Clinical Neurophysiology, in press

Optimizing PCA Methodology for ERP Component Identification and Measurement: Theoretical Rationale and Empirical Evaluation

Jürgen Kayser*, Craig E. Tenke

Department of Biopsychology, Unit 50, New York State Psychiatric Institute, New York, USA

Received 25 March 2003; revised 29 May 2003; accepted 8 July 2003

Abstract

Objective: To determine how specific methodological choices affect "data-driven" simplifications of event-related potentials (ERPs) using principal components analysis (PCA). The usefulness of the extracted component measures can be evaluated by knowledge about the variance distribution of ERPs, which are characterized by the removal of baseline activity. The variance should be small before and at stimulus onset (across and within cases), but large near the end of the recording epoch and at ERP component peaks. These characteristics are preserved with a covariance matrix, but lost with a correlation matrix, which assigns equal weights to each sample point, yielding the possibility that small but systematic variations may form a factor.

Methods: Varimax-rotated PCAs were performed on simulated and real ERPs, systematically varying extraction criteria (number of factors) and method (correlation/covariance matrix, using unstandardized/standardized loadings before rotation).

Results: Conservative extraction criteria changed the morphology of some components considerably, which had severe implications for inferential statistics. Solutions converged and stabilized with more liberal criteria. Interpretability (more distinctive component waveforms with narrow and unambiguous loading peaks) and statistical conclusions (greater effect stability across extraction criteria) were best for unstandardized covariance-based solutions. In contrast, all standardized covariance- and correlation-based solutions included "high-variance" factors during the baseline, confirming findings for simulated data.

Conclusions: Unrestricted, unstandardized covariance-based PCA solutions optimize ERP component identification and measurement.

Keywords: Event-related Potential (ERP); Principal Components Analysis (PCA); Covariance matrix; Correlation matrix; Varimax rotation; Extraction criteria

1. Introduction

1.1. Conceptualizations of ERP components

Event-related potentials (ERPs) reflect activity patterns of neuronal generators, such as the modality-specific, sequential activation in afferent and central pathways evoked by transient sensory stimuli, which sum and volume conduct to scalp electrodes (e.g., Rose and Woosley, 1949). The component structure of ERPs incorporates the transfer properties of these neuronal processes in their time course and topography (e.g., Kraut et al., 1985; Schroeder et al., 1991; Tenke et al., 1993), although much of this information is masked in the awake, behaving organism. Thus, ERP components are classically conceived as an electrophysiologic correlate of the underlying neuronal generators associated with information processes.

* Corresponding author. Tel.: +1-212-543-5169; fax:+1-212-543-6540. *E-mail address*: kayserj@pi.cpmc.columbia.edu (J. Kayser) 1051 Riverside Drive, New York, NY 10032, USA

The linkage between ERP waveforms and the underlying biophysics is less obvious in complex cognitive paradigms, which typically focus on differences between overlapping ERP components (e.g., Friedman, 1990; Kutas and Hillyard, 1980; Näätänen and Picton, 1987; Simson et al., 1976; Squires et al., 1975; Sutton et al., 1965). Although ERP studies appear to share a common nomenclature, based on polarity, sequence or timing, and attributed function (e.g., N1/N100, P3/P300), a prominent peak or trough in a waveform may not be informative or important in the context of a complete topography or a particular paradigm. For instance, a peak may invert in polarity across the scalp topography, or appear entirely different when converted to a different reference (see discussion in Kayser et al., 2003, for an example and implications; cf. also Dien, 1998b).

A frequently neglected issue in ERP studies is the difference between a *conceptual* (theoretical) and a *technical* (observational) definition of an ERP component (see, e.g., Donchin et al., 1977, 1978, 1997; Fabiani et al., 2000; Näätänen and Picton, 1987; van Boxtel, 1998).

Cognitive ERP research has extended the traditional conceptualization of an ERP component, being a construct of the underlying neuronal generators, by operationally defining endogenous ERP components as condition-related differences of distinct ERP deflections (e.g., MMN, N2b, P3b, Novelty P3), which may markedly differ between various populations (e.g., age, gender, handedness, clinical diagnosis). Paradoxically, all practical estimates of these generic ERP components, ranging from simple baseline-to-peak measures to more sophisticated waveform decomposition methods, are formally unrelated to the neuronal origin of the ERP. Instead, they are based on an evaluation of the temporal and/or spatial structure of the ERP signal and its statistical properties, which is not a trivial problem considering the richness of the data source with its high temporal resolution in the millisecond range (e.g., Chapman and McCrary, 1995; Picton et al., 2000). The analysis and interpretation of ERPs requires an effective method for measuring components of interest; however, the significance and heuristic value of any component measure must be judged by insights gained about anatomical, physiological, and/or cognitive information processes.

Recent technological advancements have made possible dense electrode arrays with a spatial resolution of 128 or more channels, calling for an overhaul of the existing 10-20 system for electrode placements (Oostenveld and Praamstra, 2001; Pivik et al., 1993), and have therefore further increased the volume of data recorded during a single experiment. Given the goal of differentiating the impact of experimental manipulations, the most direct, if simplistic, approaches to quantify this information have continued to measure the "obvious" characteristics of ERPs, namely the amplitude and latency of waveform peaks and troughs, even in the absence of a theoretical rationale that these conspicuous deflections are always meaningful (e.g., see Chapman and McCrary, 1995; Donchin and Heffley, 1978; Picton et al., 2000). However, as convincingly exemplified by Donchin and Heffley (1978), the simplicity of traditional ERP peak and area measures is deceiving, as these measures are subject to experimenter bias in specifying peaks in noisy waveforms (a problem not removed by using a computer algorithm), or in determining area integration limits for deflections that invert and shift across scalp recording locations.

1.2. Principal Components Analysis (PCA) of ERP data

To overcome the limitations of traditional peak and area measures, principal components analysis (PCA) has been advanced as a heuristic tool to determine "data-driven" ERP component measures (e.g., Donchin 1966; Donchin and Heffley, 1978; Glaser and Ruchkin, 1976; Möcks, 1988a, 1988b), and since has widely been used as an effective linear reduction method for multivariate ERP data (e.g., Chapman and McCrary, 1995; Dien, 1998a; Duffy et al., 1992; Picton et al., 2000; van Boxtel, 1998). In this approach, ERP waveforms are conceptualized as an ordered sequence of recorded scalp potentials, typically using either time (temporal) or topography (spatial) as the ordering domain.[1] The basic assumption of a temporal PCA is that a collection of ERP waveforms (i.e., the cases, their number being determined by multiplying participants, recording locations, and experimental conditions), each consisting of a finite number of surface potentials recorded at discrete time points (i.e., the variables, their number being determined by the sampling rate and the length of the recording epoch), can be decomposed into a linear combination of principal component coefficients (i.e., the factor loadings) and associated weights (i.e., the factor scores). The component coefficients represent, with respect to the cases, invariant loading patterns across the ordered variables (i.e., time points), often referred to as 'component waveforms' (Chapman and McCrary, 1995), which are orthogonal to each other (for a brief description of principle axis rotation, e.g. Hunt, 1985; Rösler and Manzey, 1981; for computational details, see Glaser and Ruchkin, 1976; Harman, 1967; Appendix). The component or factor scores represent the relative contribution or weight of each loading pattern for each case (i.e., each ERP), and may therefore subsequently be used to compare the different ERPs for a given loading pattern, for instance, by employing an analysis of variance (ANOVA) reflecting the study design.

ERP waveforms may be linearly decomposed in an infinite number of ways (e.g., Glaser and Ruchkin, 1976). An initial PCA factor extraction likewise provides only one nonunique solution set, unless a single component fully describes the data (e.g., van Boxtel, 1998). The maximum number of PCA components that can be extracted from a cases-by-variables data matrix is determined by the smaller of the number of rows (cases) or columns (variables), although fewer components may be sufficient to fully account for the variance of a linearly dependent matrix. For

[1] It should be noted that the PCA algorithm is blind to any underlying data organization (i.e., temporal, spatial, spectral, etc.), and identifies systematics solely by intercorrelations of data variables. Randomly shuffling the variable sequence will not change the PCA solution, but will make it very difficult for a researcher to interpret the component structure, unless the original order is restored. It is the *a priori* ordered, sequential nature of ERPs (temporal and spatial) that distinguishes these data sets from other unstructured data sets to which PCA is commonly applied.

a temporal PCA, the number of variables depends on the duration of the recording epoch considered for analysis and the sampling rate, whereas the number of observations (cases) equals the number of ERP waveforms submitted (i.e., subjects, conditions, and scalp sites). To obtain a more stable PCA solution, it is generally suggested that the number of observations should exceed the number of variables (e.g., Chapman and McCrary, 1995), usually several times more, which may require adjusting epoch length, sampling rate, or both. However, this general rule has been challenged by the findings of Guadagnoli and Velicer (1988), who demonstrated that absolute sample size, magnitude of component loadings, and to a lesser degree the number of variables defining a component were by far more important to attain a stable solution.

Apart from selecting input variables (e..g., epoch length, sample rate, trials used for averaging ERPs), any application of PCA to ERP data requires a choice between several specific methodological alternatives, including the type of association matrix, whether (and how) factors are rotated, and the criterion for the number of components to be extracted (e.g., Chapman and McCrary, 1995; Donchin and Heffley, 1978; Picton et al., 2000). Frequently, the raw data matrix is initially transformed into the correlation matrix (i.e., the standardized variance-covariance matrix), which has the advantage that all variables have the same variance and that the extracted factors are therefore invariant under scaling of the original variables (cf. Chapman and McCrary, 1995; Donchin and Heffley, 1978). However, it has been argued that when all variables are measured in the same units, which is usually the case for surface potentials calibrated in microvolts, using either a cross-products or a covariance matrix is preferable over using a correlation matrix (e.g., Donchin and Heffley, 1978). Moreover, there is no a priori reason to assume that the standard deviations of the input variables are the same (cf. van Boxtel, 1998), and, as argued below, exactly the opposite may be expected for baseline-corrected ERP data. The main advantages of a cross-products matrix are that PCA components will (1) likely be related to large ERP deflections, and (2) factor loadings and scores can be directly interpreted with respect to the original data, as the sign of the factor scores reflects the polarity of the underlying ERP component (e.g., van Boxtel, 1998). These two advantages are also valid for a covariance-based PCA, although for slightly different reasons, as factors reflect the ERP variance around the grand mean waveform, which is removed by this procedure. The extracted factors should still closely relate to prominent ERP deflections, because ERP variance is likely centered around or in close temporal and/or spatial proximity of these peaks, and the polarity of the associated ERP variance *relative* to the grand mean can be inferred from the sign of the component scores. As stressed by van Boxtel (1998), a covariance-based PCA extracts components only “if there is variation across electrodes, conditions, and subjects, and that is exactly what researchers are looking for” (p. 92). Still, it is widely believed that there is little, if any, difference between PCA solutions based on the covariance or correlation matrix when applied to typical ERP data (e.g., van Boxtel, 1998), and any advantages of the covariance matrix, such as the close relation to the original metric, can be easily resolved for the correlation matrix by multiplying each loading value with the standard deviation of its input variable (e.g., Chapman and McCrary, 1995). For these reasons, Chapman and McCrary (1995) “can find no reason for any vehement preferences” (p. 294), and van Boxtel (1998) suggests that practical considerations, such as the availability of an extraction procedure in a statistical package, may be used as a guide to choose between these alternatives. Unfortunately, there is considerable confusion about how extraction procedures are implemented and used in different statistical software packages, particularly in combination with other methodological choices. For instance, when a covariance matrix is factored, the loadings may by default be standardized *after* extraction, that is, each value of the component waveform is divided by the standard deviation of the raw variables (across cases).

