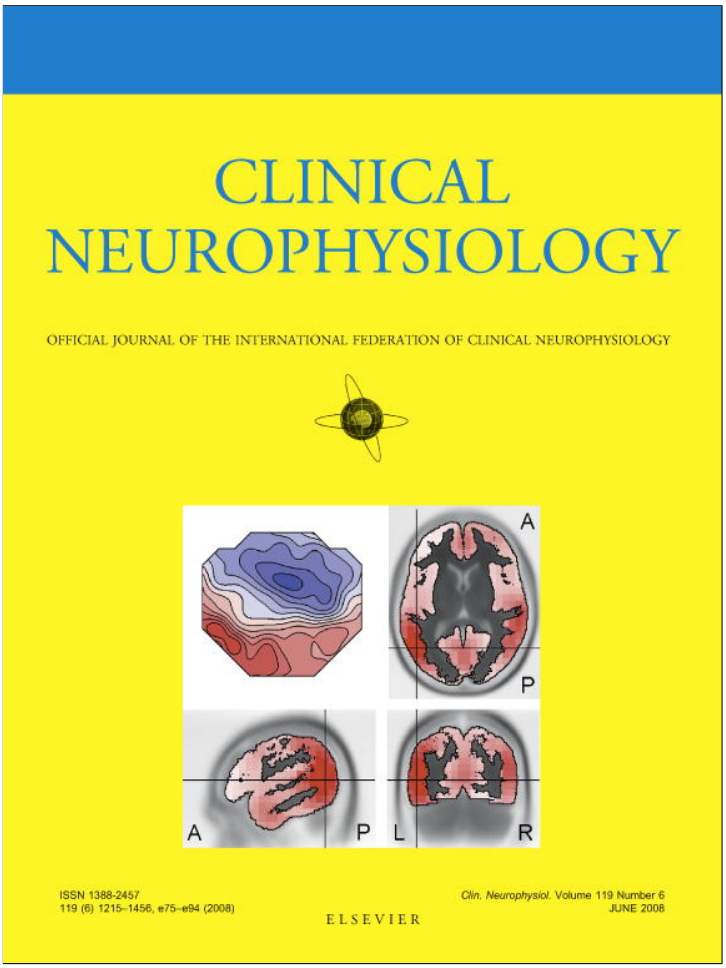

ELSEVIER

Clinical Neurophysiology 119 (2008) 1262–1270

www.elsevier.com/locate/clinph

Establishing correlations of scalp field maps with other experimental variables using covariance analysis and resampling methods

Thomas Koenig[a,*], Lester Melie-García[b], Maria Stein[a], Werner Strik[a], Christoph Lehmann[a]

[a] *Department of Psychiatric Neurophysiology, University Hospital of Psychiatry, Bolligenstr. 111, 3000 Bern 60, Switzerland*
[b] *Neuroimaging Department, Cuban Neuroscience Center, Calle 25 esq.158, Cubanacan, Playa, Havana, Cuba*

Accepted 26 December 2007

Abstract

Objective: In EEG/MEG experiments, increasing the number of sensors improves the spatial resolution of the results. However, the standard statistical methods are inappropriate for these multivariate, highly correlated datasets. We introduce a procedure to identify spatially extended scalp fields that correlate with some external, continuous measure (reaction-time, performance, clinical status) and to test their significance.
Methods: We formally deduce that the channel-wise covariance of some experimental variable with scalp field data directly represents intracerebral sources associated with that variable. We furthermore show how the significance of such a representation can be tested with resampling techniques.
Results: Simulations showed that depending on the number of channels and subjects, effects can be detected already at low signal to noise ratios. In a sample analysis of real data, we found that foreign-language evoked ERP data were significantly associated with foreign-language proficiency. Inverse solutions of the extracted covariances pointed to sources in language-related areas.
Conclusions: Covariance mapping combined with bootstrapping methods has high statistical power and yields unique and directly interpretable results.
Significance: The introduced methodology overcomes some of the 'traditional' statistical problems in EEG/MEG scalp data analysis. Its application can improve the reproducibility of results in the field of EEG/MEG.

Keywords: ERP; Topography; Correlation; Statistics; Randomization; Inverse solution

1. Introduction

Technical improvements have made it possible that in EEG/MEG and ERP/ERF experiments, the number of sensors on the scalp could be substantially increased. This higher spatial sampling frequency is a relevant factor for the quality of results obtained in electrophysiological experiments. First, the increase in spatial information reduces spatial aliasing (Gevins, 1996; Luu et al., 2001) and improves the sensitivity and specificity of the results. Second, the accuracy of EEG/MEG and ERP/ERF inverse solution improves significantly when high-density electrode arrays are being used (see Michel et al., 2004, for a review).

However, the increase of number of sensors results in an increasingly multivariate dataset that is increasingly correlated in space. When scalp field data are used to study the effects of some experimental conditions, such data requires adequate statistical treatment. In many studies, the statistical approaches chosen to analyze multi-channel scalp field data do not take the relations between sensors properly into account and disregard the physical basis of the signals to be analyzed. Namely, many studies employed strategies where some sensors or groups of sensors are selected a-pri-

* Corresponding author. Tel.: +41 31 930 9369; fax: +41 31 930 9961.
E-mail address: thomas.koenig@puk.unibe.ch (T. Koenig).

1388-2457/$34.00
doi:10.1016/j.clinph.2007.12.023

ori. Then, standard univariate statistics are used and the different recording sites are considered as repeated (within-subject) measures. The additional information obtained with higher spatial sampling is thus often poorly exploited by the statistics applied.

The univariate approach is furthermore problematic from a physics point of view: the function that relates the activity of a given point source in the brain to a measurable electric and/or magnetic field on the scalp (the so-called leadfield) implies that any source in the brain produces a field extending over the entire scalp surface. With several sources simultaneously active, the measured scalp field becomes the sum of the scalp fields produced by those sources (Mosher et al., 1999).

In terms of statistics of EEG/MEG and ERP/ERF data, this implies that the basic entity for analysis should be the scalp electric field. Furthermore, since effects of intracerebral generators on the scalp fields are additive, the effects of a difference in processing in two experimental conditions are directly reflected as the difference field between the scalp fields evoked by those two conditions. The difference field is thus exactly the field produced by those intracerebral generators that account for the difference between the conditions. Accordingly, difference maps have been employed routinely in EEG and ERP studies (e.g. Duffy et al., 1981; Steger et al., 2000).

