Nevertheless, in studies where there is a strong a priori hypothesis best captured by explicit contrasts, it would be reasonable to run an unrotated version of PLS to test the hypothesis, followed by a rotated version to further explore the dominant effects in the data. We should note that where univariate analyses are conducted using the same set of contrasts, there will be similarities in the spatial patterns identified by PLS, but because of the sensitivity to correlations among voxels, PLS will always have greater power (see McIntosh et al., *in press* for explicit comparison). Lukic et al. (2002) have also demonstrated the improved signal detection properties of multivariate analyses similar to PLS over conventional univariate approaches.

Conclusions

Starting from the basic emphasis on explaining the relation between two or more blocks of data, PLS can address questions ranging from identifying task-specific patterns of activity to extracting neural activity patterns that predict behavior. The flexibility of PLS, especially when used in concert with other analytic methods, enables more thorough testing of specific hypotheses and development of neurocognitive theory.

As noted in the introduction, PLS is part of a family of multivariate data analyses, providing direct means of ascertaining distributed patterns of activity that contribute to task differentiation, predict behavior, or identify particular patterns of functional connectivity. We have found PLS to be very powerful in this regard, revealing patterns that are not easily derived from other forms of analysis, yet at the same time complementing them. The application of task PLS allowed for the identification of general memory networks, on top of which were superimposed collections of regions that were recruited for more specific memory operations (Nyberg et al., 1996). We were also able to demonstrate memory retrieval-dependent changes in functional connectivity that were not detectable either by task PLS or univariate activation analysis (Nyberg et al., 2000). In studies of learning and awareness, we combined behavior PLS and with functional connectivity analysis (seed PLS) to identify interregional interactions involving left prefrontal cortex that were specific to learning with awareness (McIntosh et al., 1999a). In a subsequent study, again using behavior and seed PLS, we showed how the hippocampus can be related to learning with and without awareness depending on the pattern of functional connectivity (McIntosh et al., 2003). These latter examples provided strong evidence for the notion of *Neural Context*, where the functional relevance of a brain area to a particular cognitive function depends on the specific pattern of interregional communication (McIntosh, 2000, 2001, *in press*).

PLS applications to group studies have mostly occurred in the study of normal aging (Grady and Craik, 2000), although there have been applications to patient populations (Mentis et al., 2002; Turner et al., 2003; Vanlancker-Sidtis et al., 2003). The focus of

these analyses has been on common task effects and the interaction of group and task effects, and several have explored used behavior and seed PLS to further characterize similarities and difference across groups (Della-Maggiore et al., 2000; Grady et al., 2003; McIntosh et al., 1999b). In all cases, PLS has been extremely important in providing a full characterization of the unique neural patterns within distinct groups, but also indicating where groups are similar to one another. This latter point is often omitted in other approaches but is vital for the complete understanding of different age and patient groups.

Finally, we must emphasize that PLS, like any analytic approach, is not a panacea. The complexity of the data that are extracted from neuroimaging methods dictates that a single analytic approach is insufficient. We concur with the position advocated by others that a pluralistic approach will provide a much better appreciation how the brain brings about human mental function (Lange et al., 1999).

Acknowledgments

This work was supported by Canadian Institutes for Health Research grants to ARM and NJL, and a JS McDonnell Foundation grant to ARM.