A rotation is commonly applied after factor extraction with the goal of obtaining simpler interpretations of the extracted components (e.g., Chapman and McCrary, 1995; van Boxtel, 1998; Picton et al., 2000; but see Rösler and Manzey, 1981, for caveats and arguments against factor rotation). If a component waveform (i.e., the factor loadings vector) consists of multiple or significant secondary loading peaks, it can be very difficult or impossible to interpret such a component. Frequently, a Varimax criterion (Kaiser, 1958) is applied to the initial PCA solution to achieve *simple structure* (Thurstone, 1947), which maximizes the loadings variance for the components retained for the rotation procedure (e.g., Bortz, 1993; Chapman and McCrary, 1995). Apart from minimizing component overlap, the Varimax rotation also maintains orthogonality of component scores (i.e., independence between components), which is a significant advantage with respect to inferential statistics usually performed on the factor scores. As the probability of Type I errors increases with the number of dependent variables (i.e., extracted factors) considered for statistical analysis, the orthogonality of the Varimax solution counteracts this undesired effect. Covariance-based, Varimax-rotated component waveforms are typically characterized by unique triangle-shaped, positive factor loadings that are 1) clustered in a narrow time range and 2) lack inverse (negative) or significant secondary loadings at different latencies. In other words, because maximum weight is assigned to the time point that coincides with the factor loadings' peak, while neighboring time points have smaller weights, and all other time points of the recording epoch have zero weights, Varimax-rotated PCA components can be conceived as a measure of weighted time window amplitude (Kayser et al., 2000a, 2001). However, as pointed out by Chapman and McCrary (1995), this is not a characteristic of the Varimax procedure, which would identify unique variance with multiple component peaks (e.g., if the epoch includes two successive stimulus presentations, an common N1 factor with two peaks is likely extracted), but rather an inherent feature of the variance distribution in ERPs.

Other rotation procedures, particularly oblique rotations, are less commonly used, although they may achieve a greater degree of simple structure (Chapman and McCrary, 1995; van Boxtel, 1998). Dien (1998a), arguing that the orthogonality criterion of the Varimax procedure (factors must be uncorrelated) may result in distortions of the component waveforms and may also be incompatible with the underlying brain processes (i.e., different ERP components and the activity of their underlying generators may be correlated), has proposed the use of Promax (Hendrickson and White, 1964) as an alternative rotation procedure. Promax loosens the strict orthogonality criterion of Varimax and allows multiple factors to share variance. However, the advantage of Promax and any oblique rotation method is at the same time a disadvantage, as the analyzed components are no longer independent.

Fava and Velicer (1992, 1996) have warned that a bias towards under- or overextraction of the true underlying number of components may produce degraded and unstable factors and thereby yield inaccurate solutions; overextraction, however, was generally found to be a less serious problem (Wood et al., 1996). A number of general guidelines have been proposed for the number of components to extract and interpret (e.g., Everett, 1983), among them the scree test (Cattell, 1966), the criterion proposed by Kaiser (1960) to use only components that explain at least the variance equal to the average variance of the original variables (i.e., components with Eigenvalues larger than one when a correlation matrix is factored; e.g., Chapman and McCrary, 1995; van Boxtel, 1998), or parallel analysis (Horn, 1965) to estimate the degree of random noise inherent in the data (Bortz, 1993; Zwick and Velicer, 1986; see Dien, 1998a, and Beauducel et al., 2000, for examples of parallel analysis in the context of ERP data). Instead of using rather arbitrary criteria to extract, retain and interpret PCA components derived from ERP data, we and others (e.g., Donchin et al., 1997; Kayser et al., 1997, 1998, 1999, 2000a, 2001; Spencer et al., 1999, 2001) reasoned that ERP expertise and a priori knowledge about the paradigm should be used as a guide to focus on components of interests. These research groups adopted a strategy that extracts as many factors as needed to account for most or even all of the data variance, but interpretation and statistical analysis is restricted to factors that either directly vary as a function of the experimental manipulation, for instance, distinguishing a novelty P3 from a classical P300 (Spencer et al., 1999), or that can unambiguously be related to known ERP components evident in the averaged waveforms (e.g., Kayser et al., 1997, 1998, 1999, 2000a, 2001). Based on similar considerations, Dien et al. (2003) related parametric measures of expectancy and meaningfulness to factors extracted in a semantic comprehension task. Fabiani et al. (1987) have also emphasized that determining the number of components to rotate, and deciding how many to interpret, are two separate choices, and strongly encouraged investigators to closely examine "the time course of the component, its scalp distribution, and its response to experimental manipulations" (p. 28) to safeguard against component overinterpretation and misindentification.

1.3. Theoretical rationale

It is unclear which of the general guidelines and concerns are of relevance to ERP component definition using PCA. Few efforts have been made to specifically develop or validate the empirical relevance of such recommendations to the problem of identifying and defining meaningful, interpretable ERP components (Chapman and McCrary, 1995; Dien, 1998a; Glaser and Ruchkin, 1976; Picton et al., 2000). Specifically, it is not obvious that methods suited to psychometric data are equally suited to summarize the topographic organization of intercorrelated time series data comprising ERPs. There may be useful, important and empirically verifiable applications of PCA which are irrelevant or inappropriate in other contexts. Similar methods have been used to identify and *remove* high-variance components from the data, for example, eliminating EEG artifacts related to blinks (Berg and Scherg, 1994; Neurosoft, 2001).

The usefulness of the extracted PCA factors can be evaluated by specific knowledge about the variance distribution of ERPs, which are typically characterized by the removal of baseline activity. The variance should be small for sample points before and shortly after stimulus onset, both across cases (i.e., for any particular sample point across waveforms) and within cases (i.e., across sample points for each waveform). In contrast, the variance should be large near the end of the recording epoch, which reflects the autocorrelation of EEG time series data. An analogy for this property is the movement of a swinging rope that has been grasped in one hand. The movements of the rope are constrained by the grasping hand, comparable to the ongoing EEG activity constrained by the baseline. Aside from the baseline, signal variance should be larger at ERP component peaks (global field power; e.g., Lehmann and Skrandies, 1980). Furthermore, ERP component amplitudes are expected to vary as a function of the experimental manipulation and/or between different subjects. Whereas the factor loadings of a PCA based on the covariance matrix preserve this information, it is lost with a correlation matrix that assigns equal weights to each sample point, yielding the possibility that small but systematic variations may form a factor. Such small and rather uninformative systematic variations are likely to occur during the baseline period, or shortly after its end, which typically coincides with stimulus onset. For instance, if the ERPs comprise a small CNV or anticipatory drift, the baseline correction (i.e., subtracting the mean of all sample points within the baseline interval from the entire waveform) will force the drift to intersect with the baseline at a fixed sample point – exactly at half the duration of the baseline period in case of a linear drift. Other small but systematic ERP variations may also occur during the recording epoch, for example, originating from particular recording characteristics such as digitization mode, filter settings, ‘random’ digitizer noise, etc. Using a correlation matrix for factor extraction will exaggerate contributions of negligible amplitude, and may therefore obscure the underlying ERP component structure, which the set of extracted factors is intended to reflect.

We evaluated these considerations with simulated and real ERP data by comparing the PCA solutions resulting from either a covariance or a correlation association matrix. Factor extraction was followed by Varimax rotation to achieve simple structure by means of minimizing the number of time points with high loadings on a factor, thereby enhancing the interpretability of the extracted factors, and at the same time avoiding the interpretational uncertainties of correlated components. As some statistical packages apply a standardization of factor loadings by default (e.g., the SPSS procedure MANOVA can be used to perform a covariance-based PCA; SPSS, 1988), effectively scaling (normalizing) PCA components *before* rotation, this specific extraction/rotation procedure was also included in the comparison of real ERP data.[2] Furthermore, we systematically varied the number of components to be extracted, ranging from the maximum limitation of one to an unrestricted solution (i.e., number of input variables), to address the potential risks of factor over- or underextraction (e.g., Everett, 1983; Fava and Velicer, 1992, 1996; Wood et al., 1996) for real ERP data, and determine the optimal number of factors to be retained before rotation.

2. Methods

2.1. Simulated ERP data

Our theoretical considerations prompted us to construct artificial data sets that would clearly distinguish the operational properties of a covariance- versus correlation-based PCA. An invariant, pseudo ERP waveform template (128 sample points, 100 samples/sec, 200 ms baseline) was used to generate two simulated data sets for 30 ‘electrode sites,’ applying the EEG montage used in our laboratory (e.g., Kayser et al., 2000a, 2001), and 20 ‘participants’ (see Fig. 1). The template consisted an early negative ‘component’ with an amplitude of -8 μV peaking at 110 ms (sample point 32), and a symmetrical, linear rise and fall spanning an interval of ±50 ms (±5 sample points), and a late positive ‘component’ with an amplitude of +12 μV peaking at 450 ms (sample point 66), and a symmetrical, linear rise and fall spanning an interval of ±210 ms (±21 sample points). A ‘topography’ was introduced by scaling the template for selected sites with a factor of 0.5 (frontopolar sites Fp1, Fp2), 0.8 (frontal coronal plane F7, F3, Fz, F4, F8), 1.2 (medial-central sites C3, Cz, C4), and 1.0 for the remaining 20 sites. In addition, a constant, low-voltage offset of -0.01 μV was applied to the pre-stimulus baseline at every other electrode, spanning an interval of -200 to -50 ms (first 15 sample points; see inset in Fig. 1A). The second simulated data set was created by adding random noise (range ±0.25 μV, uniform distribution) to each sample point in each waveform of the first simulated data set. As the noise exceeded the amplitude of the constant offset, the offset was completely obscured in the second simulated data set (see inset in Fig. 1B), while the two components of the template and its topography were preserved.

Each simulated data set was submitted to two temporal

[2] To clearly distinguish this extraction method from the regular use of the covariance matrix, we will refer to these two procedures as standardized and unstandardized covariance-based PCA solutions. However, when these two covariance-based procedures are not directly compared, we have omitted this additional descriptor to improve readability. In this case, a covariance-based solution always refers to the unstandardized covariance-based PCA.

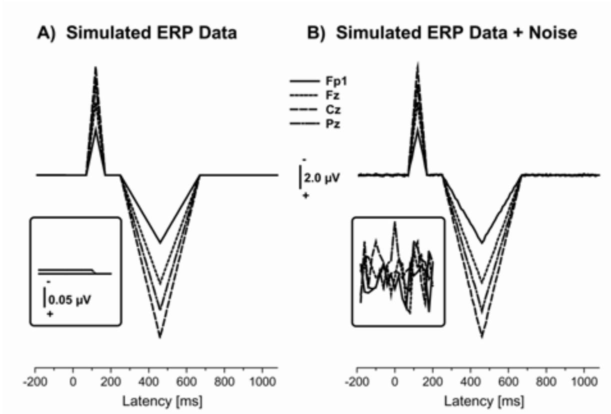

Fig. 1. Grand mean waveforms at four 'electrode sites' (Fp1, Fz, Cz, Pz) for simulated ERP data **(A)** without and **(B)** with random noise (range ±0.25 µV) added to each sample point. Waveforms were created from an invariant waveform template (128 sample points, 100 samples/sec, 200 ms baseline). A 'topography' was introduced by scaling the template for selected sites with a factor of 0.5 (e.g., Fp1), 0.8 (e.g., Fz), 1.2 (e.g., Cz), or 1.0 (e.g., Pz). At every other electrode, a constant, low-level voltage offset (-0.01 µV) was systematically applied to portions of the pre-stimulus baseline (i.e., from -200 ms to -50 ms). Insets show enlargements of the baseline period (-200 ms to 0 ms) at a different scale, revealing that the constant, low-level baseline offset at 'sites' Cz and Pz is present for the noise-free simulation, but lost for the simulated data set with added noise.

PCAs using either the covariance matrix or the correlation matrix for factor extraction. BMDP statistical software (4M; Dixon, 1992) was used for all PCA computations. Variables consisted of 128 sample points (-200 to 1,070 ms), representing the data matrix columns, while the data matrix rows consisted of 600 cases stemming from the product of 'participants' (20) and 'electrode sites' (30). Both covariance- and correlation-based PCA were followed by unrestricted Varimax rotation. Factor loadings of the four PCA solutions were plotted and compared to the original template.