In order to establish the statistical significance of such difference maps, one can either use the standard multivariate statistical approaches such as MANOVA (Vasey and Thayer, 1987). However, a MANOVA requires that there are more observations than sensors, a condition that becomes increasingly difficult to meet with increasing number of sensors. Furthermore, since the data is spatially correlated, the degrees of freedom are much lower than the number of sensors suggests. Therefore, it has become increasingly popular to use multivariate randomization statistics to establish the significance of a difference map measured between two conditions (Karniski et al., 1994; Galan et al., 1997; Maris, 2004; Greenblatt and Pflieger, 2004). These procedures are computationally more expensive, but require very little assumptions and have a high statistical power. The procedure to compute such a randomization statistics for scalp field data is simple and straightforward. Assume we have scalp field maps of a series of subjects recorded under two conditions. In a first step, the scalp field maps are averaged separately for both conditions. Then, the difference map is computed. The total amplitude of this difference map is an indicator of the strength of the difference and can be easily measured using the Global Field Power (i.e. the standard deviation across sensors, Lehmann and Skrandies, 1980). Next, the amplitude of the difference map under the null-hypothesis is established. This is achieved by randomly shuffling the two conditions (either within subject, for paired designs, or across subjects, for unpaired designs). The field maps of the two conditions are again averaged across subjects, and the difference map between the condition mean maps is computed. The total amplitude of this randomly obtained difference therefore represents a value obtained under the null-hypothesis. By repeating this randomization many times, one can therefore obtain a good estimate of the distribution of the total map difference under the null hypothesis. The probability that the total amplitude of the difference map obtained from the two real conditions is random is then defined as the percentage of observations where this amplitude was smaller than the amplitude of the randomly obtained difference maps. This randomization approach has the advantage that it is fully multivariate, that it is based on a (realistic) additive model of scalp fields, and that it does not require any a-priori assumptions about the distribution of the variables. It has been used in a series of studies (e.g. Kondakor et al., 1995; Strik et al., 1998).

Now how can one proceed if one does not have two discrete conditions, but some continuous external variable? The additivity of EEG/MEG fields implies that if there is a set of sources with activation that is linearly related to the external variable, this will result in a single scalp field that is added to the measurements at the sensors, proportionally to the external variable. Since the absence of the effect of the external variable implies the absence of the field generated by that effect, one can further impose that the regression line crosses the origin. Such a relation can easily be assessed using the covariance, across observations, of all the single sensor signals with the external variable. In order to establish whether such a covariance scalp field is of statistical significance, one can again use resampling methods.

The aim of the current paper is to (a) introduce the methodology to extract such covariance scalp fields and to test their statistical significance using resampling methods, (b) illustrate the utility of such a method using foreign-language evoked potentials in subjects with varying language proficiency and (c) relate the method to other methods such as partial least squares (PLS, see McIntosh et al., 1996; Lobaugh et al., 2001).

2. Methods

Notation: In this paper, bold symbols denote a column vector or matrix and non-bold symbols, a scalar magnitude. Superscript T denotes transpose. The notation $N(\boldsymbol{\mu},\boldsymbol{\Sigma})$ represents a normal distribution with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$. The symbol $\sim$ denotes "distributed as", e.g. $\mathbf{x} \sim N(\boldsymbol{\mu},\boldsymbol{\Sigma})$ means the random variable $\mathbf{x}$ is normally distributed with parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. Symbol '*tr*' denotes the trace of a matrix. Symbol $\mathbf{1}_n$ denotes a column vector of length n with all elements with value 1.

We assume a linear relation between the current density strength in a voxel and the behavioral variable (e.g. reaction-time):

$$\mathbf{j}_i = X_k \boldsymbol{\theta}_i + \mu \mathbf{1}_3 \quad (1)$$

where $\mathbf{j}_i = \begin{bmatrix} j_{ix} \\ j_{iy} \\ j_{iz} \end{bmatrix}$ is the current density vector in the voxel 'i' and $\boldsymbol{\theta}_i = \begin{bmatrix} \theta_{ix} \\ \theta_{iy} \\ \theta_{iz} \end{bmatrix}$ is the vector of linear regression coefficients, X_k is the behavioral variable for subject 'k' and μ is the mean effect term.

For all N_g current density sources in the brain we obtain the following equation:

$$\mathbf{J}_k = X_k\boldsymbol{\theta} + \mu\mathbf{1}_{3N_g} \tag{2}$$

where $\mathbf{J}_k = \begin{bmatrix} \mathbf{j}_1 \\ \vdots \\ \mathbf{j}_{N_g} \end{bmatrix}$ and $\boldsymbol{\theta} = \begin{bmatrix} \boldsymbol{\theta}_1 \\ \vdots \\ \boldsymbol{\theta}_{N_g} \end{bmatrix}$ The mathematical relation between primary current density and the voltage recorded in the electrode array on the scalp is obtained through the leadfield or gain matrix $\mathbf{K}_{N_e\times 3N_g}$. Then the voltage measured on the scalp using can be expressed by:

$$\mathbf{v}_k = \mathbf{K}\mathbf{J}_k = X_k\mathbf{K}\theta + \mu\mathbf{K}\cdot\mathbf{1}_{3N_g} \tag{3}$$

If we defined the term $\boldsymbol{\beta} = \mathbf{K}\cdot\boldsymbol{\theta}$ as a voltage regression coefficient, and $\boldsymbol{\mu} = \mu\mathbf{K}\cdot\mathbf{1}_{3N_g}$ as the voltage mean effect coefficient, we could rewrite the previous equation as:

$$\mathbf{v}_k = X_k\boldsymbol{\beta} + \boldsymbol{\mu} \tag{4}$$

where $\mathbf{v}_k = [v_{1k},\ldots,v_{N_ek}]^{\mathrm{T}}$ is a voltage vector for the N_e electrodes for the subject 'k'.

Then for N_s subjects the equation (4) becomes:

$$\mathbf{V} = \mathbf{X}\cdot\boldsymbol{\beta}^{\mathrm{T}} + \mathbf{1}_{N_s}\cdot\boldsymbol{\mu}^{\mathrm{T}} \tag{5}$$

where $\mathbf{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_{N_s} \end{bmatrix}$ is a matrix that encloses the behavioral variable value for all N_s subjects. The voltage matrix is organized as $\mathbf{V} = \begin{bmatrix} \mathbf{v}_1^{\mathrm{T}} \\ \vdots \\ \mathbf{v}_{N_s}^{\mathrm{T}} \end{bmatrix} = \begin{bmatrix} v_{11} & \ldots & v_{1N_e} \\ \vdots & \ddots & \vdots \\ v_{N_s1} & \ldots & v_{N_sN_e} \end{bmatrix}$. One row '$i$' of the matrix $\mathbf{V}$ is a topography for the subject 'i'.One column'j' is the voltage of the electrode 'j' for all subjects.

