ELSEVIER

Clinical Neurophysiology 116 (2005) 1747–1753

www.elsevier.com/locate/clinph

Editorial

Trusting in or breaking with convention: Towards a renaissance of principal components analysis in electrophysiology

See Article, pages 1808–1825

1. Historical perspective on the concept of an electrophysiologic component

Event-related potential (ERP) measures have an intimate and historic association with basic neuroanatomy and neurophysiology. The sensory projection systems were originally mapped by the pioneers using the simplified evoked potential waveforms of deeply anesthetized animals, from which functional activity is easily quantified by direct amplitude and latency measurements (e.g. Marshall et al., 1937, 1941). The clinical evoked potentials and cognitive ERPs that descended from these classical studies rely on signal averaging methods (e.g. Dawson, 1954) to reveal time-locked activity that reflects the transfer and processing of information by the nervous system in awake human subjects. The variability and complexity of the resulting waveforms provide a considerably greater challenge for identifying and quantifying functional activity of interest. The construct of an ERP component (or any equivalent construct; e.g. Lehmann and Skrandies, 1984) is used to decompose and understand ERP waveforms according to both their intracerebral origin (i.e. underlying neuronal generators) and their dependence on specific experimental manipulations (e.g. Picton et al., 2000).

2. The challenge: signal identification and quantification

Time-locked electrophysiological activity measured during cognitive tasks is complex on many levels. Different aspects of information processing partially overlap in time and all contribute through principles of superposition and volume-conduction to the observed ERP waveform at any recording location. Since multi-channel EEG with at least a moderate number of scalp placements (i.e. 20 or more) has become the de facto recording standard, a single ERP signal easily consists of thousands (e.g. 21 channels, 100 samples/s) or even hundreds of thousands of data points (e.g. 128 channels, 1000 samples/s). Temporal (i.e. neighboring sample points) and spatial (i.e. neighboring recording sites) proximity warrants highly intercorrelated signals. Adding to this complexity are physiological and non-physiological artifacts, such as movement, eye, muscle, or electrical field, all also contributing to the ERP signal through the very same biophysical properties (i.e. superposition and volume-conduction). How can one reduce this overwhelming data multi-dimensionality to meaningful ‘constituent elements’ or components, that provide an effective description and summary for a limited number of temporal patterns of activity, each originating from coordinated activity in neuroanatomical structures, and corresponding to a specific manner of information processing by the brain (cf. Picton et al., 2000)?

The mere recognition of the ‘obvious’ ERP characteristics, that is, prominent deflections in ERP waveforms, and, subsequently, simple quantification of their peak amplitude and latency, is an intuitive approach remarkably similar to the one used in classic evoked potential studies and has been surprisingly useful for ERP research. For instance, identifying an N2/P3 complex in target ERPs recorded at midline sites during an oddball task, the prototypical paradigm of cognitive ERP research, may be characterized as a ‘pop-out’ phenomenon, which can easily be recognized with only moderate experience. However, more subtle variations are not immediately evident in an ERP, even after thorough inspection of the waveforms. Certain systematic effects, which are only observable with more inclusive scalp coverage and/or more sophisticated paradigms, would qualify under the construct of an ERP component (e.g. low-amplitude experimental effects superimposed on the rising or falling slopes of prominent peaks, waveform shifts across the montage resulting from corresponding peaks and troughs at similar latencies). The separation and measurement of such meaningful ERP elements continues to be a prominent challenge in ERP research, particularly with the advent of high-density EEG. Paradoxically, ERP researchers frequently meet this ‘new’ challenge by analyzing conventional ERP measures (e.g. peak amplitude, peak

1388-2457/$30.00
doi:10.1016/j.clinph.2005.03.020

latency, area measurement) at selected recording sites or pooled across multiple sites, and sometimes refer to the latter as "regions-of-interest" measures in loose analogy to functional imaging methods with high anatomical resolution. These strategies are incongruous with the basic premise of dense-electrode arrays (i.e. improved ERP topographies), considering that the signal is either spatially smeared or needlessly impoverished, and subject to experimenter bias in the selection or grouping of electrodes. Similar limitations of traditional ERP peak and area measures affect the temporal properties of the signal, emphasizing that the intuitive simplicity of these approaches is deceiving (e.g. Donchin and Heffley, 1978). The continued use of these familiar methods, which is largely guided by the convenience of convention, impedes the implementation of more appropriate ones.

