Psychophysiology, 38 (2001), 517–530. Cambridge University Press. Printed in the USA.

Spatiotemporal analysis of experimental differences in event-related potential data with partial least squares

NANCY J. LOBAUGH, ROBERT WEST, AND ANTHONY R. MCINTOSH
Rotman Research Institute, Baycrest Centre for Geriatric Care and University of Toronto, Canada

Abstract

One challenge in the analysis of event-related potentials (ERPs) is to identify task-related differences in scalp topography. The multivariate Partial Least Squares (PLS) analysis was used to identify the spatiotemporal distribution of ERP differences related to experimental manipulations. Two simulations included latency shifts and amplitude changes at peaks with temporal overlap. PLS identified effects only at modeled timepoints and electrodes. In contrast, principal components analysis identified differences at most timepoints. We also demonstrated that PLS identified combinations of waveform differences, not isolated sources. ERP components in an auditory oddball task were also assessed with PLS. The primary distinction was between ERPs on hit and correct rejection trials, expressed at multiple timepoints and electrodes. PLS provides a mechanism to describe experimental differences in ERP waveforms, simultaneously across the head.

Descriptors: Partial least squares, Principal components analysis, Singular value decomposition, Event-related potentials

Event-related potential (ERP) data provide both spatial and temporal measures of brain activity related to cognitive processing. Typically, epochs analyzed include about 100 ms prestimulus activity and 1,000 ms poststimulus, at 30 or more electrode sites across the scalp. These are large, multivariate datasets, generally containing more than 6,000 datapoints per subject in every condition. Until recently, univariate analyses (ANOVA, *t* tests) have been most frequently used to extract effects related to experimental manipulations from the data. To reduce the impact of multiple comparisons, analysis is typically restricted to latencies of particular interest and a small number of electrode locations. The choice of electrode sites as well as the interval of interest are either based on a priori hypotheses, or selected after visual inspection of the ERP data. Both approaches have problems and limitations. Analysis restricted to a priori selected intervals and/or electrodes may overlook new information present in the dataset. Analysis based on post-hoc selection of electrodes and intervals is influenced by the experimenter's own biases. The need for more objective statistical analysis of these large datasets has prompted the application of a number of linear multivariate techniques such as principal components analysis (PCA; Donchin & Heffley, 1978), independent components analysis (ICA; Makeig, Juny, Bell, Ghahremani, & Sejnowski, 1997; Makeig et al., 1999), and spatiotemporal modeling (STM; Achim & Bouchard, 1997).

One primary assumption made in the application of linear multivariate analyses to study ERPs is that the scalp recordings reflect a linear combination of electrically active sources within the brain. PCA has been the most frequently used, and can provide separate analysis of the spatial or temporal relations in ERP data. Most often, it has been used to examine the underlying temporal patterns elicited by different experimental conditions (Picton et al., 2000). Temporal PCA provides orthogonal vectors, principal components (PC), reflecting patterns of scalp amplitudes across the time interval. The PCA solution is then typically rotated (Varimax, Promax) to maximize the signal within individual PCs. Newer methods such as ICA remove the orthogonality restriction of PCA (Makeig et al., 1997), examine spatial PCA separately (Dien, 1998) or combine spatial and temporal decomposition into a single analysis (Achim & Bouchard, 1997; Spencer, Dien, & Donchin, 1999). These techniques have generated interest because the solutions are represented as multiple "virtual" or "spatiotemporal" ERPs containing the spatial expression of a temporal pattern. It is important to note that the components identified by *all* of these techniques are not just related to the experimental manipulations. In the PCA approach, in order to examine which PCs are relevant to experimental manipulations, component scores are analyzed by ANOVA. Because PCA extracts commonalities as well as differ-

This research was supported by grants from the Medical Research Council of Canada to M. J. Taylor and N. J. Lobaugh, and to A. R. McIntosh. N. Lobaugh was supported by a research fellowship from the Rotman Research Institute. R. West was supported by a postdoctoral fellowship from the National Institute of Aging (RO1 Ag13845-01) and is now at the Department of Psychology, University of Notre Dame. The ERP data were graciously provided by Dr. C. Alain, Rotman Research Institute. The authors thank Drs. C. K. Van Petten, M. J. Taylor, C. Alain, and R. Staines, and two anonymous reviewers for their comments on this manuscript. Lively discussions with Drs. F. Bookstein, T. W. Picton, and A. Achim were also instrumental in the development of this method.

Address reprint requests to: Nancy J. Lobaugh, Ph.D., Imaging/Bioengineering Research & Cognitive Neurology Unit, Sunnybrook and Women's College Health Sciences Centre, 2750 Bayview Ave., Room S604, Toronto, ON M4N 3M5, Canada. E-mail: nlobaugh@sten.sunnybrook.utoronto.ca.

ences among conditions (Widaman, 1993), a single PC may reflect both of these sources of variance.

Typically, the goal of multivariate techniques such as PCA, ICA, and STM is ultimately to describe important differences in the ERP waveform that are related to the experimental manipulations. We present here the application of a new multivariate tool, partial least squares (PLS; Wold, 1975), to the analysis of ERP datasets. PLS can be used to describe the relation between one set of measures, like the experimental design or behavioral responses, and a large set of dependent measures, in this case scalp potentials. The primary advantage over other multivariate techniques is that it is designed to identify where, simultaneously in space and time, the strongest experimental effects are expressed. PLS has been used extensively for one-dimensional images from spectrographs, as in chemometrics or remote sensing (e.g., Heise, Marbach, Janatsch, & Kruse-Jarres, 1989; Hellberg, Sjostrom, & Wold, 1986). Its first application in psychology was an examination of the relations between multiple measures of maternal drinking behavior and offspring outcome in a study on fetal alcohol syndrome (Streissguth, Bookstein, Sampson, & Barr, 1993). PLS has recently been adapted for functional neuroimaging analysis to identify unique relations between (a) experimental design and brain activity (McIntosh, Bookstein, Haxby, & Grady, 1996), (b) brain activity and behavioral responses (McIntosh, Lobaugh, Cabeza, Bookstein, & Houle, 1998), and (c) single brain regions and the rest of the brain (McIntosh, Rajah, & Lobaugh, 1999).

To illustrate the utility of the PLS technique as it is applied to the full ERP dataset, we first present the results from two simulation studies. For these studies, activity in three neural sources was combined and projected onto the scalp. The first simulation demonstrates how PLS identifies experimental effects expressed as latency shifts at a single neural source. The second simulation was designed to test the limits of PLS by incorporating multiple effects at simultaneously active sources. Because temporal PCA is the most frequently used multivariate analysis in the literature, we contrasted the PLS results from this simulation with those from a temporal PCA. Finally, to illustrate the relevance of PLS for real ERP data, we applied the PLS analysis to an ERP dataset from an auditory oddball paradigm.

To anticipate the results, in all cases, PLS identified stable differences in the waveforms only at timepoints where they were modeled. Where the scalp potentials did not differ across conditions, the PLS results indicated that those timepoints did not contribute to the results. Interpretation of the first simulation was straightforward, as the PLS results mapped onto the topography of the manipulated source. The results from the second simulation demonstrated that PLS identifies only the specific *combinations* of waveform differences that distinguish conditions. In the case where differences in scalp activity deriving from two simultaneously active sources best defined an experimental effect, this combination was revealed on a single dimension. These findings indicate that PLS is a sensitive tool for detecting the spatiotemporal scalp distribution of ERP waveform differences.

Methods

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. The part is the "cross-block" correlation between the exogenous and dependent measures. PLS is similar to PCA or eigenimage analysis (Friston, Frith, Liddle, & Frackowiak, 1993; Moeller, Strother, Sidtis, & Rottenberg, 1987), but one important feature of PLS is that we constrain the solutions to the part of the covariance structure attributable to experimental manipulations. Moreover, PLS is ideal for datasets where the measures within a block are highly correlated (e.g., scalp potentials) because items within a block are not adjusted for these correlations (cf., canonical correlation). Figure 1 provides a graphical outline of the major steps in the PLS analysis, which is described in detail below.

Data matrix. The data matrix consists of one row of data per subject, blocked by condition. The spatial and temporal information is maintained by stringing together the amplitudes at all timepoints for each electrode. For example, in a study with 32 electrodes having 200 timepoints at each electrode, each row of data would contain 6400 datapoints (Figure 1A, right).

Design matrix. PLS examines the relations between the data of interest and some exogenous source thought to influence or other-

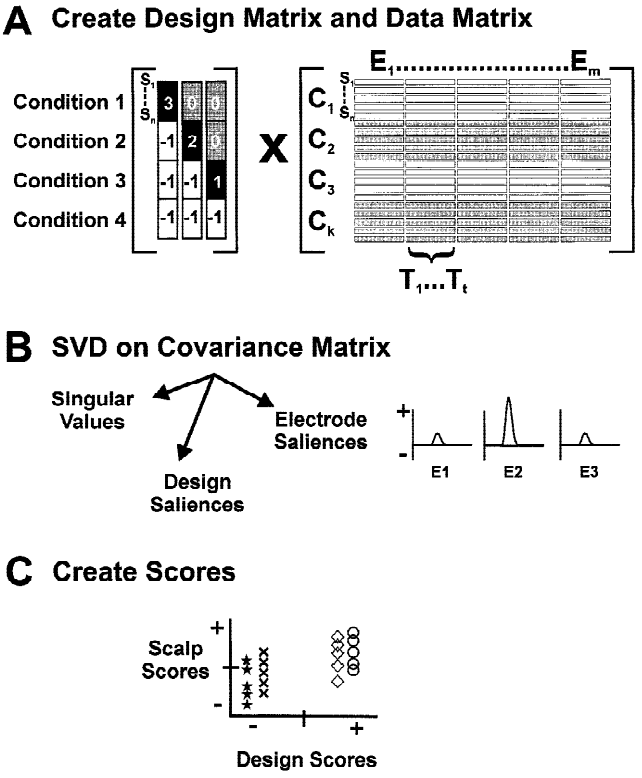


Figure 1. Steps in the design-brain PLS. A: A matrix containing design contrasts and the data matrix are created. The design matrix can be any orthonormal set of vectors defining the degrees of freedom in the experiment. Helmert contrasts are depicted in this illustration (e.g., Condition 1 versus all others, Condition 2 versus 3 and 4, Condition 3 versus 4). Within a condition, all subjects ($S_1 \ldots S_n$) have the same values. The data matrix is organized such that a single row contains all timepoints ($T_1 \ldots T_t$) for a single subject for all electrodes ($E_1 \ldots E_m$) within a condition. B: Singular value decomposition on the cross-block covariance matrix of design and data generates two sets of vectors for each latent variable: design saliences and electrode saliences as well as its singular value. The spatiotemporal distribution of the electrode saliences for a LV is plotted here for three electrodes. C: Scalp scores and design scores are obtained by matrix multiplication of the electrode saliences with the data matrix, and the design saliences with the design matrix.

wise relate to the measures in the dataset. In this study, the exogenous block is a set of contrasts defining the experimental design. Four conditions were generated for each simulation, providing three degrees of freedom, and thus three contrasts. A set of orthonormal vectors was generated, contrasting the mean of the first condition with the average of the next three conditions; the mean of the second with the average of the next two, and the mean of the third condition with the mean of the fourth (Helmert contrasts; Figure 1A, left). Any set of three orthonormal contrasts could be used without changing the analytic outcome (McIntosh et al., 1996).

Cross-block matrix. Either the cross-block correlation or cross-block covariance matrix can be examined with PLS. To maintain the ERP amplitude information, the cross-block covariance between orthonormal design contrasts and each timepoint at each electrode in the ERP dataset was used.

Singular value decomposition. Singular value decomposition (SVD) was conducted on the cross-block covariance matrix (Figure 1B). SVD re-expresses the cross-block covariance matrix as a set of orthogonal singular vectors or *latent variables* (LVs), the number of which is equal to the number of contrasts. The LVs are analogous to eigenvectors in PCA, and account for the covariance in the matrix in decreasing order of magnitude. This magnitude is indicated by a third vector containing the *singular values* for each LV (eigenvalues). The singular values are used to calculate the proportion of cross-block covariance accounted for by a LV.

Each LV consists of a pair of vectors that reflect a symmetrical relationship between those components of the experimental design (i.e., the contrasts) most related to amplitude measures on one hand, and the optimal (in the least-squares sense) spatiotemporal pattern of ERP amplitudes related to the identified design components on the other. The numerical weights at each timepoint/electrode location combination are called *electrode saliences* (Figure 1B, far right). The electrode saliences identify the collection of timepoints that, as a group, are most related to the design effects expressed in the LV. The *design saliences* indicate the degree to which each contrast is related to the identified pattern of scalp amplitude differences.

Additional information obtained from the PLS analysis are *scalp scores* and *design scores* for each LV, which are similar to factor scores. Scalp scores indicate how strongly individual subjects express the patterns on the LV. The scalp scores are the dot product of subject's measured amplitudes and the electrode saliences on a particular LV. Although they derive from electrode saliences, we use the term scalp scores rather than electrode scores to reflect the fact that the value reflects all electrodes. Design scores are calculated in a similar fashion using the design saliences. (Note that within a condition or task, all subjects will have the same design score). Plotting scalp scores for each condition by the design scores (Figure 1C) provides a visual depiction of the experimental effect on a LV, showing which conditions are maximally distinguished.

Mathematical description of partial least squares analysis. PLS is presently implemented using MATLAB code, and the following equations describe the procedure. For the analysis of experimental effects on ERP waveforms, the m electrodes, measured across t time points, are made into single vectors for each of n subjects, who are measured in k conditions (see Figure 1A). Thus, the data matrix $\mathbf{M}$ has $n * k$ rows and $m * t$ columns. A matrix of orthonormal contrasts $\mathbf{C}$ is constructed 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. When the data matrix $\mathbf{M}$ is zeroed relative to the grand mean, the operation:

$$\mathbf{C}^{\mathrm{T}} * \mathbf{M}/(n * k - 1)$$

yields a $k - 1 \times m * t$ matrix $\mathbf{Y}$ containing the covariance of each time point for each electrode with each contrast in $\mathbf{C}$ (superscript T represents a matrix transpose).

