BRAIN AND COGNITION 27, 288–310 (1995)

EP Component Identification and Measurement by Principal Components Analysis

ROBERT M. CHAPMAN AND JOHN W. MCCRARY

Center for Visual Science and Department of Psychology, University of Rochester

Between the acquisition of Evoked Potential (EP) data and their interpretation lies a major problem: What to measure? An approach to this kind of problem is outlined here in terms of Principal Components Analysis (PCA). An important second theme is that experimental manipulation is important to functional interpretation. It would be desirable to have a system of EP measurement with the following characteristics: (1) *represent* the data in a concise, *parsimonous way;* (2) *determine EP components* from the data without assuming in advance any particular waveforms for the components; (3) extract components which are *independent* of each other; (4) *measure* the amounts (contributions) of various components in observed EPs; (5) use measures that have greater *reliability* than measures at any single time point or peak; and (6) identify and measure conponents that *overlap* in time. PCA has these desirable characteristics. Simulations are illustrated. PCA's beauty also has some warts that are discussed. In addition to discussing the usual two-mode model of PCA, an extension of PCA to a three-mode model is described that provides separate parameters for (1) waveforms over time, (2) coefficients for spatial distribution, and (3) scores telling the amount of each component in each EP. PCA is compared with more traditional approaches. Some biophysical considerations are briefly discussed. Choices to be made in applying PCA are considered. Other issues include misallocation of variance, overlapping components, validation, and latency changes.

Between the acquisition of Evoked Potential (EP) data and their interpretation lies a major problem: What to measure? The data acquired often exhibit a richness that provides the opportunity for many different measures. The appearance of peaks and valleys tempts one to measure

Preparation of this article was supported in part by MH40703 and EY01319 from USPHS. We gratefully thank John A. Chapman for extensive computing help and Geoffrey Nowlis and the editors for helpful comments. Correspondence and reprint requests should be addressed to Robert M. Chapman, Center for Visual Science, University of Rochester, Rochester, NY 14627. E-mail: rmc@cvs.rochester.edu or to mccr@cvs.rochester.edu.

0278-2626/95 $6.00

their amplitudes and latencies. This may carry with it the implicit assumption that each such peak and valley represents a separate underlying process. Is such an assumption warranted and how can the reasonableness of this assumption be assessed? A possible approach to this kind of problem is outlined here in terms of Principal Components Analysis (PCA).

A second theme is intimately interwoven: experimental manipulation is important to functional interpretation. *Function* carries many meanings and operationally can be assessed by measuring changes related to experimental variables of interest. Thus, the function associated with some EP measure is dependent on the particular experimental variables whose variation produces changes in that EP measure. This points up the importance of skillful experimental design in order to sharpen the functional interpretation.

The EP data acquired are multivariate and are usually assumed to be only partly signal, the remainder being noise. Thus, one problem is to separate signal from noise. What if the EP data contain several signals, as would be expected of data obtained from sources as complex as the brain. This possibility leads to the problem of separating the various signals or, said another way, separating the various components of the signal. It would be desirable to identify and measure these components in each EP observation. One approach to this problem is to apply multivariate procedures to the set of observations. Here we discuss a particular multivariate procedure, PCA. Some uses of PCA, and its properties, are discussed in other places, e.g., Chapman, McCrary, Bragdon, and Chapman (1979), Donchin and Heffley (1978), Freeman (1987), Glaser and Ruchkin (1976), and Mocks (1988).

It would be desirable (Chapman, 1974; Chapman et al., 1979) to have a system of EP measurement with the following characteristics: (1) *represent* the data in a concise, *parsimonious* way; (2) *determine EP components* from the data without assuming in advance any particular waveforms for the components; (3) extract components which are *independent* of each other; (4) *measure* the amounts (contributions) of various components in observed EPs; (5) use measures that have greater *reliability* than measures at any single time point or peak; and (6) identify and measure components that *overlap* in time. PCA has these desirable characteristics. PCA's beauty also has some warts that we will discuss below.

GENERAL TWO-MODE MODEL: PRINCIPAL COMPONENTS ANALYSIS (PCA)

A general two-mode model may be formulated for EP data, $x(t, l, m, i)$ collected as measures at T time points from L electrodes for M condi-

tions from I subjects, as follows:

$$\underbrace{x(t, l, m, i)}_{\text{Data}} = \sum_{k=1}^{K} [\underbrace{a_k(l, m, i)}_{\text{Mode 1}} \underbrace{c_k(t)}_{\text{Mode 2}}] + \text{Grandmean}(t) + \text{error} \qquad \text{(Model 1)}$$

where K is number of components, $a_k(l, m, i)$ are scores telling the weight of the kth component at the lth electrode, mth condition, and ith subject, and $c_k(t)$ are component functions of time.

It is assumed that there are a few, K, underlying components each of which has a constant waveform, $c_k(t)$, that is different from the waveforms of the other components. These component waveforms retain their shapes but vary in amplitude, $a_k(l, m, i)$ from EP to EP (cases or observations). For each EP, then, each component is represented by multiplying its waveform at each time point, $c_k(t)$, by the amplitude value, $a_k(l, m, i)$. The amplitude-adjusted component waveforms are summed over the K components and added to the Grandmean waveform to reproduce the original EP (minus the error).

Principal components is sometimes referred to as a Karhunen-Loeve transformation or filter in the engineering literature. The latter method is for continuous functions rather than discrete sample values but the distinction is not critical in applications. Principal Components Analysis, Principal Factor Analysis, Eigen analysis, and Singular Value Decomposition produce essentially the same results.

If the component waveforms (basis functions) were known, then only their amplitudes would need to be estimated from the data. One approach would be to assume a particular set of component waveforms. Sine waves, for example, are assumed as component waveforms in Fourier analyses. However, many workers would not accept sine waves as being very physiological and usually the number of such components (K) needed to account for the same percentage of the data is much greater than Principal Components (PCs) need. An alternative to assuming particular component waveforms (basis functions) is to derive the underlying component waveforms from the data. This is quite a trick because both the waveforms, $c_k(t)$, and their amplitudes, $a_k(l, m, i)$ are to be estimated from the data set. Given so few constraints, the question of a unique solution arises. Often a rotation criterion (e.g., Varimax criterion) is applied to the PCA to achieve a unique solution that has simple structure. Other useful aspects of rotated PCA outputs are amount of data variance accounted for by each component and correlations between components.

PCAs correspond to an easily formulated least squares method and statistical methods are available for comparison of PCA across groups and between conditions within subjects (e.g., Guthrie, 1990).