From the early days of cognitive ERP research, principal components analysis (PCA) was already proposed as a linear, multivariate data-reduction approach to address this challenge (Donchin, 1966). PCA can serve as an objective, empirical tool to determine 'data-driven' ERP component measures (e.g. Chapman and McCrary, 1995; Donchin and Heffley, 1978; Glaser and Ruchkin, 1976; van Boxtel, 1998). By identifying and grouping unique variance patterns for a given set of ERP waveforms, PCA decomposes the variance structure of the observed data into components (i.e. factor loadings) and their associated weights (i.e. factor scores), which may be interpreted as observational definitions of ERP components. Early attempts at using PCA to operationally define ERP components were quite promising (e.g. Chapman et al., 1978; Friedman et al., 1981; John et al., 1973; Molfese, 1984; Ruchkin et al., 1980; Squires et al., 1977), and the Zeitgeist made the PCA-Varimax strategy (i.e. factor extraction followed by orthogonal factor rotation to simplify factor interpretation) the fashionable and primary analytic tool for many ERP researchers (cf. Gaillard and Ritter, 1983). Some disagreements were limited to how to take full advantage of this new methodology, such as the separability of components (Donchin and Heffley, 1979; Wastell, 1979) or the use of a rotation procedure (Rösler and Manzey, 1981; Wastell, 1981). A less obvious, but no less crucial problem of that period, however, was that PCA required considerable computational resources to process the data from even the small recording montages available. In line with these constraints, PCA implementations were limited in scope (i.e. number of input variables and cases, output of extracted factors), providing variable and unwieldy solutions that were anything but intuitive (e.g. Donchin and Heffley, 1978; Pritchard, 1984). In retrospect, one could also argue that the limited, spatially undersampled ERP topographies of that time were inadequate to characterize the underlying neuronal generators, and, consequently, to arrive at a stable PCA factor structure, which was primarily inferred from variance arising from conditions and subjects, rather than topographies.

When it became clear that PCA was not a panacea, ERP applications of PCA began to be met with increased skepticism (e.g. Hunt, 1985). It was in this environment that Wood and McCarthy (1984) published an influential and now classic study, which explored the ability of PCA to reconstruct ERP component prototypes (i.e. their shapes) used to simulate ERP data, and the degree to which the ensuing univariate analysis of variance performed on the associated factor scores would correctly attribute a condition effect introduced to only one component. Their findings indicated that when prototypes overlapped, a significant condition effect was often falsely attributed to the unaffected prototype. This inflation of Type I error rates was termed misallocation of variance. Although the authors stressed that variance misallocation and misinterpretation of experimental effects are equally possible for other (i.e. more conventional) ERP component measures, the implications of these cautious conclusions were by-and-large overlooked. Shortly thereafter, Möcks and Verleger (1986) clarified that the Varimax rotation step is responsible for the misallocation of variance, rather than the PCA extraction itself. Varimax may even produce slow-wave-like factors (Verleger and Möcks, 1987). Furthermore, the extracted factors may not even be a 'true' representation of the underlying ERP components, since infinitely many sets of prototypes (or generators) may render an identical association matrix used for factor extraction. The adequacy of using PCA as a tool for ERP analysis became an open controversy (e.g. Collet, 1989; Donchin, 1989; Möcks, 1989), and reasonable suggestions to improve this methodology (Möcks, 1988a,b) did not resonate within the research community. Reluctance to apply what was perceived to be an uncertain technique that required enormous caution led many to abandon PCA in favor of traditional, intuitive methods (i.e. peak and window amplitudes) for all but descriptive purposes (e.g. to summarize data and to select time windows for obvious components).