References

- Cox, R.W., 1996. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res.* 29, 162–173.
- Della-Maggiore, V., Sekuler, A.B., Grady, C.L., Bennett, P.J., Sekuler, R., McIntosh, A.R., 2000. Corticolimbic interactions associated with performance on a short-term memory task are modified by age. *J. Neurosci.* 20, 8410–8416.
- Duzel, E., Habib, R., Schott, B., Schoenfeld, A., Lobaugh, N., McIntosh, A.R., Scholz, M., Heinze, H.J., 2003. A multivariate, spatiotemporal analysis of electromagnetic time-frequency data of recognition memory. *NeuroImage* 18, 185–197.
- Edgington, E.S., 1980. *Randomization Tests*. Marcel Dekker, New York.
- Efron, B., Tibshirani, R., 1986. Bootstrap methods for standard errors, confidence intervals and other measures of statistical accuracy. *Stat. Sci.* 1, 54–77.
- Friston, K.J., Ashburner, J., Frith, C.D., Pline, J.-B., Heather, J.D., Frackowiak, R.S.J., 1995. Spatial registration and normalization of images. *HBM* 2, 165–189.
- Good, P., 2000. *Permutation Tests: A Practical Guide to Resampling Methods for Testing Hypotheses*. Springer, New York.
- Grady, C.L., Craik, F.I., 2000. Changes in memory processing with age. *Curr. Opin. Neurobiol.* 10, 224–231.
- Grady, C.L., McIntosh, A.R., Bookstein, F., Horwitz, B., Rapoport, S.I., Haxby, J.V., 1998. Age-related changes in regional cerebral blood flow during working memory for faces. *NeuroImage* 8, 409–425.
- Grady, C.L., McIntosh, A.R., Beig, S., Keightley, M.L., Burian, H., Black, S.E., 2003. Evidence from functional neuroimaging of a compensatory prefrontal network in Alzheimer's disease. *J. Neurosci.* 23, 986–993.
- Hasegawa, K., Kimura, T., Miyashita, Y., Funatsu, K., 1996. Nonlinear partial least squares modeling of phenyl alkylamines with the monoamine oxidase inhibitory activities. *J. Chem. Inf. Comput. Sci.* 36, 1025–1029.
- Hay, J.F., Kane, K.A., West, R., Alain, C., 2002. Event-related neural activity associated with habit and recollection. *Neuropsychologia* 40, 260–270.