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

$$[\mathbf{USV}] = SVD[\mathbf{Y}^{\mathrm{T}}],$$

where

$$\mathbf{U} * \mathbf{S} * \mathbf{V}^{\mathrm{T}} = [\mathbf{Y}^{\mathrm{T}}].$$

From the decomposition, $\mathbf{U}$ is an $m * t \times k - 1$ orthonormal matrix containing the electrode saliences, $\mathbf{V}$ is a $k - 1 \times k - 1$ orthonormal matrix of design saliences and $\mathbf{S}$ is a diagonal matrix of the $k - 1$ nonzero singular values. (MATLAB code for PLS is available through anonymous FTP at ftp.rotman-baycrest.on.ca/pub/Randy/pls/erp_pls).

Assessment of significance. The arbitrary decisions regarding the number of LVs to retain (e.g., scree plots) and which of the 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 electrode 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 electrode saliences from the bootstrap tests are used to assess the reliability of the nonzero saliences on significant LVs. These two tests are described below.

Statistical significance of each LV was assessed by means of a permutation test using 1,000 permutations (Braun et al., 1998; Edgington, 1980; McIntosh et al., 1996). This was accomplished using sampling without replacement to reassign the order of conditions for each subject. PLS was recalculated for each sample, and the number of times the permuted singular values exceeded the observed singular values was calculated. Exact probabilities are presented for all LVs. With the use of the permutation test probabilities, the threshold for determining the number of LVs to be considered can be justified on a statistical basis.

To determine the stability of the maximal electrode saliences identified on the LVs, the standard errors of the saliences were estimated through 200 bootstrap samples (Braun et al., 1998; Efron & Tibshirani, 1986; Fabiani, Gratton, Corballis, Cheng, & Friedman, 1998). Bootstrap samples were generated using sampling with replacement, keeping the assignment of experimental conditions fixed for all subjects. PLS was recalculated for each bootstrap sample. A salience whose value depends greatly on which subjects are in the sample is less precise than one that remains stable regardless of the sample chosen (Sampson, Streissguth, Barr, & Bookstein, 1989). The ratio of the salience to the bootstrap standard error is approximately equivalent to a z score if the normality of the bootstrap distribution is valid (Efron & Tibshirani, 1986). Those timepoints where the salience was greater than twice the standard error are indicated above the plots of the

electrode saliences. The primary purpose of the bootstrap is to determine those portions of the ERP waveforms that show reliable experimental effects across subjects; thus no corrections for multiple comparisons is necessary because 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.

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

Interpretation of PLS results. As described above, the PLS analysis provides pairs of patterns defining experimental differences in the recorded waveforms. Figure 2 contains simplified results from a PLS analysis. We show how the pieces would be put together for a single electrode from a two-condition experiment (based on a 31-electrode simulation). In the top panel, two grand average ERP waveforms are plotted, clearly showing Condition 1 latency is longer than Condition 2 latency for the early peak (points 10 to 25). Note that some small differences are also visible during the slow-wave portion of the waveforms.

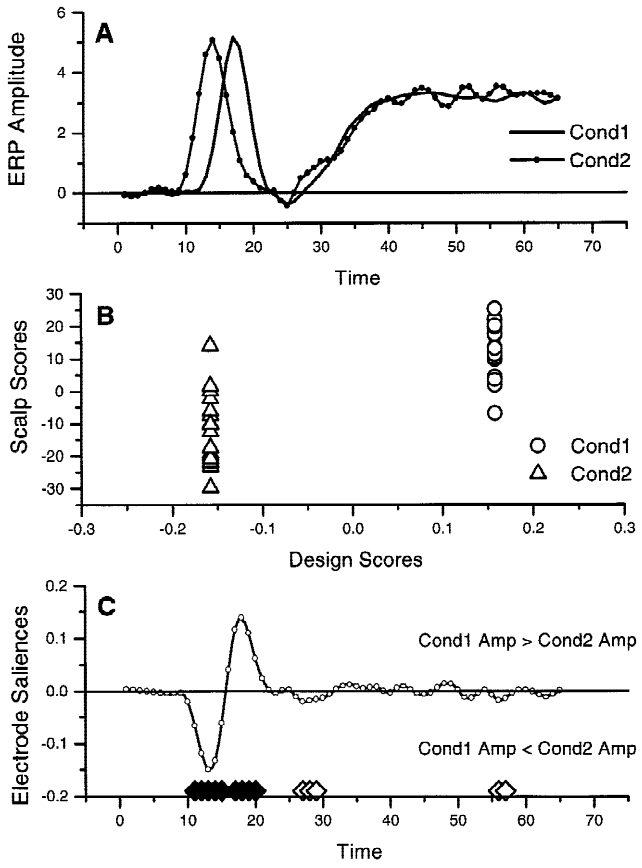


Figure 2. Interpreting PLS results. A: Grand average waveforms for two conditions at one electrode. B: If the latent variable is significant, scalp scores are plotted. This identifies the contrasts among conditions. C: Electrode saliences identify the timepoints where the contrast is strongly expressed. Reliable saliences by bootstrap are indicated at the bottom of the plot.

Following a significant statistical assessment using permutation tests, the pattern expressed on the LV can be examined. This starts with inspection of the scalp scores and design scores (Figure 2, middle panel). The pattern of scores indicates that Condition 1 amplitudes are generally higher (positive scores) than Condition 2 amplitudes (negative scores) *across the whole head*. To identify this pattern of amplitude differences, the electrode saliences (Figure 2, bottom panel) are examined in relation to the grand averages. Saliences greater than zero indicate portions of the waveform where Condition 1 amplitude is higher than Condition 2 amplitude. Saliences less than zero indicate that Condition 1 amplitude is lower than Condition 2 amplitude. The strongest saliences are for the timepoints at the first peak. They perfectly map onto the latency shift seen in the grand averages: saliences are negative from points 10 to 15, and positive from points 16 to 21. The bootstrap results indicated that nine of these amplitude differences were reliable across subjects (solid marks at the bottom of the plot).

Figure 2 also illustrates that small saliences can be identified by PLS. The onset of the slow wave is somewhat earlier in Condition 2, and there are some fluctuations at the end of the recording. Those deviations are reflected in the saliences as belonging to the pattern discriminating the two conditions. However, only small portions of those timepoints were stable by bootstrap (open marks at the bottom of the plot). When small stable findings such as these are due to artifact or random noise, they are typically associated with a single electrode or few timepoints. In most cases, the pattern of electrode saliences extends across multiple electrode sites. If so, the pattern would be interpreted to mean that the differences between Condition 1 and Condition 2 were due to a large latency shift at the first peak *combined with* small amplitude differences later in processing. The electrode saliences can be used then to identify new time periods or electrode locations for examination with more traditional analyses of latency and amplitude, eliminating guesswork regarding the significance of waveform differences. More importantly, PLS provides a scalp-wide assessment of amplitude differences, and can indicate that multiple differences in waveform amplitudes are related to experimental manipulations.

Simulation Datasets

Source waveforms. The data consisted of three, 65-point sources set along the midline (Figure 3A, B). The Brain Electromagnetic Source Analysis (BESA 2.2) software package was used to simulate the ERP and autocorrelated noise data for a full scalp topography containing 33 data channels. The simulated epoch began with an initial 11-point period with no systematic ERP activity, followed by Source 1 (points 12–20), Source 2 (points 21–36), and Source 3 (points 21–65). Source 1 reached its maximum at point 16, and returned to zero at point 20, before the onset of Sources 2 and 3. Sources 2 and 3 began at the same time. Source 2 peaked at point 29, and had returned to zero at the point where Source 3 reached its maximum (point 36). The maximal amplitude of Source 3 persisted over the remainder of the epoch. The dipoles for the three sources (Figure 3A) were set to an eccentricity of 40%; Source 1 was maximally distributed over the frontocentral region (theta −60°, phi −90°); Source 2 was maximally distributed over the parietal region (theta 40°, phi −80°); and Source 3 was maximally distributed over the fronto-polar region (theta −60°, phi −90°). The noise-free fundamental topography is presented in Figure 3C for 9 of the 33 channels.

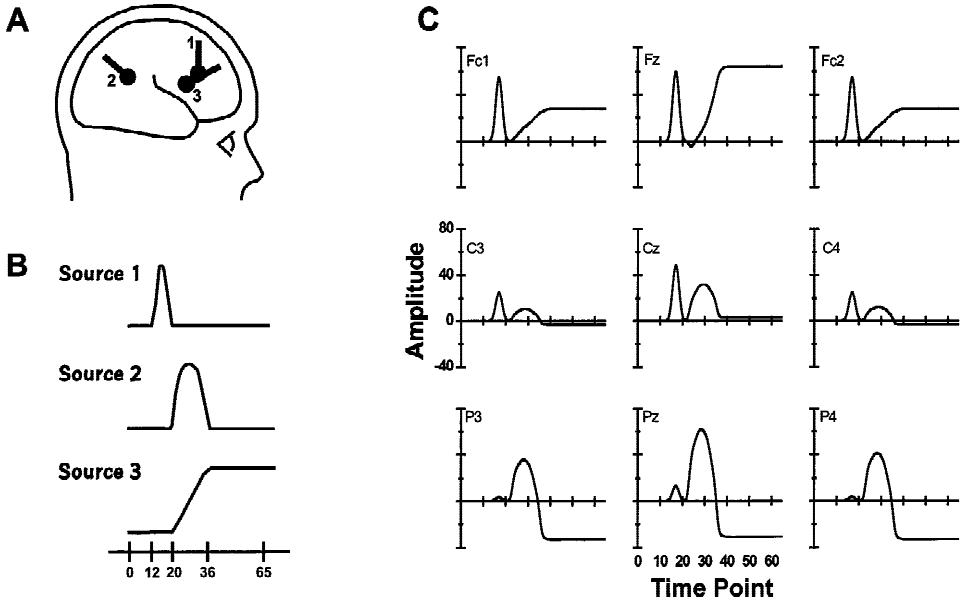


Figure 3. Simulation data. A: Location and orientation of dipoles. All three sources were placed along the midline. B: Source waveshapes corresponding to each dipole. C: Scalp topography for the baseline condition for 9 of 29 electrodes. Note that Fc1 and Fc2 show effects from Source 1 and Source 3 activity, whereas Cz shows primarily effects from Source 1 and Source 2 activity. All other electrodes express some combination of all three sources.

Experimental effects. Six experimental conditions were constructed for 20 simulated subjects. The first two conditions (B1 and B2) represented baseline conditions reflecting the initial values for the three sources. To generate these conditions, amplitude values for each source were selected from a Gaussian distribution ($M = 200$, $SD = 50$). The third condition reflected a doubling in the amplitude of Source 1 (S1*a*), selected from a Gaussian distribution ($M = 400$, $SD = 50$) for each subject. The fourth condition (S1*al*) reflected both a doubling of the amplitude of Source 1 ($M = 400$, $SD = 50$) and a 25% decrease in its onset latency (on average from point 12 to 8). The fifth condition was a doubling of the amplitude of Source 2 (S2*a*, $M = 400$, $SD = 50$). The sixth condition (S2*a*S3*a*) simultaneously increased the amplitude of Source 2 ($M = 400$, $SD = 50$) and decreased the amplitude of Source 3 ($M = 100$, $SD = 50$). The inclusion of this condition permitted a consideration of whether PLS could separately identify these two sites of changing activity. Where no experimental effects were modeled, any differences between conditions resulted from chance variation.

Subject variability. Between-subject temporal jitter in source activity onset was accomplished by varying the onset latency of the sources for each of the 20 simulated subjects. The onset latency was based upon a Gaussian distribution ($M = 0$, $SD = 5$) and this value was added or subtracted from the fundamental source activity onset. For Sources 1 and 2, the termination of the source was also decreased or increased by the same amount. This was not necessary for Source 3, as its activity persisted over the epoch. The individual simulated subject source-onset points were maintained across all simulated experimental conditions.

Spatially and temporally autocorrelated noise, designed to reflect ongoing EEG, was also simulated. The BESA regional-source model (RS4.par) was chosen, with four sources placed in the left and right anterior and posterior quadrants. The same noise level was used for each of the 12 dipoles making up the regional sources, and was based upon a Gaussian distribution ($M = 2.5$, $SD = 1.25$). The noise level was maintained for each simulated subject across the six simulated conditions, and unique noise was generated for each condition. Spatially and temporally uncorrelated noise, designed to reflect artifact arising from recording equipment and other nonbiological sources, was simulated using a random numbers generator ($M = 0$, $SD = 0.5$). As with the autocorrelated noise, separate noise files were generated for each simulated subject and condition. The uncorrelated noise files were average referenced to meet the convention of BESA before adding them to the autocorrelated noise and source waveform files.

Subject data files were constructed by adding the appropriate unique autocorrelated noise file and uncorrelated noise file to the source waveform file for each subject in each condition. For plotting convenience, two lateral electrode positions in the simulation montage were dropped from the dataset (FC5, FC6), leaving 31 channels for analysis.

Design of simulations. Two simulations with four experimental conditions in each were generated by selectively combining the six different experimental effects. Baseline 1 (B1) and Baseline 2 (B2) were included in each of the simulations to reflect situations where an experimental manipulation does not result in ERP differences. The remaining two conditions in each simulation reflected experimental effects on ERP amplitudes. In Simulation 1, the experimental effects were restricted to Source 1. Here we incorporated an amplitude increase in one condition (S1*a*) and a combined amplitude increase and latency decrease in another (S1*al*). For Simulation 2, effects at temporally overlapping sources were examined. This simulation examined an amplitude increase at Source 2 (S2*a*) in one condition, and the combination of a simultaneous amplitude increase at Source 2 and decrease at Source 3 in another (S2*a*S3*a*).

PCA

As the primary expression of the source activity was at the midline, temporal PCA was conducted on three midline electrodes from Simulation 2 (Fz, Cz, and Pz). The columns of the data matrix were the 65 timepoints, with electrodes and conditions for each subject represented in the rows. SVD was applied to the 65×65 covariance matrix. PCs explaining the first 95% of the covariance matrix were retained for further analysis.

The significance of the PCA results was assessed in two ways. First, for each retained PC, scores for each experimental condition were created for each subject, averaged across electrodes. Significant differences among the score means were identified by paired *t* tests (corrected for multiple comparisons, $p_{crit} = .008$).

Permutation tests were also conducted on this dataset. This allowed us to use the same statistical criteria to compare the PCA results with that of PLS. The statistic assessed on the PCA results was the R^2 obtained from regressing a set of orthonormal design contrasts on the PCs from the permuted datasets. PCs having an R^2 probability less than .05 were considered to be significant. Bootstrap estimates for the component weights were also generated to identify which timepoints on the PC reflected stable factor loadings.