3. Principal components analysis: a renaissance

Over the last two decades, computation has become powerful, cheap, and accessible to a degree that was previously unimaginable. Likewise, when high-density recordings, based on extensions of the 10–20 system (e.g. Oostenveld and Praamstra, 2001) or evenly spaced geometry (e.g. Tucker, 1993), provided improved topographic information, an urgent need arose to simplify the massive amount of data in a physiologically meaningful way, warranting a renewed effort at PCA. Conservative skepticism was replaced by a new understanding of both the strengths and the weaknesses of PCA, when several simulation studies revisited the problem of misallocation of variance, again modeling the component prototypes introduced by Wood and McCarthy (1984), but using new perspectives and comparisons not previously considered

(Achim and Marcantoni, 1997; Beauducel and Debener, 2003; Chapman and McCrary, 1995; Dien, 1998). For instance, Chapman and McCrary (1995) noted that the misallocation of variance, or component leakage, is likely due to correlated prototypes rather than their overlap, and may spread from components producing a large effect size to neighboring components yielding a much smaller effect size, suggesting that a critical review and comparison of analysis of variance (ANOVA) effects across *all* components will confine the risk of accepting false positives. Furthermore, if there is only *one* specific ANOVA effect on *one* component, there is no ambiguity. Conversely, the authors also demonstrated that when the original component prototypes were slightly altered, there was no leakage among PCA components; however, a peak amplitude measure resulted in a perfect 100% misallocation of variance.

Recently, Beauducel and Debener (2003) more directly addressed the importance of effect size and topography for simulated ERP data analyzed with PCA and base-to-peak measures. When using a more realistic (i.e. smaller) test power and component topography, two aspects not covered by Wood and McCarthy (1984), misallocation of variance was greatly reduced for PCA-derived component measures, and baseline-to-peak measures were equally, or even more, subject to this problem (Beauducel and Debener, 2003). These results provided a more formal theoretical basis for earlier observations of real ERP data, which showed superior statistical properties (i.e. larger effect sizes, better reliabilities) for PCA than conventional ERP measures (e.g. Beauducel et al., 2000; Kayser et al., 1998).

Although established PCA methods became the new orthodoxy, modifications began to emerge. In particular, much recent research has focused on the assumption that an ERP component is also, and perhaps primarily, characterized by its spatial (i.e. topographic) rather than its temporal configuration. However, the traditional two-dimensional ERP data matrix arrangement, consisting of observations (electrodes, conditions, subjects) by time points, does not explicitly use topographic information as a constraint for factor extraction. These limitations could be overcome by a topographic component model (Field and Graupe, 1991; Möcks, 1988a,b; Wang et al., 2000), which has the added benefit of being less prone to misallocation of variance and latency jitter (Achim and Bouchard, 1997; Achim and Marcantoni, 1997; Möcks, 1986). Dien (1998), using an extended simulation model, has argued that spatial PCA as a complement to temporal PCA, in which the ERP data matrix is rearranged to observations (time points, conditions, subjects) by electrodes, together with parallel analysis (Horn, 1965) to identify noise factors, and oblique rotation to allow for correlated factors, can address several of these and other PCA shortcomings, such as retaining the correct number of extracted factors and correlated prototypes. A combined spatiotemporal approach has since successfully been applied to real ERP data concisely separating known ERP components (e.g. Spencer et al., 1999, 2001). Independent component analysis (e.g. Makeig et al., 1997) and partial least square analysis (e.g. Lobaugh et al., 2001) are more recent multivariate developments to decompose ERPs, which also formally integrate the concept of spatiotemporal invariance of ERP components.

4. Acceptance and beyond: unorthodox evaluations of methodological norms

Modeling ERP components through prototypes has the advantage of knowing the true underlying data structure, and the second wave of simulation studies using this approach has rehabilitated the usefulness of PCA for ERP data analysis (Achim and Marcantoni, 1997; Beauducel and Debener, 2003; Chapman and McCrary, 1995; Dien, 1998). The flip side of this control is the inability to systematically vary and evaluate all aspects relevant for real ERP data, and therefore the inability to decide whether the modeled findings apply to any given data set. For real ERP data, the true underlying component structure is unknown, which leaves the ERP researcher with the challenge to evaluate and interpret the PCA solution. Given the wide variety and range of methodological choices one has when submitting ERP data to PCA, it seemed reasonable to ask whether any of these choices would be more helpful for the goal of identifying and measuring ERP components.