2.2. Real ERP data

For this systematic comparison, we reanalyzed ERP data of healthy adults previously collected in our laboratory. ERPs were from 30 scalp locations using a nose tip reference with an Fpz ground, and impedances maintained at 5 kΩ or less (Kayser et al., 2000a, 2001). EEG gain was 10,000. Data were sampled at 100 Hz with a .01 to 30 Hz band pass. Recording epochs of 1,280 ms (200 ms pre-stimulus baseline) were extracted off-line and digitally low pass filtered at 20 Hz.

Visual ERPs of 16 right-handed control participants (7 men; mean age = 27.3 ±8.6 years) were collected during a hemifield paradigm, in which negative and neutral stimuli (medical textbook pictures showing face areas before or after surgical treatment of dermatological disorders) were briefly exposed for 250 ms to either the left or right visual field. Stimuli were presented with variable interstimulus intervals (range 12 to 18 s, mean = 15 s). Participants were instructed to attend to the lateralized stimulus presentations while maintaining fixation, however, no manual response or any specific cognitive operation was required (for methodological details, objective and rationale, see Kayser et al., 2000a). Fig. 2 reveals the ERP topography and component structure of this paradigm, in which the surface potentials are characterized by: (1) a central N1 with a peak latency of about 150 ms; (2) a strongly right-lateralized P1/N2 complex over inferior-parietal sites (cf. P8 and P10), peaking at 120 and 220 ms, respectively; and (3) a prominent mid-parietal P3 peaking at site Pz at 480 ms. Fig. 2B also shows enhanced P3 amplitude for negative compared with neutral stimuli, which was greatest over right lateral parietal sites.

Auditory ERPs of 32 right-handed volunteers (18 men; mean age = 35.3 ±11.9 years) were collected during tonal and phonetic oddball tasks. Participants listened to a series of either two complex tones or two consonant-vowel syllables (250 ms stimulus duration, fixed interstimulus interval of 1,750 ms), and responded to infrequent target stimuli (20% probability assigned to one of the two stimuli in the series) with either a left or right button press (for methodological details, objective and rationale, see Kayser et al., 1998, 2001). The ERP component structure generated by these auditory tasks was fundamentally different from that found with the visual-half field paradigm. For target stimuli, the most distinctive components were: (1) a central N1 peaking at approximately 100 ms; (2) an N2 peaking around 220 ms, which was characterized by task-dependent regional hemispheric asymmetries (right-larger-than-left over frontotemporal sites for tonal stimuli, left-larger-than-right over parietotemporal sites for phonetic stimuli); and (3) a mid-parietal P3b peaking between 340 and 420 ms (waveforms are given in Kayser et al., 2001).

By changing the computational instructions[3] of BMDP program 4M (Dixon, 1992), three extraction methods and

[3] The critical BMDP syntax consisted of command variations in the /FACTOR paragraph of program 4M. Principal components analysis (METH = PCA) was combined with the request for factoring either the covariance matrix and to use unstandardized loadings (FORM = COVA), the correlation maxtrix (FORM = CORR), or the covariance matrix and to use standardized loadings (FORM = COVA and LOAD = CORR). The maximum number of factors obtained was determined by the number command (NUMB = #). All other BMDP statements were identical, including the request for Varimax rotation (METH = VMAX) in the /ROTATE paragraph (Dixon, 1992).

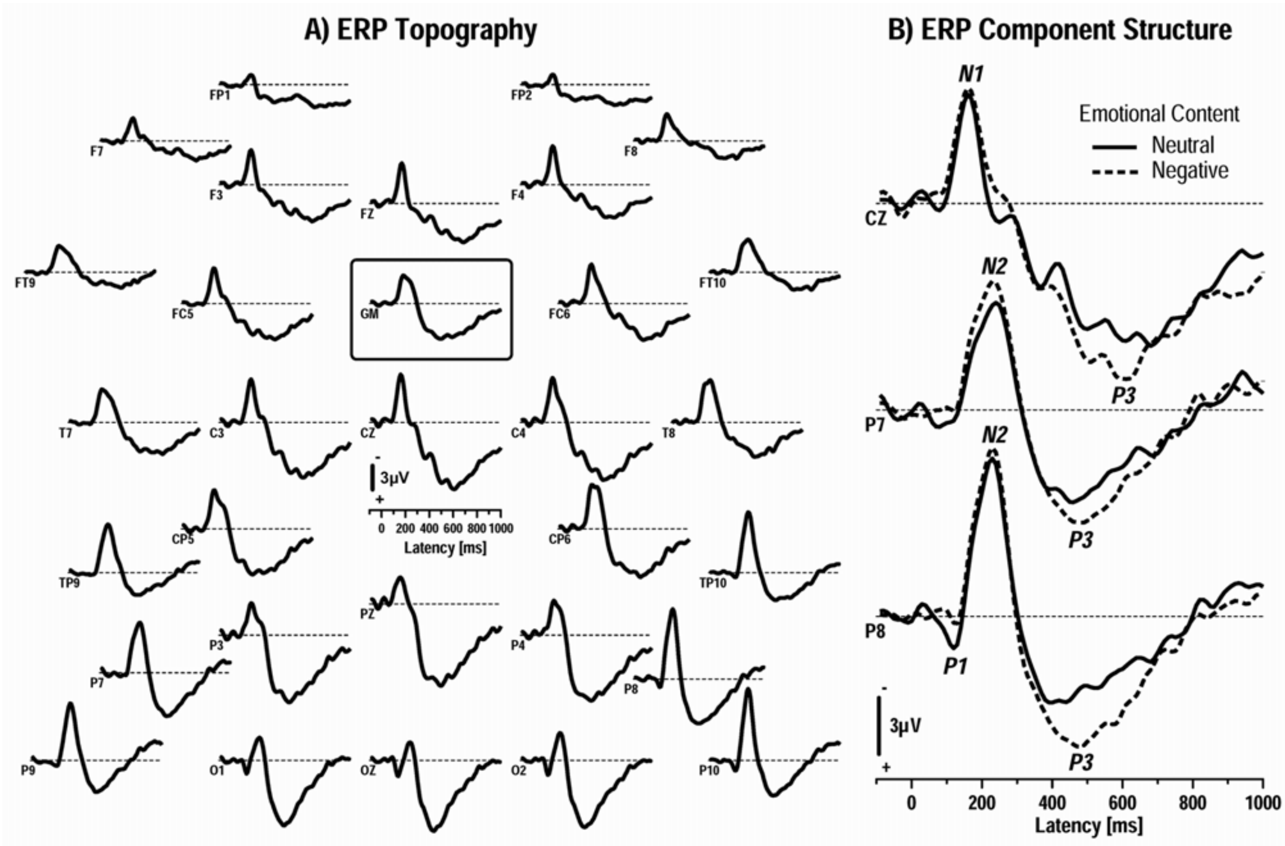

Fig. 2. Grand mean ERP waveforms to visual stimuli for 16 healthy adults at 30 scalp locations **(A)**, averaged across all experimental conditions (data from Kayser et al., 2000a). The inset shows the grand mean (GM) across all recording sites. To clarify the ERP component structure, waveforms were enlarged and separately plotted for negative and neutral stimuli at representative sites Cz, P7, and P8 **(B)**, where P1, N1, N2, and P3 were prominent when using a nose reference. ERP differences of emotional content, the main objective of this previous study, were clearly evident between 400 and 700 ms, and larger over right (P8) than left (P7) lateral-parietal sites. N2, the most distinct ERP component at these sites, revealed a marked right-larger-than-left hemispheric asymmetry.

110 extraction/retention criteria were systematically combined to perform a total of 330 temporal PCAs on each real data set. Factors were extracted and Varimax rotated using either: (1) the covariance matrix and unstandardized factor loadings; (2) the correlation matrix; or (3) the covariance matrix and standardized factor loadings. The number of components to be extracted and retained for rotation was varied between 1 (maximum restriction) and 110 (unrestricted solution only limited by the number of input variables). Data matrix columns (variables) consisted of 110 sample points (-100 to 990 ms), whereas data matrix rows consisted of 1920 or 3840 cases (visual and auditory data sets, respectively) resulting from the combination of participants (16/32), conditions (4/4) and electrode sites (30). For all PCA solutions, factors were described by the time courses of their factor loadings and the topographies of their factor scores. Hence, all factor loadings and all factor score topographies were plotted and compared to the corresponding grand mean waveforms, applying common ERP knowledge and the theoretical considerations outlined in the introduction.

3. Results and discussion

3.1. Simulated ERP data

Fig. 3 directly compares the PCA solutions for the two simulated ERP data sets using either a covariance or a correlation matrix for factor extraction. As can be seen from Fig. 3A, the first covariance-based factor (solid line) explained effectively all the data variance for the simulated, noise-free data, which is an almost exact representation of the true variation introduced to this data set by jointly scaling the ERP template. Moreover, the shape of the component waveforms also reflects the shape of the template. It is important to note that the shape of the loadings vector is *not* a simple copy of the template, as the grand mean is removed when factoring the covariance matrix, but rather reflects the variance around that template. As no

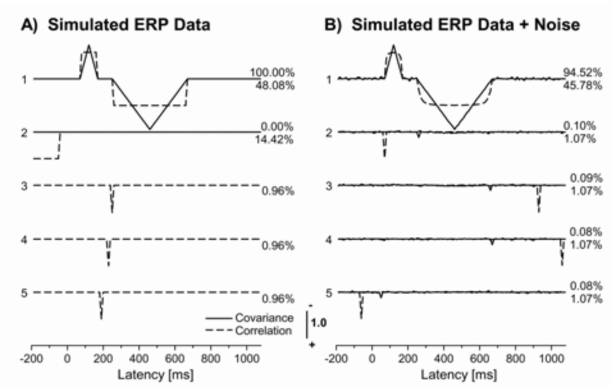

Fig. 3. Time courses of factor loadings for the initial PCA factors extracted (up to five) from the covariance (solid lines) or correlation (dashed lines) matrix for simulated ERP data without **(A)** and with noise **(B)**. The percentage of explained variance after Varimax rotation is indicated at the right of each loadings vector (values for covariance-based loadings are plotted above the baseline, correlation-based loadings are plotted below). Note that only two factors were extracted in the covariance-based PCA for simulated ERP data without noise, but 41 factors in the correlation-based PCA using the same data. For simulated ERP data with noise, 78 factors were extracted for both extraction methods.

additional noise was included in this data set, all variance stems exclusively from the 'topographic' differences between selected electrodes, which were introduced by linear scaling of the template. As a direct consequence of the linear scaling, differences from the original template are larger in absolute terms at 'component' peaks, and gradually decrease along the rising and falling slopes.

In contrast, the correlation-based PCA produced a first component (dashed line) that merely indicates the direction but not the variance shape, resulting in two loading rectangles: the first rectangle had a constant amplitude of -1.0 and spanned exactly the time period of the simulated component N1, whereas the second rectangle had a constant amplitude of +1.0 and spanned exactly the time period of the simulated component P3. Likewise, a second factor had a constant amplitude of -1.0 and an onset and duration that exactly matched the constant, low-level offset introduced to the baseline. A correlation-based PCA does not distinguish between small and large variations of different variables (time points), provided they are linearly related.

The amount of variance explained by the first two correlation-based factors was 48.08% and 14.42%, respectively. This approximates the proportion of the number of variables used for simulating N1 and P3 (50) and the low-level baseline offset (15) with respect to the total number of submitted variables (128), but disproportionally reflects the absolute amount of variance introduced by these two independent variations. In addition, the correlation-based PCA extracted another 39 factors, the first three are shown in Fig. 3A as factors 3 to 5, each explaining 0.96% variance and having a loading of +1.0 that spanned exactly one variable, apparently randomly assigned to variables outside the loading range of factor 1 and 2. It is obvious that these additional factors are a mere artifact of the need to explain another 37.5% variance which remains unexplained after extraction of the first two factors.