COMPARISONS OF PCA WITH MORE TRADITIONAL APPROACHES

How appropriate is this Principal Components Analysis model? Nearly all would agree that EPs are made of components, and that these components may occur at different times with respect to a stimulus, overlap to varying extents and have a variety of waveforms and durations. Furthermore, it is generally accepted that the component amplitudes depend on experimental conditions and electrode placement. However, the particular components and their characteristics are more controversial, partly due to the kinds of measures used. Various distinctive features within the EP waveform, especially positive and negative peaks, are called components when they occur within some latency range, exhibit some regular pattern of amplitudes over the scalp, and relate to particular experimental factors. Much has been learned from this approach. To illustrate a problem with this approach, consider measuring the peaks or other features of the EPs in Fig. 1 (top). The early peak is easy to see, but what about the later one(s)? These EPs were simulated data made from random combinations of the three prototypes shown in Fig. 1 (middle). Confidently deciding whether there are two or three underlying components by only examining the EP data set (Fig. 1, top) is questionable and measuring each of them seems even more of a problem. Some limitations of this traditional approach are (1) the problem of identifying and agreeing upon the peaks and features to measure, (2) relatively few of the time points in the EP contribute to the measures, limiting measurement stability, (3) when increasing numbers of measures within EPs are made, there is increasing likelihood that the measures will be found to be intercorrelated and cannot be assumed to be independent, and (4) the possibility of component overlap suggests that peak or feature measures are not pure measures of a single underlying component.

PCA identifies components by a systematic approach that analyzes the variations in all the EPs in the analysis set. This analysis starts by computing the correlations (or covariances) of each of the time points with each of the other time points (across all EPs in the data set). The structure is computed from these correlations, with the key idea that variables (time points) that are correlated belong to the same underlying component. Often when applied to EP data the time points whose values are correlated tend to be within limited poststimulus time zones leading the PCA to result in components whose loadings are high in different time zones. However, there is nothing inherent in the PCA procedure that encourages this local temporal result; it is the consistencies in the EP data that are local temporally. A different outcome that occurs for some components (e.g. CNV, late slow-wave, etc.) is a much broader time zone of relatedness and the PCA results show correspondingly long-lasting

components. It is conceivable that a single component could exhibit an early and a late part with nothing between; any waveform is possible. The useful idea pursued by PCA is that, from one EP to another, parts that wax and wane together are taken as evidence that they belong to the same component. A part of the EP may be partly correlated with two other parts of the EP that are not correlated with each other; this can result in principal components that overlap each other in time. What about a part of the brain activity that one thinks should be considered a separate component but is always present with the same amplitude in every EP in the data set being analyzed? PCA cannot find a "component" that does not vary from EP to EP; it's waveform contributes to the grandmean. Extending this idea, only the amount of a brain component that varies will be measured as a PCA component, the constant part of that brain component will contribute to the grandmean of all EPs. Thus, *PCA looks where the action is to find components, rather than looking for peaks.*

This measurement idea nicely fits the scientific approach of experimentally manipulating conditions and measuring what effects occur. If experimental conditions were simply held constant for a set of EPs from one electrode from one subject, only trial-to-trial variations can be analyzed for components. Perhaps some of the same components that would have been activated by different experimental conditions will be "spontaneously" activated and found by PCA, but interpreting their functional significance is difficult without any evidence of what events contributed to a component's activation in a given trial (measured by component score). Therefore, attempting to influence the activity of the EP components by varying experimental conditions provides a chance to attach functional interpretations to the components. Some kinds of variables that influence the activity of the components are experimental conditions, electrode positions, subjects, and natural (unknown) variation. All, or some, of these variables, may be used to improve reliability and determinacy of the measurements. One can choose which variables to be studied by a PCA by what is allowed to vary within the set of EPs submitted for the analysis.

The discussion above indicates that in contrast to peak measurement: (1) PCA identifies what to measure by a systematic approach, (2) all the time points in the EP contribute to the PCA measures, enhancing measurement stability, (3) the number of components and their independence are an integral part of the PCA approach, and (4) overlapping components are separated allowing purer measures of each underlying component (although not to everyone's satisfaction, see below). An example below will show poor recovery of prototype variation by peak measures, when the recovery by PCA was quite good.

SOME BIOPHYSICAL CONSIDERATIONS AND PCA

Some biophysical considerations are briefly discussed next (see also Mocks, 1988b). The potential at a scalp location at a certain instant of time is due to a number of current sources and sinks located somewhere in the brain (e.g., Nunez, 1981). The complicated relationships between the internal generators and the surface measurements that depend on geometric and electric properties of the skull and tissue layers are not relevant for the present purpose. Changes in generator current are proportionately related to changes in surface voltage with the proportionality coefficients remaining constant at each electrode provided that these geometric and electric properties remain fixed. The same generator viewed from a different electrode may be represented by a different proportionality coefficient while its waveform retains the *same* shape; this fits the PCA model well in that $a_k(l, m, i)$ of Model 1 depends on electrode, l, whereas $c_k(t)$ does not. Furthermore, the effects of different generators seen at an electrode may be said to simply sum at the electrode. This, too, fits the PCA model in summing different components.

PCA: LETTING THE DATA TALK

Let us consider an experiment in which EP observations have been obtained in response to a stimulus. Evoked Potentials (EPs) obtained in this kind of experiment are multivariate observations. Each EP might contain measures at 200 time points. In any collection of such EPs, the parameters would involve not only the 200 means and 200 variances of the amplitudes, but also the 19,900 covariances among the measures at different time points. Techniques may be used that incorporate these important parameters into the analyses. These tallies become even larger for EPs based on more than 200 time points.

Varimaxed PCA may be used to identify a component structure for the EPs and to obtain measures of the components. In many sets of data a relatively small number of component measures, perhaps ten, do a reasonably good job of representing the large number of measures, perhaps several hundred, obtained during data acquisition. In fact, PCA is the most effective linear method of data reduction, accounting for the most variance with the fewest number of components.

Although PCA, like any other technique, is not without problems, we find it quite useful in EP research, where the assumption of linear summation of components is consonant with volume conduction considerations (Helmholtz's principle of superposition). Surface voltages that are produced by different current sources summate linearly. We are trying to find nature's seams so that the underlying pieces can be used to describe and understand brain functions in this complex machine.

CHOICES TO BE MADE IN APPLYING PCA

There are technical choices to be made in applying any technique, and often biases in their selection. In the case of PCA, these include, among others, the time epoch, the association matrix, the criterion for the number of components to be rotated, and the rotation criterion.