Equation (5) can be rewritten in a more compact way as:

$$\mathbf{V} = \widetilde{\mathbf{X}}\cdot\tilde{\boldsymbol{\beta}} \tag{6}$$

where $\widetilde{\mathbf{X}} = [\mathbf{X} \quad \mathbf{1}_{N_s}]$ and $\tilde{\boldsymbol{\beta}} = \begin{bmatrix} \boldsymbol{\beta}^{\mathrm{T}} \\ \boldsymbol{\mu}^{\mathrm{T}} \end{bmatrix}$ The observation equation is obtained from (6) by including an additive noise term:

$$\mathbf{V} = \widetilde{\mathbf{X}}\cdot\tilde{\boldsymbol{\beta}} + \boldsymbol{\varepsilon} \tag{7}$$

The equation (7) is a standard multivariate regression model, where $\widetilde{\mathbf{X}}$ is the design matrix and $\tilde{\boldsymbol{\beta}}$, the regression coefficients.

The $\boldsymbol{\varepsilon}$ term is the experimental noise, and most applications assume that it has mean $\mathbf{0}$ and unknown covariance matrixes $\boldsymbol{\Sigma}_e$ and $\boldsymbol{\Sigma}_s$ such that $\boldsymbol{\varepsilon} \sim N(\mathbf{0},\boldsymbol{\Sigma}_s,\boldsymbol{\Sigma}_e)$, where $\boldsymbol{\varepsilon}$ is independent of $\mathbf{X}$. The term $\boldsymbol{\Sigma}_e$ provides the covariance structure between EEG sensors. The matrix $\boldsymbol{\Sigma}_s$ expresses the covariance structure between subjects.

Under this noise assumption, the log-likelihood function for the data matrix $\mathbf{V}$ in terms of the parameters $\tilde{\boldsymbol{\beta}}$, $\boldsymbol{\Sigma}_s$ and $\boldsymbol{\Sigma}_e$ is given by:

$$l(\tilde{\boldsymbol{\beta}},\boldsymbol{\Sigma}) = -\frac{1}{2}N_s\log|2\pi\boldsymbol{\Sigma}_s| - \frac{1}{2}N_e\log|\boldsymbol{\Sigma}_e| - \frac{1}{2}\mathrm{tr}\left(\mathbf{V}-\widetilde{\mathbf{X}}\tilde{\boldsymbol{\beta}}\right)\boldsymbol{\Sigma}_e^{-1}\left(\mathbf{V}-\widetilde{\mathbf{X}}\tilde{\boldsymbol{\beta}}\right)^{\mathrm{T}}\boldsymbol{\Sigma}_s^{-1} \tag{8}$$

In order to estimate the parameters in (7), we made use of the Bayesian formalism. A summary of the basis of the Bayesian Inference Theory and the derivations of the parameter estimators are summarized in the Supplementary Appendixes B and C.

The design matrix $\widetilde{\mathbf{X}}$ in our case is full rank. If no prior information for regression parameter $\tilde{\boldsymbol{\beta}}$ is considered a maximum likelihood estimator is obtained. The $\tilde{\boldsymbol{\beta}}$ estimator is unique and given by the expression:

$$\hat{\boldsymbol{\beta}} = \left(\widetilde{\mathbf{X}}^{\mathrm{T}}\boldsymbol{\Sigma}_s^{-1}\widetilde{\mathbf{X}}\right)^{-1}\widetilde{\mathbf{X}}^{\mathrm{T}}\boldsymbol{\Sigma}_s^{-1}\mathbf{V} \tag{9}$$

If one assumes that there is no covariance across subjects as $\boldsymbol{\Sigma}_s = \xi\cdot\mathbf{I}$, ($\xi$ is a variance term) one obtains that

$$\hat{\boldsymbol{\beta}} = \left(\widetilde{\mathbf{X}}^{\mathrm{T}}\widetilde{\mathbf{X}}\right)^{-1}\widetilde{\mathbf{X}}^{\mathrm{T}}\mathbf{V} \tag{10}$$

(Mardia et al., 1979, Note that this estimator does not depend on the covariance matrix $\boldsymbol{\Sigma}_e$): Term $\left(\widetilde{\mathbf{X}}^{\mathrm{T}}\widetilde{\mathbf{X}}\right)^{-1}$ is a 2×2 symmetric matrix that can be represented by:

$$\mathbf{A} = \left(\widetilde{\mathbf{X}}^{\mathrm{T}}\widetilde{\mathbf{X}}\right)^{-1} = \left(\begin{bmatrix}\mathbf{X}^{\mathrm{T}} \\ \mathbf{1}_{N_s}^{\mathrm{T}}\end{bmatrix}[\mathbf{X} \quad \mathbf{1}_{N_s}]\right)^{-1} = \begin{bmatrix}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1} & 0 \\ 0 & N_s^{-1}\end{bmatrix} \tag{11}$$

Without loss of generality, we can center the column of $\mathbf{X}$ to have mean $\mathbf{0}$,which is obtained by calculating the mean and subtracting it from the $\mathbf{X}$ column. It is convenient to separate the effect of the mean from the other independent variables. Then $\mathbf{A}$ is a diagonal matrix because $\mathbf{X}^{\mathrm{T}}\cdot\mathbf{1}_{N_s} = 0$. The term $(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}$ is a scalar magnitude.

Using the result of equation (11) the voltage regression coefficient and the voltage mean effect have the following expressions:

$$\boldsymbol{\beta}^{\mathrm{T}} = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{V} = c\cdot\mathbf{X}^{\mathrm{T}}\mathbf{V} \tag{12}$$

$$\boldsymbol{\mu}^{\mathrm{T}} = \left[\tfrac{1}{N_s}\sum_{i=1}^{N_s}V_{i1}, \quad \ldots \quad ,\tfrac{1}{N_s}\sum_{i=1}^{N_s}V_{iN_e}\right] \tag{13}$$

where $c = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}$.

Each element of the vector $\boldsymbol{\mu}$ is the mean for each electrode through all subjects.