The PCA solutions performed on the simulated data with added noise when factoring either the covariance- or correlation matrix are compared in Fig. 3B. Each of the two unrestricted solutions extracted a total of 78 factors. Again, the first covariance-based factor explained most of the data variance (94.52%), closely matching the introduced variation, and its time course closely approximated the introduced variance shape. However, no distinct second covariance factor was extracted that could be related to the low-level baseline offset. Evidently, this systematic variation was completely obscured by the added noise. Rather, each of the additional 77 factors explained less than 0.1% variance, gradually decreasing to 0.06%, and had loadings peaks of small amplitude scattered across the recording epoch (see solid lines for factors 2 to 5 in Fig. 3B).

In contrast, the first correlation-based factor explained 45.78% variance, with a time course that resembled that of the simulated, noise-free data (dashed line in Fig. 3B). However, rising and falling edges of the rectangles were notable curved, which appeared to be a logical consequence of those time periods when the random noise exceeded or equaled the signal. In other words, the noise smoothed the loading shape and feigned a true loading peak, and this effect would be exaggerated by further lowering the signal-to-noise ratio. The remaining 77 factors explained 1.07% or less variance, gradually decreasing to 0.01%, and had again single, isolated loadings with an amplitude of approximately +1.0, dispersed over the recording epoch (see dashed lines for factors 2 to 5 in Fig. 3B).

3.2. Real ERP data (visual modality)

3.2.1. Component waveforms of PCA solutions

An overview of the different Varimax-rotated PCA solutions derived from real ERP data collected during the visual half-field paradigm is given in Fig. 4. The left panel (Fig. 4A) shows up to the first ten component waveforms that were obtained after factoring the covariance matrix of the raw data, ordered across columns in the sequence of extraction. Ordered across rows are the solutions derived by systematically varying the number of factors to be extracted (using this systematic, individual component waveforms will be referred to by their row and column number).

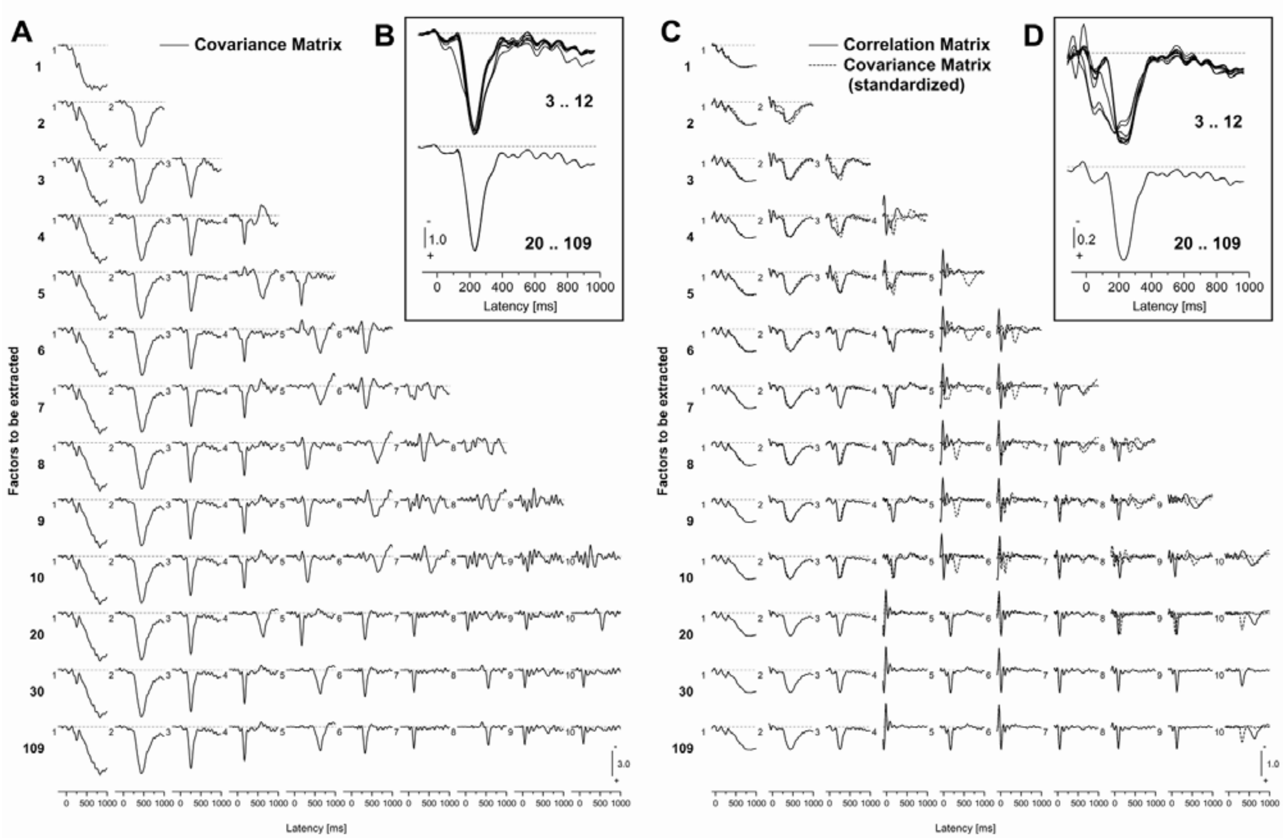

Fig. 4. Overview of PCA solutions for real ERP data (visual task). Varimax-rotated factor loadings are plotted in their extraction sequences (along columns and up to the first 10 factors extracted) for PCA solutions based on the **(A)** covariance matrix using unstandardized loadings, and **(C)** the correlation matrix (solid lines) or the covariance matrix using standardized loadings (dashed lines). PCA solutions are ordered (along rows) by the extraction criterion (i.e., the number of factors to be extracted and retained before rotation). Note that for this data set 109 factors were sufficient to completely explain the data variance produced by 110 input variables. Insets show overlaid factor loadings of a single component (factor 3, approximate peak latency 250 ms) for restricted ($\leq$12) and liberal ($\geq$20) extraction criteria for **(B)** covariance-based and **(D)** correlation-based solutions.

A closer look at the first column reveals that all solutions extracted a high variance factor that extended over a relatively long time period with higher loadings towards the end of the recording epoch. However, the shape of the loading vector of this first factor changed depending on the extraction limit, which is particularly obvious when comparing conservative solutions (e.g., see component waveforms 1-1, 2-1, and 3-1 in Fig. 4A). In a similar manner, the second factor revealed, for all solutions, loading peaks at approximately 450 ms, which changed shape for more restrictive solutions (see component waveforms 2-2, 3-2, and 4-2 in column 2 of Fig. 4A). The third factor peaked for all solutions at approximately 250 ms. Fig. 4B (left panel inset) compares the different solutions for factor 3 by overlaying all component waveforms for extraction limits of 3 to 12 total factors (i.e., more restricted solutions), and also overlaying all component waveforms for extraction limits of 20 or more total factors (i.e., more liberal solutions). As can be seen, there is considerable variation between more restricted solutions, whereas the component waveforms of factor 3 are practically identical for more liberal solutions.

For low variance components (i.e., for factors extracted after the first three components), fluctuations of shape *and* position within the extraction sequence were observed. Despite this inconsistency, some components could be identified without difficulty by their unique loading peak across the different solutions (e.g., one distinct factor with an approximate peak latency of 170 ms can be recognized in component waveforms 5-5, 20-5, and in column 4 for the remaining solutions shown in Fig. 4A). By-and-large, component waveforms tended to become more stable with more liberal extraction criteria. This was notably the case for all of the first ten components when applying an extraction criterion of 30 or more factors – there were no differences in shape or extraction sequence between these

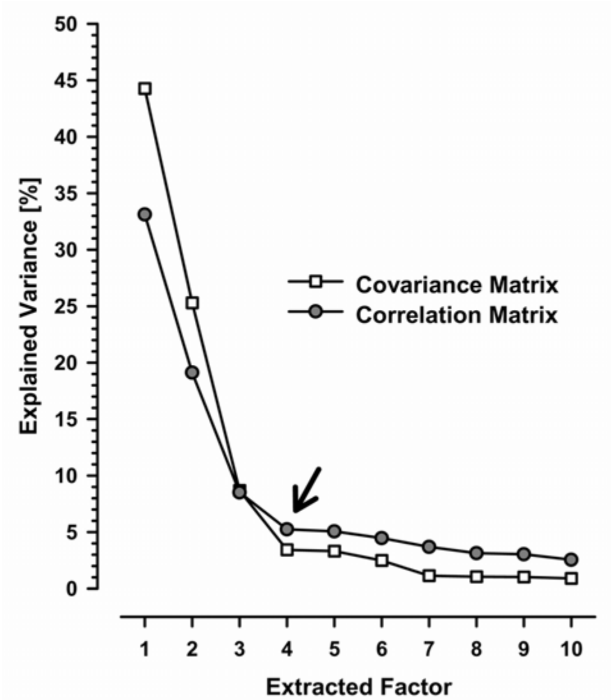

Fig. 5. Eigenvalues for the first ten factors extracted from the unrestricted covariance- or correlation-based PCA solutions, plotted as fractions of overall variance after Varimax rotation. The arrow indicates the Scree-test cutoff-criterion for the corresponding slope of Eigenvalues of the covariance matrix before rotation.

solutions. A total of 109 components was sufficient to fully account for the data variance of this particular data set with 110 input variables (i.e, the PCA allowing the extraction of 110 components extracted exactly the same 109 factors).

Fig. 4C shows the equivalent PCA solutions when either the correlation matrix was factored and rotated (solid lines), or the covariance matrix was factored but standardized loadings were rotated (superimposed dashed lines). On the whole, there were only marginal differences between the PCA solutions derived from these two extraction/rotation methods: almost none were found for liberal extraction criteria (30 or more factors extracted), and the few notable differences were mainly observed for more conservative extraction criteria and towards the end of the extraction sequence (e.g., see component waveforms 2-2, 3-3, 4-4, and 5-5).

By comparing the correlation-based PCA solutions (Fig. 4C) to their covariance-based counterparts (Fig. 4A), it becomes clear that the first three factors target the same variance: all extracted a high-variance component as factor 1, followed by two factors with approximate peak latencies of 450 and 250 ms. However, when compared to the covariance-based solutions, the correlation-based components always had a wider, less focused loading peak, which was frequently accompanied by significant secondary loadings. These undesirable characteristics are prominent for factor 3, particularly for more conservative solutions (Fig. 4D top, as compared to 4B top). Nevertheless, correlation-based PCA solutions also become more stable with more liberal extraction criteria. For example, the shape of the component waveforms of factor 3 did not change for an extraction criterion of 20 or more factors (Fig. 4D bottom).

The most striking difference between the unstandardized covariance-based PCA solutions and the correlation- or standardized covariance-based solutions was the presence of factors with multiple and inverted loading peaks of short-latency, some of which occurred during the baseline period (cf. component waveforms in columns 4 to 6 of Fig. 4C). These factors were even present in the most stable, unrestricted solution (component waveforms 109-4 and 109-6 of Fig. 4C), and also explained a considerable proportion of variance (i.e., 5.25% and 4.46% for factors 4 and 6, respectively, after Varimax rotation; see Fig. 5).

