## Consensus on PCA for ERP data, and sensibility of unrestricted solutions

In our recent invited editorial (Kayser and Tenke, 2005), we placed the contributions of a simulation study by Dien et al. (2005) in the broader context of what we perceive as a renaissance of PCA methodology in electrophysiology. This editorial also addressed several misrepresentations and apparent misunderstandings of Dien et al. (2005) regarding the research strategy and recommendations given in our previous article (Kayser and Tenke, 2003). In his response to our critical explanations, Dien (2006) shows a much better understanding of our recommendations than when his simulation study was published. Indeed, a number of statements in his letter now paraphrase our own, although the need for such a picturesque repetition eludes us. While we prefer a more scientific language appropriate to the nature of the data, we suspect that some of his argumentation and style were and are merely rhetorical. Unfortunately, even though some of his points are well taken, he still attributes conclusions or recommendations to us that are incorrect, and we gladly accepted the invitation to respond to his letter.

First and foremost, we are in complete agreement with Dien on the value of PCA for ERPs. Promoting more objective approaches to the analyses of ERP data, PCA methodology being just one example, appears to be a common goal to which we will continue to contribute. Secondly, Dien (2006) identifies two areas of disagreement, namely factor scaling and factor restriction, and also provides additional comments regarding the Promax rotation. We will underscore our consensus on the specific methodological issue of scaling choices. However, we will argue for the sensibility of using unrestricted PCA solutions in light of empirical evidence, and restate our reasons and provide additional considerations for not enthusiastically endorsing Promax rotation for empirical ERP data at this time.

### 1. Scaling choices

Dien (2006) stresses that a scaling choice (covariance or correlation) has to be made at the extraction, rotation and presentation step. Dien goes to great lengths in soliciting personal communications from the SPSS and SAS technical support staff, verifying that Varimax rotation is performed on the 'standardized loadings' regardless of the association matrix used for factor extraction.[1] Thus, the implementation of this scaling choice is rather nebulous in SPSS and SAS, and if it cannot be altered, it is actually worse than a default setting. In contrast, BMDP is fairly clear on the separation of factor extraction and rotation, which involves separate statements for extraction (i.e. FORM=COVA or FORM=CORR for factoring the covariance or correlation matrix; default is CORR) and rotation (i.e. LOAD=COVA or LOAD=CORR for using covariance or correlation loadings; cf. Dixon, 1992, p. 355). As with SPSS and SAS, BMDP's Varimax rotation (ROTATE METH=VMAX) uses Kaiser's normalization as the default setting unless NO NORMAL is specified (cf. Dixon, 1992, p. 357). Our original publication explicitly stated the BMDP syntax used (Kayser and Tenke, 2003, footnote 3, p. 2312), and we reiterated in our editorial that BMDP with this syntax was used for all factor analyses (Kayser and Tenke, 2003, p. 1750).[2]

With respect to the choice of association matrix for factor extraction, one of the two main purposes of the Kayser and Tenke (2003) study, our theoretical rationale for not using a correlation matrix was straightforward: ERP data variance should be small during inactive periods (i.e. baseline) but large during active periods (i.e. components). Fig. 1 graphs covariance and correlation matrices for the two ERP data sets used by Kayser and Tenke (2003). Due to the 'restriction' and 'inflation' of highly intercorrelated neighboring variables to one, the data space of a correlation matrix (Fig. 1C–D) appears by comparison less suited for factorization than that of a covariance matrix (Fig. 1A–B). As this reasoning is strongly supported by the empirical evidence (e.g. Dien et al., 2005; Kayser and Tenke, 2003), we, too, considered this point settled.