The time epoch is limited by practical constraints. The most serious of such constraints are the number of time points (variables) in relation to the number of EPs (observations or cases). The number of variables should be less than the number of observations; the smaller of these two numbers is the mathematical upper limit to the number of principal components. Furthermore, the number of variables may be limited by the particular computer program. Although most modern programs can compute a PCA with fewer variables than observations, the solution is more stable when more observations contribute to the PCA. This epoch, and the time samples within it, may be adjusted depending upon the degree of interest in, or expected relevance of, early and later components. It is not necessary that each EP be obtained under different experimental conditions nor that all EPs be obtained from the same experimental condition. In the latter case, however, the components could not be interpreted in terms of experimental variations, because no experimental variations would have been used.

An association matrix must be selected. Typically, we transform the data matrix into the standardized variance-covariance (correlation) matrix. Principal components of the correlation matrix are invariant under separate scaling of the original variables and those of the covariance matrix are not (Gnanadesikan, 1977). When all variables are measured in the same units, some researchers prefer the covariance matrix (e.g., Donchin & Heffley, 1978). However, the method is extremely dependent upon the total variance of the original variables. Harman's latest revision (1976) states again, "it is customary to express the variables in standard form, i.e., select the unit of measurement for each variable so that its sample variance is one. Then the analysis is made on the correlation matrix, with the total variance equal to n (the number of variances). For such a matrix (symmetric, positive definite) all n principal components are real and positive" (p. 134). This "customariness" is often simply taken for granted by many specialists (e.g., Cooley & Lohnes, 1971; Mulaik, 1972). Rummel (1970) points out that almost all factor analyses (not merely the specific PCA) have started with a correlation matrix, but cites Sir Cyril Burt (1941) in particular as having pointed out some advantages of the covariance matrix. We have frequently compared PCA results from covariance and correlation matrices and can find no reason for vehement preferences. When, as is often the case, analytic results are desired to be expressed in the original microvolt metric, the restoration of

the metric to the components based on the correlation matrix is straightforward (first described in Chapman et al., 1979).

The waveforms of the components can be displayed in the original metric by the following procedure when the correlation matrix is used as the association matrix (and when the output loadings relate to the standardized variables, even though the covariance matrix is factored, as some programs do). The loading for each point is multiplied by the standard deviation for that time point, where the standard deviations are based on the set of EPs submitted to the PCA. In matrix notation, the matrix of loadings is multiplied by a diagonal matrix of the standard deviations.

Number of components to be rotated. The number of components is K in Model 1. Theoretically, K could be as large as the number of time points, T, (often in the hundreds). For EP data K is typically much less, perhaps in the neighborhood of 10, but it ought to depend on experimental conditions, electrodes, etc. In practice there are several possible approaches to reaching this important decision. Rummel (1970) presents an unusually comprehensive review. Since PCA extracts the components in the order of the amount of variance they contribute (indicated by their eigenvalues), from most to least, the components not used account for relatively little of the data. This takes some of the sting out of deciding exactly where to make the cut, although it is possible that a component is important in relation to experimental conditions while being small in size. Those components not retained may be considered to be in the error term of Model 1. We generally use the widely employed criterion originally proposed by Kaiser (1960). This is often referred to as the eigenvalue = one rule: retain and rotate only those components which have associated eigenvalues equal to or greater than one (or equal to or greater than the average variance of the original variables when the covariance matrix is factored). This recommendation has been supported by convergence of algebraic derivations, psychometric reliability, and interpretability. It also has the intuitive appeal and plausibility of excluding components not accounting for at least the total variance of a single variable. We use this rule in initial computations and then apply the scree test (Cattell, 1966) to confirm or question whether the appropriate number has been selected. Where the rules disagree both are used, and the desired rotated result selected on the basis of judgments of compacted components (e.g., too few variables in the hyperplane) and fissuring of components (e.g., too few substantial loadings). Velicer (1976) has shown that the average of squared partial correlations decreases as components are extracted until a "unique" or "specific" component (high loadings on only one original variable, i.e., time point) is partialed out, then, the average begins to increase. Under Velicer's Minimum Average Partial

(MAP) method, no more components are retained after the average has reached its minimum. This rule bears an obvious relationship to the intuitive concept of parsimony involved in the eigenvalue = one rule. Simulation studies (Zwick & Velicer, 1982, 1986) have examined alternative stopping rules.

Rotation criterion. The purpose of the rotation is usually to obtain simpler interpretations of the components. The problem is that there is a basic indeterminacy in factor analysis—an infinitude of factorizations of a correlation matrix may account for the observed data equally well. Which of these makes the best sense from whatever else we know about the research topic? Often a Varimax criterion is applied to the PCA to achieve a unique solution that has simple structure and retains orthogonality of component scores. The simplicity of a component is defined as the variance of its squared loadings (the loadings in Model 1 are functions of time, $c_k(t)$) and the Varimax criterion maximizes this loading variance for all the components retained. Varimax finds components whose loadings tend to be large or small, not intermediate. For the EP application this is saying that the components tend to be either large or small at any time point and this tends to minimize overlap of different components. This idea leans toward the idea of measuring EPs by amplitudes or areas, which carries the implicit assumption that only a single underlying component is being measured. Measuring the average amplitude within some time zone of the EP is similar to a component that has loadings of 1.0 at time points within that time zone and loadings of 0.0 at all other time points. The Varimaxed PCA components are not typically as extreme, in that more than one component may contribute to any time point. Because EP data are a time series where adjacent time points tend to be more correlated, the PCA tends to group them together in identifying components. Since a major EP task is separating this time series into meaningful components it seems reasonable, especially at early stages of understanding, to segment this time record into as separate time pieces as possible while considering all the data. The Varimax rotation, it seems to us, is likely to do this. Following this logic leads us to suggest that Varimax tends to give narrow waveforms to the components; perhaps we could think of Varimax (and even more so oblique rotations) as providing a kind of lower limit estimate to the width of the true component waveforms that may have more overlap and be more spread out over time. However, the Varimax PCA is not looking for one-peak components (as might be misread in Mocks and Verleger, 1986); if the data had underlying components of multiple peaks that waxed and waned together, then the Varimax PCA would find this multiple-peak component (an example is the unmistakably sinusoidal component related to power line frequency found by Arbuckle, 1970). Model 1 applies equally well before or after rotation as a general

model; only the values of the parameters may change. With a new set of rotated component functions of time, $c_k(t)$, there will be a corresponding new set of component scores, $a_k(l, m, i)$ and a corresponding new set of variances accounted for by each component, although the total variance for the K components remains constant. After orthogonal rotations, such as Varimax, the component scores, $a_k(l, m, i)$, for the various components remain uncorrelated, while the component functions of time may be correlated. A further relaxation of criteria may be obtained in oblique rotations that allow the component scores to be correlated. Oblique rotations may result in "simpler" loadings than those for an orthogonal rotation. With oblique rotation there is a greater tendency for each variable (time point) to be associated with a single component. The method of oblique rotation generally recommended is the *direct quartimin* (Jennrich & Sampson, 1966). An example is briefly discussed below. An important objective of the rotational problem is to achieve factorial invariance. The Varimax criterion tends to lead to factorial invariance (Harman, 1976, p. 298). Whatever the rotation decisions, the results are simply linear transformations of the data. PCA does not create effects which are not in the data.