Actual Event-Related Potential Data

Participants and design. Data from two single-feature conditions in a feature-conjunction auditory-oddball paradigm were used to assess PLS in actual ERP data. Participants were 14 young adults (23 ± 4 years, 8 females) who were participating as controls for a larger patient study. Participants attended to loudness-matched tones (e.g., 500 and 1500 Hz, 100-ms duration with 5-ms rise/fall time) presented randomly to left or right ear via headphones. In separate blocks, the target dimension was either pitch (high or low) or location (left or right ear). Target dimension was counterbalanced across participants. In both conditions, 720 trials were presented, and targets occurred on 25% of all trials ($n = 180$).

Hits were defined as button presses between 200 and 1,000 ms after target onset. Correct rejections were defined as "no response" 200 to 1,000 ms following a distractor. False alarms were button presses occurring outside the hit window. Misses corresponded to no response between 200 and 1,000 ms after target presentation.

ERP recording. The electrophysiological signals were digitized continuously (256 Hz sampling rate per channel; bandpass 0.15–40 Hz) from 32 electrodes using NeuroScan Synamps and software (v. 3.1), and archived for off-line analysis after each session. Eye movements were recorded with electrodes at the outer canthi and at the superior and inferior orbit. All electrodes were referenced to Cz during the recording; the data were re-referenced to an average reference off-line and digitally filtered using a 32 Hz lowpass filter.

ERP analysis. Before averaging the ERPs, the ocular artifacts associated with blinks were corrected using an ocular source component approach (NeuroScan). Averaging occurred off-line, following the computerized automated rejection of trials contaminated by excessive peak-to-peak deflection, or amplifier saturation (± 100 μV). The epoch included a 200-ms prestimulus baseline and 800-ms poststimulus. In each condition, the ERPs were averaged separately for hits, correct rejections, false alarms, and misses.

For the PLS analysis, subject grand averages for hits and correct rejections from 28 electrodes were used (the four eye movement channels were not included). The epoch analyzed consisted of 205 timepoints starting at stimulus onset (0–800 ms).

Results

Simulation 1: Latency and Amplitude Effects at a Single Source

The grand averages from the four conditions (B1, B2, S1*a*, and S1*al*) are shown in Figure 4A (9 of the 29 electrodes are plotted). The first and second LVs accounted for 67.1% and 31.1% of the cross-block covariance matrix, respectively, and both were significant by permutation test ($p = .00$). The third LV accounted for the remaining 1.8% of variance, and was not significant ($p = 1.00$).

LV1. Plots of the scalp scores by design scores for LV1 are shown in Figure 4B, left panel. The values for the design scores (DS) indicate the primary distinction on this LV was between the S1*al* (DS = +0.19) condition and the other three conditions (DS < −0.06). The electrode saliences for this LV (Figure 4C, thick lines), indicate the direction and the magnitude of the differences, simultaneously at all scalp locations. To reiterate, the electrode saliences simply indicate which timepoints and electrodes showed amplitude differences related to the contrast shown in the design scores. The deflections should not be interpreted as separate waveform components.

Visual inspection of the saliences indicates the ERP differences were restricted to those timepoints where Source 1 was active. The shape of the saliences identifies both shorter latency (positive followed by negative saliences) and the increased amplitude (stronger positive saliences than negative) in condition S1*al* relative to the other three conditions. This pattern was expressed most strongly at frontal and central electrodes. The nonzero saliences stable by bootstrap are indicated at the top of each of the plots. The bulk of the stable timepoints were located in the 10 to 20 point range, consistent with the simulation. Thus, the experimental effect of a combined increase in amplitude and latency shift at Source 1 was depicted by a reversal of the saliences within the time period the source was active.

LV2. Plots of the scalp scores by design scores for LV2 are shown in Figure 4B, right panel. The design score weights indicated the primary difference was between the S1*a* condition (DS = +0.18) and the two baseline conditions (DS = −0.07, −0.11). Note the weights for the S1*al* condition were very close to zero (DS = −0.01), indicating its contribution to this LV was minimal. Although the saliences for this LV (Figure 4C, thin lines) resemble a "peak," what they indicate is that the amplitude in the S1*a* condition was higher than in the baseline conditions. The topography of the saliences matched the topography for Source 1, and the bootstrap results indicated the effect was most stable between points 10 to 20.

In summary, the PLS on this dataset distinguished the two experimental effects. The strongest effect was the combination of an increase in amplitude and a latency decrease in condition S1*al*. Any experimental manipulations that produce combined amplitude and latency differences will have a shape that resembles LV1. If both aspects are reliable by bootstrap (as in the present simulation), then one can confidently claim the two effects occurred. The second LV identified the second experimental effect, the increase in amplitude in condition S1*a*. At all other timepoints, the saliences were essentially zero, indicating the other peaks in the recorded waveform did not differ among conditions.

Simulation 2: Spatiotemporal overlap

The grand averages for the four conditions (B1, B2, S2*a*, and S2*a*S3*a*) are shown in Figure 5A. Due to the orientation of the

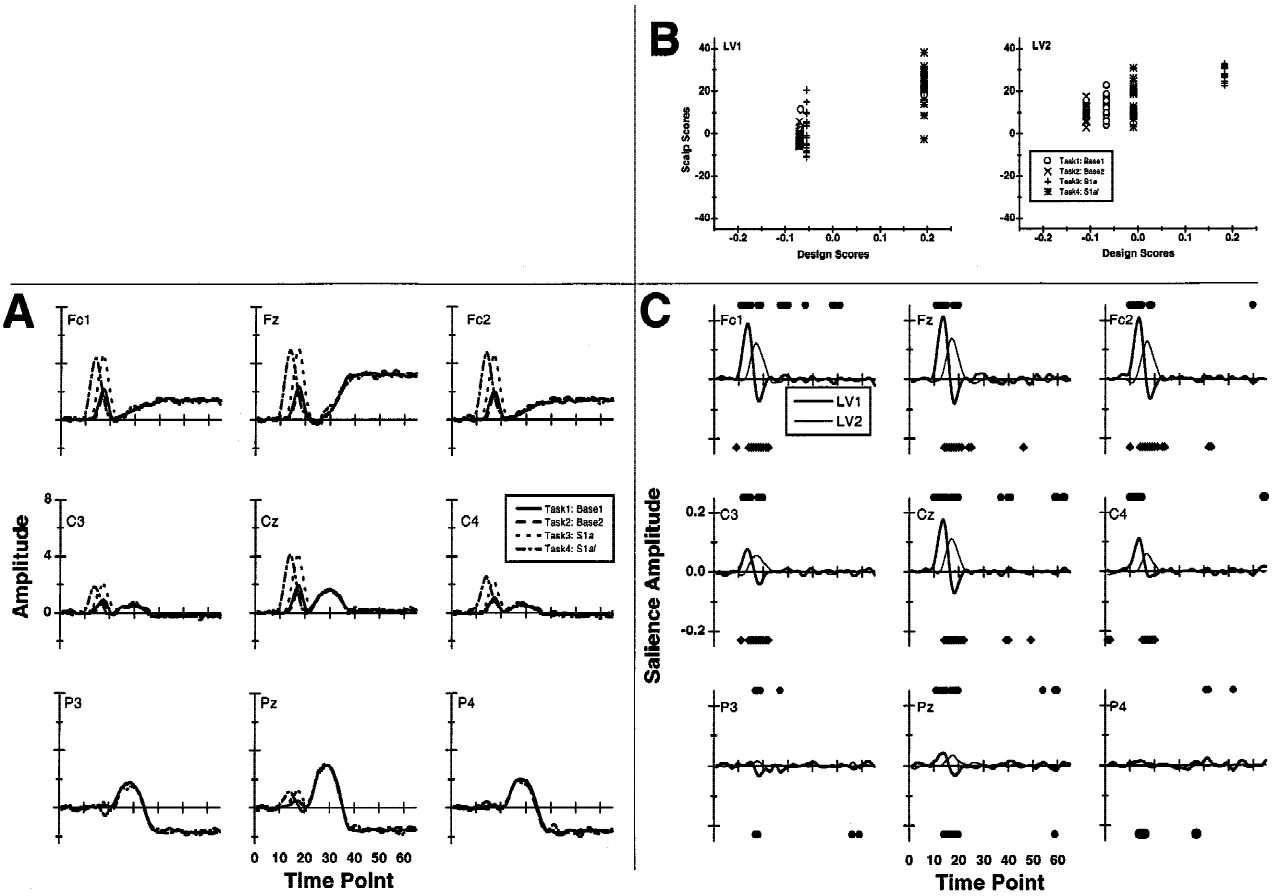


Figure 4. Simulation 1: Latency and amplitude effects at a single source. A: Grand averages for 9 of 29 electrodes. Both effects are clearly seen in the S1*al* condition (– · –), while the S1*a* condition (– –) contains only the increased amplitude. B: Left panel: Scatterplot of scalp scores by design scores for LV1. The weight of the design scores for each condition can be used to identify the contrasts being expressed on the LV. LV1 design scores indicates the S1*al* condition (∗) is different from the other three conditions. The overall positivity of the scalp scores in the S1*al* condition indicates the electrode saliences are likely to be mostly positive. Right panel: Scatterplot of scalp scores by design scores for LV2. The weight for the S1*al* condition (∗) is zero, indicating it is not contributing to this contrast. Thus, this LV is contrasting the S1*a* condition (+) with the baseline conditions (×,○). The effect is largest for the difference between the S1*a* condition and the second baseline condition. This reflects small differences between the two baseline conditions resulting from the random assignment of peak amplitudes. C: Electrode saliences for LV1 (▬) and LV2 (—), rescaled by the LV singular value. Stable peak saliences are identified at the top of each plot for LV1 (●) and at the bottom of each plot for LV2 (♦). For LV1, the saliences identify latency and amplitude effects, expressed most strongly at the frontal electrodes. The weights are primarily positive, which when combined with the positive loading on the LV1 design scores, correctly identifies the increased amplitude and shorter latency in the S1*al* condition. For LV2, the peak saliences are also positive, again indicating a positive relation with the design contrasts. This correctly identifies the higher amplitudes in the S1*a* condition compared to the baseline conditions.

sources, the scalp potentials from Source 2 and Source 3 also had spatial overlap at most electrode sites (cf. Fz, Pz). However, this overlap did not occur at Fc1 and Fc2, where only effects of Source 3 were expressed, nor at C3, Cz, and C4, where ERP differences related to Source 2 activity were expressed.

The first and second LVs accounted for 84.1% and 14.9% of the cross-block covariance matrix, respectively, and both were significant by permutation test ($p = .00$). The third LV accounted for the remaining 1.0% of variance, and was not significant ($p = .34$).

LV1. The S2*a*S3*a* condition consisted of two simultaneous effects, an increase in amplitude of Source 2 and a decrease in amplitude of Source 3. This contrasts with the single effect in condition S2*a*, which had only the increased amplitude at Source 2. It would be expected that the experimental effect differing the most from the two baseline conditions would be the S2*a*S3*a* condition. Plots of the scalp scores by design scores for LV1, shown in Figure 5B, left panel, indicated a more complex pattern of differences among conditions, but supported this expectation. The design scores showed strong positive weights for the S2*a*S3*a* condition (DS = 0.16) and equal, strong negative weights for the two baseline conditions (DS = −0.11). This indicated the primary difference being expressed on LV1 was between the S2*a*S3*a* condition and the two baseline conditions. However, the positive score for the S2*a* condition (DS = 0.06) indicated it also differed from baseline. Because PLS groups similar differences together, LV1

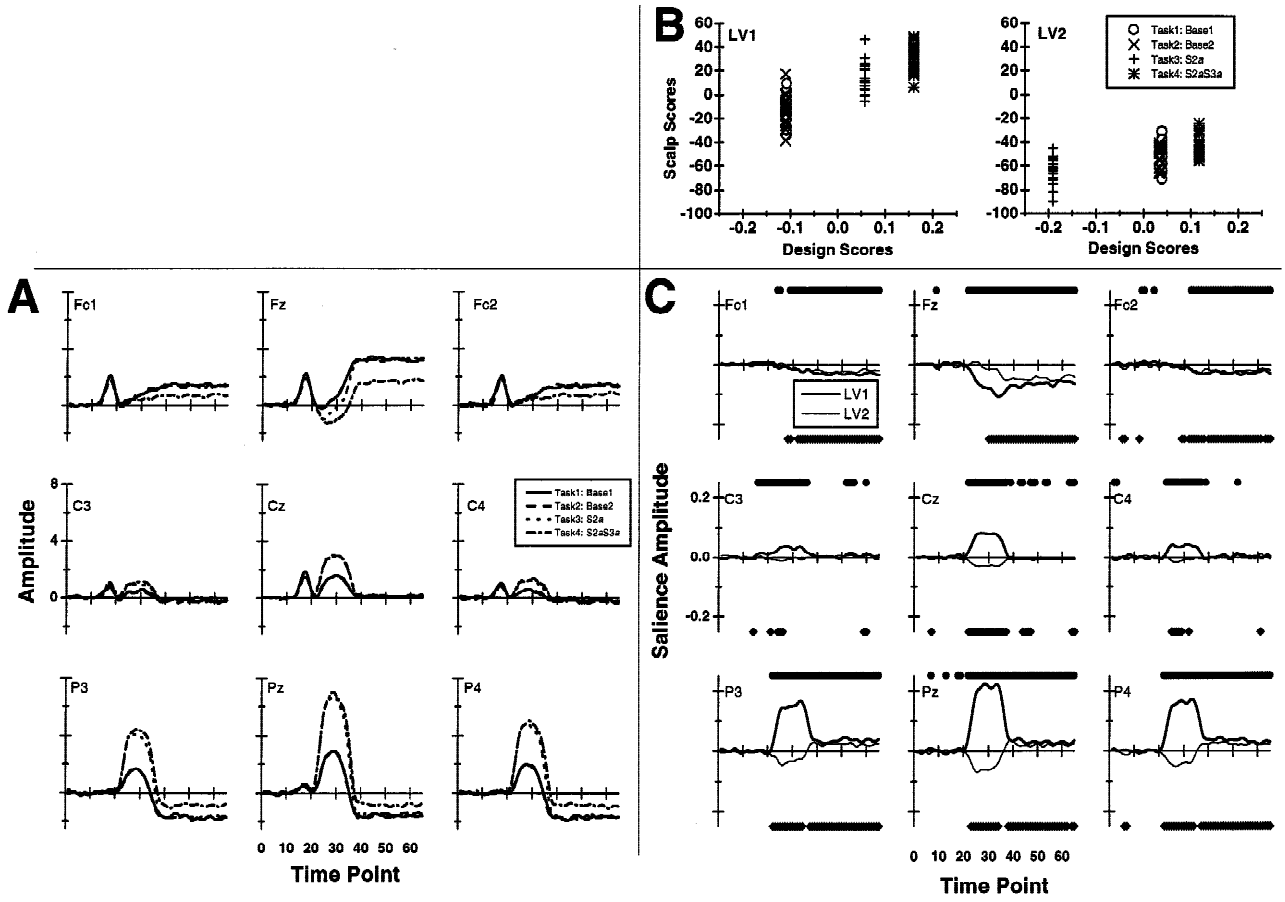


Figure 5. Simulation 2: Experimental effects at sources with spatiotemporal overlap. A: Grand averages. The increased amplitude at the second peak in the S2*a* condition (– –), and the combined effect of increased amplitude at Source 2 and decreased amplitude at Source 3 in the S2*a*S3*a* condition (– · –) are shown. B: Left panel: Scatterplot of scalp scores by design scores for LV1. The largest discrimination on this LV is between the S2*a*S3*a* condition (∗) and the two baseline conditions (×,○). This LV also indicates that the S2*a* condition (+) differs from the two baseline conditions. Right panel: Scatterplot of scalp scores by design scores for LV2. The largest discrimination on this LV is between the S2*a*S3*a* condition (∗) and the S2*a* condition (+), with additional contributions from the difference between the two baseline conditions (×,○) and the S2*a* condition (+). C: Electrode saliences for LV1 (—) and LV2 (—), rescaled by the LV singular value. Stable peak saliences are as in Figure 3. For LV1, the saliences indicate stable peak effects at the second peak and at the slow wave. For LV2, the effects are also expressed at the second peak and at the slow wave, but the relationships are reversed for the second peak. As was seen in the grand means, Fc1 and Fc2 only show stable saliences in the slow wave, whereas C3, C4, and Cz show strong saliences only at the second peak. All other electrodes reflect the combination of ERP amplitude differences.