As for the scaling choice during Varimax rotation, Dien et al. (2005) clarified that the use of Kaiser's normalization cancels any differences as long as the set of submitted factor loadings is identical. However, given our rationale with respect to correlation loadings, we have argued that low variance factors may attain prominence when the original data space (in microvolts) is reexpressed in standardized form (i.e. the total variance is equal to the number of variables; cf. Fig. 1). When the factor's variance is computed as the sum of squares for the elements of the corresponding column in the factor loading matrix, the explained variance will differ (i.e. not only in absolute terms but also as a fraction of the total variance). BMDP expresses the variance contribution of a factor this way (i.e. using appropriate mathematics), which conforms to the decision of submitting either correlation or covariance loadings.[3] As a result of factor rotation, the

[1] The term 'standardized loadings' is evidently common language among SPSS, SAS and BMDP representatives. However, in complete consensus with Dien et al. (2005), we feel that the terms 'correlation loadings' and 'covariance loadings' are preferable.

[2] Our lab, as many others, first applied PCA to ERP data using BMDP on mainframe computers long before desktop implementations were available. While BMDP remained largely unchanged, competing software packages using different approaches underwent drastic changes in scope over recent years (e.g. changes of the procedure FACTOR from SPSS versions 6.1–13.0). Although it is beyond the scope of this letter to account for all these differences and changes, we fully agree with Dien that inconsistent access to different 'standard' statistical packages constitutes a problem for the field.

[3] It must be noted that this difference in expression of variance (i.e. using either the original or standardized data space) is at the core of our original argumentation and apparently also the reason for the misunderstandings with Dien, as this difference will result in what we called 'erroneous' high-variance factors (i.e. a high-variance factor in standardized data space but a low-variance factor in original data space). In his letter, Dien elaborates on assigning equal weight to relative inactive time points, apparently unaware that he is merely reiterating our original argument.