Another kind of rotation criterion could be based on other assumed characteristics, for example scalp distributions that are related to particular neural generators (e.g., Maier, Dagnelie, Spekreijse, & Van Dijk, 1987; Vokey, 1989).

WARTS: PRINCIPAL COMPONENTS ANALYSIS OF EPS IS NOT WITHOUT PROBLEMS OR CRITICS

The plausibility of the mathematical assumptions behind PCA of EPs was questioned by Hunt (1985), who analyzed some simulations. He concluded that his results "indicate that PCA is a surprisingly robust technique for estimating component waveform shapes and the relative size of component forms, across records, even when some of the assumptions are not met."

An influential example of a critical analysis of PCA use with EPs is the simulation study by Wood and McCarthy (1984). Their objective was to investigate the ability of PCA, Varimax rotation, and univariate ANOVA (a) to reconstruct component wave shapes, (b) to allocate variance correctly across components, and (c) to identify the component to which a simulated treatment effect was applied. Simulated EPs were constructed using three 64-point prototype components (Fig. 1, middle panel) "not intended to simulate any particular set of EP results, but were designed to capture in a general way some of the main features of the EP components discussed in recent experiments." Each simulation set of EPs consisted of randomly weighted combinations of these prototypes, corresponding

to a 2 × 2 × 10 factorial repeated-measures design with 20 subjects. We illustrate these simulations with our own replication of Wood and McCarthy (1984) in Fig. 1. Random weights were selected for each component from normal distributions with standard deviations of 50 and means of 100, except for one component for which a main effect was introduced by treatment levels with means of 100 and 200. In addition, a small additive noise term was chosen independently for each of the 64 time points from a normal distribution with a mean of zero and a standard deviation of 2.

Samples of our simulated EP data are shown in the top panel (Fig. 1) (only 6, randomly selected from the 800 EPs in that data set are shown). Covariance PCAs, Varimax rotation and univariate ANOVAs were calculated for several hundred simulations to observe outcomes with (a) no input effects on any component for any treatment and (b) with an input treatment effect for each of the prototypes, one at a time. Wood and McCarthy showed and concluded that the "wave shapes of the simulated components were reconstructed reasonably well, although not completely, by the rotated principal component (PC) loadings." This agrees with our experience, a sample of which is in Fig. 1 (bottom). Two of the prototype components overlapped and a quantitative comparison of rotated PC scores with the random weights used to generate the simulated EPs showed that PCAs "misallocated" some of the variance across these overlaps.This was found to have dramatic consequences on the ANOVAs' assessment of treatment effects. With *no* simulated treatment effect on any component, F ratios with $p < .05$ closely approximated the 5% type I error rate. When treatment effects were introduced on either of the overlapping components, type I error rates for ANOVAs on the other component were inflated to levels of 70 to 80%. Wood and McCarthy cautioned against overgeneralization of the results, but the outcomes, as presented, illustrated that the PCA–Varimax–ANOVA strategy *can* result in misinterpretation of treatment effects, especially if the analysis is on a component-by-component univariate basis. The authors are careful to point out that the misallocation of variance problem is not peculiar to PCA: "Other approaches to ERP analysis, measurement of peak amplitudes and latencies for example, are no less subject to the problem of component overlap than PCA; they simply make it easier to ignore by not representing it explicitly. Misallocation of variances and misinterpretation of experimental effects are just as possible using such techniques as they are with PCA" (p. 258). This paper has been the subject of frequent discussion and, unfortunately, some informal summaries have lost the temperate and qualified aspects of the authors' conclusions.

The experience with PCA in our laboratory led us to believe that additional simulations might aid understanding the basis for this "misalloca-

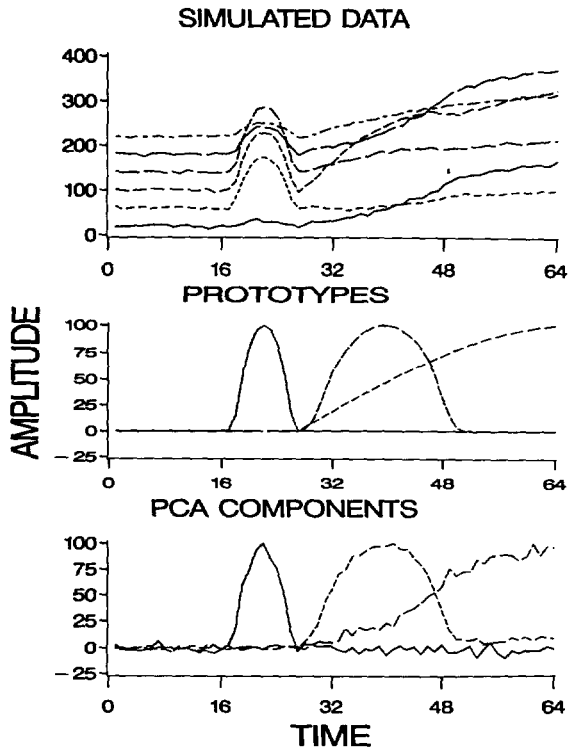

FIG. 1. Principal Components Analysis (PCA) of simulated data. The pseudo-EPs are samples of our simulated EP data (top) made from random combinations of the three prototypes (middle) used by Wood and McCarthy (1984). The six EPs are arbitrarily displaced vertically for viewing here. Varimaxed PCA from a simulated data set of 800 EPs (like those in the top panel) extracted three components whose loadings, $c_k(t)$, are plotted (bottom). These PCA component functions of time reconstruct the prototypes reasonably well, although not completely. Prototypes 1 (solid line), 2 (broad-dash line), and 3 (dash line) had peak amplitudes at early, middle, and late times, respectively.

tion" of variance. The validity of the orthogonal Varimax rotation assumes that the underlying component structure is orthogonal. If this assumption is not at least approximated, then the Varimax procedure is limited in recovering the original prototypes. We wished to check this with the same three prototype waveforms used by Wood and McCarthy (graciously provided by them).

Using these and other prototypes, our simulations and those of Mocks and Verleger (1986) have provided some insights into the "misallocation" of variance problem.