indicated that the S2*a*S3*a* and S2*a* conditions shared some common differences from the baseline conditions (that is, the amplitude effect at Source 2).

The pattern for the electrode saliences for this LV reflected this combination of differences among conditions (Figure 5C, thick lines). Interestingly, where Sources 2 and 3 combined at the scalp, both signals were incorporated into the LV. Where the sources had little or no overlap (e.g., Cz for Source 2, Fc1 and Fc2 for Source 3), the saliences reflected effects of the single source.

The strongest saliences were associated with the increased amplitude of Source 2 from points 20–36. These were seen as positive saliences at the posterior electrodes (e.g., Pz), which reversed at frontal-polar sites (not plotted). The remaining portions of the electrode saliences reflected differences in the slow wave. At electrodes where the recorded scalp activity reflected activity from Source 3 but not Source 2 (Fc1, Fc2), PLS identified the experimental effect at the slow wave. As the design scores were positive for the contrast between the S2*a*S3*a* and baseline conditions (S2*a*S3*a* < baseline), the negative saliences for these points indicated the effect of the decreased amplitude at Source 3.

LV2. The scores for LV2 are shown in the right panel of Figure 5B. As for LV1, the scores for LV2 present a complex picture. The primary distinction was between the S2*a* (DS = −0.19) and S2*a*S3*a* (DS = 0.12) conditions, which at the level of their sources, differed only in the slow-wave amplitude at Source 3. However, the scores for the two baseline conditions were also positive (DS = 0.03), indicating they were also different from the S2*a* condition. This pattern of scores suggested that this LV indexed differences between S2*a*S3*a* and S2*a* as well as similar differences between the S2*a* condition and the baseline conditions that were not accounted for in LV1.

The electrode saliences for LV2, shown in Figure 5C (thin lines), reflected a combination of differences at the second peak and during the slow wave. The negative saliences at the central electrodes, taken with the negative design scores for the S2*a* condition, identified the modeled *increased* amplitude in S2*a* relative to the baseline conditions (e.g., points 22–37 at Cz). The centroparietal and parietal electrodes expressed a combination of effects: negative saliences during the second peak, and positive saliences during the slow wave. The frontal electrodes expressed differences primarily reflecting the slow-wave activity.

This again demonstrates that when multiple effects in the scalp amplitudes define differences among experimental conditions, the LV will contain those multiple effects. When the topography allows a dissociation of those effects, they can be identified on the plots of the saliences, offering a suggestion to the researcher that sources with different scalp topography may be responsible. Interestingly, note that the saliences for LV1 and LV2 at all electrodes were essentially identical during the slow-wave time period. We interpret this to indicate the differences between the S2*a*S3*a* condition and the two baseline conditions (the primary distinction on LV1) were of the same magnitude as the differences between the S2*a*S3*a* and S2*a* conditions (the primary distinction on LV2).

Comparison with PCA

To determine if temporal PCA and PLS would produce similar conclusions, data from three midline electrodes (Fz, Cz, Pz) from Simulation 3 were analyzed with both techniques. The findings for the PLS analysis on the reduced dataset were equivalent to those shown in Figure 5C for the full dataset. The first two PCs accounted for 93.9% and 5.7% of the covariance matrix, respectively. Only the first two PCs were significant as assessed by permutation tests, and are shown in Figure 6A and B.

Because PCA solutions are not constrained to identify a particular source of variance, they may reflect differences or commonalities among conditions, or some combination of these effects. This was evident in PC1, where effects from all three sources can be seen (Figure 6A). Comparisons of the scores for PC1 indicated the baseline conditions ($M = -0.65, -0.66$) differed from the two experimental conditions ($M = 1.74$ for S2*a* and $M = 2.78$ for S2*a*S3*a*) and the two conditions differed from each other. Bootstrap estimates of the PC weights, plotted under the PC, indicated all nonzero timepoints after point 16 were stable. This included the weights associated with Source 1 activity. PC2 (Figure 6B) indicated effects resulting from amplitude differences in Source 2 and Source 3. As for PC1, the two baseline conditions ($M = 5.87, 5.93$) differed from the two conditions with added effects ($M = 10.16$ for S2*a* and $M = 8.72$ for S2*a*S3*a*), and the two conditions with effects differed from each other. Bootstrap estimates indicated that all points after point 21 were stable.

The shape of the second PC was most similar to the PLS results, in that it only identified differences during the second peak and the slow wave. However, the statistical comparison of the scores for both PC1 and PC2 indicated significant differences among conditions. Thus, the experimental effects would have been incorrectly assigned to the early peak as well as to the later peak and slow wave.

Because the first PC accounted for 94% of the variance, the two vectors would typically not be rotated. After Varimax rotation (Figure 6C), Source 1 activity was seen on both PC1 and PC2. [This was also true when the third, nonsignificant PC (0.27% of variance) was included.] Analysis of the scores from the rotated vectors again indicated significant effects for both PCs. For rotated PC1, S2*a*S3*a* differed from all other conditions, and mapped onto the modeled differences. For rotated PC2, S2*a* and S2*a*S3*a* differed from the two baseline conditions, but not from each other. The point here is that although rotation extracted the modeled experimental manipulations as simple effects, some of that difference would be attributed to amplitude differences during Source 1 activity, where no differences were modeled.

Actual ERP Data

Analysis of the behavioral measures indicated performance was similar in the location and pitch conditions. For the location condition, the proportion of hits (*SD*) was 97% (2.5%), and the mean number of false alarms was 3.6 (2.5), and mean reaction time was 327 (49) ms. In the pitch condition, the mean number of hits was 98% (1.6%), with a mean of 3.2 (2.1) false alarms, and 313 (38) ms mean reaction time.

The grand averages for ERPs in the four conditions for the auditory oddball task are presented in Figure 7. Visual examination of these data identified a large positive deflection from 200 to 600 ms, most strongly expressed at Pz in the two hit conditions

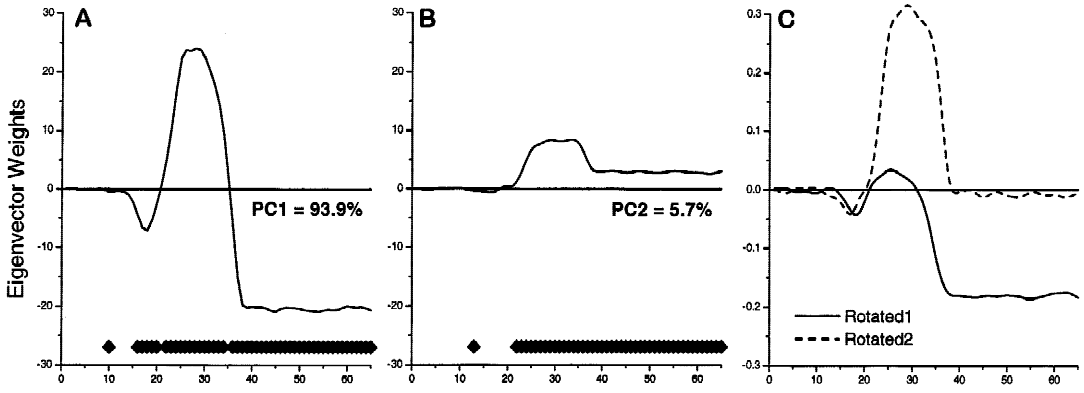


Figure 6. Simulation 3: Temporal PCA results. Points stable by bootstrap are indicated below the waveform (♦). A: PC1 accounted for 93.9% of the covariance and reflected effects at all three peaks. B: PC2 accounted for 5.7% of the covariance and reflected effects related to the second peak and the slow wave. C: Results of varimax rotation of the first two principal components. Rotated PC1 reflected effects at the first and third peaks; rotated PC2 reflected effects at the first and second peaks.

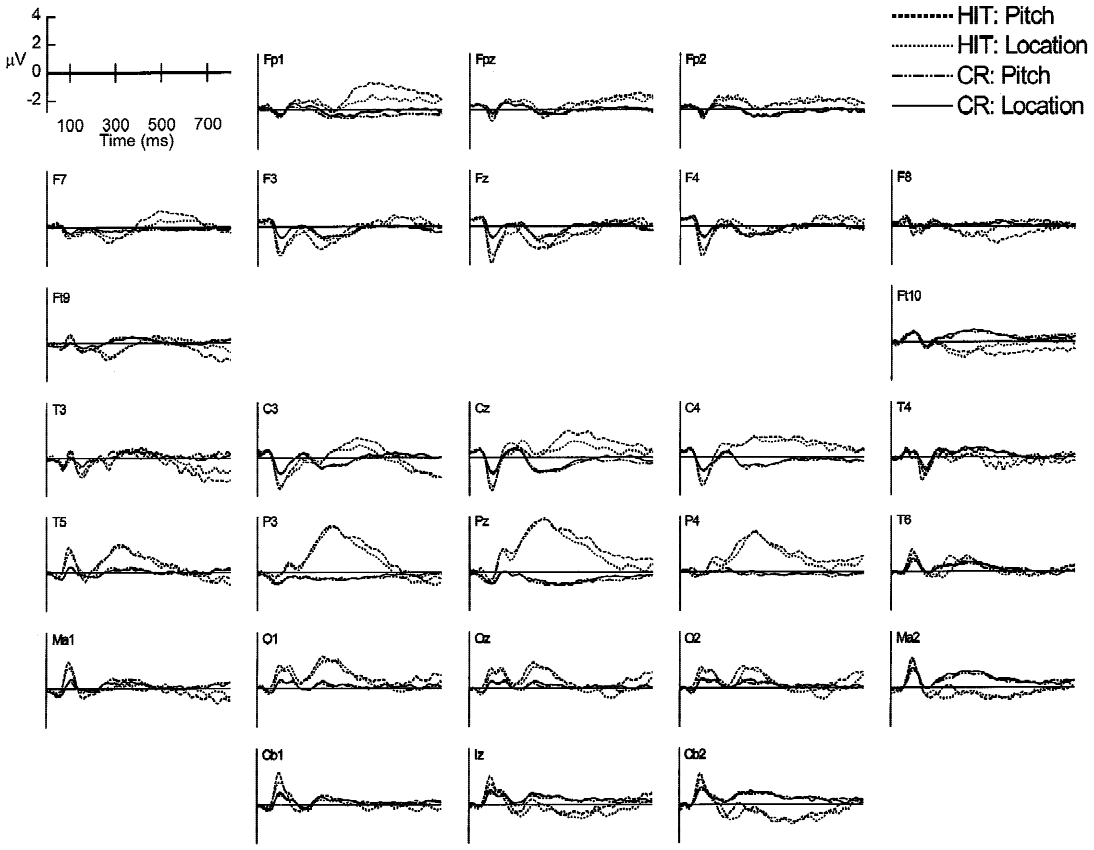


Figure 7. Grand averages for actual ERP data: Auditory Oddball task. Targets were distinguished from nontargets based on pitch or location. Positive is up. Data from four conditions are plotted: Hits for Pitch (– – –) and Location (- - -); and Correct Rejections for Pitch (– · · –) and Location (—).

compared to the two correct rejection conditions. Additionally, there were differences at 100 ms at a number of electrodes, and slow-wave differences after 400 ms at frontal and posterior electrodes. The goal of the PLS analysis, then, is to identify that particular combination of scalp amplitudes related to specific distinctions among the four conditions. The full sweep from 0 to 800 ms was analyzed, and three Helmert contrasts were used to create the covariance matrix. The first LV accounted for 92.5% of the cross-block covariance, and was significant by permutation test ($p = .000$). The second and third LVs accounted for 6.8% and 0.7% of the cross-block covariance, respectively, but were not significant ($ps = .714$; 1.00), so they were not considered further.

LV1. The design scores (Figure 8A), indicated the primary contrast in this dataset was between the two correct rejection conditions ($DS = -0.13$) and the two hit conditions ($DS > +0.11$). The smaller range in scalp scores for the correct rejection conditions indicated the ERP amplitudes were more similar among subjects in those two conditions compared to their ERP amplitudes in the two hit conditions. (Examination of the scalp scores can also be used to identify potential outliers.) A combination of factors may have contributed to the differences in variability between hit and correct rejection scores. First, there were fewer target trials compared to nontarget trials (180 versus 540). Secondly, it is likely that the ERPs to target stimuli reflected not only the perceptual response to the deviant stimulus, but also some aspects of the preparation and execution of the motor response.