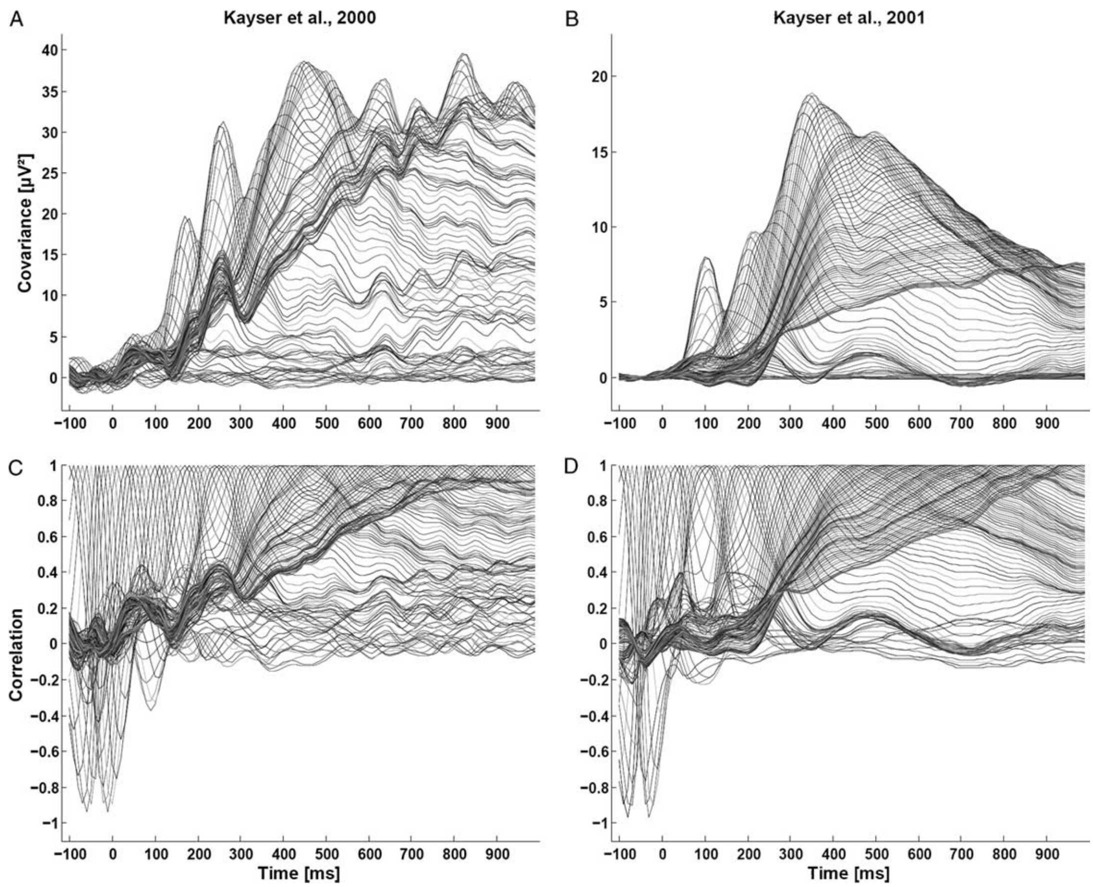


Fig. 1. Covariance (A,B) and correlation (C,D) association matrices of two typical ERP data sets consisting of healthy adults (Kayser et al., 2000; 2001; cf. Kayser and Tenke, 2003). The underlying ERP/PCA component structure is intuitively more obvious for a covariance matrix, which does not limit the variance range to one. While the diagonal of a correlation matrix is characterized by ones, correlations close to one are also observed for neighboring time points (variables), which is particularly evident during the baseline period (−100 to 0 ms).

variance explained by a factor will change. To conform to this change in explained variance, BMDP resorts the factors accordingly *after* rotation, which appears to be reasonable and not 'an arbitrary decision.' Moreover, it is consistent with the overall purpose of PCA as a means of systematic data reduction, and is implicit in the rationale for various approaches to factor restriction.

With regard to covariance-based solutions using correlation loadings, we have to give Dien credit for his persistence. Although he still incorrectly refers to our Fig. 3 (Kayser and Tenke, 2003) as displaying standardized loadings derived from a covariance matrix (cf. our previous comments on this error in Kayser and Tenke, 2005, footnote 2), his argument partly holds for our Fig. 4 which indeed also presented these data. We had assumed that BMDP would consistently sort factors also *before* rotation, which could result in different sets of factors submitted to rotation; however, we have now verified that this is not the case. Rather, it appears that the original order resulting from the covariance-based extraction is maintained, resulting in the possibility that the sequence of explained variance is out of order (i.e. now using standardized data space). Dien is therefore correct that, in this case, identical sets of factors stemming from covariance-based solutions are rotated. These factors differ only in terms of variance explained (and hence in their order after rotation), which could be reconciled by rescaling (and recomputing their variance contribution, as this is not necessarily provided by standard statistical packages). With this acknowledgment, we should have reached a consensus on the implication of the scaling choices.

## 2. Restricted solutions

Turning to the issue of factor retention, the other main purpose of the Kayser and Tenke (2003) study, Dien claims that our editorial differs from our original paper by newly 'advocating a threshold approach' and that this is a question of 'analytical philosophy.' We wholeheartedly disagree on both counts. We have never suggested that all factors be considered for statistical analyses, any more than we'd recommend submitting all possible ERP time points (or windows) to inferential statistics. It's not about a variance threshold, but using the available tool (e.g. time window average or PCA factor) as an empirical measure of the concept of an ERP component (i.e. apply meaning to a factor). Variance alone is insufficient for separating