"MISALLOCATION" OF VARIANCE

Using the Wood and McCarthy prototypes and procedures in simulations with the PCA–Varimax–ANOVA approach, we obtained essentially the same outcomes. In order to assess whether the assumption of orthogonality is appropriate for the underlying structure, an oblique rotation—usually direct quartimin or "rotation to simple loadings" (Jennrich and Sampson, 1966)—can be employed to check it. Using a PCA-direct quartimin approach with simulations using the Wood–McCarthy prototypes produced components which correlated about +.22. The "misallocation" of variance may be in part due to underlying non-orthogonality.

Under those circumstances where variance misallocation or "leakage" has in fact occurred, the problem of increased type I error rates exists only when the research interest is focused upon the results of univariate analyses of individual components. Even in these cases, there are ways to identify and attempt to cope with the problem.

In the absence of other knowledge about the structure underlying the data, routine inspection of the results of both the orthogonal and the oblique rotations can be helpful in providing information about (1) the degree to which the assumption of orthogonality is approximated by the underlying component structure and (2) which, if any, components may likely have variance "shared" with other components.

If univariate ANOVAs have been computed separately for each of the components, then the overall pattern of the complete set of test results can be examined for indications of variance misallocation or "leakage." If two or more components have the same pattern of ANOVA outcomes and the wave shapes of the components overlap, then the occurrence of variance misallocation, "sharing" or "leakage" can be considered as one of the possible explanations of the similarity. This explanation becomes more probable if the ANOVA effect sizes are very large for one of these components compared to the other(s). Such differences in magnitudes of effects where the ANOVA significance patterns are the same may be used to identify the component whose variance may be partly misallocated to another component and to judge the implications for risking type I error rates if univariate ANOVA results are interpreted without qualification and additional control analyses.

The risk is primarily associated with uncritical acceptance of ANOVA outcomes for the component(s) with the smaller effect sizes. The ANOVA outcomes for the component with the much larger effect sizes, restricting the comparison to the same ANOVA effect, are valid. In our replication of the Wood and McCarthy simulations, we discovered these differences in effect sizes were always quite large; the F for the real component was often 20 times larger than the F for the component dis-

playing misallocation. Comparison of F values is reasonable here because the degrees of freedom are always the same when the same ANOVA design is used with different components. A special case where interpretation is simple occurs *when only one component has a significant, specific, ANOVA effect:* such an outcome is clearly valid.

Where the research questions are more concerned with EP effects in general than with single components, analytic approaches can be used which circumvent the problem completely. Under these circumstances, PCA is used primarily as an EP measurement system. For example, the scores for all components might be used as the dependent variables in a Multivariate Analysis of Variance (MANOVA). This procedure could determine whether there are statistically reliable differences among linear combinations of these scores as a function of experimental manipulations without encountering threats to inference due to misallocations of variance.

If, as in some of our research on semantic processing (Chapman, 1979; Chapman, McCrary, Chapman, & Bragdon, 1978; Chapman, McCrary, Chapman, & Martin, 1980), Discriminant Analysis is used, possible problems due to variance misallocation are also circumvented. The technique is used to develop linear combinations of the PCs whose scores best reveal differences among EPs separated into groups according to an appropriate criterion (different experimental treatments, different types of eliciting stimuli, etc.). These combinations of PC scores (discriminant functions) then serve as the basis for classifying EPs into these groups. The resulting classification accuracy (proportion of correct classifications) is easily evaluated on a preliminary basis for apparent magnitude and reliability. Cross-validation of the discriminant functions is always important to avoid capitalizing on chance. This is carried out by applying the functions to EPs not used in the computation of them (either a "hold-out" group or subset of the EPs or a newly collected set of EPs) and statistically evaluating the resulting classification success rates obtained with the "new" EPs. Reaching decisions about differences among the EPs based on such cross-validation is added protection against a more generally encountered source of type I error.

OVERLAPPING COMPONENTS NEED NOT RESULT IN MISALLOCATION OF VARIANCE

We performed additional simulations directed at disentangling the matter of component *overlap* from the matter of component *correlations* in "misallocation" of variance across components.

We prepared a set of three 64-point prototype components that approximated orthogonal simple structure while retaining as much of the form and appearance of the Wood-McCarthy prototypes as possible. Plots of

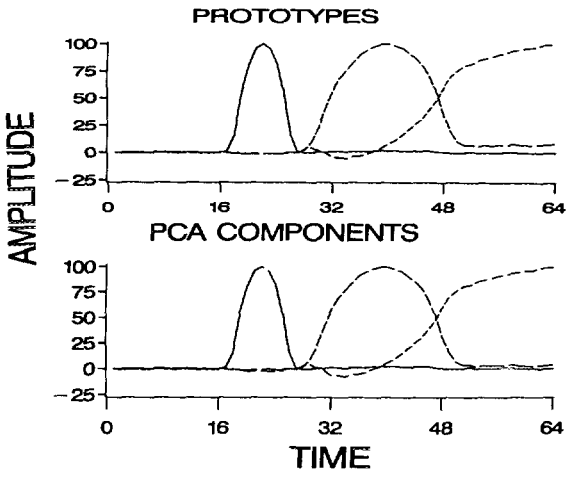

FIG. 2. PCA recovering particular prototypes. Three new, orthogonalized prototypes (top) were used to construct simulation sets of EPs. Varimax rotated PC loadings from a representative example of 100 PCA simulations (bottom). The orthogonal PCA loadings (bottom) correspond closely to the waveforms of the prototypes (top). In these simulations, no leakage of variance among components was found. Prototypes 1′ (solid line), 2′ (broad-dashed line) and 3′ (dash line) had peak amplitudes at early, middle, and late times, respectively.

these components are presented in Fig. 2 (top). The second and third prototypes overlap about as extensively as their Wood–McCarthy counterparts. These prototypes were used with the Wood and McCarthy procedures to construct simulation sets of EPs.

In the first set of 100 simulations no treatment effect was applied to any of the three components. Fig. 2 (bottom) displays plots of the Varimax rotated PC loadings from a representative example of these 100 simulations. The rotated PCs can be seen to correspond closely to the wave shapes of the prototypes.

An additional 100 simulations were evaluated to test for leakage among components. In these, we used the procedures of Wood and McCarthy to apply a single main treatment effect of their same large size on one of our overlapping components. The effect was applied to Prototype 3′ because it most closely resembles their Prototype 3 that resulted in their "worst case" leakage of variance into the overlapping second component to produce the 81% level of type I error rates. No treatment effect was applied to either the first or the second component. The PCA–Varimax–ANOVA strategy was applied in these 100 simulations. All 100 ANOVAs on PC scores for component 3 were significant at $p < .05$. Neither of the other two components—including the completely overlapping component

2—showed any indication whatsoever of any "misallocation" of variance. Type I error rates for these components were in both cases less than 5%. Thus, there was *no leakage among components.*

This may be contrasted with the *leakage obtained by peak measures* on the same simulation set. For example, a peak measure of the response maximum near time 40 (where the second prototype had its maximum) received sufficient leakage from the third prototype (the one to which the effect was actually applied) to produce significant F values ($p < .05$) in all 100 ANOVAs. This poor performance of peak measures suggests that they are not suitable for validating PCA results, as some would suggest (see below).