When inter-subject covariance is modeled, the expression for the voltage regression coefficient and voltage mean effect is given by equations (13) and (14) of the Supplemen-

tary Appendix C. In this case it is necessary to estimate the covariance matrices $\mathbf{\Sigma}_e$ and $\mathbf{\Sigma}_s$ given by equations (17) and (23) of the Supplementary Appendix C. The estimation algorithm is defined in Supplementary Appendix C as well.

The amplitudes of the estimated covariance map $\boldsymbol{\beta}$ depend on the variance of $\mathbf{V}$, on the variance of $\mathbf{X}$, and on the strength of the relation between $\mathbf{V}$ and $\mathbf{X}$. (The reference of $\boldsymbol{\beta}$ is identical to the reference of $\mathbf{V}$; given formula (12), post-multiplying $\mathbf{V}$with any matrix that defines a reference will result in $\boldsymbol{\beta}$ being multiplied by the same matrix).

In order to obtain a global (across electrodes) measure of the size of the estimator $\boldsymbol{\beta}$, we calculated the Global Field Power measure (GFP, Lehmann and Skrandies, 1980) using the following equation:

$$d = \mathrm{GFP}(\boldsymbol{\beta}) = \sqrt{1/N_e (\boldsymbol{\beta} - \bar{\boldsymbol{\beta}})^{\mathrm{T}} (\boldsymbol{\beta} - \bar{\boldsymbol{\beta}})} \tag{14}$$

where $\bar{\beta} = \frac{1}{N_e}\sum_{j=1}^{N_e}\beta_i$, $\bar{\boldsymbol{\beta}} = \bar{\beta} \cdot \mathbf{1}_{N_e}$

Apart from the constant scaling factor 'c' and the elimination of the spatial baseline, GFP is identical to the singular value used by Lobaugh et al. (2001). The GFP has the advantage that it is reference- independent (Lehmann and Skrandies, 1980).

In order to establish the significance of such a covariance scalp field $\boldsymbol{\beta}$, i.e. in order to estimate the probability that a covariance field with d can be obtained by chance, one randomizes the sequence of $\mathbf{X}$, resulting in an $\mathbf{X}^*$. The strength d^* of the covariance field resulting from $\mathbf{X}^*$ is then, using equations (12) and (14) (where no covariance between subjects is assumed), computed as follows:

$$d^* = \mathrm{GFP}(c \cdot \mathbf{V}^{\mathrm{T}}\mathbf{X}^*) \tag{15}$$

yielding a d^* based solely in the null-hypothesis. When repeating this randomization n times and n is sufficiently large, the random distribution of d^* is approximated by d_n^*. Since the randomization destroys a possible physiological relation between $\mathbf{V}$ and $\mathbf{X}$, the value of d_n^* depends again on the variance of $\mathbf{V}$, on the variance of $\mathbf{X}$, and on a now randomly obtained strength of the relation between $\mathbf{V}$and $\mathbf{X}$. Since the variance of $\mathbf{V}$ and $\mathbf{X}$ are not affected by randomizing and can thus be considered as scaling factors that are constant for d and all d^*, differences between d and d^* indicate solely differences in the strength of the relation between the real and the randomized datasets. Thus, the one-tailed probability p that the original d is part of the distribution of d_n^* is given by the percentage of randomization runs where d_n^* is larger than the original d.

$$p = \frac{1}{n}\sum_{i=1}^{n} 1_{\{d_i^* > d\}} \tag{16}$$

In correlation statistics, one is often not only interested in the significance of the correlation, but also in the correlation coefficient, its confidence interval, and the fraction of common variance. In order to compute the correlation coefficient, the strength of the covariance map $\boldsymbol{\beta}$ in the ERP data $\mathbf{V}$ has to be computed case-wise. This strength $\mathbf{s}$ is defined as

$$\mathbf{s} = \mathbf{V} \cdot \boldsymbol{\beta} \tag{17}$$

Using (12) we thus obtain

$$\mathbf{s} = c \cdot \mathbf{V}\mathbf{V}^{\mathrm{T}}\mathbf{X} \tag{18}$$

The correlation r is then defined as the Pearson Correlation Coefficient of $\mathbf{s}$ and $\mathbf{X}$; the percent common variance is equal to r^2. The confidence interval of r can be estimated using bootstrapping methods: From the original sample of N_s observations, m new subsets of the same size are randomly drawn by sampling with replacement (Efron and Tibshirani, 1993). For each sub-sample, r^* is computed. The distribution of these r^* values is corrected for non-normality using the Fisher transformation ($\mathrm{atanh}(x) = 1/2^*\ln((1+x)/(1-x))$, see Davison and Hinkley, 1997). In the Fisher transformed distribution of r^*, the so- called studentized bootstrap confidence intervals are constructed from the bootstrap replicates (see Davison and Hinkley, 1997 for details). These studentized bootstrap confidence intervals finally are back-transformed to the original scale. Studentized confidence intervals are known to give the best coverage overall.

Similar to the ANCOVA, the method presented here is a general linear model with one continuous variable. (In the ANCOVA, the continuous variable may or may not be a confounding factor of no interest). Furthermore, it will be shown below that the method is an extension of a method of comparing ERP topographies between groups that have been labeled topographic analysis of variance (TANOVA, Strik et al., 1998, see Wirth et al., in press for a multifactorial implementation), although so far, the TANOVA has only been used for comparison of two conditions. The TANOVA is thus similar to the ANOVA in the sense that it used categorical independent variables, while the present method is similar to the ANCOVA by using one continuous independent variable. Accordingly, the method will be called TANCOVA.

3. Simulations

In order to test the sensitivity of the proposed model to noise, a series of simulations were computed. For each simulation, a dataset was generated consisting of an either 19 or 64 channel, zero-mean, normally distributed random map and a vector of either 12 or 50 also normally distributed random values that served as an external variable. The random map was multiplied by the random external variable to obtain simulated data that are compatible with the model outlined in formula (4). To these data, uncorrelated random noise was added that was scaled against the simulated data to obtain different signal-to-noise ratios (SNRs). Using these noisy random simulated data and the random external variable, the p-value was computed for each simulation run. Furthermore, using the squared correlation coefficient (r^2, which indicates the amount of common variance), the covariance map extracted by the simulation was compared to the random map used to