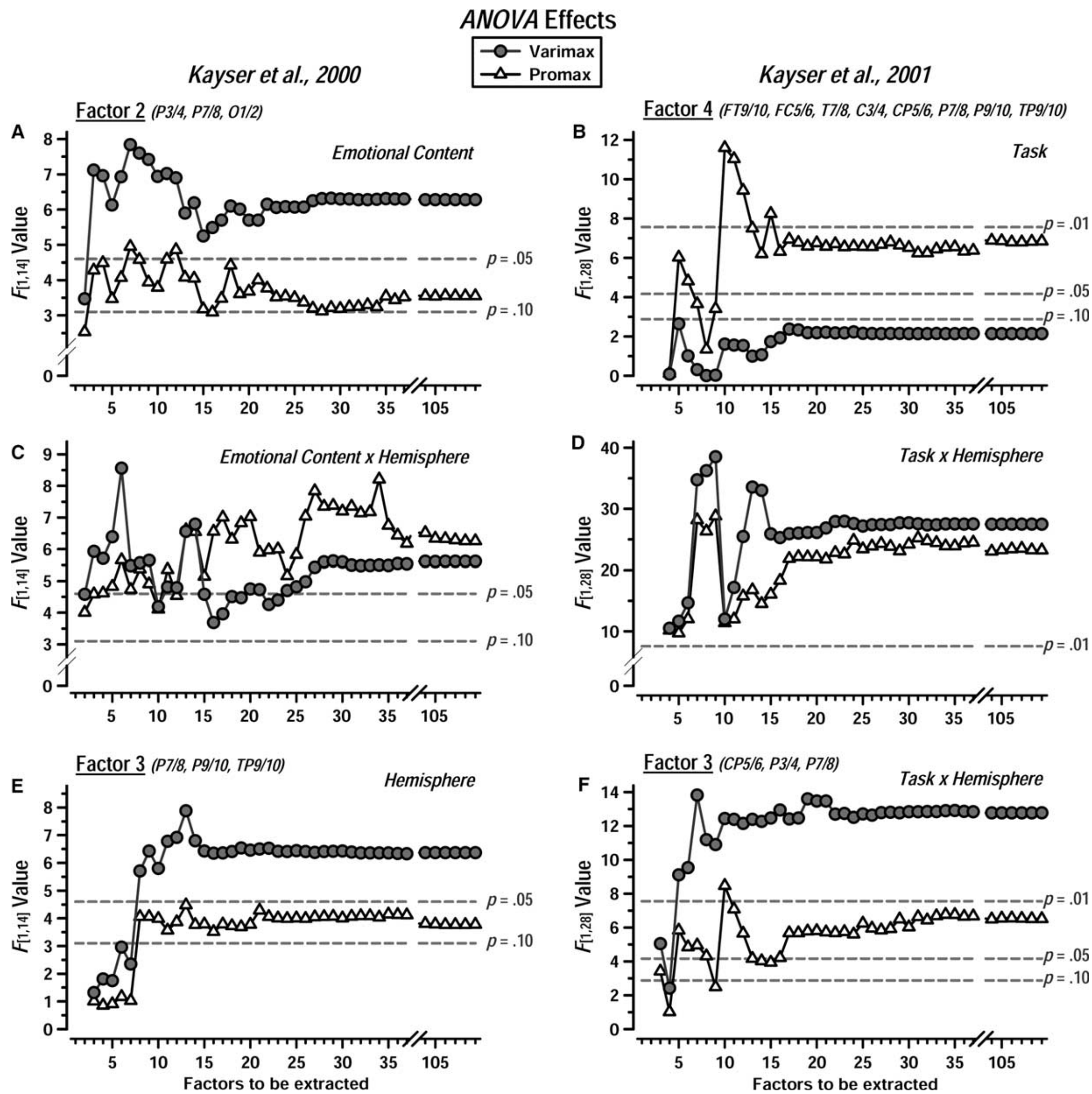


Fig. 2. Statistical effects ($F$ values) of analyses of variance (ANOVA) performed on PCA factor scores derived from covariance-based solutions followed by Varimax (circles) or Promax (triangles) rotation of covariance loadings as a function of extraction criterion (number of factors to be extracted). Representative examples are provided for data of two previous studies (cf. Kayser and Tenke, 2003). For all analyses presented in the left column (Kayser et al., 2000), gender was a between-subjects factor ($N=16$, 7 male), and emotional content (2), visual field (2), hemisphere (2), and site (3 homologous electrode pairs as indicated in parentheses) were within-subjects factors. Shown are critical $F$ tests (all $df=1,14$): for factor 2 (P3, peak latency 460 ms), the emotional content main effect (A) and the Emotional Content × Hemisphere interaction effect (C); for factor 3 (N2, peak latency 260 ms), the hemisphere main effect (E). For all analyses presented in the right column (Kayser et al., 2001), gender and response hand were between-subjects factors ($N=32$, 18 male, 18 right press), and task (2), hemisphere (2), and site (8 or 3 homologous electrode pairs as indicated in parentheses) were within-subjects factors. Shown are critical $F$ tests (all $df=1,28$): for factor 4 (N2, peak latency 210 ms), the task main effect (B) and the Task × Hemisphere interaction effect (D); for factor 3 (P3, peak latency 320 ms), the Task × Hemisphere interaction effect (F). Dashed lines indicate $F$ values corresponding to conventional significance levels of $p=.01$, .05 and .10. Factor numbers refer to Varimax-rotated solutions, which were paired with the corresponding factor of the Promax-rotated solutions (closest factor loadings peak latency within the first six factors). Note that with more conservative extraction criteria, the $F$ statistic of any given effect is highly variable, particularly for Promax-rotated solutions. In contrast, statistical effects become very stable with more liberal extraction criteria, particularly for Varimax-rotated solutions. The abscissa break between 35 and 105 extracted factors indicates the omission of largely redundant $F$ values.

physiologically meaningful components and noise. For example, a PCA factor with a restricted topography corresponding to visual P100 may be appropriately quantified, despite the fact that it accounts for a only a small proportion of the total variance. Conversely, a high-variance factor may be considered noise if it clearly relates to a known artifact (Kayser and Tenke, 2003).