In a previous report published in this journal (Kayser and Tenke, 2003), we studied the variability of PCA solutions and their ensuing inferential statistics (ANOVA) for real ERP data sets as a function of (1) the factor retention criterion, systematically comparing all possibilities for the number of components to be extracted, and (2) the type of association matrix used for factor extraction, comparing the commonly used correlation and covariance matrices believed to produce similar results (e.g. Chapman and McCrary, 1995; van Boxtel, 1998); all other PCA parameters, including a normal orthogonal factor rotation (i.e. Varimax with Kaiser's normalization), were fixed. The results were unambiguous, and confirmed our impressions gained over many years of experience in ERP research (e.g. Kayser et al., 1997, 1998). First, all restricted PCA solutions continued to become more similar with more liberal retention criteria, and eventually converged on the unrestricted solution, independent of the association matrix. Second, unrestricted factor extraction improved the interpretability of high-variance factors and yielded stable *F*-statistics, whereas conservative retention criteria resulted in highly variable ANOVA effects. Third, correlation-based solutions lead to spurious high-variance factors stemming from irrelevant but systematic variations in the ERP waveforms (e.g. baseline intersections). From these findings, we concluded that unrestricted, covariance-based PCA solutions are a superior choice for identification and measurement of physiologically meaningful ERP components. As it turns out, the recommendation of unrestricted PCA solutions was like declaring that the earth is not flat.

5. Pioneering a new frontier: redefining convention in uncharted territory

In this issue, Dien et al. (2005) provide an informative comparison of several methodological choices, arranged in factorial fashion, which include the extraction matrix type, loading weighting method, rotation method, and factor retention criteria on PCA solutions using simulated prototypes. The ambitious objective of this report is to resolve previous conflicting recommendations and to formalize a standard protocol for PCA with ERP data. One important methodological advancement is the construction of high-density ERP topographies derived from idealized dipole generators at predetermined intracranial locations, and their simplified reconstruction via dipole localization (software by Patrick Berg; http://www.besa.de), thereby mimicking idealized ERP components with considerable realism. As previously, two prototypes were constructed with a condition effect imposed on one of them, and misallocation of variance was investigated by comparing ANOVA effects derived from many simulations. Of particular interest are comparisons of an oblique factor rotation method (Promax; e.g. Hendrickson and White, 1964) against the conventional orthogonal rotation (Varimax). The good performance of Promax for these simulated data sets is impressive, and emphasizes the need for a more thorough understanding of both the advantages and disadvantages of this promising approach for real, physiological data.

The findings by Dien et al. (2005) are in close agreement with those of Kayser and Tenke (2003) where methods overlap, documenting the preferability of the covariance over the correlation association matrix for restricted PCA solutions, and the convergence of covariance and correlation matrix results for unrestricted solutions. This corroboration is particularly striking in view of the marked methodological and conceptual differences between these two reports, one following the convention of simulated prototypes, statistical tallies, and operationalized measures for reconstruction fit, the other providing an unorthodox, yet thorough exploration of effects for real ERP data, driven by the pragmatic goal of attaining interpretable and stable solutions. It is unfortunate that these theoretical and methodological differences were apparently misunderstood by these authors, who categorically criticized our methods, interpretations, and recommendations. We will therefore address important misconceptions stemming from the misrepresentation of central aspects of our rationale and conclusions that are in striking conflict with the recommendations given by Dien et al. (2005).

At the heart of any PCA is the variance explained by each extracted factor in proportion to the overall data variance. The explained variance is used to rank the factors, prioritizing factors with higher variance. Both the overall and relative variance proportion of each factor are affected by the type of association matrix. Small, but systematic noise variance is exaggerated in its importance by correlation-based extractions. In addition, factor rotation also reorganizes their variance, which may affect their ranking. Standard statistical packages (BMDP, SPSS, SAS) typically rotate the factor loadings as extracted and weighted. Although microvolt-scaled factor loadings are preferable when interpreting PCA solutions because the shape of the loadings are altered by their scaling, this necessarily requires an additional recalculation of the overall and explained variance contributions, and an additional reranking of the factors. Given the premise that correlation loadings can only be directly compared to covariance loadings (and ERP waveforms) after being rescaled to μV units, it seems almost reasonable to become upset by our decision to plot the factor loadings of the correlation- and covariance-based solutions in their respective data units. However, we simply adopted the pragmatic perspective of an applied researcher without programmers on staff, using the default output provided by a standard statistical package, in our case BMDP[1] (cf. program 4M syntax given in footnote 3 of Kayser and Tenke, 2003; Dixon, 1992). More importantly, plotting the factor loadings in the common units used for extraction and rotation provides an intuitive explanation for the percentage of explained variance assigned to any given factor, which is, as explained in our paper, directly related to the weight assigned to each variable.