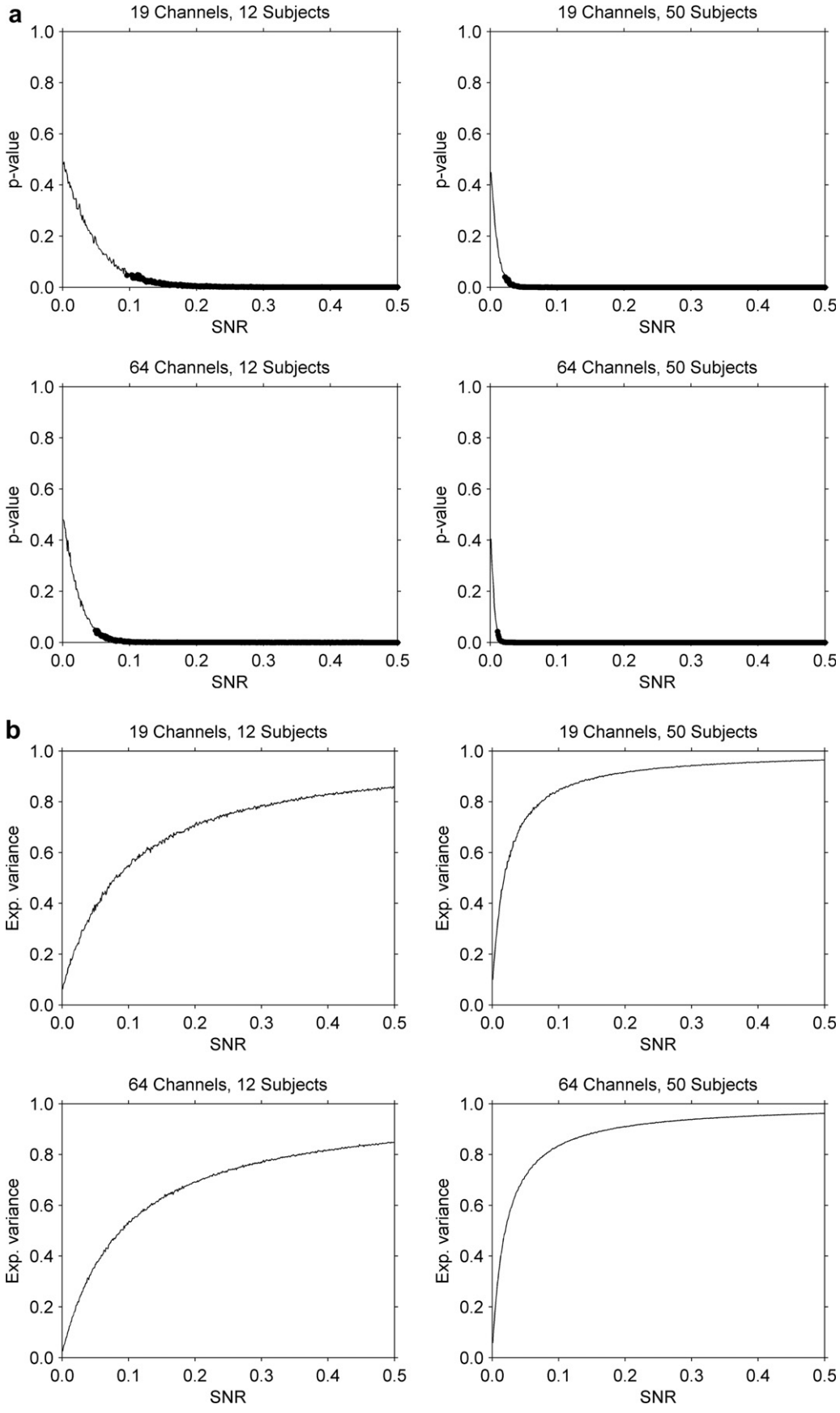

Fig. 1. Results of the simulations with 19 and 64 channel EEG data and 12 and 50 subjects. (a) The vertical axis indicates the obtained *p*-value as a function of the SNR, plotted on the horizontal axis. Bold lines indicate *p*-values below .05. (b) Mean r^2-values as function of the SNR. The sensitivity of the method clearly improves with more electrodes on the scalp and an increasing number of subjects.

generate the simulated data. Such simulations were run for SNRs ranging from 0.001 to 0.5, in steps of 0.001. For each step of the SNR ratio, 500 simulations were computed, each with 500 randomization runs. For each step of the SNR, the mean p-value and the mean r^2-value were then computed. The graphs of the mean p-values and the mean r^2-values against the SNR are shown in Fig. 1. The figure indicates that with 12 subjects and 19 channels of EEG, significant ($p < .05$) effects can be detected at SNRs above about 10%. Increasing the number of subjects or electrodes improves the sensitivity of the method, and effects can be detected at lower SNRs. The r^2-values as a function of the SNRs, however, indicate that at low SNRs, the obtained covariance maps may contain a considerable amount of noise (up to 50%) while the p-value still indicates a significant effect.

4. Example

In the following section, we show an example from an ongoing study. Seventy-four channel ERPs were collected in 10 English-speaking exchange students to Switzerland while reading single-German words. The ERPs were collected after the students had spent about 3 months in Switzerland and had already acquired some proficiency in German. After recording the ERPs, students underwent two language tests developed by "Inlingua International" language schools (http://www.inlingua.com); one was a-multiple choice test where sentences had to be completed, the second one was a vocabulary test, where 40 German words had to be translated into English. To combine the two language proficiency tests into a single (more accurate) language proficiency score, percentages of correct answers from each test were averaged for each individual (for details, see Stein et al., 2006). The study was approved by the Canton of Bern's ethical committee.

The topography of the map series was correlated to a combined score of both language tests. TANCOVA of the ERP data and the test score was applied (with estimation of the noise covariance across subjects) time instant by time instant (4 ms gap), and the p and r values with their confidence intervals were plotted (Fig. 1). The figure indicates that effects of language proficiency are found in a time window that has often been associated with language processing (Kutas and Hillyard, 1980). The covariance map at the most significant time point (448 ms post stimulus) is shown, together with a distributed inverse solution of that map (Fig. 2). The obtained covariance map had a typical N400 topography, and the inverse solutions showed maximal current density values in the left parietal regions that are often associated with language processing (see Fig. 3).

Fig. 4 shows the mean across time of the estimated noise covariance matrix across subjects. The matrix is almost diagonal, which indicates that there is little covariance of noise across subjects. The assumption of uncorrelated noise thus appears to be a good approximation.