The electrode saliences are shown in Figure 8B, with stable peak saliences based on bootstrap at the top of each electrode plot. Positive electrode saliences indicated that the difference expressed on this LV was stronger amplitudes to hits compared to correct rejections. The pattern of electrode saliences indicated that a combination of differences in scalp potentials distinguished hits and correct rejections. The strongest saliences were at those electrodes expressing the P300 effect (i.e., Pz, P3, P4, and C4). The bootstrap results indicated these saliences were stable for a large portion of the 200–600 ms time period at those electrodes. The early peak at 100 ms also showed stable saliences across the head. These saliences were positive at the back of the head and negative at the front of the head, and the shape of the saliences indicates an amplitude difference, which is confirmed by examining the grand averages (e.g., T5). The magnitude of this difference was somewhat larger on the left than the right side (e.g., T5, Ma1 versus T6, Ma2). Differences were also seen for the second early peak at 150 ms at O1, Oz, and O2, and at P3, Pz, and P4. Slow-wave saliences were stable for the central and right posterior as well as right temporal electrode sites (e.g., Iz, Cb2, Ma2, and Ft10), but this effect was not as strong at frontal electrodes (Fp1, Fpz, and Fp2). The late negativity (>700 ms) at left temporal sites in the hit conditions was also identified (Ft9, T3, T5, Ma1).

LV2. This LV was not significant by permutation test, and we are commenting on those results only to demonstrate that PLS can be used to discount apparent differences in the wave-

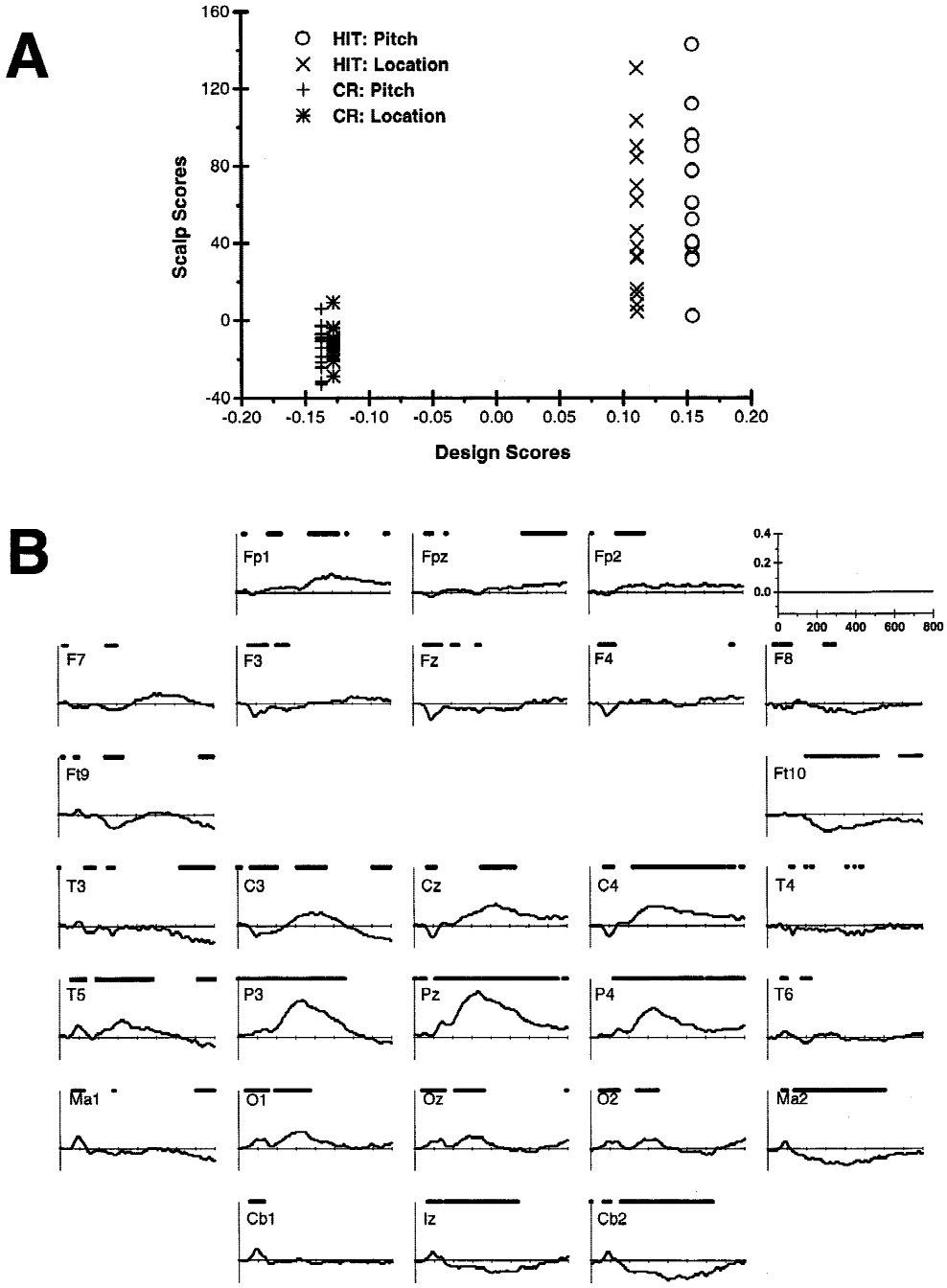


Figure 8. PLS results for actual ERP data. A: Scatterplots of scalp scores by design scores for LV1. This LV contrasts the two correct rejection conditions (+, ∗) from the two hit conditions (×,○). Note the tighter clustering of the scalp scores in the correct rejection conditions compared to the hit conditions, indicating more subject variability on hit trials. B: Electrode saliences for LV1, with stable peak saliences shown at the top of each plot (●). This LV expressed a combination of effects across the scalp. The strongest saliences are seen at the central-parietal electrodes, reflecting the P300 differences. Two different kinds of slow-wave differences were seen, one at posterior electrodes (e.g., Ma2, Cb2) and the other at frontotemporal electrodes (e.g., Ft9, T3, T5). Early peak differences at 100 and 150 ms also distinguished hits from correct rejections. This distinction was seen as a single peak at temporal sites and as overlapping peaks at occipital electrodes.

forms. The design scores on LV2 indicated the primary difference was between the two hit conditions. In fact, some differences between pitch and location hits appear in the grand averages (e.g., Cz). However, because this LV did not reach statistical criterion for significance, we conclude that there are no reliable amplitude differences when detecting targets based on pitch compared to location in this sample. ANOVA on the amplitude and latency of the P300 at Pz also indicated no significant differences between the two hit conditions, $F(1,13) = 0.129$ and 0.018, respectively.

Discussion

This paper introduced the application of the multivariate Partial Least Squares technique to the full spatiotemporal extent of event-related potential data. PLS shares some features with other multivariate tools, such as PCA, in that it makes use of the singular value decomposition algorithm to extract information from the dataset. It differs from PCA in that the solutions are constrained to reflect the part of the variance in the dataset associated with amplitude differences due to experimental manipulations. The resulting pairs of vectors (design and electrode saliences) define the commonalities between the ERP measures and the experimental design. Additionally, we provide mechanisms to assess the significance of the LVs as a whole (permutation tests), and subsequently assess the stability of the maximal saliences (bootstrapping). This type of statistical assessment could equally be applied to PCA (Braun et al., 1998).

To show the range of effects PLS can identify, the results of two simulations were presented, as well as a direct comparison with PCA. In all cases, the amplitude differences identified by PLS were restricted to those electrodes and timepoints that reflected the experimental differences. The first simulation applied two separate experimental effects to a single peak. PLS separated differences due to increased amplitude from those due to a combination of increased amplitude and decreased latency. The second simulation was designed to demonstrate that PLS identifies differences in the scalp recordings, rather than differences in source waveforms. In this simulation, experimental effects were placed on portions of the waveform where the scalp topographies reflected the temporal overlap of two underlying sources. Here, the differences identified by PLS reflected combined effects, rather than singularly identifying the differences in the source amplitudes.

Temporal PCA on this same dataset produced a PC (PC2; Figure 6B) that had a shape very similar to that of the electrode saliences identified by PLS. Statistical comparisons of the scores on PC2 identified the correct set of differences among experimental conditions. However, PC1 also identified reliable effects (Figure 6A) that corresponded to all nonzero portions of the waveform (this was true even when more stringent criteria were applied). As no experimental effect was placed on the early peak, this PC likely reflected some combination of commonalities and differences among the four conditions. Statistical analysis of the scores indicated that the two conditions with experimental effects differed from each other and from the two baseline conditions. Thus, experimental effects would be incorrectly attributed to all measured peaks. Varimax rotation of the PCA solution identified the two simple effects, but did not eliminate the erroneous weights associated with the first peak.

Application to Actual ERP Data

The subject, electrode, and peak onset variability in the simulation studies was added to the experimental differences to approximate intersubject variability that would be found in actual ERP data. The simulation studies indicated that PLS was not affected by this variability. The real power of tools such as PLS is in their ability to explain differences in actual scalp recordings. We chose to do this on data from an auditory oddball task, with large differences between the experimental conditions. These particular data were part of a larger experiment designed to extend the findings in the literature by directly contrasting the effects of attention to pitch with attention to location. As is typically seen in oddball paradigms, the largest amplitude differences were between hits and correct rejections at central/parietal electrodes—the "P300" effect. PLS identified this effect as the major source of variance in the cross-block covariance matrix (LV1), as the electrode saliences were strongest at this time period. The discrimination between hits and correct rejections resulted from amplitude differences at multiple scalp locations and timepoints. Thus, in addition to the differences related to the P300, PLS identified early amplitude differences at 100 and 150 ms, as well as additional slow-wave differences at 300 to 600 ms, and late effects at 700 to 800 ms. The distinct spatial topographies of the electrode saliences indicated that multiple neural sources contributed to the observed waveform differences. Differences on hit trials when attending to pitch versus location were not statistically significant (LV2).

For this initial demonstration, we purposely selected an ERP dataset with clear distinctions in the waveforms. It is important to note that PLS would be expected to identify any effect that would be typically be seen in more traditional analyses of peak amplitudes and latencies. One advantage of PLS is that because all electrodes are assessed, new information may be obtained that would not be seen otherwise. A second, and equally important feature, is that the PLS results indicate which differences in the ERP waveforms "hang together." In the present dataset, it would not be sufficient to describe the results in terms of simple differences in the P300 or P100 components because the differences between hits and correct rejections were in fact a combination of differences at multiple timepoints and electrodes.

The PLS analysis, as presented in this paper, can be considered to conduct an omnibus test of the major relations between the experimental design and the measured amplitudes, and as such, we have used Helmert contrasts to reflect the experimental design. One of the strengths of PLS is that it is also an excellent tool for exploratory data analysis (Martens, Izquierdo, Thomassen, & Martens, 1986), so if one had specific questions regarding relations between task conditions, the design and data matrices can be generated to address those questions. So, for example, many ERP effects sensitive to attentional manipulations are seen in the early peaks. In this case, the researcher may decide to limit the analysis to this time period. Another question raised in ERP research concerns laterality effects, in which case the data and design matrices can be organized to contrast both experimental condition and hemisphere. In fact, there is a suggestion in the present results that the differences between hits and correct rejections are more strongly expressed over the left hemisphere. We have conducted PLS analyses on this dataset to examine both types of questions, and those data will be presented more fully in a future paper.

Data Considerations

As is the case for any analysis technique, PLS works best when the data are relatively clean. In situations where there is large variability within condition across the recorded waveforms (e.g., either peak amplitudes or latencies), PLS will be less able to detect significant experimental effects in the scalp waveforms. This is likely to be especially true in patient populations. In fact, our first application of PLS to ERP data was in a study comparing ERP measures of visual selective attention in two groups of children. A group diagnosed with Attention Deficit Hyperactivity Disorder was compared to a group of control children (Lobaugh, Taylor, & McIntosh, 1997). The high variability in the patient group produced results that were difficult to interpret, and stimulated the current simulation study.

Relations of Scalp Differences to Source Activity

It has been known for some time that linear decomposition models, such as PCA, ICA, and SVD, cannot distinguish underlying sources contributing to separate PCs without added constraints or assumptions (e.g., Borgen & Kowalski, 1985; Sylvestre, Lawton, & Maggio, 1974). PLS also does not do this. What it does do is identify distinct combinations of scalp potentials that define contrasts between two or more experimental conditions. Conceptually, the LV electrode saliences frequently represent *weighted* difference waveforms, which are assessed statistically and simultaneously. In the rare case where only one part of the scalp topography differs between conditions, PLS will identify that single topography. This was seen in Simulation 1, where the topography accurately reflected the activity of a single neural source. In most cases, ERP differences related to experimental manipulations will be the result of changes in multiple neural sources. So for example, Simulation 2 contained two experimental effects (S2*a* and S2*a*S3*a*) having both spatially and temporally overlapping scalp topographies. Any PLS contrast that simultaneously identified these two effects would by definition also have electrode saliences expressing these two topographies. Thus, both LVs reflected *combinations* of amplitude effects, and we were able to state that both the slow-wave portion of the waveform and the second peak were affected by the experimental manipulations.

Extensions and Future Directions

The present simulations and application to real ERP data were applied to a single experimental group with multiple conditions. PLS is also applicable to designs where two groups are contrasted (Cabeza et al., 1997; Grady et al., 1998; Lobaugh et al., 1997). Additionally, if the research questions of interest relate to the strength of a *specific* set of contrasts, these can be used in lieu of the Helmert contrasts presented here. The only restriction is that they must be orthonormal. We note though, that any set of orthonormal contrasts that account for all of the degrees of freedom will produce the same outcome with PLS (McIntosh et al., 1996).