Dien et al. (2005) are puzzled by the difference of the factor waveforms plotted in our Figs. 3 and 4, and contend that this is entirely due to employing different scales.[2] While rescaling would indeed increase the similarity of corresponding waveforms, and in fact result in identical waveforms for the unrestricted solutions, the main point is that the sequence of the extracted factors differs due to differences in explained variance. As shown in our Fig. 4 for real ERP data, and in our Fig. 3 for an illustrative example constructed to explain the underlying principle, correlation loadings can result in erroneous high-variance factors. This has profound implications for the proof presented by Dien et al. (2005) in their footnote 1, showing that Kaiser's normalization effectively

[1] Highly comparable default outputs, yielding identical factor loadings and sequences, were obtained with SPSS and SAS for corresponding extraction and rotation procedures. To overcome limitations of the abandoned, DOS-based BMDP software, a MatLab code was later developed (appendix of Kayser and Tenke, 2003) to replicate the BMDP algorithms for common use and larger ERP data sets.

[2] Further confusion is evident from Dien et al.'s incorrect assertion that our Fig. 3 depicts unconfounded "…correlation and covariance loadings (termed by Kayser and Tenke as the standardized and unstandardized loadings) with the covariance matrix." As clearly stated in the caption, this figure compares the factor loadings of PCA solutions using a correlation or covariance matrix for factor extraction, which by default results in correlation or covariance loadings with BMDP statistical software.

equalizes the rotation of covariance and correlation loadings if otherwise identical factors are submitted. However, if a retention criterion is applied before rotation, a non-identical set of factors may be submitted to the rotation procedure, resulting in the computation of different communalities, which leads to a different Kaiser's normalization, and finally different rotated loadings. Only the unrestricted case guarantees comparable communalities for weighting the factor loadings using Kaiser's normalization, in which case rescaling the rotated correlation loadings to covariance (μV) units will indeed result in identical waveforms, but not necessarily in the same order of extraction. It is a remarkable insight that, by equalizing corresponding factor loadings using Kaiser's normalization, not only do solutions derived from the covariance matrix using either covariance or correlation loadings during rotation fully converge for unrestricted solutions, but they are also identical, other than being out of order, with unrestricted solutions derived from the correlation matrix (see factor score topographies given in our Fig. 7).

Dien et al. (2005) claim that we proposed to retain all factors "… on the grounds that underextraction (retaining too few factors) degrades solution quality (Fava and Velicer, 1992; Wood et al., 1996)…", and in turn use this claim to argue that retaining and analyzing all factors results in a multiple comparison problem. Instead, our recommendation for using unrestricted solutions is based on our empirical findings using real ERP data, which clearly show that (1) all restricted solutions eventually converge on the unrestricted solution, and (2) that the representations of high-variance factors (i.e. the very first factors to be extracted) are improved in the sense that F statistics are stable and do not randomly depend on the factor retention criterion. The crucial point is that in either case (restricted or unrestricted solution) only a limited number of factors are considered for statistical analyses, essentially considering corresponding high-variance factors. In practice, we have never considered analyzing factors explaining less than 1% of the overall data variance, and these factors constitute by far the majority of factors in an unrestricted solution (the variance explained typically drops below 1% before the 10th extracted factor in a temporal PCA, even when using a 128-channel ERP data set; Kayser et al., 2000). Although, considering such a low-variance factor is not completely out of the question, it would require a very strong a priori rationale for analyzing and interpreting it. Unlike a simulation study, in which the true prototype components are known (in the Dien et al. article, this knowledge is used to select two out four factors), real ERP data do not provide the luxury of knowing what factors correspond to the underlying ERP components.