5. Discussion

The presented methodology provides an extension of the currently available methods for multi-channel randomization statistics of EEG and MEG topographies. It can be considered as a special case of PLS, where the design matrix has only one column and corresponds to the behavioral variable **X**: As in PLS (Lobaugh et al., 2001), **X** is multiplied with the data to obtain the covariance matrix (Eq. 12). Since **X** has only one column, the corresponding covariance matrix is a single column/row vector. Applying a singular value decomposition (in its economic form) to this covariance vector yields the following: a) a vector of unity length and with the same orientation as the covariance vector (corresponding to the singular images or electrode saliences in the PLS formulation), b) a scalar that indicates the length of the covariance vector and corresponds to the singular values in PLS and c) always 1 for the design salience.

The method allows for a continuous variable to be accounted for by a linear contribution of a topographic map. When there is a good reason to assume that the relation between the external variable and the scalp field data is not directly linear, but follows some other, monotonic relation such as a logarithmic or exponential one, the external variable can easily be transformed accordingly to fit the linear framework and the method can be used based on the transformed external variable.

Comparing the randomization methodology for statistics with groups (Karniski et al., 1994; Galan et al., 1997; Maris, 2004; Greenblatt and Pflieger, 2004) and the methodology for the correlation statistics described here, it becomes apparent that the statistics for the two groups can be considered as a special case of the correlation statistics, where the external variable represents the group assignments (One group would be coded by values of 1, the other group would be coded by values of -1). It has indeed been shown that experimental designs can be reformulated as general linear models (James, 2002).

A continuous variable is likely to contain more information than a binary one. One can therefore expect that the correlation approach provides more statistical power than when the observations are assigned to groups based on a (often artificial) discretization of an initially continuous external variable (e.g. by median split). This may become useful to better assess the effects of continuous external variables such as reaction-time, test-experience, test-performance, ratings, age, drug doses and many more. The usefulness of the approach is also illustrated by the example given, where r values of up to 0.83 were found, which correspond to a common variance of nearly 70%.

By using a single, continuous external variable, the method thus has an intermediate position between statistical tests of topographic differences between two conditions, as implemented in the TANOVA, and statistical tests designed to account for several potentially continuous external variables (PLS). Its main advantages are that by using a continuous external variable rather than a binary one, it may

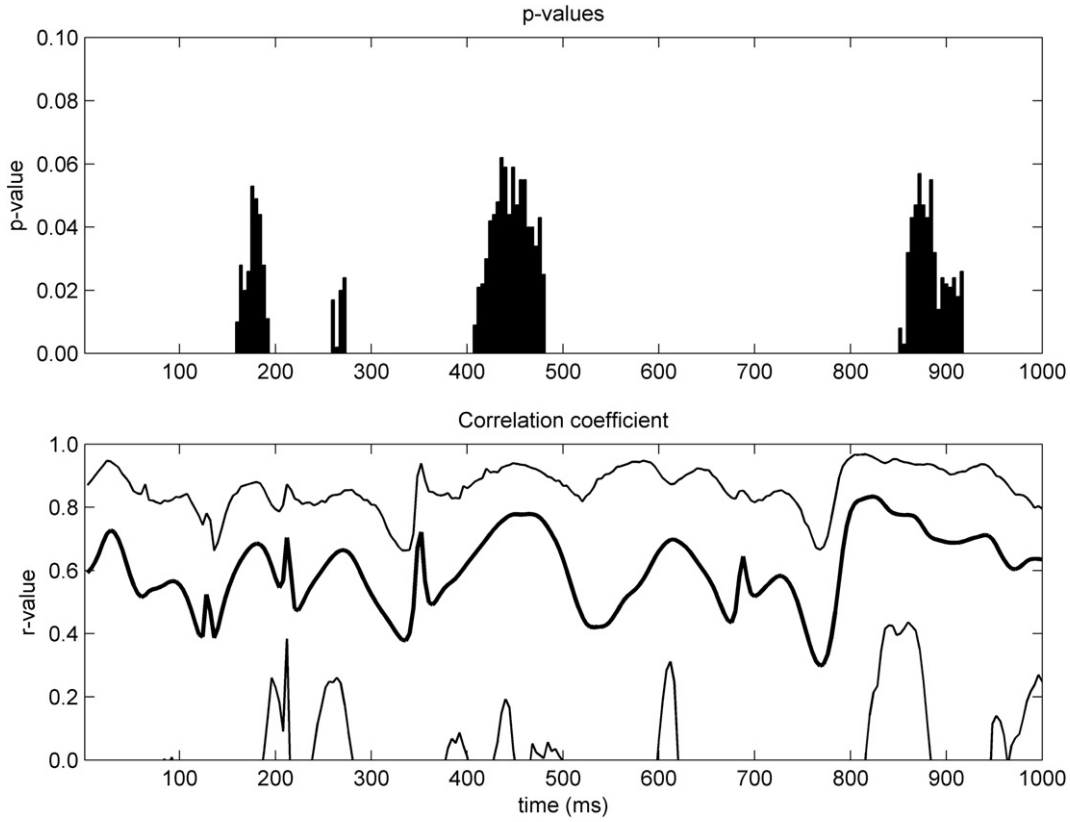

Fig. 2. Moment by moment TANCOVA results obtained when correlating 74 channel visual, German-language evoked ERP data with German proficiency scores. The data were obtained in an ongoing study investigating American students learning German. The significance (vertical, upper graph) and the correlation coefficient with its 95% studentized confidence interval (vertical, lower graph) are shown as function of time (horizontal axis). There was a significant effect of language proficiency between 400 and 500 ms and around 900 ms post-stimulus. Note that during periods of non-significance, the confidence interval of *r* often includes zero, which is closely related to the null hypothesis we tested, namely that there may be no relation. Regarding the width of the confidence interval, the simulations shown in Fig. 1 indicate that one may be able to reject the null-hypothesis while the estimation of the covariance map (and thus the correct *r*-value) might contain considerable noise and thus be quite uncertain.

provide more statistical power than the TANOVA. By limiting the number of external variables to one, no rotation of the covariance matrix is necessary, and the resulting covariance map remains directly interpretable in terms of intracerebral sources associated with the external variable.

In order to overcome the problem of multiple testing across time, one may employ procedures to identify ERP components (i.e. Pascual-Marqui et al., 1995; Michel et al., 2001) and average the data across time periods belonging to the same component before the statistics are computed. Alternatively, one may employ techniques for the correction of multiple comparisons that are commonly employed in functional neuroimaging (Nichols and Holmes, 2002; Carbonell et al., 2004).