PLS is not restricted to pairwise comparisons, and frequently the LV results will indicate multiple differences in scalp potentials. This is problematic only when one wishes to assign a specific part of the contrasts being expressed (e.g., A > B versus A > C when the LV indicates A > B > C) to specific segments of the electrode saliences (i.e., derive component waveforms). For example, LV1 in Simulation 2 (Figure 5) expressed two contrasts. The strongest effect was the difference between condition S2*a*S3*a* and baseline, with a smaller difference between condition S2*a* and baseline. We are currently examining whether further post hoc analysis of the PLS solutions can be used to separate multiple effects on a LV. One approach operates on the design scores to determine how strongly specific aspects of a contrast are expressed on the LV. In this case, one can regress post hoc contrasts against the design scores to verify the visual impression of the strength of the effects being expressed on a LV. This approach has been used successfully in brain imaging data to distinguish age and delay effects in working memory (Grady et al., 1998). The post hoc regression is done as part of the permutation test, allowing both an assessment of the significance of the LV as well as hypothesis-driven contrasts. Another alternative is to conduct a "post hoc" SVD on the LV saliences to produce separate temporal and electrode saliences. This approach appears to be most successful when the PLS results identify single effects such as those found in Simulation 1 (Lobaugh, West, Taylor, & McIntosh, 1999). However, the important feature of PLS as it is presently implemented is its ability to isolate only those aspects of waveform differences that are related to experimental manipulations.

PLS can also be extended to examine relations between behavioral responses and scalp potentials, in a fashion similar to that used by O'Donnell and colleagues (1999) to examine relations between brain measures and ERP peak amplitudes. In that study, PLS was used to examine the relations between measures of brain morphology and the N200 and P300 components of the ERP in patients with schizophrenia. Peak amplitudes of the N200 and P300 were identified for each subject, and PLS was conducted on the two amplitude measures from each electrode, rather than the full waveform. In this case, PLS would identify those aspects of behavior (or morphology) that most strongly correlated with the measured ERPs.

The current paper has described a two-block PLS model, with design and ERP amplitude as the two blocks under consideration. Any multiblock path model can be examined using the PLS approach (Frank & Kowalski, 1985). The multiblock approach has been successful in behavioral teratology (Bookstein, Sampson, Streissguth, & Barr, 1996) and PET brain imaging (McIntosh et al., 1998), so the ability to describe simultaneous spatiotemporal relations among task variables, behavioral responses, and measured scalp potentials is a very enticing future possibility.

REFERENCES

Achim, A., & Bouchard, S. (1997). Toward a dynamic topographic components model. *Electroencephalography and Clinical Neurophysiology, 103,* 381–385.

Bookstein, F. L., Sampson, P. D., Streissguth, A. P., & Barr, H. M. (1996). Exploiting redundant measurement of dose and developmental outcome: New methods from the behavioral teratology of alcohol. *Developmental Psychology, 32,* 404–415.

Borgen, O. S., & Kowalski, B. R. (1985). An extension of the multivariate component-resolution method to three components. *Analytica Chimica Acta, 174,* 1–26.

Braun, A. R., Balkin, T. J., Wesensten, N. J., Gwadry, F., Carson, R. E., Varga, M., Baldwin, P., Belenky, G., & Herscovitch, P. (1998). Dissociated pattern of activity in visual cortices and their projections during human rapid eye movement sleep. *Science, 279,* 91–95.

Cabeza, R., Grady, C. L., Nyberg, L., McIntosh, A. R., Tulving, E., Kapur, S., Jennings, J. M., Houle, S., & Craik, F. I. M. (1997). Age-related differences in neural activity during memory encoding and retrieval: A positron emission tomography study. *Journal of Neuroscience, 17,* 391–400.

Dien, J. (1998). Addressing misallocation of variance in principal components analysis of event-related potentials. *Brain Topography, 11,* 43-55.

Donchin, E., & Heffley, E. F. (1978). Multivariate analysis of event-related potential data: A tutorial review. In D. Otto (Ed.), *Multidisciplinary perspectives in event-related brain potentials research* (pp. 555–572). Washington DC: Environmental Protection Agency.

Edgington, E. S. (1980). *Randomization tests*. New York: Marcel Dekker.

Efron, B., & Tibshirani, R. (1986). Bootstrap methods for standard errors, confidence intervals and other measures of statistical accuracy. *Statistical Science, 1,* 54–77.

Fabiani, M., Gratton, G., Corballis, P. M., Cheng, J., & Friedman, D. (1998). Bootstrap assessment of the reliability of maxima in surface maps of brain activity of individual subjects derived with electrophysiological and optical methods. *Behavior Research Methods, Instruments, & Computers, 30,* 78–86.

Frank, I. E., & Kowalski, B. R. (1985). A multivariate method for relating groups of measurements connected by a causal pathway. *Analytica Chimica Acta, 167,* 51–63.

Friston, K. J., Frith, C. D., Liddle, P. F., & Frackowiak, R. S. J. (1993).

Functional connectivity: The principal-component analysis of large (PET) data sets. *Journal of Cerebral Blood Flow and Metabolism*, *13*, 5–14.

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.

Heise, H. M., Marbach, R., Janatsch, G., & Kruse-Jarres, J. D. (1989). Multivariate determination of glucose in whole blood by attenuated total reflection infrared spectroscopy. *Analytical Chemistry*, *61*, 2009–2015.

Hellberg, S., Sjostrom, M., & Wold, S. (1986). The prediction of bradykinin potentiating potency of pentapeptides. An example of a peptide quantitative structure-activity relationship. *Acta Chemica Scandinavica. Series B. Organic Chemistry and Biochemistry*, *40*, 135–40.

Lobaugh, N. J., Taylor, M. J., & McIntosh, A. R. (1997). Event-related potentials reveal visual search processing differences in Attention Deficit Hyperactivity Disorder. *Society for Neuroscience Abstracts*, *23*, 300.

Lobaugh, N. J., West, R., Taylor, M. J., & McIntosh, A. R. (1999). Spatiotemporal analysis of task-related components in event-related potentials: Simulation results and application to a visual selective attention task. Abstract presented at the Cognitive Neuroscience Society annual meeting, April 11–13, 1999.

Makeig, S., Juny, T.-P., Bell, A. J., Ghahremani, D., & Sejnowski, T. J. (1997). Blind separation of auditory event-related brain responses into independent components. *Proceedings of the National Academy of Sciences*, *94*, 10979–10984.

Makeig, S., Westerfield, M., Jung, T.-P., Covington, J., Townsend, J., Sejnowski, T. J., & Courchesne, E. (1999). Functionally independent components of the late positive event-related potential during visual spatial attention. *Journal of Neuroscience*, *19*, 2665–2680.

Martens, H., Izquierdo, L., Thomassen, M., & Martens, M. (1986). Partial least squares regression on design variables as an alternative to analysis of variance. *Analytica Chimica Acta*, *191*, 133–148.

McIntosh, A., Lobaugh, N., Cabeza, R., Bookstein, F., & Houle, S. (1998). Convergence of neural systems processing stimulus associations and coordinating motor responses. *Cerebral Cortex*, *8*, 648–659.

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. (1999). Interactions of prefrontal cortex in relation to awareness in sensory learning. *Science*, *284*, 1531–1533.

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

O'Donnell, B. F., McCarley, R. W., Potts, G. F., Salisbury, D. F., Nestor, P. G., Hirayasu, Y., Niznikiewicz, M. A., Barnard, J., Shen, Z. J., Weinstein, D. M., Bookstein, F. L., & Shenton, M. E. (1999). Identification of neural circuits underlying P300 abnormalities in schizophrenia. *Psychophysiology*, *36*, 388–398.

Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R., Jr., Miller, G. A., Ritter, W., Ruchkin, D. S., Rugg, M. D., & Taylor, M. J. (2000). Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology*, *37*, 127–152.

Sampson, P. D., Streissguth, A. P., Barr, H. M., & Bookstein, F. L. (1989). Neurobehavioral effects of prenatal alcohol: Part II. Partial least squares analysis. *Neurotoxicology and Teratology*, *11*, 477–491.

Spencer, K. M., Dien, J., & Donchin, E. (1999). A componential analysis of the ERP elicited by novel events using a dense electrode array. *Psychophysiology*, *36*, 409–414.

Streissguth, A. P., Bookstein, F. L., Sampson, P. D., & Barr, H. M. (1993). *The enduring effects of prenatal alcohol exposure on child development: Birth through seven years, a partial least squares solution*. Ann Arbor, Michigan: The University of Michigan Press.

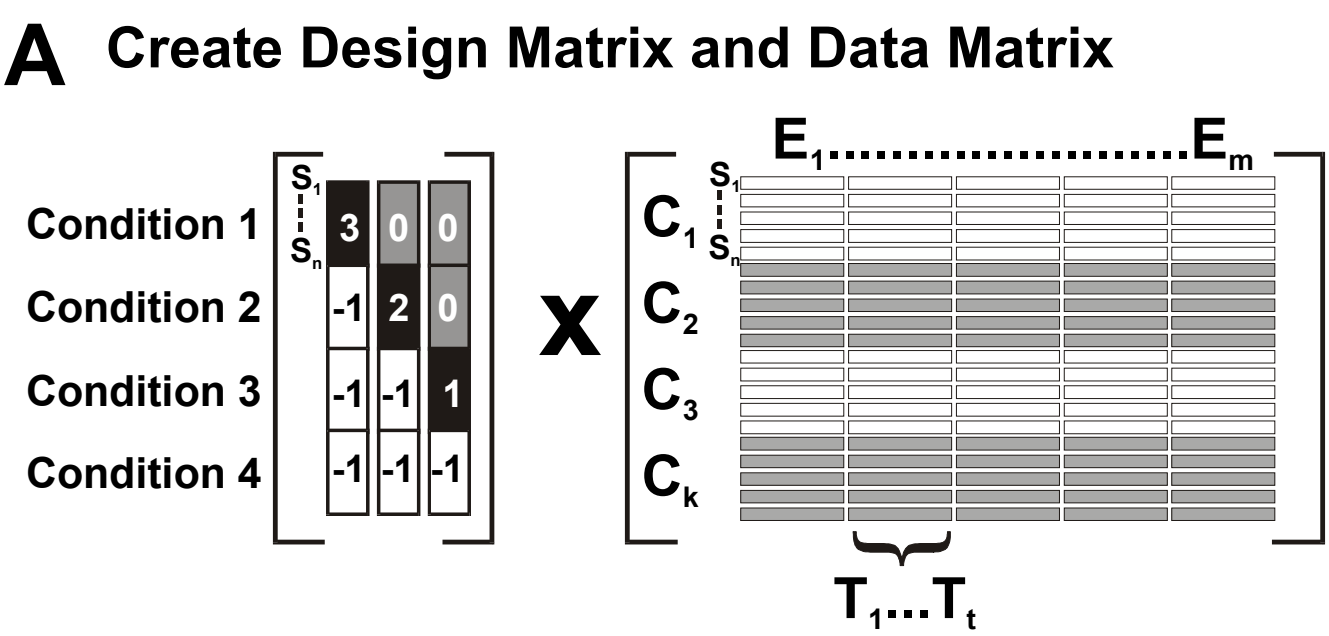
Sylvestre, E. A., Lawton, W. H., & Maggio, M. S. (1974). Curve resolution using a postulated chemical reaction. *Technometrics*, *16*, 353–368.

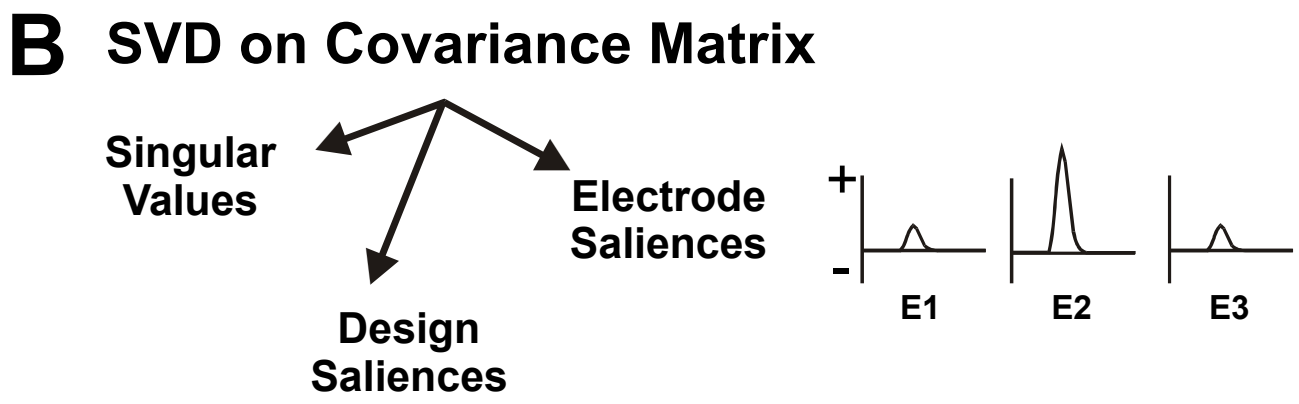
Widaman, K. F. (1993). Common factor analysis versus principal components analysis: Differential bias in representing model parameters? *Multivariate Behavioral Research*, *28*, 263–312.

Wold, H. (1975). Path models with latent variables: The NIPALS approach. In H. M. Blalock, A. Aganbegian, F. M. Borodkin, R. Boudon and V. Cappecchi (Ed.), *Quantitative sociology: International perspectives on mathematical and statistical modeling* (pp. 307–357). New York: Academic Press.

(Received September 15, 1999; Accepted September 29, 2000)

Fig 1


A
Create Design Matrix and Data Matrix
Condition 1
Condition 2
Condition 3
Condition 4
S1
Sn
3 0 0
-1 2 0
-1 -1 1
-1 -1 -1
X
E1.....Em
C1
C2
C3
Ck
T1...Tt


B
SVD on Covariance Matrix
Singular Values
Electrode Saliences
Design Saliences
+
-
E1
E2
E3

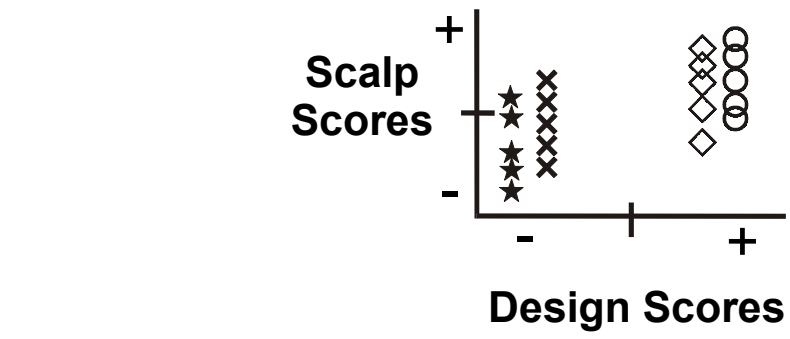

C
Create Scores
Scalp Scores
+
-
-
+
Design Scores

Fig 2.