With these underlying components, two of them overlapping considerably, the PCA–Varimax–ANOVA strategy resulted in excellent recovery of the component wave forms, no misallocation of variance among components and no inflated type I error levels.

Wood and McCarthy (1984) and Mocks and Verleger (1986) do not propose that (a) these issues prohibit applying PCA followed by some rotation, nor (b) that methods like measuring peaks or areas should be used instead. The peak and area methods assume that components do not overlap (Donchin and Heffley, 1978). If there is indeed no overlap, then PCA, Varimax rotated or not, provides the correct solution (Mocks & Verleger, 1986). However, if there is some overlap, the simpler peak or area approaches inevitably render biased measures, i.e., some kind of "misallocation." In contrast, PCA does allow for overlap and shows whether it is present at all, but has difficulties in that it cannot resolve overlap in a way that necessarily agrees with the true prototypes.

The overall ability of PCA to parsimoniously disentangle components has not yet been equaled by any other procedure. The Varimaxed Principal Components method has performed well in achieving maximally parsimonious descriptions of a wide variety of actual data sets obtained from differing scientific areas (Chapman, McCrary, & Tuttle, 1981; Thorndike & Weiss, 1970) where other factor analytic methods sometimes fail.

The importance of the rotation after PCA was emphasized by Mocks and Verleger (1986) who further extended analyses of the Wood and McCarthy simulations. Mocks and Verleger provided evidence that the "misallocation of variance" does not originate from the component extraction step used by PCA. This was demonstrated by the fact that a rotation could be defined which perfectly reconstructed the set of prototypes. However, it is not possible to find this rotation in a real data situation, they indicated, since it requires knowledge of the true prototypes, which, is available in simulations only. To the extent that the true prototypes satisfy the Varimax criterion they will be recontructed by PCA Varimax procedures (see Fig. 2). Furthermore, if there is no overlap between true components, then PCA, Varimax rotated or not, provides

the correct solution (see prototype 1 in Fig. 1). They wrote, "The very best to expect from a procedure like PCA is that it comprises the information in a few components, without considerable loss or distortion, and, when applying some rotation criterion, that it ends up with rotated versions of the underlying true brain potentials." The situation might be improved by formulating more definite data models by using more information or assumptions (e.g., see Three-Mode Models below). Mocks and Verleger concluded that "PCA does not distort or lose information when extracting components" and "In sum, it appears to us that despite its difficulties PCA may be applied to advantage on ERP data."

VALIDATION OF COMPONENTS

The acceptance of PCA as a measurement method for EPs can be attributed, at least in part, to some Principal Components looking similar to what researchers expected from simpler measurement approaches such as viewing EPs, measuring peaks, subtracting one EP from another, etc. For example, in an early application of PCA (Chapman, 1974; Chapman, McCrary, Bragdon, and Chapman, 1979) Component 1 looked (temporal waveform obtained from loadings) *and* behaved (amplitude, obtained from component scores, changes with experimental conditions) like CNV (Grey Walter, Cooper, Aldridge, McCallum, & Winter, 1964) and Component 2 looked *and* behaved like P300 (Chapman & Bragdon, 1964; Chapman, 1965). In a sense, the PCA results were validated against other measures of EPs (Chapman, 1973). Such comparisons with other measures seem useful, especially at early stages of using a new method. Some workers suggest that PCA results should *always* be validated by other EP measures. Before subscribing to such a mandate, further elaboration of the implications of such comparisons may be helpful. When agreement between PCA and other EP measures is found, then there is no problem. However, when agreement is not found, which measure is to be taken as the criterion for validating the other? Some investigators want to "see" the component in the EP traces and imply that visual inspection is more valid than PCA. First, it is not always appreciated that PCA does not invent data, but is a view of the data. Summing the components with the proper weights (all specified in the PCA results) can reproduce *any* of the EPs (e.g., Chapman et al., 1979, Fig. 2). Thus, it is theoretically possible to "see" any PC in the EPs after one knows where to look. They need not appear necessarily as peaks in the EPs, but may appear as shoulders or thickenings of peaks or as longer lasting raises or falls, etc. This may be a fun exercise, but can not help validate the PC because invalidation is not possible; not "seeing" a PC may be the problem of the viewer and the difficulty of the viewing task. Peak measurements as validation criteria are flawed if overlap of components is

permitted theoretically. Second, is a logical point: if passing an "other-measure" test is required in order to accept a PC, then why not simply use the "other-measure"? The other side of this coin suggests that the value of PCA may lie not in those PC's that are similar to other measures, but rather in those PC's that are not so similar and thus more difficult to validate by the "other-measure" unless the "other-measure" is, in fact, very similar to the PC. Although the method of validation by the "other-measure" for some of the components was useful in early applications of PCA to EPs because it lent credence to the method, requiring such validation for "new" components, we suggest, is not as useful now. Other forms of validation may be more useful, e.g., comparisons to non-EP measures both behavioral and physiological, as well as assessing the mapping of the component scores onto the experimental conditions (discussed above).

IDENTIFICATION OF LATENCY CHANGES BY PCA

If there is variability in the time domain within EPs then there will be variability in amplitudes at the various EP time points beause of the shifting back and forth of EP features. Since PCA is based on amplitude measures, the method accepts latency variability as amplitude variability. Donchin and Heffley (1978) appropriately caution that "results should be interpreted with great care whenever latency variability is substantial across ERPs" (p. 567). If there are consistent variations of sufficient size to permit resolution within the EP sampling rate, the PCA can produce a component which represents the latency shift. Unfortunately, this prospect has led some critics to assert that variability in latency is a major reason why PCA yields misleading component structure. Careful examination of the data can bring insights about such effects. Donchin and Heffley (1978) have illustrated how this representation occurs. For a specific example, systematically shifting a constant-amplitude component, they computed the PCA and displayed the resulting biphasic component. Such a component may be called a "latency-adjusment component." The name identifies its "function" in the PC structure, i.e., an adjustment to a component's latency that may be so identified, related to specific experimental manipulations, and subjected to further analysis when appropriate. If the identification and analysis of such latency-adjustment components is viewed as undesirable or irrelevant to an overall analysis of experimental outcomes, then an alternative procedure might be used to exclude them. As suggested by Donchin and Heffley, an adaptive filtering algorithm proposed by Woody (1967) (see also Steeger, Hermann, & Spreng, 1983; Kramer & Donchin, 1987; Ruchkin, Sutton, & Stega, 1980; Mocks, Kohler, Gasser, & Pham, 1988) might be applied to adjust component latencies prior to PCA. Presumably the time shift indicated for a

trial by this filtering algorithm would be applied to all channels on that trial. Woody filtering generally tends to focus on a single component with less concern for other components which may be shifted as a consequence.