The method implies several assumptions that need some consideration here: First, it is assumed that scalp fields are additive (which is well founded in the physics of EEG and MEG), and that the relation between the external variable and the generators that cause the variance in the external variable is linear (this may be adapted by transforming the external variable). By basing the analysis on a vector of scalp potentials that is scaled for each subject (Eq. 5), it is furthermore assumed that across subjects a common vector of scalp potentials corresponds to the same distribution and orientation of current density. Theoretically, this is not proven, because the inverse problem of the EEG (deducing the current density distribution from a scalp potential vector) is in general non-unique (Eqs. (2) and (3) imply Eq. (5), but Eq. (5) does not generally imply Eqs. (2) and (3)). However, the same theoretically not generally proven assumption is made when averaged ERPs are computed. Since ERP averaging has produced an abundant mass of meaningful results and often served as the basis for the computation of convincing inverse solutions, we think that practically, this assumption is very plausible. (In addition, the opposite hypothesis that substantially different individual current density distributions produce a consistent effect upon a single common vector of scalp potentials is not very convincing.)

Taken together, the assumptions made here are thus well justified, general, simple and closely related to the assumptions made when applying standard ERP averaging. The method is therefore ideal to establish, on a global, model-free level whether there is a statistical effect in the

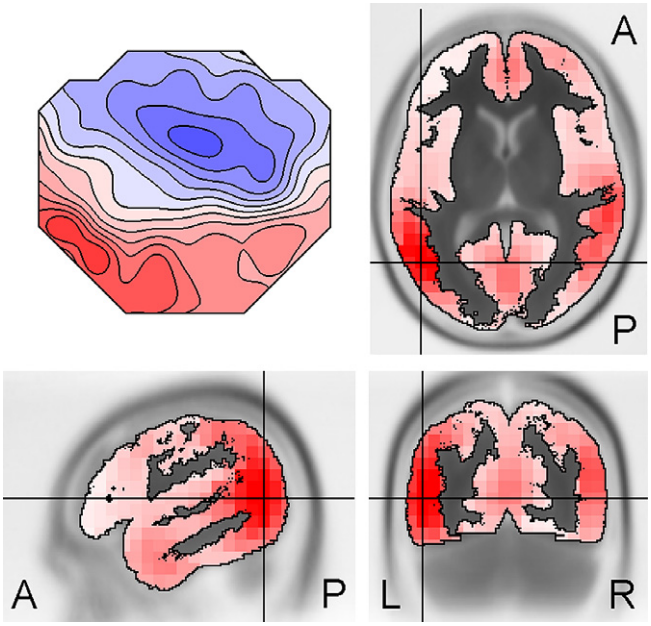

Fig. 3. Covariance map (upper-left graph, seen from above, with isopotential contour lines) of the data used in Fig. 2 at the moment with minimal p-value (448 ms post stimulus), and LORETA inverse solution of the covariance map. The LORETA solution was computed in 2394 cortical voxels ($7 \times 7 \times 7$ mm) of the digitized brain atlas of the Montreal Neurological Institute (MNI). The crosshair line indicates the point of maximal current source density, which was at Talairach coordinates $x = -52$, $y = -67$, $z = 8$, close to left BA 37, 19 and 39.

data. Since the GFP used to compute the size of the relation is reference-independent, the obtained significance of the effect is also reference-independent. Significant effects can then of course be further investigated with more model-driven approaches such as voxel-based statistics based on distributed inverse solutions. This sequence of statistical analysis has previously been used successfully for group statistics (e.g. Strik et al., 1998; Lehmann et al., 2005). Care should, however, be taken in the interpretation of negative results, since they might indicate either that an effect is indeed absent, or that some of the above-mentioned assumptions do not hold.

Another benefit of the method is that the obtained covariance map is directly interpretable. Given that the assumptions hold, it represents the scalp field of intracerebral generators that account for the effect observed in the external variable with the full spatial resolution of the measured data. This is justified by the fact that both the statistical model and the fields generated by different generators are additive. The covariance map can therefore also directly be submitted to a source localization procedure, or it can be used as a spatial filter that outputs the strength of the generators related to the external variable.

There is no general answer to the question whether to use global scalp field statistics or voxel-wise statistics of estimated intracerebral current density; the methods emphasize different features of the data, and make different

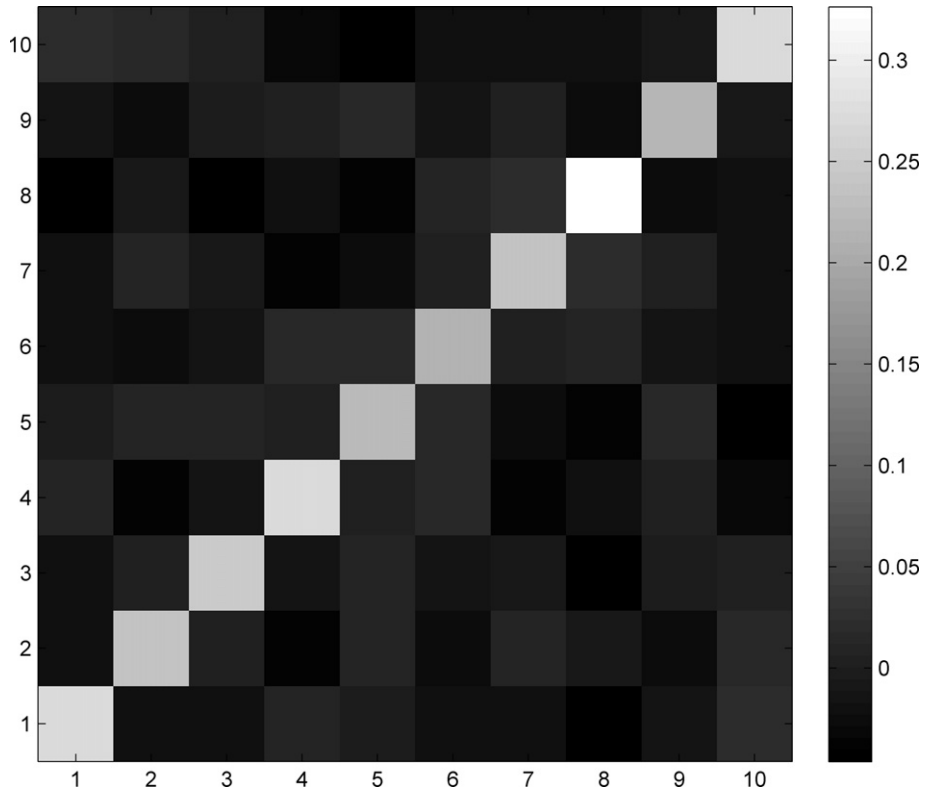

Fig. 4. Mean covariance matrix of the noise across subjects estimated from the data shown in Fig. 3. Ninety seven percent of the variance of the matrix is on the diagonal, indicating little covariance of noise across subjects.