3.2.2. Component topographies of PCA solutions

To compare the associated factor score topographies of related components for each of the different PCA solutions, both within extraction methods (covariance or correlation matrix) as well as across the two methods, squared differences between factor scores of two solutions were summed across all cases. As illustrated for the scores of

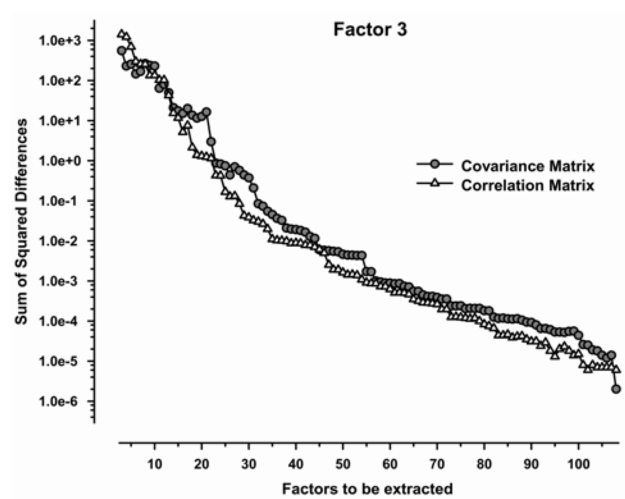

Fig. 6. Similarity of factor scores as a function of extraction criterion (number of factors to be extracted) for covariance- and correlation-based PCA solutions, illustrated for factor 3. Similarity is expressed as the sum of squared differences from the unrestricted solutions (109 extracted factor), summed across all cases (1920) for visual ERP data. Data range from 1,414.23 to 0.000002, and are shown on a common logarithmic scale. Factor scores continue to become more stable and more similar with more liberal extraction criteria, until they converge on the unrestricted solution.

factor 3, Fig. 6 shows this similarity measure as a function of the number of extracted factors for the within-methods comparison, computed using the differences from the unrestricted solutions. By plotting the slope of this function on a common logarithmic scale, it becomes clear that factor scores continue to become more stable and more similar with more liberal extraction criteria for both extraction methods, until they converge on the unrestricted solution. However, for all practical purposes, differences in factor scores are negligible (10^3 reduction) for this data set after the extraction of 20 or more factors. Comparable functions were observed for all factors, revealing the same pattern of convergence. Moreover, the similarity functions *between* extraction methods (i.e., using the differences between covariance- and correlation-based solutions for each extraction criterion) also indicated an increase of similarity with more liberal extraction criteria, until differences between methods stabilized on a negligible (10^5 reduction) level after the extraction of 40 or more factors.

Fig. 7 shows the factor score topographies for the first ten components of the unrestricted PCA solutions, along with superimposed plots of their factor loadings. As a direct result of the increased similarity, the most remarkable observation is that eight of these initial ten components revealed almost identical factor score topographies across extraction methods (Fig. 7A and 7B), although correspon-ding factors differed in peak latency (Fig. 7C and 7D), explained variance (Fig. 5), and position within the extraction sequence (Fig. 4). For example, factor 4 of the covariance matrix and factor 5 of the correlation matrix were both characterized by a peak latency of 170 ms, and a mid-central negativity paired with an occipital positivity (Fig. 7A and 7B), thereby closely corresponding to a central N1 which inverted at occipital sites (Fig. 2). Analogously, factors 1 to 3 revealed topographies closely corresponding to a mid-central positive slow wave, a posterior P3, and a

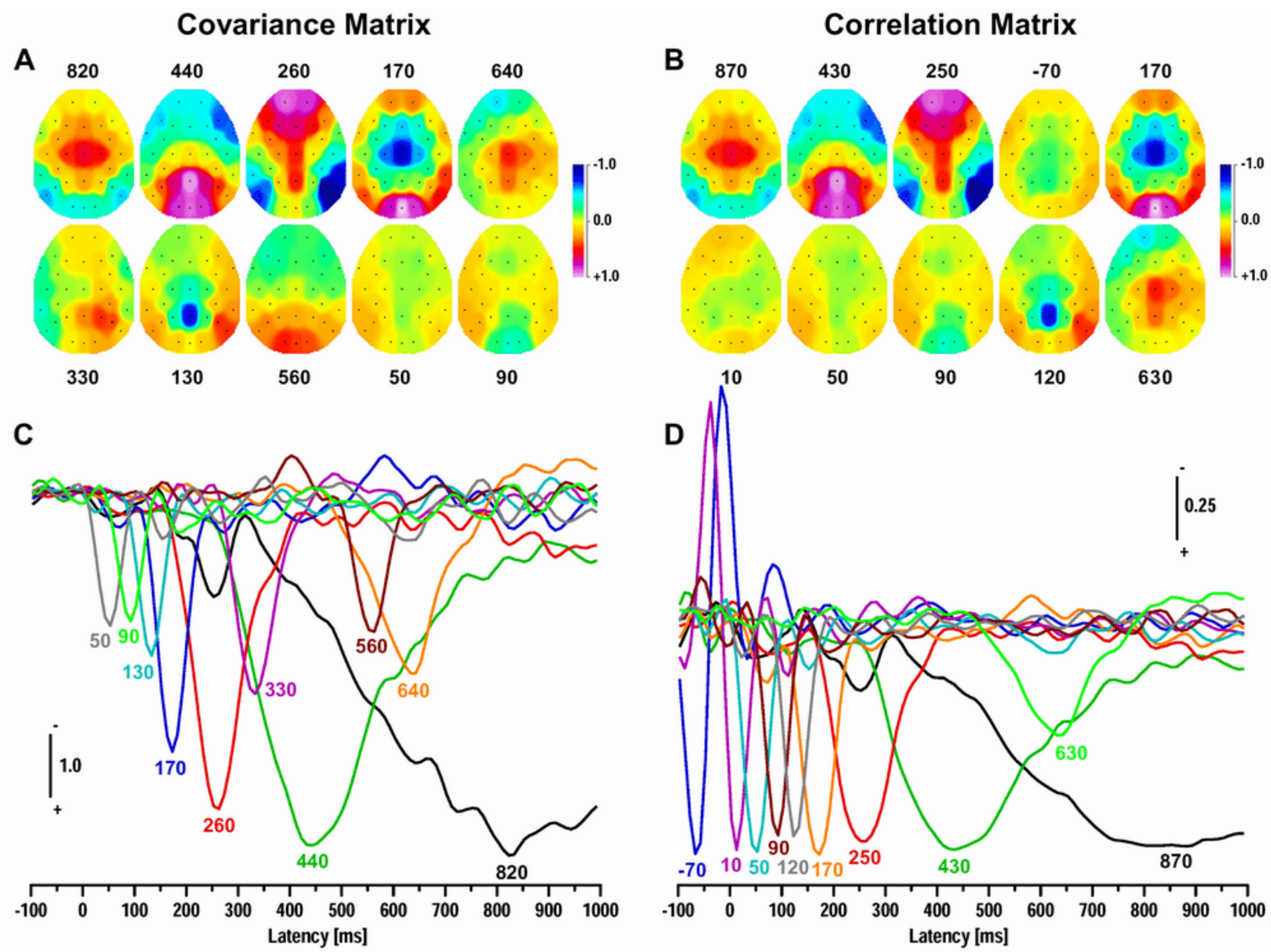

Fig. 7. Topographies of factor scores **(A, B)** and superimposed factor loadings **(C, D)** for the first ten components of the unrestricted PCA solutions using real ERP data recorded from 16 healthy adults during a visual half-field paradigm (Kayser et al., 2000a). Shown are the solutions based on the covariance matrix (A, C) and the correlation matrix (B, D). Factors are labeled using the peak latency of the loadings, but topographies of factor scores (top view, nose upwards) are ordered according to their extraction sequence (left to right, top to bottom).

right-lateralized posterior visual N2 for both extraction methods, despite differences in shape and peak latency (Fig. 7C and 7D). Even low-variance factors were paired across extraction methods, some of them apparently related to various phases of P3 activity (covariance peak latencies of 330, 560, and 640 ms), although not all were among the first ten components extracted. Evidently, by enforcing simple structure, the Varimax algorithm will rotate factors of different extraction methods to a unique solution, if the number of vectors to rotate is sufficiently large to reallocate the loadings accordingly. Once identical or almost identical component waveforms are established, irrespective of their origin, the same associated factor scores are generated.

Factor 7 of the covariance matrix (peak latency 130 ms) and factor 9 of the correlation matrix (120 ms) were characterized by a distinct right-larger-than-left posterior positivity paired with posterior midline negativity, which was clearly related to early P1 activity (cf. sites P9 and P10 in Fig. 2). It is important to note, however, that this distinct activity was not extracted as a unique component unless 20 or more factors were extracted and rotated (Fig. 4A and 4C). Thus, most common rules to determine the number of factors to extract, retain, or interpret would have precluded the formation of this factor. Applied to the correlation matrix, the Eigenvalue-larger-than-one rule would only retain the first 14 factors. The arrow in Fig. 5 indicates the Scree-test cut-off criterion for the covariance-based solution, revealing that only the first 3 factors would have been retained, therefore also missing the N1 factor.

The most glaring problem of the correlation-based solutions is the extraction of irrelevant, high-variance components, exemplified for the current data set by two factors with positive and negative loading peaks of maximum amplitude before or immediately after stimulus onset (see factors labeled -70 and 10 in Fig. 7D). As demonstrated by the correlation-based PCA of the simulated ERP data, the most likely cause is systematic, low-amplitude signal around stimulus onset, which in these cases inverted in polarity.

To fully comprehend the spatiotemporal dynamics of the extracted components, one needs to simultaneously appreciate the topographic information provided by the factor scores with regard to the duration the overlapping component waveforms. While this mental transformation can be accomplished by close scrutiny of their static two-dimensional representations (e.g., Fig. 7A and 7C), the spatiotemporal dynamics become intuitively obvious by virtue of the linear decomposition of the association matrix, which allows a time series of factor score topographies to be reconstructed by simply multiplying the factor loadings vector with the mean of the associated factor scores. An animation of such a time series created for the covariance data presented in Fig. 7A and 7C reveals the distinctive spatiotemporal sequence of P1, N1 and N2, and further clarifies the temporal overlap, however, topographic specificity, of factors corresponding to various phases of the late positive complex.[4]

3.2.3. Statistical implications of different extraction criteria

The most important question for practical research purposes is whether the observed differences for identified components, which may be quite small at times, have an impact on statistical tests commonly used to evaluate variation across electrodes, conditions, and subjects. Using high-variance factors corresponding to ERP components of critical interest for the Kayser et al. (2000a) study (i.e., N2 and P3), we systematically computed separate ANOVA reflecting the study design for each factor and each covariance- and correlation-based PCA solution. The resulting F statistics were compared for representative experimental effects crucial for the objective of the study. As an illustration, Fig. 8 gives the F values of four effects as a function of extraction criterion. The direction of these effects did not change with any extraction method, and can be inferred from Fig. 2B. The greater P3 amplitude to negative than neutral stimuli, as measured by factor 2, was significant ($F_{[1,14]} > 4.60$, $p < .05$) for covariance-based solutions (filled circles in Fig. 8A) for all but the most restricted extraction. In contrast, this main effect was insignificant for correlation-based solutions (open triangles circles in Fig. 8A) for strict conservative extraction criteria (for 4 or less factors extracted, all $F_{[1,14]} < 3.10$, $p > .10$), only marginally significant for somewhat relaxed extraction criteria (for 5 to 9 factors extracted, all $F_{[1,14]} < 4.60$, $p < .10$), and equaled or just exceeded a conventional significance level for liberal extraction criteria (for 10 or more factors extracted, all $F_{[1,14]} \geq 4.60$, $p \leq .05$). The N2 asymmetry, as measured by factor 3 (Fig. 8B), found no statistical support with conservative extraction criteria (for 7 or less factors extracted, all $F_{[1,14]} < 3.10$, $p > .10$), but was significant with more liberal criteria (for 13 or more factors extracted, all $F_{[1,14]} > 4.60$, $p < .05$) for solutions of either association matrix. However, intermediate extraction criteria (8 to 12 factors extracted) yielded significant results for covariance-based solutions (all $F_{[1,14]} > 4.60$, $p < .05$), but insignificant or marginally significant results for

[4] This animation can be viewed and obtained at the following URL: *http://psychophysiology.cpmc.columbia.edu/cn2003.html*