DON'T BE A WART HOG: SHARE THE TROUBLE

All techniques have problems. They are easier to see in simulation studies where the "truth" is known. In most of these simulation studies the distortions from this "truth" were small (e.g., 6–10% of total variance in Wood and McCarthy, 1984), even acceptable by the National Bureau of Standards for some of their physical measures. It is well recognized that even a small amount of biased error can regularly pass the decision rule of inferential statistical tests. This can explain how principal components whose waveforms are just a little off from the prototypes can make non-linear decision jumps to many "wrong" rejections of null hypotheses. Other measures also are subject to the same problem. In choosing measures, it may be helpful to have comparisons under the same conditions (e.g. see comparison of PCA and amplitude measures above in which the PCA measures were considerably closer to the "truth").

Carroll (1992) said: "Although it has sometimes been fashionable to criticize factor analysis as being a technique with many problems—for example, the problem of the number of factors to extract, the problem of whether principal components or principal factors should be extracted (Velicer & Jackson, 1990), and the problem of indeterminate rotations to simple structure—I am persuaded that these problems can be and have been satisfactorily resolved to the point that at least for well-designed sets of variables, reasonable and replicable solutions can confidently be arrived at (Carroll, 1985)."

AN EXTENSION OF PCA: THREE-MODE MODELS

Extensions of the PCA approach have been proposed, some of which seem to hold special promise. One of these will be summarized here. Mocks (1988a) suggested that attempting to treat three-mode data, $x(t, l, i)$, by some two-mode decomposition (e.g. Model 1 above) is bound to emphasize some aspect more than another. He suggested that a component be defined by two properties: a fixed time course and a topographic pattern independent of time. Neither of these two properties needs to be known in advance of the analysis that searches for them in a data set by estimating their parameters. He suggested a trilinear decomposition of EPs through a set of component functions, $c_k(t)$, with associated electrode coefficients, $b_k(l)$, both common to all subjects and also a set of scores, $a_k(m, i)$, telling the weight of the kth component in the mth condition for

the ith subject. Mock's Topographic Components Model (TCM), slightly enlarged to include experimental conditions, is

$$\underbrace{x(t, l, m, i)}_{\text{Data}} = \sum_{k=1}^{K} [\underbrace{a_k(m, i)}_{\text{Mode 1}} \quad \underbrace{b_k(l)}_{\text{Mode 2}} \quad \underbrace{c_k(t)}_{\text{Mode 3}}] + \text{Grandmean}(t) + \text{error} \qquad \text{(Model 2)}$$

where K is number of components, $a_k(m, i)$ are scores telling the weight of the kth component in the ith subject and mth condition, $b_k(l)$ are coefficients of strength at the lth electrode, and $c_k(t)$ are component functions of time.

Model 2 has three ingredients (scores, electrode coefficients, and loadings), whereas Model 1 has two (scores and loadings). Although Model 2 may look more complicated, it is simpler in that fewer parameters (unknowns) are involved (due to combinations of 3 things, rather than 2 things). One advantage is that Model 2 has unique solutions for a fixed number of components, K. This has the favorable consequence of not needing to choose a rotation as with Model 1. More importantly, the assumptions needed are rather weak (see also Kruskal, 1976, 1977): (1) The K component functions differ in their pattern, and the same holds for the K vectors of scores (i.e., the matrices $(c_k(t))$ and $(a_k(m, i))$ have full rank K); (2) the topographical distributions pertaining to any two of the components are not the same (i.e., any two of the L-vectors $b_k(l)$ are not collinear). These assumptions are less restrictive than in PCA. Mocks noted that there is nothing prescribed about orthogonality or the like. This means that scores of two components can correlate with each other and that these correlations are determined as well. As for the component functions it means that there may occur almost any overlap in time, without affecting the unique decomposition. A three-mode model like Model 2 (Mocks, 1988a, 1988b; Pham & Mocks, 1992) has been considered before in other frameworks (Harshman, 1970; Harshman & Lundy, 1984; Carroll & Chang, 1970) under the names PARAFAC (parallel factors analysis) and CANDECOMP (canonical decomposition). However, Model 2 differs from the better known model of three-mode factor analysis (Tucker, 1966). The three-mode Model 2 can be extended towards a four-mode analysis that separates parameters for experimental conditions and subjects (Mocks, 1988a). It is also possible to generalize it by including individual latency parameters of the component functions, along the lines of Mocks (1986).

The favorable properties of the new model have their costs in terms of problems on the algorithmic side. More experience with these problems seems worthwhile and could lead to substantial methodological progress. Aside from the mathematical virtues, of greatest importance is whether the results make sense for the research problem being investigated. Here

again, we suggest interpreting the components by relating their measures, the component scores, $a_k(m, i)$, to experimental conditions. The other parts of the model provide parameters that estimate the time courses of the components, $c_k(t)$, and the scalp distributions of the components, $b_k(l)$, both of which lead to physiological predictions.

REFERENCES

Arbuckle, J. L. (1970). Multivariate analysis with stationary error: An application to auditory evoked potentials. *Educational Testing Service Research Bulletin*, RB-70-36.

Burt, C. L. (1941). *The factors of the mind: An introduction to factor-analysis in psychology*. Macmillan: NY.

Carroll, J. D., & Chang, J. J. (1970). Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition. *Psychometrika*, **35**, 283–319.

Carroll, J. B. (1985). Exploratory factor analysis: A tutorial. In D. K. Detterman (Ed.), *Current topics in human intelligence: Vol. 1. Research methodology*. Ablex: Norwood, NJ. Pp. 25–58.

Carroll, J. B. (1992). Cognitive abilities: The state of the art. *Psychological Science*, **3**, 266–269.

Cattell, R. B., (Ed.) (1966). *Handbook of multivariate experimental psychology*. Rand McNally & Co.: Chicago.

Chapman, R. M., & Bragdon, H. R. (1964). Evoked responses to numerical and non-numerical visual stimuli while problem solving. *Nature*, **203**, 1155–1157.