assumptions. Scalp field maps are very sensitive to differences in source orientation, which is usually disregarded in inverse solutions. On the other side, some changes in source localization may produce disproportional small differences in scalp field configuration, and source orientation may be noisy, such that appropriate inverse solutions may be considerably more sensitive to some changes compared to scalp field maps. Another difference is that voxel-wise testing typically needs corrections for multiple testing, which is not the case with the proposed methodology. Finally, both if results of voxel-wise statistics of some inverse solution or results of an inverse solution of an extracted scalp field map are shown, these results depend of course on the correctness of and limitations of the inverse solution employed. For LORETA (which was used in the example shown), these limitations are that the data should contain little noise, that the real (and unknown) current density distribution does not contain high spatial frequencies, and that the real current density distribution fits into the solution space of the inverse solution. Since the real current density distribution is unknown, it remains difficult to say whether the data match these limitations.

Appendix A. Supplementary data

In order to facilitate the usage of the method, a platform-independent Matlab version is made available in the Supplementary Appendix A (doi:10.1016/j.clinph.2007.12.023) or for download (http://www.puk.unibe.ch/tk2/Software.htm).

References

Carbonell F, Galan L, Valdes P, Worsley K, Biscay RJ, Diaz-Comas L, et al. Random field-union intersection tests for EEG/MEG imaging. Neuroimage 2004;22:268–76.

Davison AC, Hinkley DV. Bootstrap methods and their application. Cambridge: Cambridge University Press; 1997.

Duffy FH, Bartels PH, Burchfiel JL. Significance probability mapping – an aid in the topographic analysis of brain electrical-activity. Electroencephalogr Clin Neurophysiol 1981;51:455–62.

Efron B, Tibshirani RJ. An introduction to the bootstrap. London: Chapman & Hall; 1993.

Galan L, Biscay R, Rodriguez JL, Perez-Abalo MC, Rodriguez R. Testing topographic differences between event related brain potentials by using non-parametric combinations of permutation tests. Electroencephalogr Clin Neurophysiol 1997;102:240–7.

Gevins A. High resolution evoked potentials of cognition. Brain Topogr 1996;8:189–99.

Greenblatt RE, Pflieger ME. Randomization-based hypothesis testing from event-related data. Brain Topogr 2004;16:225–32.

James GM. Generalized linear models with functional predictors. J Roy Stat Soc B 2002;64:411–32.

Karniski W, Blair RC, Snider AD. An exact statistical method for comparing topographic maps, with any number of subjects and electrodes. Brain Topogr 1994;6:203–10.

Kondakor I, Pascual-Marqui RD, Michel CM, Lehmann D. Event-related potential map differences depend on the prestimulus microstates. J Med Eng Technol 1995;19:66–9.

Kutas M, Hillyard SA. Reading senseless sentences: brain potentials reflect semantic incongruity. Science 1980;207:203–5.

Lehmann D, Faber PL, Galderisi S, Herrmann WM, Kinoshita T, Koukkou M, et al. EEG microstate duration and syntax in acute, medication-naive, first-episode schizophrenia: a multi-center study. Psychiatry Res 2005;138:141–56.

Lehmann D, Skrandies W. Reference-free identification of components of checkerboard-evoked multichannel potential fields. Electroencephalogr Clin Neurophysiol 1980;48:609–21.

Lobaugh NJ, West R, McIntosh AR. Spatiotemporal analysis of experimental differences in event-related potential data with partial least squares. Psychophysiology 2001;38:517–30.

Luu P, Tucker DM, Englander R, Lockfeld A, Lutsep H, Oken B. Localizing acute stroke-related EEG changes: assessing the effects of spatial undersampling. J Clin Neurophysiol 2001;18:302–17.

Mardia KV, Kent JT, Bibby JM. Multivariate analysis. London: Academic Press; 1979.

Maris E. Randomization tests for ERP topographies and whole spatiotemporal data matrices. Psychophysiology 2004;41:142–51.

McIntosh AR, Bookstein FL, Haxby JV, Grady CL. Spatial pattern analysis of functional brain images using partial least squares. Neuroimage 1996;3:143–57.

Michel CM, Murray MM, Lantz G, Gonzalez S, Spinelli L, Grave dP. EEG source imaging. Clin Neurophysiol 2004;115:2195–222.

Michel CM, Thut G, Morand S, Khateb A, Pegna AJ, Grave dP, et al. Electric source imaging of human brain functions. Brain Res Brain Res Rev 2001;36:108–18.

Mosher JC, Leahy RM, Lewis PS. EEG and MEG: forward solutions for inverse methods. IEEE Trans Biomed Eng 1999;46:245–59.

Nichols TE, Holmes AP. Nonparametric permutation tests for functional neuroimaging: a primer with examples. Hum Brain Mapp 2002;15:1–25.

Pascual-Marqui RD, Michel CM, Lehmann D. Segmentation of brain electrical activity into microstates: model estimation and validation. IEEE Trans Biomed Eng 1995;42:658–65.

Steger J, Imhof K, Steinhausen H, Brandeis D. Brain mapping of bilateral interactions in attention deficit hyperactivity disorder and control boys. Clin Neurophysiol 2000;111:1141–56.

Stein M, Dierks T, Brandeis D, Wirth M, Strik W, Koenig T. Plasticity in the adult language system: a longitudinal electrophysiological study on second language learning. Neuroimage 2006;33:774–83.

Strik WK, Fallgatter AJ, Brandeis D, Pascual-Marqui RD. Three-dimensional tomography of event-related potentials during response inhibition: evidence for phasic frontal lobe activation. Electroencephalogr Clin Neurophysiol 1998;108:406–13.

Vasey MW, Thayer JF. The continuing problem of false positives in repeated measures anova in psychophysiology – a multivariate solution. Psychophysiology 1987;24:479–86.

Wirth M, Horn H, Koenig T, Razafimandimby A, Stein M, Mueller T, Federspiel A, Meier B, Dierks T, Strik W. The early context effect reflects activity in the temporo-prefrontal semantic system – evidence from electrical neuroimaging of abstract and concrete word reading. Neuroimage, in press.