It is ironic that Dien et al. (2005) criticize us for providing no formal guideline for choosing factors for analysis, as they themselves encounter the very same problem by having to select the 'right' factors for analysis out the set of retained factors, which may be as many as 15 or 20 when applying a Scree test, even by their own account. Furthermore, these authors also agree with our contention that a priori knowledge about component latency and scalp topography should be applied to restrict statistical analysis to the factors of interest. The bottom line is that high-variance factors are improved with unrestricted solutions, and Figs. 2 and 3 of Dien et al. (2005) resoundingly underscore this conclusion. Given the high degree of prototype reproducibility accomplished for all PCA choices studied, it really seems counterintuitive to argue for restricted solutions. However, it would to be extremely easy to establish a formal guideline or 'systematic objectivity' to exclude the majority of factors, for instance, by simply applying a Scree test *after* rotation[3], or by setting a variance threshold (e.g. more than 1%) for the factors to be considered for analysis and interpretation. For all of these reasons, we continue to advocate postponing factor restriction until after rotation, explicitly breaking with this convention.

As it is true that in certain scenarios reasonable restrictions for factor retention *before* rotation will result in similar or almost identical high-variance factors compared to unrestricted solutions (e.g. as demonstrated for the two prototypes employed by Dien et al., 2005), it is wrong to conclude that unrestricted solutions are unnecessary. Instead, the logic of the argument would require evidence that unrestricted solutions are inferior to restricted solutions, and that researchers will always be able to select the 'optimal' cut-off point for real ERP data sets. Since information about an optimal cut-off point, if it exists, is not independently available for real ERP data, we have strongly argued for instead using the unrestricted solution as the natural stable end-point in a continuum of possible solutions. Dien et al. (2005) dismiss this reasoning by suggesting that between-subject variance may yield 'individual' factors collectively representing a component of interest, and, referring to the Monte-Carlo studies by others, that overextraction may result in the distortion of major factors and/or the creation of uninterpretable, unreplicable minor factors. Being extremely skeptical ourselves, we have systematically evaluated these possibilities, but found no evidence that they pose a problem for real ERP data, and, paradoxically, concluded that the improvement of major factors is due to the extraction and

[3] By pointing to our Fig. 5, Dien et al. (2005) imply that we incorrectly implemented the logic of the Scree test, that is, not looking for the "last bend counting from the right," suggesting that "more bends are present in the eigenvalues past that shown in the figure." Firstly, Dien et al. overlook the fact that this graph depicts the Eigenvalues *after* rotation, as stated in the figure caption, and can therefore not be used to evaluate the Scree criterion for the Eigenvalues *before* rotation. It should be noted that the slope and possibly the factor sequence change as a result of reassigning the variance through rotation. Secondly, selecting the 'appropriate' bend by moving from low- to high-variance factors necessarily requires a threshold criterion to decide when Eigenvalues increase beyond data noise levels, since irregularities in the slope of Eigenvalues (i.e. bends) are present at different scales throughout the range of extracted factors. Ironically, the ambiguous subjectivity of selecting among multiple and subtle bends when applying the Scree test has been acknowledged by the first author himself (Dien, 1998).

removal of noise or otherwise uninterpretable factors. In contrast to our pragmatic empirical approach, Dien et al. (2005) present no evidence to support their contention that overextraction degrades solutions for ERP data sets, despite the fact that this was evidently not a problem in their unrestricted solution using two simulated components, or in our unrestricted solutions for two real ERP data sets. Apart from the unrestricted solution, it remains to be seen whether other conservative retention criteria (i.e. any other number than four) would have affected the findings reported by Dien et al. (i.e. reconstruction of waveform time course and spatial topography; misallocation of the condition effect). In the absence of such evidence contesting our systematic findings of extremely variable *F* statistics for a wide range of conservative retention criteria, the conclusion that unrestricted solutions are unnecessary must be considered presumptuous. As their own pioneering work suggests otherwise, it is a final irony that Dien et al. (2005) prefer trusting in the norm of using restricted solutions, although their premise is to reform ERP-PCA conventions stemming from statisticians having primarily psychometric data in mind.