- Itier, R.J., Taylor, M.J., 2002. Inversion and contrast polarity reversal affect both encoding and recognition processes of unfamiliar faces: a repetition study using ERPs. *NeuroImage* 15, 353–372.
- Itier, R., Taylor, M.J., Lobaugh, N.J., 2004. Spatiotemporal analysis of event-related potentials to upright, inverted, and contrast-reversed faces: effects on encoding and recognition. *Psychophysiology* 41, 643–653.
- Lange, N., Strother, S.C., Anderson, J.R., Nielsen, F.A., Holmes, A.P., Kolenda, T., Savoy, R., Hansen, L.K., 1999. Plurality and resemblance in fMRI data analysis. *NeuroImage* 10, 282–303.
- Lin, F.H., McIntosh, A.R., Agnew, J.A., Eden, G.F., Zeffiro, T.A., Belliveau, J.W., 2003. Multivariate analysis of neuronal interactions in the generalized partial least squares framework: simulations and empirical studies. *NeuroImage* 20, 625–642.
- Lobaugh, N.J., West, R., McIntosh, A.R., 2001. Spatiotemporal analysis of experimental differences in event-related potential data with partial least squares. *Psychophysiology* 38, 517–530.
- Lobaugh, N.J., Chevalier, H., Batty, M., Taylor, M.J., submitted for publication. Accelerated and amplified neural responses in visual discrimination: Two features are processed faster than one.
- Lukic, A.S., Wernick, M.N., Strother, S.C., 2002. An evaluation of methods for detecting brain activations from functional neuroimages. *Artif. Intell. Med.* 25, 69–88.
- Martin, Y.C., Lin, C.T., Hetti, C., DeLazzer, J., 1995. PLS analysis of distance matrices to detect nonlinear relationships between biological potency and molecular properties. *J. Med. Chem.* 38, 3009–3015.
- McIntosh, A.R., 2000. From location to integration: how neural interactions form the basis for human cognition. In: Tulving, E. (Ed.), *Memory, Consciousness, and the Brain: The Tallinn Conference*. Psychology Press, Philadelphia.
- McIntosh, A.R., 2001. Towards a network theory of cognition. *Neural Netw.* 13, 861–876.
- McIntosh, A.R., in press. Contexts and Catalysts: a resolution of the localization and integration of function in the brain. *Neuroinformatics*.
- McIntosh, A.R., Bookstein, F.L., Haxby, J.V., Grady, C.L., 1996. Spatial pattern analysis of functional brain images using partial least squares. *NeuroImage* 3, 143–157.
- McIntosh, A.R., Rajah, M.N., Lobaugh, N.J., 1999a. Interactions of prefrontal cortex related to awareness in sensory learning. *Science* 284, 1531–1533.
- McIntosh, A.R., Sekuler, A.B., Penpeci, C., Rajah, M.N., Grady, C.L., Sekuler, R., Bennett, P.J., 1999b. Recruitment of unique neural systems to support visual memory in normal aging. *Curr. Biol.* 9, 1275–1278.
- McIntosh, A.R., Rajah, M.N., Lobaugh, N.J., 2003. Functional connectivity of the medial temporal lobe relates to learning and awareness. *J. Neurosci.* 23, 6520–6528.
- McIntosh, A.R., Chau, W., Protzner, A.B., in press. Spatiotemporal analysis of event-related fMRI data using partial least squares. *NeuroImage*.
- McKeown, M.J., Makeig, S., Brown, G.G., Jung, T.P., Kindermann, S.S., Bell, A.J., Sejnowski, T.J., 1998. Analysis of fMRI data by blind separation into independent spatial components. *HBM* 6, 160–188.
- Mentis, M.J., McIntosh, A.R., Perrine, K., Dhawan, V., Berlin, B., Feigin, A., Edwards, C., Mattis, P., Eidelberg, D., 2002. Relationships among the metabolic patterns that correlate with mnemonic, visuospatial, and mood symptoms in Parkinson's disease. *Am. J. Psychiatry* 159, 746–754.
- Milan, L., Whittaker, J., 1995. Application of the parametric bootstrap to models that incorporate a singular value decomposition. *J. R. Stat. Soc., Ser. C, Appl. Stat.* 44, 31–49.
- Nichols, T.E., Holmes, A.P., 2002. Nonparametric permutation tests for functional neuroimaging: a primer with examples. *HBM* 15, 1–25.
- Nyberg, L., McIntosh, A.R., Cabeza, R., Habib, R., Houle, S., Tulving, E., 1996. General and specific brain regions involved in encoding and retrieval of events: what, where, and when. *Proc. Natl. Acad. Sci. U. S. A.* 93, 11280–11285.
- Nyberg, L., Habib, R., Tulving, E., Cabeza, R., Houle, S., Persson, J., McIntosh, A.R., 2000. Large scale neurocognitive networks underlying episodic memory. *J. Cardiovasc. Nurs.* 12, 163–173.
- Sampson, P.D., Streissguth, A.P., Barr, H.M., Bookstein, F.L., 1989. Neurobehavioral effects of prenatal alcohol: Part II. Partial least squares analysis. *Neurotoxicol. Teratol.* 11, 477–491.
- Strother, S.C., Anderson, J., Hansen, L.K., Kjems, U., Kustra, R., Sidtis, J., Frutiger, S., Muley, S., LaConte, S., Rottenberg, D., 2002. The quantitative evaluation of functional neuroimaging experiments: the NPAIRS data analysis framework. *NeuroImage* 15, 747–771.
- Turner, R.S., Grafton, S.T., McIntosh, A.R., DeLong, M.R., Hoffman, J.M., 2003. The functional anatomy of parkinsonian bradykinesia. *NeuroImage* 19, 163–179.
- Vanlancker-Sidtis, D., McIntosh, A.R., Grafton, S., 2003. PET activation studies comparing two speech tasks widely used in surgical mapping. *Brain Lang.* 85, 245–261.
- Wold, H., 1982. Soft modeling: the basic design and some extensions. In: Wold, K.G.JaH (Ed.), *Systems Under Indirect Observation: Causality–Structure–Prediction, Part II*. North-Holland Publishing Company, Amsterdam, pp. 1–54.