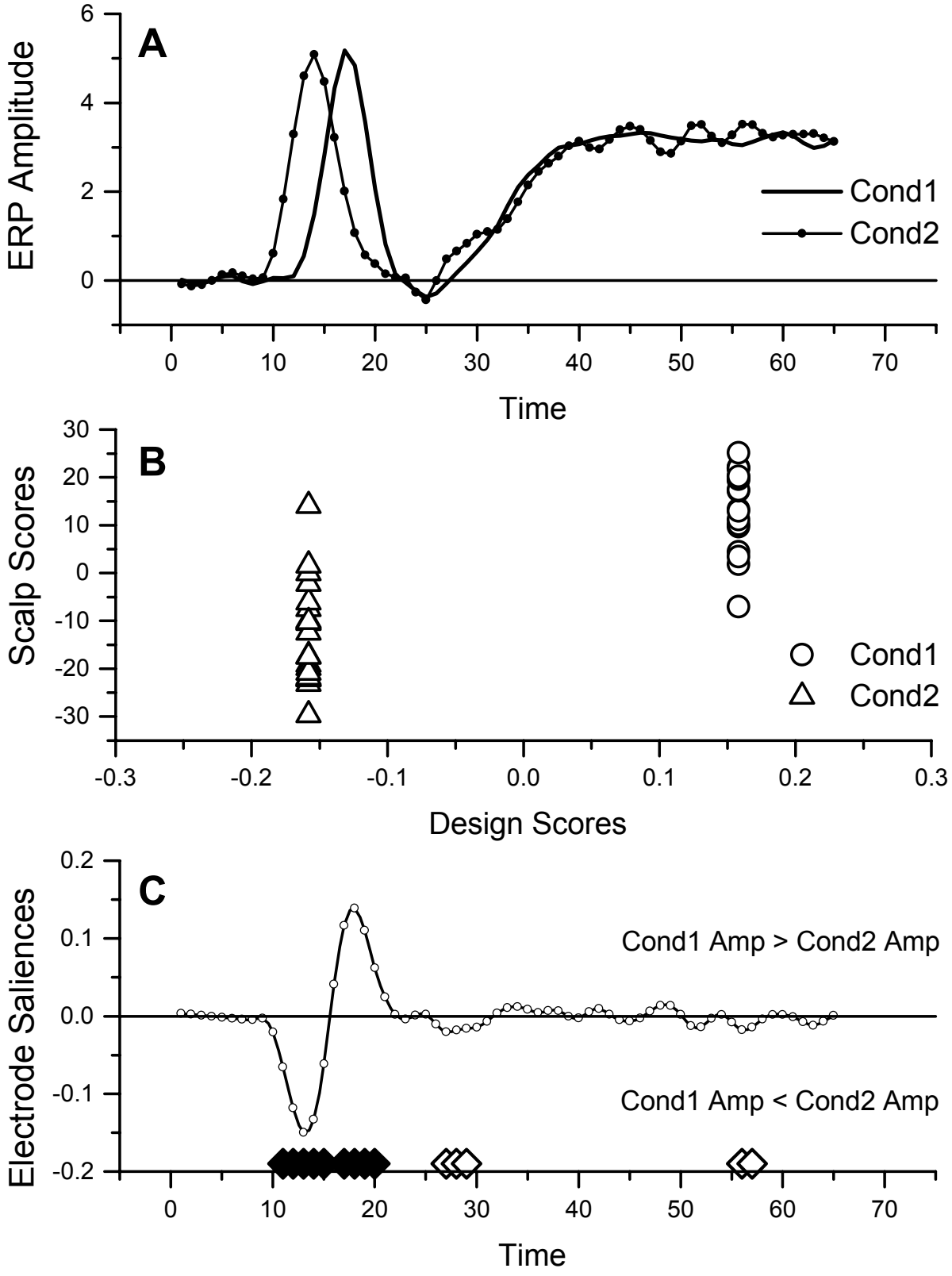

A
ERP Amplitude
6
4
2
0
0
10
20
30
40
50
60
70
Time
Cond1
Cond2
B
Scalp Scores
30
20
10
0
-10
-20
-30
-0.3
-0.2
-0.1
0.0
0.1
0.2
0.3
Design Scores
Cond1
Cond2
C
Electrode Saliences
0.2
0.1
0.0
-0.1
-0.2
Cond1 Amp > Cond2 Amp
Cond1 Amp < Cond2 Amp
0
10
20
30
40
50
60
70
Time

Fig 3.

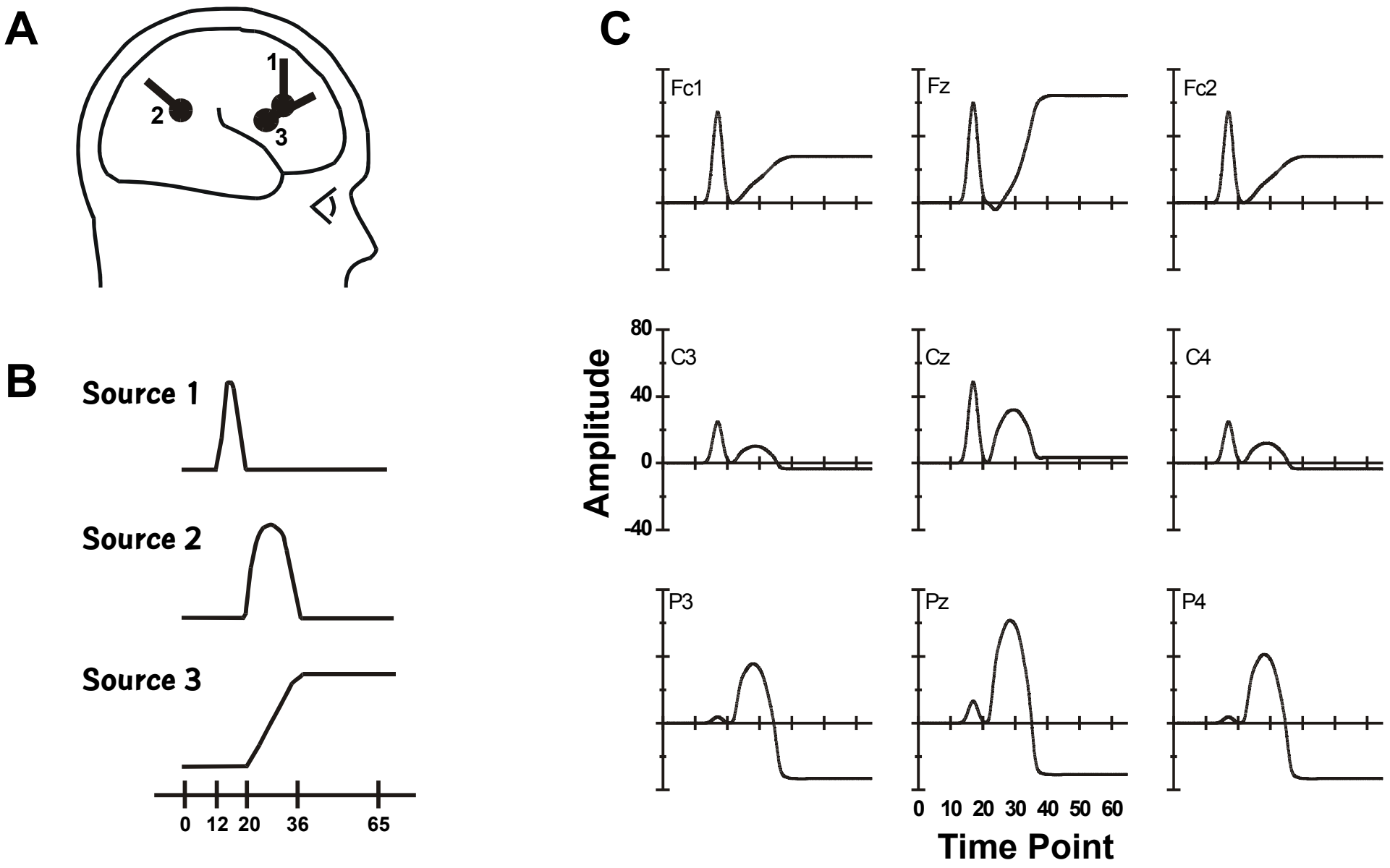

A
1
2
3
B
Source 1
Source 2
Source 3
0 12 20 36 65
C
Fc1
Fz
Fc2
C3
Cz
C4
P3
Pz
P4
80
40
0
-40
Amplitude
0 10 20 30 40 50 60
Time Point

Fig 4.

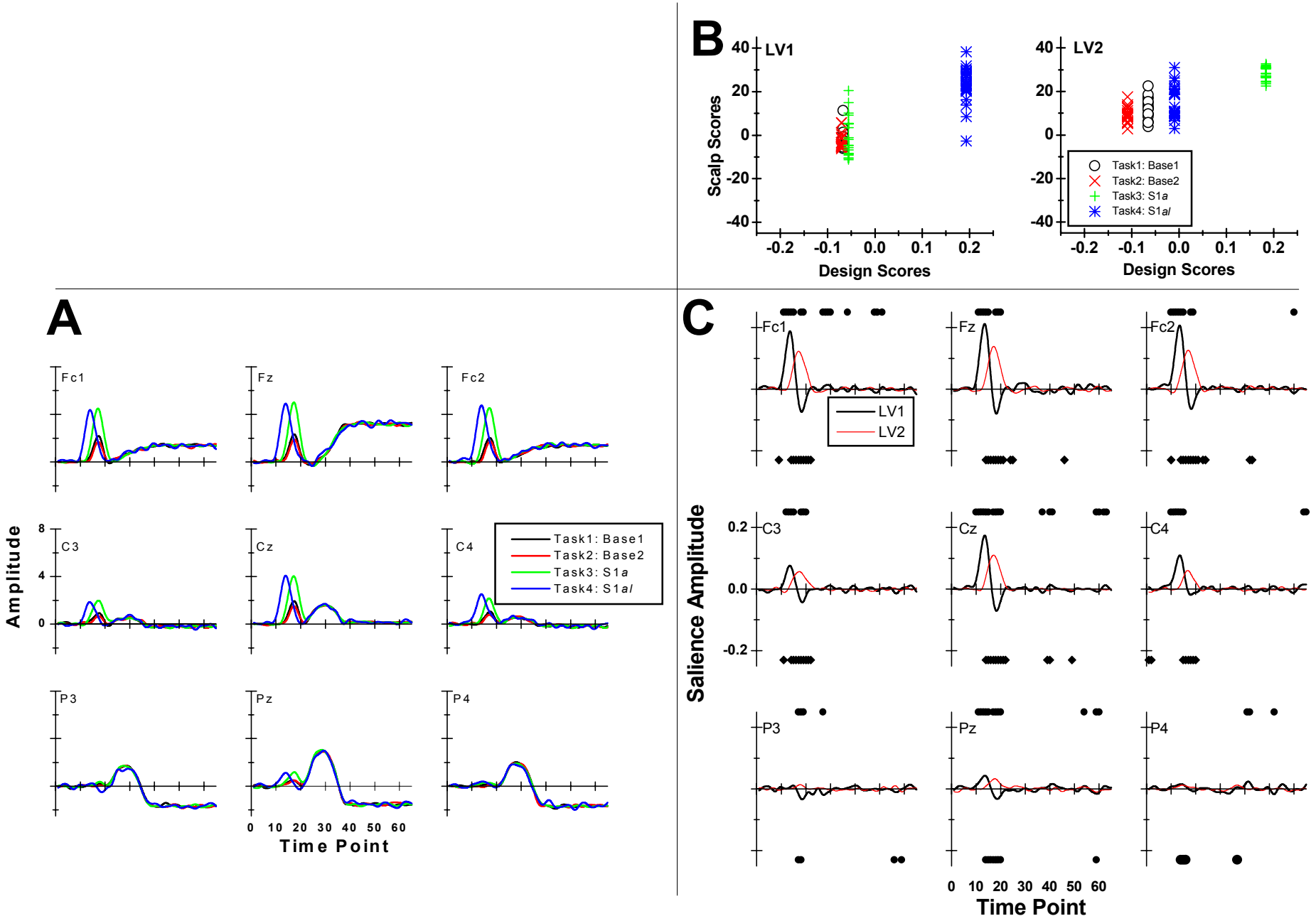

A
Fc1
Fz
Fc2
C3
Cz
C4
P3
Pz
P4
Amplitude
8
4
0
0 10 20 30 40 50 60
Time Point
Task1: Base1
Task2: Base2
Task3: S1a
Task4: S1al
B
LV1
LV2
Scalp Scores
40
20
0
-20
-40
-0.2 -0.1 0.0 0.1 0.2
Design Scores
Task1: Base1
Task2: Base2
Task3: S1a
Task4: S1al
C
Fc1
Fz
Fc2
LV1
LV2
C3
Cz
C4
P3
Pz
P4
Salience Amplitude
0.2
0.0
-0.2
0 10 20 30 40 50 60
Time Point

Fig 5.