Dien is still quite mistaken in his assumption that our rationale for recommending unrestricted solutions is based on the significance levels of the $F$-tests reported in our 2003 study, which happened to yield 'more significance with the unrestricted solution,' evidently leading to the erroneous assertion that we were 'equating size of an effect with meaningfulness.' Our argument has nothing to do with the *size* of an $F$ value but rather with its *stability* as a function of factor restriction. Our particular concern was the high variability of $F$ values for restricted solutions (i.e. extracting less than 30 factors), which almost arbitrarily varied across conventional significance levels for neighboring solutions (i.e. extracting $n$ or $n+1$ factors). In contrast, no changes in $F$ values (i.e. large or small) were observed with more liberal extraction criteria, as clearly shown in Fig. 8 of our 2003 study (please note that the abscissa in this figure omitted identical $F$ values for solutions from 38 to 103 extracted factors). This principle was more or less independent of the association matrix used for factor extraction.

Dien is concerned that 'without knowing the true state of affairs it is unclear whether' these $F$ values represent 'an advance away from Type II error or a retreat into Type I error.' He suggests that a replication of 'this analysis with a reliable and well-characterized effect such as an oddball task' would be helpful for an evaluation of this principle. As our original article did not report the detailed findings for our auditory oddball ERP data due to space limitations, we are presenting these data here. Furthermore, we took the opportunity to systematically compare Varimax and Promax rotation with respect to the stability of $F$ values as a function of factor retention, which was regrettably not included in the Dien et al. (2005) study. This also enabled us to indirectly compare the BMDP PCA-Varimax algorithms, which can be emulated by our published Matlab code (appendix of Kayser and Tenke, 2003; available at http://psychophysiology.cpmc.columbia.edu/erpPCA.html), with those implemented in Dien's Matlab PCA toolbox (v. 1.09). Using ERPs of healthy adults stemming from two previous studies (Kayser et al., 2000; 2001; cf. details given in Kayser and Tenke, 2003, Section 2.2), these data were repeatedly submitted to the main function of the PCA toolbox (doPCA), varying only the instructions for factor rotation (Varimax or Promax) and retention (extracted factors ranged from 1 to 109), whereas the choices of using a covariance association matrix for extraction and Kaiser's normalization during rotation were kept constant.