Our report did not evaluate the effects of factor rotation (i.e. Varimax was uniformly applied). However, Dien et al. (2005) take issue with our comment that caution must be exercised when conceptualizing ERP component measures that may or may not be linearly dependent. Temporal PCA reduces a set of ERP waveforms to a set of independent (orthogonal) factors. The Varimax procedure rotates the resulting factor loadings to produce a new set of waveforms that are also orthogonal, but have a simpler relationship to the underlying data (factor scores). Varimax rotation produces solutions that are generally more interpretable than the unrotated solutions, and has proven itself to be empirically useful for ERP data, while its limitations are well documented (e.g. Beauducel and Debener, 2003; Chapman and McCrary, 1995; Glaser and Ruchkin, 1976; Verleger and Möcks, 1987). Although real ERP components, as well as their underlying neurophysiological generators, are likely to be correlated, an orthogonal solution is parsimonious in that it preserves independence of variance contributions summarizing ERP waveforms. This is a desirable property (van Boxtel, 1998), although it may be at the expense of exact reconstruction. As demonstrated by Dien (1998) and Dien et al. (2005), an oblique rotation, such as Promax, which relaxes the orthogonality criterion, has the potential of a more accurate reconstruction of simulated component prototypes and their topographies produced by idealized dipole generators, for which factor correlation may be traced to a known origin (i.e. electrodes, subjects, experimental conditions). However, for real ERP data, the exact source of factor correlation is unknown, giving rise to the concern that a certain proportion of the variance is overrepresented, which may or may not affect the interpretability of the findings. This concern can be relieved by comparing orthogonal and oblique rotations performed on the same sets of real ERP data, showing when improved interpretability of the extracted factors based on psychophysiological knowledge and reason can be achieved (cf. Chapman and McCrary, 1995). Nevertheless, the mostly superior performance of Promax over Varimax for simulated data sets is exciting, and may ultimately lead us to break with the convention of using Varimax and instead embrace Promax as the method of choice for factor rotation.

We welcome the recent efforts to improve on PCA methodology as a classic tool for ERP analysis. If for nothing else, their common denominator is that ERP researchers using PCA are well-advised to not blindly accept established conventions. These may be encountered in the form of default settings in canned statistical software, or may come as traditional norms adopted out of context. Specific parametric choices may significantly influence the meaning, effectiveness, and value of PCA as a data-reduction tool. Some choices can be shown to be superior to others for a given purpose, and their preference should be determined by empirical insight rather than dogmatic principle. The challenge of the new empirical landscape demands nothing less. The reappraisal and recent advancements in PCA methodology are therefore encouraging, and promise to enrich the field of electrophysiology.

References

Achim A, Bouchard S. Toward a dynamic topographic components model. Electroencephalogr Clin Neurophysiol 1997;103(3):381–5.

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):103–12.

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):226–40.

Chapman RM, McCrary JW. EP component identification and measurement by principal components analysis. Brain Cogn 1995;27(3):288–310.

Chapman RM, McCrary JW, Chapman JA. Short-term memory: the 'storage' component of human brain responses predicts recall. Science 1978;202(4373):1211–4.

Collet W. Doubts on the adequacy of the principal component varimax analysis for the identification of event-related brain potential components: a commentary on Glaser and Ruchkin, and Donchin and Heffley. Biol Psychol 1989;28(2):163–72.

Dawson GD. A summation technique for the detection of small evoked potentials. Electroencephalogr Clin Neurophysiol 1954;6:65–84.

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

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

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):131–9.

Donchin E. On why Collet's doubts regarding the PCA are misplaced. Biol Psychol 1989;28(2):181–6.

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. pp. 555–72.

Donchin E, Heffley EF. The independence of the P300 and the CNV reviewed: a reply to Wastell. Biol Psychol 1979;9(3):177–88.

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

Field AS, Graupe D. Topographic component (parallel factor) analysis of multichannel evoked potentials: practical issues in trilinear spatiotemporal decomposition. Brain Topogr 1991;3(4):407–23.

Friedman D, Vaughan Jr HG, Erlenmeyer-Kimling L. Multiple late positive potentials in two visual discrimination tasks. Psychophysiology 1981; 18(6):635–49.

Gaillard AWK, Ritter WK. Tutorials in event-related potential research: endogenous components. Amsterdam: North-Holland Publishing Company; 1983.

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

Hendrickson AE, White PO. Promax: a quick method for rotation to oblique simple structure. Br 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):179–85.

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

John ER, Walker P, Cawood D, Rush M, Gehrmann J. Factor analysis of evoked potentials. Electroencephalogr Clin Neurophysiol 1973;34(1):33–43.

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

Kayser J, Tenke 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):414–26.

Kayser J, Tenke CE, Bruder GE. Dissociation of brain ERP topographies for tonal and phonetic oddball tasks. Psychophysiology 1998;35(5): 576–90.