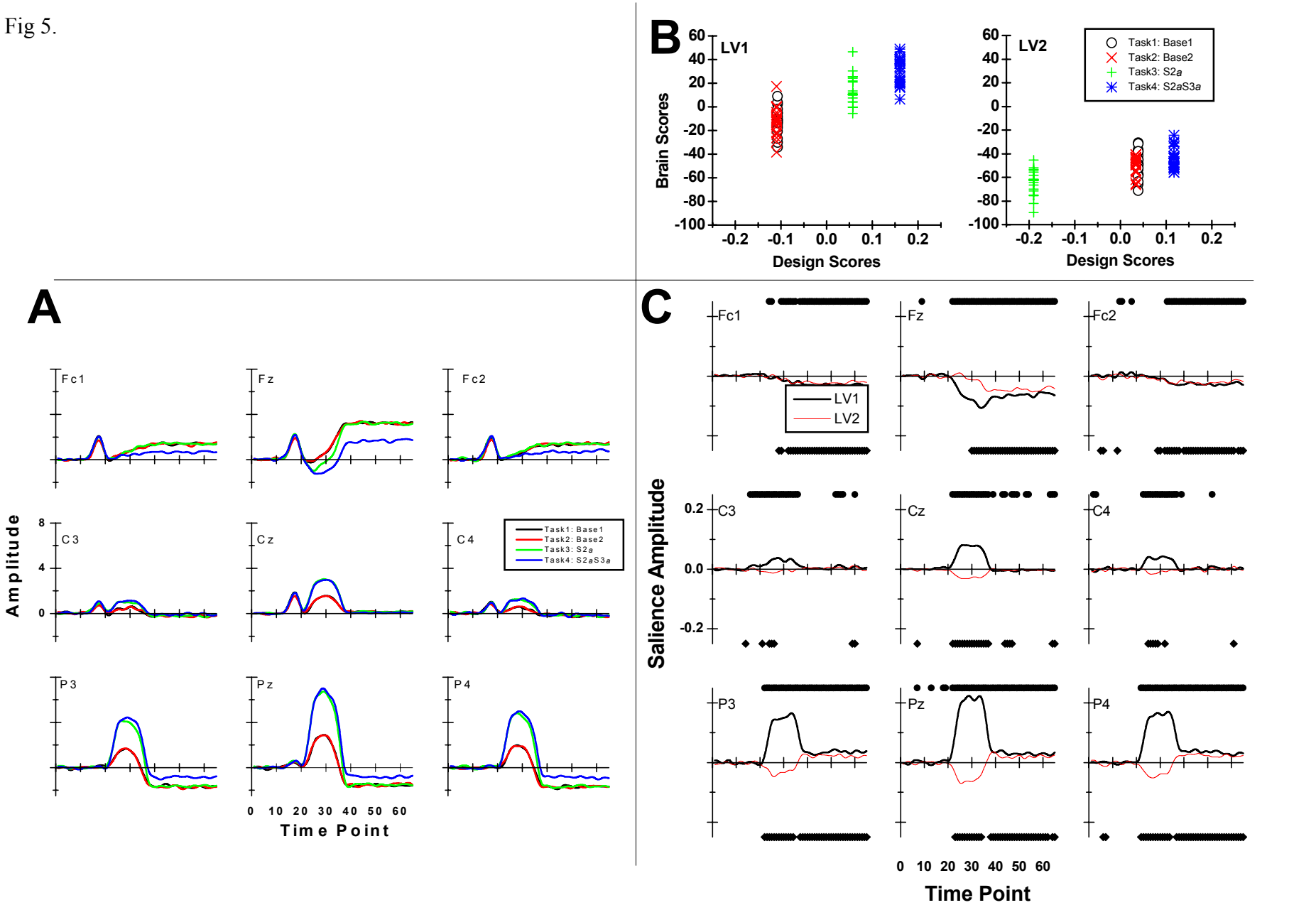

A
Fc1
Fz
Fc2
C3
Cz
C4
P3
Pz
P4
Amplitude
Task1: Base1
Task2: Base2
Task3: S2a
Task4: S2aS3a
0 10 20 30 40 50 60
Time Point
B
LV1
LV2
Brain Scores
Design Scores
Task1: Base1
Task2: Base2
Task3: S2a
Task4: S2aS3a
C
Fc1
Fz
Fc2
LV1
LV2
C3
Cz
C4
P3
Pz
P4
Salience Amplitude
0 10 20 30 40 50 60
Time Point

Fig 6.

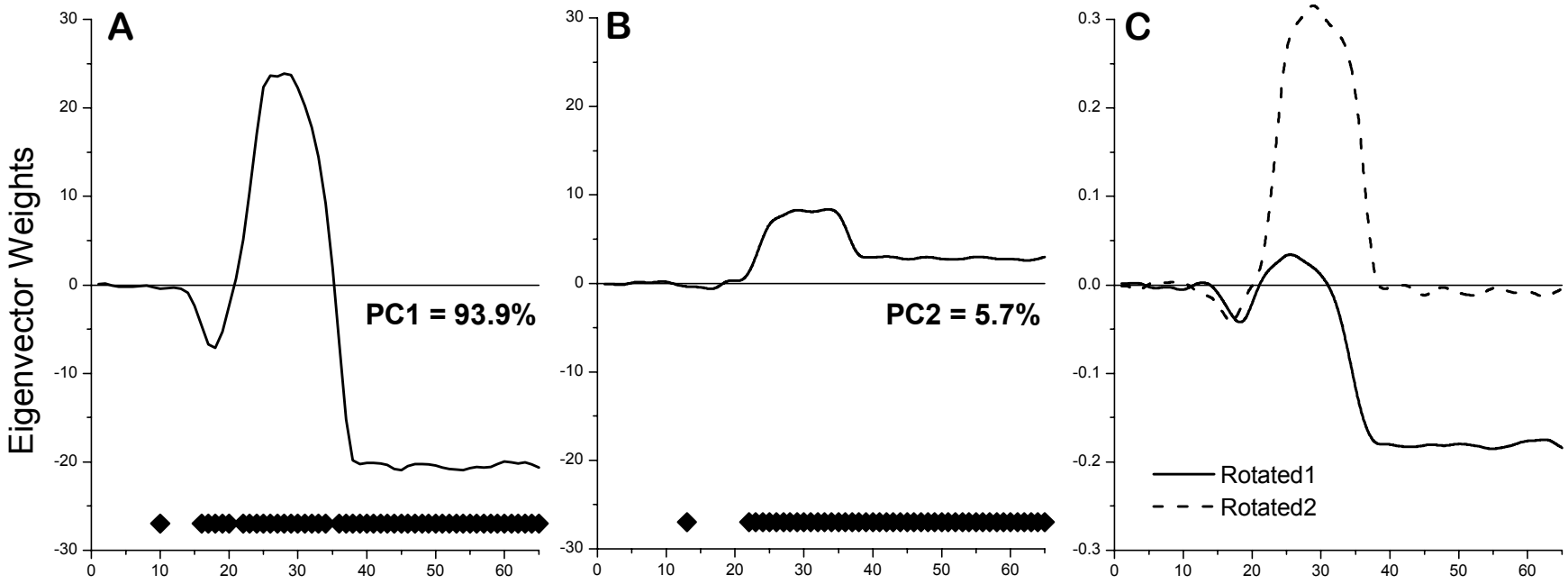


Fig 7.

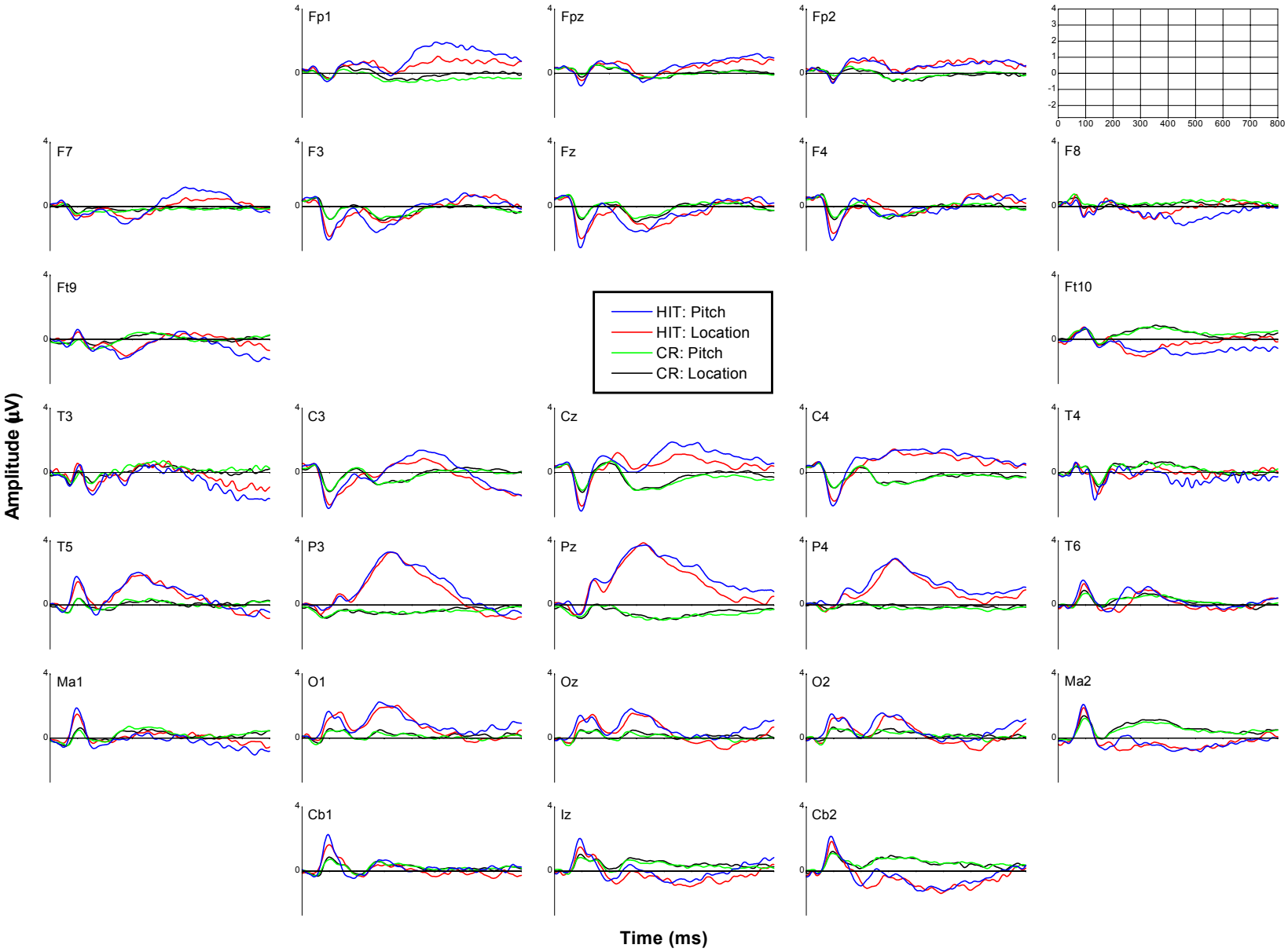
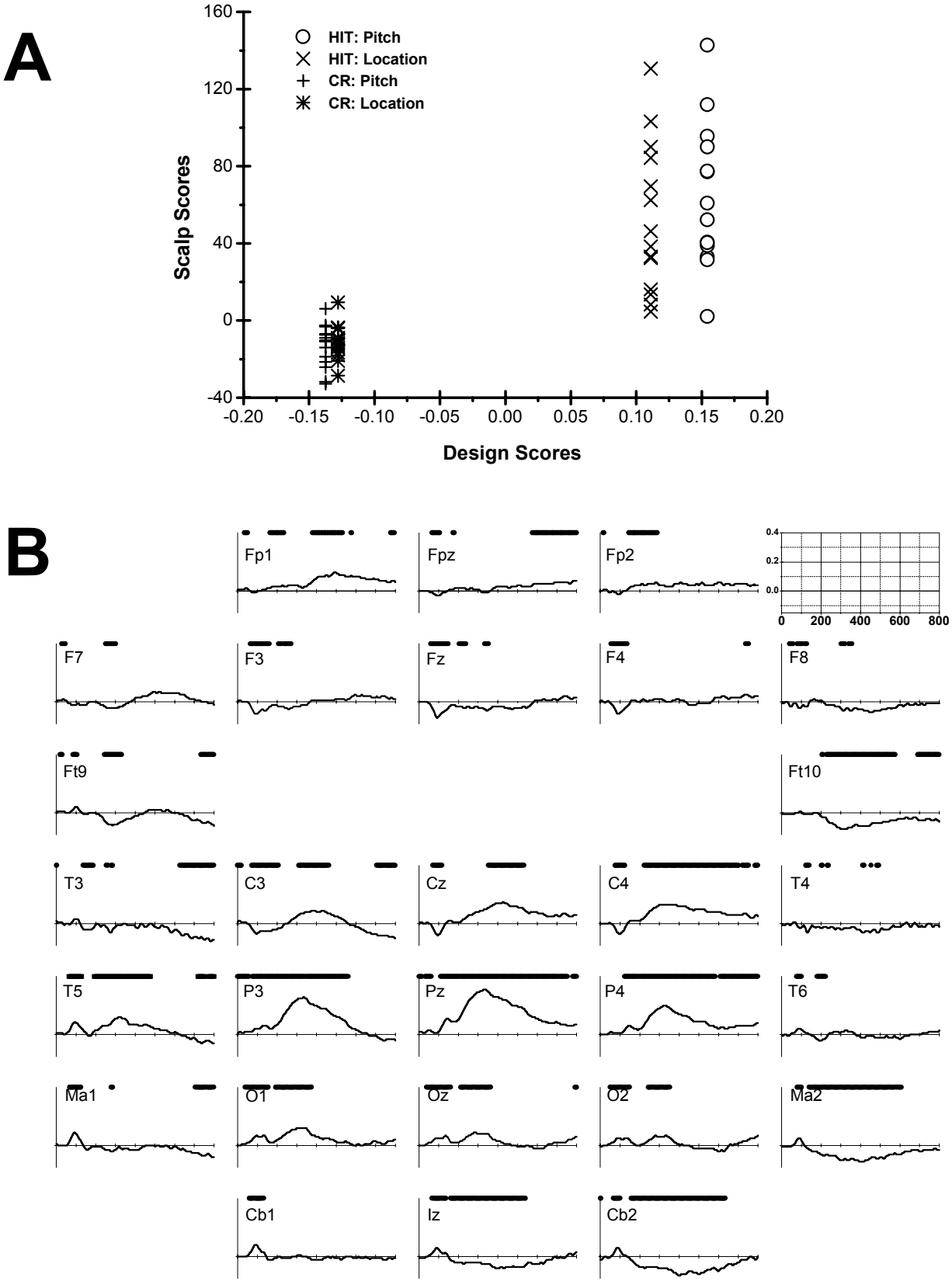

Fp1
Fpz
Fp2
F7
F3
Fz
F4
F8
Ft9
Ft10
T3
C3
Cz
C4
T4
T5
P3
Pz
P4
T6
Ma1
O1
Oz
O2
Ma2
Cb1
Iz
Cb2
HIT: Pitch
HIT: Location
CR: Pitch
CR: Location
Amplitude (μV)
Time (ms)

Fig 8


A
HIT: Pitch
HIT: Location
CR: Pitch
CR: Location
Scalp Scores
Design Scores
160
120
80
40
0
-40
-0.20
-0.15
-0.10
-0.05
0.00
0.05
0.10
0.15
0.20
B
Fp1
Fpz
Fp2
F7
F3
Fz
F4
F8
Ft9
Ft10
T3
C3
Cz
C4
T4
T5
P3
Pz
P4
T6
Ma1
O1
Oz
O2
Ma2
Cb1
Iz
Cb2
0.4
0.2
0.0
0
200
400
600
800