Fig. 2 shows representative ANOVA effects for these two data sets (i.e. those crucial to the study's objective), comparing solutions with Varimax or Promax rotation as a function of extraction criterion (number of factors to be extracted). Varimax and Promax factors were carefully matched by the peak latencies of their covariance loadings and by visual inspection of their factor score topographies. First, the sequences of $F$ values stemming from Varimax solutions for the Kayser et al. (2000) data (Fig. 2A,C,E, circles) are indeed almost identical with those previously reported with BMDP software (cf. Kayser and Tenke, 2003, Fig. 8 A,C,B, circles), which is reassuring. Second, the sequence of $F$ values stemming from Varimax solutions for the Kayser et al. (2001) oddball data (Fig. 2B,D,F, circles) follows the same principle, that is, stable $F$ values with liberal solutions, but variable $F$ statistics with restricted solutions (larger $F$ values were notably observed for some more restricted solutions, for example, for the task×hemisphere interaction; Fig. 2 D,F). Third, although clearly more variable than Varimax solutions, Promax solutions by-and-large also followed the same principle, yielding highly variable $F$ statistics with more restricted solutions, but fairly stable $F$ values with liberal solutions (Fig. 2 A–F, triangles).

Further adding to this evidence are the simulation findings reported by Dien et al. (2005), which clearly showed improved prototype reconstruction for *all* unrestricted solutions. The point is that liberal solutions improve the quality of the high-variance components ultimately considered for statistical analysis, and unrestricted solutions as the final end point along a continuum do not degrade these better solutions. In short, our recommendation for using unrestricted PCA solutions is *not* based on 'analytical philosophy,' but is instead mandated by the available empirical evidence and the scientific method, because it can be reasonably justified for real ERP data, in which 'the correct answer' is not known.

## 3. Promax rotation

After applying Promax to our data, we can now offer more insightful comments and qualifications of our encouraging remarks regarding the usefulness of Promax for the simulation of Dien et al. (2005). Due to factor intercorrelations after Promax rotation, it is anything but straightforward to rank the factors by variance in the extended data space, which may contribute to our disagreements. Using the unrestricted Promax solutions, for instance, 10 factors (ranked 5–14 after sorting by explained variance) accounted for 10.2–20.6% of the original ERP variance for the Kayser et al. (2000) data (562.2% total variance), and 18 factors (ranked 5–22) accounted for 10.1–17.8% variance for the Kayser et al. (2001) data (710.2% total). In contrast, unrestricted orthogonal rotations obviously still account exactly for 100% of the ERP variance, or a fraction thereof if fewer factors are rotated. Although Dien's program reports unique variance contributions of the factors which may also be used to evaluate and rank Promax factors, this does not overcome the problem of deciding which factors best represent the physiologic processes of interest. Simply put, the Promax factor structure is redundant and not concise. On the other hand, Promax-rotated factor loadings were by comparison sharper than Varimax-rotated loadings, and had effectively zero secondary loadings for unrestricted solutions.

Dien asserts that Promax has already been shown to be superior to Varimax using real ERPs, given divergent P300 source localization results (Dien et al., 2003). However, Dien disregards the fact that a factor retention criterion was used in this study, leaving the appropriate comparison untested. The findings of the Dien et al. (2005) simulation study strongly suggest a more favorable outcome for unrestricted Varimax solutions.

Dien's comments on our discussion of parsimony concludes with the observation that the brain is not orthogonal. We heartily agree, but would again like to note that certain models are inherently more parsimonious than others. Dien's suggestion of a single component accounting for an entire ERP data set as an example of failed parsimony misses the mark. If the ERP variance can be accounted for by a single factor, the PCA-Varimax approach will conform to the data, and this is not only parsimonious but also accurate (e.g. see our constructed example with artificial data in Kayser and Tenke, 2003, Figs. 1A and 2A). Another example is when a single PCA factor accounts for the repeated occurrence of a prominent ERP feature, such a consecutive N1 amplitudes in S1–S2 paradigms (e.g. Kayser et al., 2005). Our standard use of the term parsimony has to do with the conciseness of summarizing data. No one would assert that the brain uses only linear processes, yet a linear, volume-conduction model of EEG and ERP data has proven to be sufficiently valuable for empirical studies (Kayser and Tenke, 2006 a,b; Tenke and Kayser, 2005). To argue for using a less tractable, nonlinear model, it must be shown to have sufficient empirical value in a particular paradigm (i.e. less ambiguous, easier to interpret).