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

Lehmann D, Skrandies W. Spatial analysis of evoked potentials in man—a review. Prog Neurobiol 1984;23(3):227–50.

Lobaugh NJ, West R, McIntosh AR. Spatiotemporal analysis of experimental differences in event-related potential data with partial least squares. Psychophysiology 2001;38(3):517–30.

Makeig S, Jung TP, Bell AJ, Ghahremani D, Sejnowski TJ. Blind separation of auditory event-related brain responses into independent components. Proc Natl Acad Sci USA 1997;94(20):10979–84.

Marshall WH, Woolsey CN, Bard P. Cortical representation of tactile sensibility as indicated by cortical potentials. Science 1937;85:388–90.

Marshall WH, Woolsey CN, Bard P. Observations on cortical somatic sensory mechanisms of cat and monkey. J Neurophysiol 1941;4:1–24.

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

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):482–4.

Möcks J. Doubting these doubts—a reply to Collet [comment]. Biol Psychol 1989;28(2):173–80.

Möcks J, Verleger R. Principal component analysis of event-related potentials: a note on misallocation of variance. Electroencephalogr Clin Neurophysiol 1986;65(5):393–8.

Molfese DL. Left hemisphere sensitivity to consonant sounds not displayed by the right hemisphere: electrophysiological correlates. Brain Lang 1984;22(1):109–27.

Oostenveld R, Praamstra P. The five percent electrode system for high-resolution EEG and ERP measurements. Clin Neurophysiol 2001; 112(4):713–9.

Picton TW, Bentin S, Berg P, Donchin E, Hillyard SA, Johnson Jr R, 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):127–52.

Pritchard WS. PCAVR: a portable laboratory program for performing varimax-rotated principal components analysis of event-related potentials. Brain Res Bull 1984;13(3):465–73.

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.

Ruchkin DS, Sutton S, Kietzman ML, Silver K. Slow wave and P300 in signal detection. Electroencephalogr Clin Neurophysiol 1980;50(1–2):35–47.

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):409–14.

Spencer KM, Dien J, Donchin E. Spatiotemporal analysis of the late ERP responses to deviant stimuli. Psychophysiology 2001;38(2):343–58.

Squires KC, Donchin E, Herning RI, McCarthy G. On the influence of task relevance and stimulus probability on event-related-potential components. Electroencephalogr Clin Neurophysiol 1977;42(1):1–14.

Tucker DM. Spatial sampling of head electrical fields: the geodesic sensor net. Electroencephalogr Clin Neurophysiol 1993;87:154–63.

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

Verleger R, Möcks J. Varimax may produce slow-wave-like shapes by merging monotonic trends with other components. J Psychophysiol 1987;1:265–70.

Wang K, Begleiter H, Porjesz B. Trilinear modeling of event-related potentials. Brain Topogr 2000;12(4):263–71.

Wastell DG. On the independence of P300 and the CNV: a short critique of the principal components analysis of Donchin et al. (1975). Biol Psychol 1979;9(3):171–88.

Wastell DG. PCA and varimax rotation: some comments on Rösler and Manzey. Biol Psychol 1981;13:27–9.

Wood CC, McCarthy G. Principal component analysis of event-related potentials: simulation studies demonstrate misallocation of variance across components. Electroencephalogr Clin Neurophysiol 1984;59(3):249–60.

Wood JM, Tataryn DJ, Gorsuch RL. Effects of under- and overextraction on principal axis factor analysis with varimax rotation. Psychol Methods 1996;1(4):354–65.

Jürgen Kayser[a,b,*]
Craig E. Tenke[a,b]
[a]*Department of Biopsychology,*
New York State Psychiatric Institute,
Unit 50, 1051 Riverside Drive,
New York, NY 10032, USA
[b]*Department of Psychiatry, College of Physicians &*
Surgeons of Columbia University, New York, USA
E-mail addresses: kayserj@pi.cpmc.columbia.edu,
tenkecr@pi.cpmc.columbia.edu

* Corresponding author. Address: Department of Biopsychology, New York State Psychiatric Institute, Unit 50, 1051 Riverside Drive, New York, NY 10032, USA. Tel.: +1 212 543 5169; fax: +1 212 543 6540.