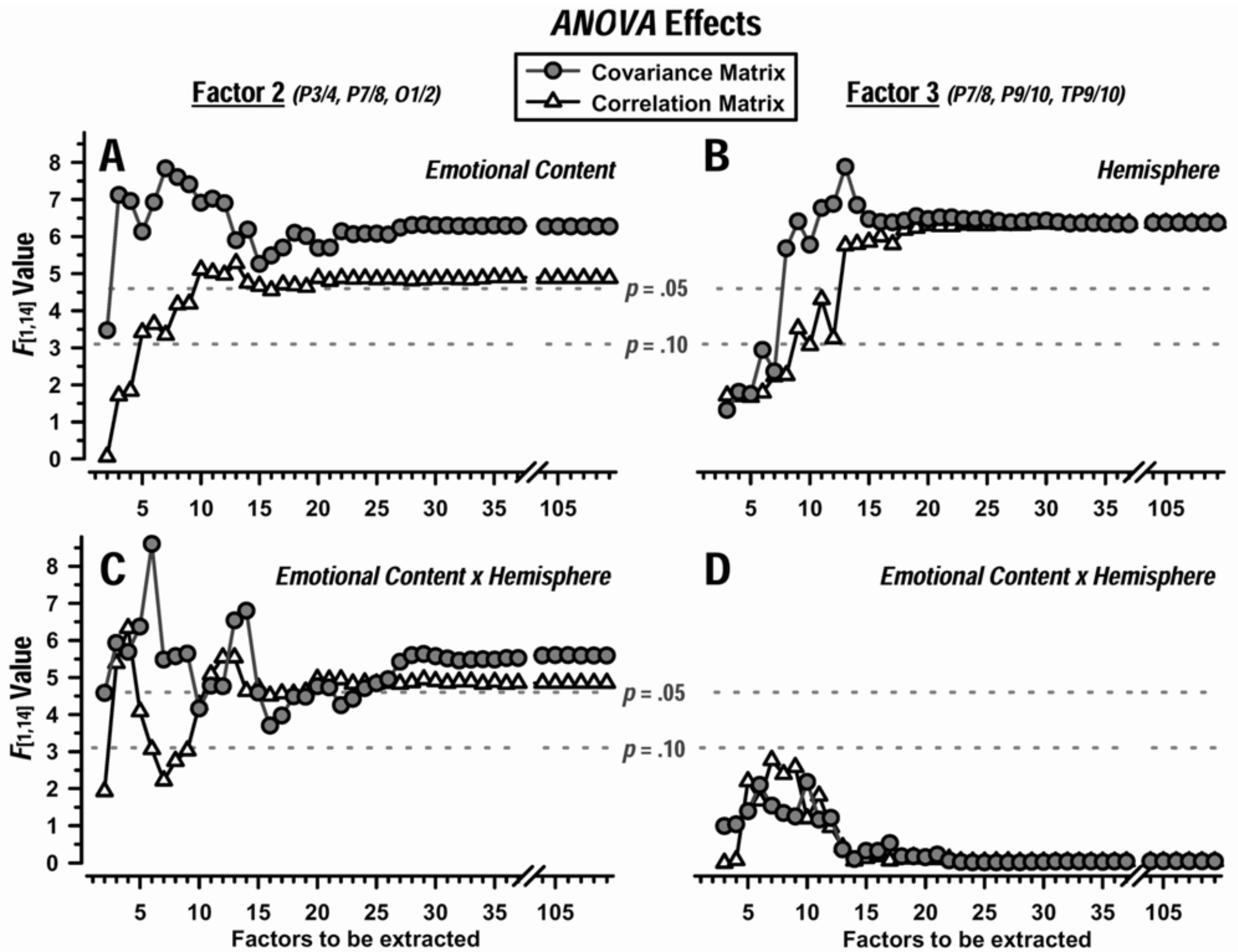

Fig. 8. Statistical effects (F values) of analyses of variance (ANOVA) performed on covariance-and correlation-based PCA factor scores as a function of extraction criterion (number of factors to be extracted). For all analyses, gender was a between-subjects factor (N = 16, 7 male), and emotional content (2), visual field (2), hemisphere (2), and site (3 homologous electrode pairs as indicated in parentheses) were within-subjects factors (cf. Kayser et al., 2000a). Shown are critical F tests (all df = 1,14): for factor 2 (P3), the emotional content main effect **(A)**; for factor 3 (N2), the hemisphere main effect **(B)**; and for both factors, the Emotional Content × Hemisphere interaction effect **(C, D)**. Dashed lines indicate F values corresponding to conventional significance levels of .05 and .10. Note that with more conservative extraction criteria, the F statistic of any given effect is highly variable and may or may not attain statistical significance, despite factors 2 and 3 representing high-variance components. In contrast, statistical effects become very stable with more liberal extraction criteria (25 or more factors extracted). The abscissa break between 35 and 105 extracted factors indicates the omission of redundant F values.

correlation-based solutions (all $F_{[1,14]} < 4.60$, $p > .05$). Likewise, the lateralized emotional content effect for P3, as measured by factor 2 (Fig. 8C), randomly crossed the conventional .05 threshold for covariance-based solutions when fewer than 27 factors were extracted, for instance, approaching a .01 significance level for 6 factors extracted ($F_{[1,14]} = 8.60$, $p = .011$), and almost dropping to .10 for the solution restricted to 16 factors, before settling on a .05 significance level for all other liberal extraction criteria. For correlation-based solutions, this important interaction even varied between significance (for 3 and 4 factors extracted, $F_{[1,14]} > 4.60$, $p < .05$) and non-significance (for 2 and 6 to 9 factors extracted, all $F_{[1,14]} < 3.10$, $p > .10$) for more conservative solutions, before also settling on a .05 significance level for more liberal extraction criteria. Conversely, an Emotional Content × Hemisphere interaction effect was effectively nonexistent in the F statistics for factor 3 (Fig. 8D) for solutions of either association matrix when more than 12 factors were extracted, but attained notable F values when extraction was limited to 6 or 10 factors for covariance-based solutions (both $F_{[1,14]} > 2.10$, $p < .17$), or to 5 and 7 to 9 factors for correlation-based solutions (all $F_{[1,14]} > 2.20$, $p < .16$).

Thus, as the factors converged towards the unrestricted solution, statistical effects also converged on a stable level, which was accomplished for the present data set when 30 or more factors were extracted, corresponding to 99.64% explained variance for the covariance-based solution (99.50% for the correlation-based solution). It should be noted that when extracting as many as 23 factors (98.96% and 98.60% data variance), the *F* statistic was still not stable (e.g., Fig. 8C). Statistical effects for the covariance-based solutions were generally stronger than those for correlation-based solutions, if these effects were also significant for the unrestricted solutions (Fig. 8A-C); analogously, statistical effects tended to be weaker for covariance- compared with correlation-based solutions if the *F* tests of the unrestricted solutions were insignificant (Fig. 8D). As these statistical analyses typically form the basis for interpreting experimental findings, this remarkable observation implies that failing to effectively explain all the data variance could lead a researcher to over- or underestimate effects of interests, missing the goal of accurately measuring relevant ERP components. This problem is move severe with correlation- than covariance-based solutions, which may result from a less-efficient description of the sequential organization of ERP variance, particularly with respect to the experimental conditions.

3.3. Real ERP data (auditory modality)

The different Varimax-rotated PCA solutions derived from auditory ERP data collected during tonal and phonetic oddball tasks revealed the very same principles. Due the completely different nature of this ERP paradigm (e.g., modality, response requirements, cognitive demands, procedural characteristics, etc.), different task-specific factors were extracted (e.g., Kayser et al., 2001). However, along the sequence created by extraction procedures and criteria, these PCA components similarly converged towards an unrestricted solution, which was again found to be the most stable one. Since the general and statistical properties of these solutions were identical to those already described for the visual ERP data set, these additional analyses for auditory data are not further detailed in this report.

4. General discussion

The implications of the reported observations for temporal PCAs using Varimax rotation to achieve simple structure are straightforward. First, factor underextraction may be a serious problem, since limiting the number of components changed the morphology of some components considerably. On the other hand, overextraction was not a concern, since more liberal or even unlimited extraction criteria did not degrade or change high-variance components, which is in close agreement with previous Monte Carlo studies (Wood et al., 1996). Instead, their interpretability was improved by more distinctive time courses with narrow and unambiguous peaks (i.e., low secondary loadings). Moreover, some physiologically meaningful ERP components, either small in amplitude, or topographically localized, or both (e.g., P1), may have a PCA equivalent, which cannot be extracted with restricted solutions due to their low overall variance contributions. Thus, whenever computational constraints are not an issue, unrestricted PCA solutions, or solutions explaining almost all the variance (> 99%), are the preferable choice.

Second, unstandardized covariance-based factors had more distinctive time courses (i.e., lower secondary loadings) than the corresponding correlation-based factors, which allowed a better interpretation of their electrophysiological relevance. Moreover, correlation-based solutions were likely to produce artificial factors that merely reflected small but systematic variations, for instance, when the ERP waveforms intersected the baseline. While such an artifact may be easily recognized as such during the baseline, similar small, systematic variations may also be present throughout the recording epoch, but then with less justification to be discounted. Including the baseline within the submitted data appears to be a reasonable methodological choice to direct attention to the extraction of spurious PCA components. Standardizing covariance-based PCA factors before rotation, a default operation in some statistical software packages, approximated correlation-based solutions, and ultimately yielded the same coefficients (factor loadings) when all components were rotated. Thus, PCA solutions based on the covariance matrix using the unstandardized components for rotation are clearly preferable over correlation-based solutions and extraction/rotation procedures that mimic the latter. Apart from merely describing the type of association matrix (Picton et al., 2000), it is strongly suggested that investigators clearly identify the procedures and statistical software used to perform the PCA.

Third, the differences of the extracted components resulting from these methodological choices can have a profound impact on the statistical analyses usually performed on the associated factor scores to test experimental hypotheses. Factor underextraction yielded volatile statistical tests, whereas increasingly liberal extraction criteria converged on a stable test statistic. The problem is aggravated with correlation-based solutions. While the reported observations clearly indicate that these problems are a real concern for the specific visual and auditory tasks considered in this report, it is currently unknown whether there are certain properties of other ERP data sets (e.g., differences in effect size, sample size, electrode montage, experimental design, etc.) that may render these concerns unnecessary – a question that may be addressed with suitable simulation studies. Meanwhile, one may be inclined to substitute the various guidelines for the number of components to extract and rotate (e.g., scree test, Eigenvalue larger than one, parallel analysis) with the graphical analysis presented in Fig. 8 to justify extracting less than all possible factors. However, not every researcher may be prepared to invest this extra effort. Since the true number of factors granting stable statistical tests is generally unknown, we would argue that employing an unrestricted PCA solution is the conservative approach.

While this report has focused on temporal PCA, similar a priori knowledge about the topographical aspects of ERP variance can be applied to evaluate component measures when using a spatial PCA, or a combination of both approaches (e.g., Achim and Bouchard, 1997; Möcks, 1988a, 1988b; Spencer et al., 2001). For a spatial PCA, the dimensions (space and time) in which an ERP component factor is defined are parsed in a complementary fashion: variables (matrix columns) consist of electrode locations included in the EEG montage, whereas matrix rows consist of time points, conditions, and participants. Factor extraction is determined by reliable variations of topography (space) rather than waveform (time). In this case, variance should be relatively small at sites closer to the recording reference, but larger at more distant scalp locations and over scalp regions used to define individual ERP components within the context of a specific paradigm. For example, variance should be larger at mid-parietal sites when a classical P3b is observed for infrequent events in an oddball paradigm (e.g., Spencer et al., 1999), or over inferior-lateral sites for visual N1 and N2 components when ERPs are referenced to nose tip (e.g., Kayser et al., 1999, 2003).