## 4. Final remarks

A word of reassurance may be appropriate for those who are unfamiliar with PCA: Dien's and our own concerns all address improvements of PCA methodology for ERP applications. We are in total accord that PCA affords superior description and quantification of ERP components when compared to other conventional methods (e.g. peak or window estimates). Although our recommendation of using unrestricted solutions is an untraditional refinement that understandably requires cautious and critical evaluation by the research community, the same considerations apply to Promax: a direct and thorough, side-by-side comparison of Promax is needed to explore and highlight both its strengths and weaknesses using real ERP data. Simulations are quite useful, but they are not sufficient. Our experience with Promax rotations using real ERP data stemming from our paradigms deepens our concerns about shared variance, and we are looking forward to a satisfactory resolution that would allow a final consensus on this issue.

## References


Dien J. Progressing towards a consensus on PCA of ERPs. Clin Neurophysiol 2006;117:699–702.

Dien J, Spencer KM, Donchin E. Localization of the event-related potential novelty response as defined by principal components analysis. Cogn Brain Res 2003;17(3):637–50.

Dien J, Beal DJ, Berg P. Optimizing principal components analysis of event-related potentials: matrix type, factor loading weighting, extraction, and rotations. Clin Neurophysiol 2005;116(8):1808–25.

Dixon WJ, editor. Statistical software manual: to accompany the 7.0 software release. Berkeley, CA: University of California Press; 1992.

Kayser J, Tenke CE. Optimizing PCA methodology for ERP component identification and measurement: theoretical rationale and empirical evaluation. Clin Neurophysiol 2003;114(12):2307–25.

Kayser J, Tenke CE. Trusting in or breaking with convention: towards a renaissance of principal components analysis in electrophysiology. Clin Neurophysiol 2005;116(8):1747–53.

Kayser J, Tenke CE. Principal components analysis of Laplacian waveforms as a generic method for identifying ERP generator patterns: I. Evaluation with auditory oddball tasks. Clin Neurophysiol 2006a; 117(2):348–68.

Kayser J, Tenke CE. Principal components analysis of Laplacian waveforms as a generic method for identifying ERP generator patterns: II. Adequacy of low-density estimates. Clin Neurophysiol 2006b;117(2):369–80.

Kayser J, Bruder GE, Tenke CE, Stewart JE, Quitkin FM. Event-related potentials (ERPs) to hemifield presentations of emotional stimuli: differences between depressed patients and healthy adults in P3 amplitude and asymmetry. Int J Psychophysiol 2000;36(3):211–36.

Kayser J, Bruder GE, Tenke CE, Stuart BK, Amador XF, Gorman JM. Event-related brain potentials (ERPs) in schizophrenia for tonal and phonetic oddball tasks. Biol Psychiatry 2001;49(10):832–47.

Kayser J, Tenke CE, Gates NA, Kroppmann CJ, Gil RB, Bruder GE. Verbal working memory (WM) in schizophrenia: event-related brain potential (ERP) findings for the Word Serial Position Test. Biol Psychiatry 2005;57:201S.

Tenke CE, Kayser J. Reference-free quantification of EEG spectra: combining current source density (CSD) and frequency principal components analysis (fPCA). Clin Neurophysiol 2005;116(12):2826–46.



Jürgen Kayser
Craig E. Tenke
*Department of Biopsychology,*
*New York State Psychiatric Institute, New York*
*Department of Psychiatry, Columbia University,*
*College of Physicians and Surgeons New York, USA*
*E-mail address:* kayserj@pi.cpmc.columbia.edu



doi:10.1016/j.clinph.2005.11.015