The *observed* ERP component structure, which is evident from the ERP waveforms, may considerably vary with different recording references (e.g., Dien, 1998b; Kayser et al., 2003), although the underlying generator activity remains the same, meaning that the *conceptual* ERP component structure is unchanged. Any change in reference will also affect PCA component extraction, since the variance is reallocated around a different grand mean waveform. However, its sensitivity to variance rather than to changes of peaks and troughs makes the PCA approach less susceptible to reference changes compared to more traditional ERP measures, such a area measurements or peak amplitudes. It has been argued that an average reference may avoid some of the pitfalls inherent to a single reference (e.g., Dien, 1998b), and suggested that an average reference best approximates a "reference-free" recording (Lehmann, 1984; Pascual-Marqui and Lehmann, 1993; however, see Desmedt and Tomberg, 1990; Tomberg et al., 1990). While an average reference may provide a reasonable compromise, it is not a universal or "reference-free' solution, since ERP waveforms are obviously unique to the specific recoding montage.

All PCA solutions are dependent on the characteristics of the raw data, which will not only change with the choice of reference but with *any* methodological variation (whether deliberate or not), including the experimental paradigm, specific procedure, stimulus modality, targeted population, sample heterogeneity, to name just a few. For any data set, the PCA approach will reveal the underlying variance structure of the raw data in a systematic, comprehensive fashion. It may be argued that by recommending use of an unrestricted solution, the main purpose of a PCA is put ad absurdum, namely to attain a systematic reduction of the data dimensionality into meaningful entities. The key issue here, however, is the term meaningful: if the extracted factors are meaningful, PCA factors are a valuable adjunct to conventional ERP techniques. It is the investigator's responsibility to justify the selection or rejection of components for further analysis by attributing a 'meaning' to these components within the context of the research. This must be guided by an understanding of the correspondence between the time course and topography of PCA factors and anticipated ERP components (e.g., N1, N2, P3, slow wave), the effects of the experimental manipulations, and interpretable sources of data artifacts.

Our findings clearly suggest that factor interpretation is improved with unrestricted, covariance-based solutions. The meaning and qualitative distinctiveness of a PCA component cannot be decided by a statistical program, because the amount of explained variance alone does not make a factor meaningful. As demonstrated here for temporal PCAs, the amount of variance depends largely on the number of time points and the number of scalp locations, which together define a component. Thus, early, rather transient ERP components with a distinct topography (e.g., N1) will necessarily explain a much smaller proportion of the overall variance than sustained, long-lasting components covering a broad region (e.g., a positive slow wave). Components that summarize slow ERP activity over a longer time period, usually towards the end of the recording epoch (sometimes also at the beginning of the epoch if no baseline correction was performed), may explain a great amount of data variance, but do not necessarily reflect meaningful ERP activity but rather variance associated with the grand mean waveform (Kayser et al., 1997, 2000a; van Boxtel, 1998; Wastell, 1981). However, as such components gather variance not systematically related to the experimental manipulations, these variance contributions are removed from other, more meaningful ERP components, thereby clarifying the interpretation of the latter factors. Depending on the experimental question, a distinct low-variance factor may be by far more important to the objective than high-variance factors, which explain unsystematic variance, or variance associated with effects of secondary interest. This separation of meaningful variance is a very desirable characteristic of the PCA, because when unsystematic variance is effectively filtered from the data, it can no longer obscure effects of primary interest.

When carefully applied with sufficient understanding of the implications for the data, PCA can become a valuable, general-purpose tool, serving filter functions such as removing noise from a waveform by reassembling it from a subset of PCA components (Sinai and Pratt, 2002), or identifying and eliminating EOG artifacts using a spatial PCA (Neurosoft, 2001). Likewise, components may be recombined in a new, meaningful fashion, for instance, by establishing a composite P3 measure, or by calculating a PCA-based equivalent for N2-P3 amplitude (Kayser et al., 1997, 1998, 2001).[5]

As with any analytic technique, ERP-PCA approaches may have pitfalls that could result in misleading conclusions, particularly when data include outlying cases, temporal or spatial jitter, or have low signal-to-noise ratios (e.g., Chapman and McCrary, 1995; Dien, 1998a; Donchin and Heffley, 1978; van Boxtel, 1998). In this case, the combined approach of PCA factor extraction, Varimax-rotation, and ANOVA performed on factor scores may result in a misallocation of variance, an issue repeatedly addressed in the literature (Achim and Marcantoni, 1997; Beauducel and Debener, 2003; Chapman and McCrary, 1995; Möcks, 1986; Möcks and Verleger, 1986; Wood and McCarthy, 1984). Since a further discussion of these issues is beyond the scope of this report, it will suffice to note that traditional component measures are also, and more severely, affected by these limitations (Beauducel et al., 2000; Beauducel and Debener, 2003; Chapman and McCrary, 1995; Donchin and Heffley, 1978; Wood and McCarthy, 1984). What is ultimately worse, traditional peak and area measures invite a more superficial data analysis, particularly when combined with an automated scoring routine, and are therefore less likely to alert the researcher that there may be a serious problem with the recorded data.

Our goal is to use PCA as a heuristic tool for 1) identifying relevant ERP components, and 2) generating efficient measurements for temporally and spatially overlapping components for a given data set. When applied appropriately, PCA-based component measures are more efficient when compared to more traditional ERP measures. We have reported better statistical properties for PCA-based component measures as opposed to area definitions (i.e., larger effect sizes; Kayser et al., 1998), and found PCA-based measures to have by far superior reliabilities than peak-determined amplitudes (Beauducel et al., 2000). In particular, PCA avoids the subjectivity of selecting time windows for components that invert across the scalp topography, or for subjects, conditions, or recording sites that reveal no distinct peaks within the time interval of interest, the most glaring problem in implementing peak amplitude and latency measures. Instead, after understanding the data variance, the investigator's expertise is required to make informed decisions regarding the interpretation of the extracted factors.

Acknowledgements

This research was supported in part by NIMH grant MH058346 awarded to the first author.

Software for ERP waveform display was developed by Charles L. Brown, III.

Summaries of various ideas and findings addressed in this report have been presented at the 40th Annual Meeting of the Society for Psychophysiological Research (SPR), October 2000, San Diego, CA, and at the 8th Evoked Potentials International Conference (EPIC XIII), July 2001, Paris.

We greatly appreciate helpful comments of two reviewers, particularly the suggestion to also evaluate the *statistical* impact of key choices when applying PCA methodology to ERP data.

[5] Just like time series data, other physiological measures characterized by sequenced data may also benefit from the proposed PCA methodology, using similar considerations tailored to the data. For instance, EEG frequency spectra contain large amplitudes at low frequencies (more variance), converge towards zero at high frequencies (less variance at 'baseline'), and include reproducible condition-related effects (eyes open and closed) in the alpha range. By submitting amplitude spectra to a *frequency* PCA, we have extracted distinct factors representing overlapping EEG components differentiating various subbands of alpha, EOG, and EMG activity (Debener et al., 2000; Kayser et al., 2000b).

References

Achim A, Bouchard S. Toward a dynamic topographic components model. Electroenceph Clin Neurophysiol 1997;103(3):381-385.

Achim A, Marcantoni W. Principal component analysis of event-related potentials: misallocation of variance revisited. Psychophysiology 1997;34(5):597-606.

Beauducel A, Debener S. Misallocation of variance in event-related potentials: simulation studies on the effects of test power, topography, and baseline-to-peak versus principal component quantifications. J Neurosci Methods 2003; 124(1):103-112.

Beauducel A, Debener S, Brocke B, Kayser J. On the reliability of augmenting/reducing: peak amplitudes and principal components analysis of auditory evoked potentials. J Psychophysiol 2000; 14(4):226-240.

Berg P, Scherg M. A multiple source approach to the correction of eye artifacts. Electroenceph Clin Neurophysiol 1994;90(3):229-241.

Bortz J. Statistik für Sozialwissenschaftler. Berlin: Springer, 1993.

Cattell RB. Scree test for number of factors. Multivar Behav Res 1966;1(2):245-276.

Chapman RM, McCrary JW. EP component identification and measurement by Principal Components Analysis. Brain Cogn 1995; 27(3):288-310.

Debener S, Kayser J, Tenke CE, Beauducel A. Principal components analysis (PCA) as a tool for identifying EEG frequency bands: II. Dissociation of resting alpha asymmetries. Psychophysiology 2000;37:S35.

Desmedt JE, Tomberg C. Topographic analysis in brain mapping can be compromised by the average reference. Brain Topogr 1990; 3(1):35-42.

Dien J. Addressing misallocation of variance in principal components analysis of event-related potentials. Brain Topogr 1998a; 11(1):43-55.

Dien J. Issues in the application of the average reference: review, critiques, and recommendations. Behav Res Methods Instrum Comput 1998b;30(1):34-43.

Dien J, Frishkoff GA, Cerbone A, Tucker DM. Parametric analysis of event-related potentials in semantic comprehension: evidence for parallel brain mechanisms. Brain Res Cogn Brain Res 2003; 15(2):137-153.

Dixon WJ, editor. BMDP Statistical Software Manual: To Accompany the 7.0 Software Release. Berkeley, CA: University of California Press, 1992.

Donchin E. A multivariate approach to the analysis of average evoked potentials. IEEE Trans Biomed Eng 1966;13(3):131-139.

Donchin E, Callaway R, Cooper R, Desmedt JE, Goff WR, Hillyard SA, Sutton S. Publication criteria for studies of evoked potentials (EP) in man. In: Desmedt JE, editor. Progress in Clinical Neurophysiology: Vol. 1. Attention, Voluntary Contraction and Event-Related Cerebral Potentials. Basel: Karger, 1977:1-11.

Donchin E, Heffley EF. Multivariate analysis of event-related potential data: a tutorial review. In: Otto DA, editor. Multidisciplinary Perspectives in Event-Related Brain Potential Research. Proceedings of the Fourth International Congress on Event-Related Slow Potentials of the Brain (EPIC IV), Hendersonville, NC, April 4-10, 1976. Washington, DC: The Office, 1978:555-572.

Donchin E, Ritter W, McCallum WC. Cognitive Psychophysiology: the endogenous components of the ERP. In: Callaway E, Tueting P, Koslow SH, editors. Event-Related Brain Potentials in Man. New York: Academic Press, 1978:349-411.

Donchin E, Spencer KM, Dien J. The varieties of deviant experience: ERP manifestations of deviance processors. In: van Boxtel GJM, Bocker KBE, Begleiter H, editors. Brain and Behavior: Past, Present, and Future. Tilburg, NL: Tilburg University Press, 1997:67-91.

Duffy FH, Jones K, Bartels P, McAnulty G, Albert M. Unrestricted principal components analysis of brain electrical activity: issues of data dimensionality, artifact, and utility. Brain Topogr 1992;4(4):291-307.

Everett J. Factor comparability as a means of determining the number of factors and their rotation. Multivar Behav Res 1983; 18(2):197-218.

Fabiani M, Gratton G, Karis D, Donchin E. Definition, identification, and reliability of measurement of the P300 component of the event-related brain potential. In: Ackles PK, Jennings JR, Coles MGH, editors. Advances in Psychophysiology (Vol. 2). Hillsdale, NJ: JAI Press, 1987:1-78.

Fabiani M, Gratton G, Coles MGH. Event-related brain potentials: methods, theory, and applications. In: Cacioppo JT, Tassinary LG, Bernston GG, editors. Handbook of Psychophysiology (2nd Edition). Cambridge University Press, 2000:53-84.

Fava JL, Velicer WF. The effects of overextraction on factor and component analysis. Multivar Behav Res 1992;27(3):387-415.

Fava JL, Velicer WF. The effects of underextraction in factor and component analyses. Educ Psychol Meas 1996;56(6):907-929.

Friedman D. ERPs during continuous recognition memory for words. Biol Psychol 1990;30(1):61-87.

Glaser EM, Ruchkin DS. Principles of neurobiological signal analysis. New York: Academic Press, 1976.

Guadagnoli E, Velicer WF. Relation of sample size to the stability of component patterns. Psychol Bull 1988;103(2):265-275.

Harman HH. Modern Factor Analysis (2nd ed). Chicago: University of Chicago Press, 1967.

Hendrickson AE, White PO. Promax: a quick method for rotation to oblique simple structure. Brit J Stat Psychol 1964;17:65-70.

Horn JL. A rationale and test for the number of factors in factor-analysis. Psychometrika 1965;30(2):179-185.

Hunt E. Mathematical models of the event-related potential. Psychophysiology 1985;22(4):395-402.

Kaiser HF. The Varimax criterion for analytic rotation in factor analysis. Psychometrika 1958;23:187-200.

Kaiser HF. The application of electronic computers to factor analysis. Educ Psychol Meas 1960;20:141-151.

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):832-847.

Kayser J, Bruder GE, Friedman D, Tenke CE, Amador XF, Clark SC, Malaspina D, Gorman JM. Brain event-related potentials (ERPs) in schizophrenia during a word recognition memory task. Int J Psychophysiol 1999;34(3):249-265.

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 2000a;36(3):211-

236.

Kayser J, Fong R, Tenke CE, Bruder GE. Event-related brain potentials during auditory and visual word recognition memory tasks. Brain Res Cogn Brain Res 2003;16(1):11-25.

Kayser J, Tenke CE, Bruder GE. Dissociation of brain ERP topographies for tonal and phonetic oddball tasks. Psychophysiology 1998;35(5):576-590.

Kayser J, Tenke CE, Debener S. Principal components analysis (PCA) as a tool for identifying EEG frequency bands: I. Methodological considerations and preliminary findings. Psychophysiology 2000b;37:S54.

Kayser J, Tenke C, Nordby H, Hammerborg D, Hugdahl K, Erdmann G. Event-related potential (ERP) asymmetries to emotional stimuli in a visual half-field paradigm. Psychophysiology 1997; 34(4):414-426.

Kraut MA, Arezzo JC, Vaughan HG, Jr. Intracortical generators of the flash VEP in monkeys. Electroenceph Clin Neurophysiol 1985; 62(4):300-312.

Kutas M, Hillyard SA. Reading senseless sentences: brain potentials reflect semantic incongruity. Science 1980;207(4427):203-205.

Lehmann D. EEG assessment of brain activity: spatial aspects, segmentation and imaging. Int J Psychophysiol 1984;1(3):267-276.

Lehmann D, Skrandies W. Reference-free identification of components of checkerboard-evoked multichannel potential fields. Electroenceph Clin Neurophysiol 1980;48(6):609-621.

Möcks J. The influence of latency jitter in principal component analysis of event-related potentials. Psychophysiology 1986; 23(4):480-484.

Möcks J. Decomposing event-related potentials: a new topographic components model. Biol Psychol 1988a;26(1-3):199-215.

Möcks J. Topographic components model for event-related potentials and some biophysical considerations. IEEE Trans Biomed Eng 1988b;35(6):482-484.

Möcks J, Verleger R. Principal component analysis of event-related potentials: a note on misallocation of variance. Electroenceph Clin Neurophysiol 1986;65(5):393-398.

Näätänen R, Picton TW. The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. Psychophysiology 1987;24(4):375-425.

Neurosoft, Inc. EDIT 4.2 - Offline analysis of acquired data (Document number 2202, Revision B). Sterling, VA: Neuroscan Labs, 2001.

Oostenveld R, Praamstra P. The five percent electrode system for high-resolution EEG and ERP measurements. Clin Neurophysiol 2001;112(4):713-719.

Pascual-Marqui RD, Lehmann D. Topographic maps, source localization inference, and the reference electrode: comments on a paper by Desmedt et al. Electroenceph Clin Neurophysiol 1993; 88(6):532-536.

Picton TW, Bentin S, Berg P, Donchin E, Hillyard SA, Johnson R Jr, Miller GA, Ritter W, Ruchkin DS, Rugg MD, Taylor MJ. Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria. Psychophysiology 2000;37(2):127-152.

Pivik RT, Broughton RJ, Coppola R, Davidson RJ, Fox N, Nuwer MR. Guidelines for the recording and quantitative analysis of electroencephalographic activity in research contexts. Psychophysiology 1993;30(6):547-558.

Rose JE, Woosley CN. Organization of the mammalian thalamus and its relationships to the cerebral cortex. Electroenceph Clin Neurophysiol 1949;1:391-404.

Rösler F, Manzey D. Principal components and varimax-rotated components in event-related potential research: some remarks on their interpretation. Biol Psychol 1981;13:3-26.

Schroeder CE, Tenke CE, Givre SJ, Arezzo JC, Vaughan HG, Jr. Striate cortical contribution to the surface-recorded pattern-reversal VEP in the alert monkey. Vision Res 1991;31(7-8):1143-1157.

Simson R, Vaughan HG, Ritter W. The scalp topography of potentials associated with missing visual or auditory stimuli. Electroenceph Clin Neurophysiol 1976;40(1):33-42.

Sinai A, Pratt H. Electrophysiological evidence for priming in response to words and pseudowords in first and second language. Brain Lang 2002;80(2):240-252.

Spencer KM, Dien J, Donchin E. A componential analysis of the ERP elicited by novel events using a dense electrode array. Psychophysiology 1999;36(3):409-414.

Spencer KM, Dien J, Donchin E. Spatiotemporal analysis of the late ERP responses to deviant stimuli. Psychophysiology 2001; 38(2):343-358.

SPSS Inc. SPSS-X User's Guide. 3rd Edition. Chicago, IL: SPSS Inc, 1988.

Squires N, Squires K, Hillyard S. Two varieties of long latency positive waves evoked by unpredictable auditory stimuli in man. Electroenceph Clin Neurophysiol 1975;38:387-401.

Sutton S, Braren M, Zubin J, John ER. Evoked potential correlates of stimulus uncertainty. Science 1965;150:1187-1188.

Tenke CE, Schroeder CE, Arezzo JC, Vaughan HG, Jr. Interpretation of high-resolution current source density profiles: a simulation of sublaminar contributions to the visual evoked potential. Exp Brain Res 1993;94(2):183-192.

Thurstone LL. Multiple factor analysis. Chicago, IL: University of Chicago Press, 1947.

Tomberg C, Noel P, Ozaki I, Desmedt JE. Inadequacy of the average reference for the topographic mapping of focal enhancements of brain potentials. Electroenceph Clin Neurophysiol 1990;77(4):259-265.

van Boxtel GJM. Computational and statistical methods for analyzing event-related potential data. Behav Res Methods Instrum Comput 1998; 30(1):87-102.

Wastell DG. On the correlated nature of evoked brain activity: biophysical and statistical considerations. Biol Psychol 1981;13:51-69.

Wood JM, Tataryn DJ, Gorsuch RL. Effects of under- and overextraction on principal axis factor analysis with varimax rotation. Psychol Methods 1996;1(4):354-365.

Wood CC, McCarthy G. Principal component analysis of event-related potentials: simulation studies demonstrate misallocation of variance across components. Electroenceph Clin Neurophysiol 1984; 59(3):249-260.

Zwick WR, Velicer WF. Comparison of five rules for determining the number of components to retain. Psychol Bull 1986;99(3):432-442.

Appendix

The proposed unrestricted, unstandardized covariance--based PCA with Varimax rotation may be compactly coded in high-level numerical languages, including the popular MatLab (The MathWorks, Inc., 2002, MatLab Version 6.5, Release 13 [*http://www.mathworks.com*]). All crucial computational steps are exemplified by the MatLab routine provided below. It is important to note that the component weights (i.e., the factor scores) are estimated from intermediate factor score coefficients (rescaled with respect to the input variables; cf. lines 25-27) and the normalized raw data (ll. 28-34). The implemented Varimax procedure applies Kaiser's normalization (i.e., the rows are normalized by their individual lengths) before rotation, which is reversed after Varimax convergence. Marginal computational differences from the BMDP-4M implementation (Dixon, 1992) are largely due to improved data precision, which affects estimates of matrix rank order (l. 8) and Varimax convergence (l. 14).

```
% erpPCA - Unrestricted, unstandardized covariance-based PCA with Varimax
%          rotation (cf. Kayser J, Tenke CE, Clin Neurophysiol, in press)
%
% Usage: [LU, LR, FSr, VT] = erpPCA( X )
%
% Generic PCA-Varimax implementation emulating the PCA agorithms used by
% BMDP-4M (Dixon, 1992) and SPSS 10.0 FACTOR (http://www.spss.com/tech/
% stat/Algorithms/11.5/factor.pdf). It expects a data matrix (X) of ERP
% waveforms, with ERPs (cases) as rows and sample points (variables) as
% columns. The routine returns the unrotated (LU) and Varimax-rotated (LR)
% factor loadings as a variables-by-factors matrix, the rotated factor
% scores (FSr) as a cases-by-factors matrix, and Eigenvalues and explained
% variance as a variables-by-variance matrix (VT), with four columns
% consisting of Eigenvalues and percentage of explained variance before
% and after rotation.
%
% erpPCA employs Varimax4M (max. 100 iterations, 0.0001 convergence criterion,
% Kaiser's normalization; MatLab code by $jk available on request), which
% emulates algorithms described by Harman (1967, pp. 304-308) as implemented
% in BMDP-4M (Dixon, 1992, pp. 602-603).
%
% Copyright (C) 2003 by Jürgen Kayser (Email: kayserj@pi.cpmc.columbia.edu)
% GNU General Public License (http://www.gnu.org/licenses/gpl.txt)
% Updated: $Date: 2003/07/08 14:00:00 $ $Author: jk $
%
function [LU,LR,FSr,VT] = erpPCA(X)
1  [cases, vars] = size(X);            % get dimensions of input data matrix
2  D = cov(X);                         % compute covariance matrix
3  [EM, EV] = eig(D);                  % determine Eigenvectors and Eigenvalues
4  UL = EM * sqrt(EV);                 % determine unrotated factor loadings
5  [u, ux] = sort(diag(EV)');          % sort initial Eigenvalues, keep indices
6  u = fliplr(u); ux = fliplr(ux);     % set descending order
7  LU = UL(:,ux);                      % sort unrotated factor loadings
8  rk = rank(corrcoef(X),1e-4);        % estimate the number of singular values
9  LU = LU(:,1:rk);                    % ... and remove all linearly dependent
10 u = u(1:rk); ux = ux(1:rk);         % ... components and their indices
11 s = ones(1,rk);                     % current sign of loading vectors
12 s(abs(max(LU)) < abs(min(LU))) = -1; % determine direction of loading vectors
13 LU = LU .* repmat(s,size(LU,1),1);  % redirect loading vectors if necessary
14 RL = Varimax4M(LU,100,1e-4,1);      % Varimax-rotate factor loadings
15 EVr = sum(RL .* RL);                % compute rotated Eigenvalues
16 [r, rx] = sort(EVr);                % sort rotated Eigenvalues, keep indices
17 r = fliplr(r); rx = fliplr(rx);     % set descending order
18 LR = RL(:,rx);                      % sort rotated factor loadings
19 s = ones(1,size(LR,2));             % current sign of loading vectors
```

```
s(abs(max(LR)) < abs(min(LR))) = -1; % determine direction of loading vectors
LR = LR .* repmat(s,vars,1);         % redirect loading vectors if necessary
tv = trace(EV);                      % compute total variance
VT = [u' 100*u'/tv ...               % table explained variance for unrotated
      r' 100*r'/tv ];                % ... and Varimax-rotated components
FSCFr = LR * inv(LR' * LR);          % compute rotated FS coefficients
FSCFr = FSCFr .* ...                 % rescale rotated FS coefficients by
   repmat(sqrt(diag(D)),1,rk);       % ... the corresponding SDs
mu = mean(X); sigma = std(X);        % compute Mean and SD for each variable
Xc = X - repmat(mu,cases,1);         % remove grand mean
FSr = zeros(cases,vars);             % claim memory to speed computations
for n = 1:cases; for m = 1:rk;       % compute rotated factor scores from
 FSr(n,m) = sum( (Xc(n,:) ./ ...     % ... the normalized raw data and
       sigma) .* FSCFr(:,m)' );      % ... the corresponding rescaled
end; end;                            % ... factor score coefficients
```