Chapman, R. M. (1965). Evoked responses to relevant and irrelevant visual stimuli while problem solving. *Proceedings of the American Psychological Association*, 177–178.

Chapman, R. M. (1973). Evoked potentials of the brain related to thinking. In F.J. McGuigan & R. Schoonover (Eds.), *Psychophysiology of Thinking*. New York: Academic Press. Pp. 66–108.

Chapman, R. M. (1974). Latent component of average evoked brain responses functionally related to information processing. In *International Symposium on Cerebral Evoked Potentials in Man, Pre-circulated Abstracts*. Brussels: Presses Universite de Bruxelles. Pp. 38–42.

Chapman, R. M., McCrary, J. W., Chapman, J. A. and Bragdon, H. R. (1978). Brain responses related to semantic meaning. *Brain and Language*, **5**, 195–205.

Chapman, R. M. (1979). Connotative meaning and averaged evoked potentials. In H. Begleiter (Ed.), *Evoked brain potentials and behavior*. New York: Plenum Press: Pp. 171–196.

Chapman, R. M., McCrary, J. W., Bragdon, H. R., & Chapman, J. A. (1979). Latent components of event-related potentials functionally related to information processing. In J. E. Desmedt (Ed.), *Progress in clinical neurophysiology*. Karger. Pp. 80–105.

Chapman, R. M., McCrary, J. W., Chapman, J. A., & Martin, J. K. (1980). Behavioral and neural analyses of connotative meaning: Word classes and rating scales. *Brain and Language*, **11**, 319–339.

Chapman, R. M., McCrary, J. W., & Tuttle, J. R. (1981). Principal components analysis of sources of variability in retinal ganglion cell responses. *Biological Cybernetics*, **42**, 45–50.

Cooley, W. W., and Lohnes, P. R. (1971). *Multivariate data analysis*. New York: Wiley.

Donchin, E. and Heffley, E. F. (1978). Multivariate analysis of event-related potential data:

A tutorial review. In D. Otto (Ed.), *Multidisciplinary perspectives in event-related brain potential research*. Washington, DC: U.S. Government Print Office. Pp. 555–572.

Freeman, W. J. (1987). Analytic techniques used in the search for the physiological basis of the EEG. In A. S. Gevins & A. Redmond (Eds.), *Methods of analysis of brain electrical and magnetic signals: EEG handbook (revised series, vol. 1)*. Amsterdam: Elsevier Science Publishers. Pp. 583–664.

Glaser, E. M., & Ruchkin, D. S. (1976). *Principles of neurobiological signal analysis*. New York: Academic Press.

Gnanadesikan, R. (1977). *Methods for statistical data analysis of multivariate observations*. New York: Wiley.

Guthrie, D. (1990). Intergroup and intrasubject principal component analysis of event-related potentials. *Psychophysiology*, **27**, 111–119.

Harman, H. H. (1976). *Modern factor analysis, third edition revised*. Chicago: The University of Chicago Press.

Harshman, R. A. (1970). Foundation of the PARAFAC procedure: Models and conditions for an "explanatory" multi-mode factor analysis. *UCLA Working Papers in Phonetics*, **16**, 1–84.

Harshman, R. A., & Lundy, M. E. (1984). The PARAFAC model for the three-way factor analysis and multidimensional scaling. In H. G. Law, C. W. Snyder, J. A. Hattie, & R. P. McDonald (Eds.), *Research methods for multimode data analysis*. New York: Praeger. Pp. 122–215.

Hunt, E. (1985). Mathematical models of the event-related potential. *Psychophysiology*, **22**, 395–402.

Jennrich, R. I. & Sampson, P. F. (1966). Rotation for simple loadings. *Psychometrika*, **31**, 313–323.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurements*, **20**, 141–151.

Kramer, A. F., & Donchin, E. (1987). Brain potentials as indices of orthographic and phonological interaction during word matching. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **13**, 76–86.

Kruskal, J. B. (1976). More factors than subjects, tests and treatments: An indeterminacy theorem for canonical decomposition and individual differences scaling. *Psychometrika*, **41**, 281–293.

Kruskal, J. B. (1977). Three-way arrays: Rank and uniqueness of trilinear decomposition with application to arithmetic complexity and statistics. *Linear Algebra and its Applications*, **18**, 95–138.

Maier, J., Dagnelie, G., Spekreijse, H., & Van Dijk, B. W. (1987). Principal components analysis for source localization of VEPs in man. *Vision Research*, **27**, 165–177.

Mocks, J. (1986). The influence of latency variations in principal component analysis of event-related potentials. *Psychophysiology*, **23**, 480–484.

Mocks, J., & Verleger, R. (1986). Principal component analysis of event-related potentials: A note on misallocation of variance. *Electroencephalography and Clinical Neurophysiology*, **65**, 393–398.

Mocks, J. (1988a). Decomposing event-related potentials: A new topographic components model. *Biological Psychology*, **26**, 199–215.

Mocks, J. (1988b). Topographic components model for event-related potentials and some biophysical considerations. *IEEE Trans. Biomed. Eng.*, **35**, 482–484.

Mocks, J., Kohler, W., Gasser, T., & Pham, D. T. (1988). Novel approaches to the problem of latency jitter. *Psychophysiology*, **25**, 217–226.

Mulaik, S. A. (1972). *The foundations of factor analysis*. New York: McGraw-Hill.

Nunez, P. L. (1981). *Electric fields of the brain*. New York: Oxford.

Pham, T. D. & Mocks, J. (1992). Beyond principal component analysis: A trilinear decomposition model and least-squares estimation. *Psychometrika*, **57**, 203–215.

Ruchkin, D. S., Sutton, S. & Stega, M. (1980). Emitted P300 and slow wave potentials in guessing and detection tasks. *Electroenceph. & Clin. Neurophysiol.*, **49**, 1–14.

Rummel, R. J. (1970). *Applied factor analysis*. Evanston: Northwestern Univ. Press.

Steeger, G.H., Herrmann, O., & Spreng, M. (1983). Some improvements in the measurement of variable latency acoustically evoked potentials in human EEG. *IEEE Trans. Biomed. Engn.*, BME-30, 295.

Thorndike, R. M., & Weiss, D. J. (1970). Parsimony of factor-analytic methods for diverse scientific areas. In *Proceedings of the 78th Annual Convention of the American Psychological Association*. American Psych. Association: Washington, DC.

Tucker, L. R. (1966). Some mathematical notes on three-mode factor analysis. *Psychometrika*, **31**, 279–311.

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, **41**, 321–327.

Velicer, W. F., & Jackson, D.N. (1990). Component analysis versus common-factor analysis: Some issues in selecting an appropriate procedure. *Multivariate Behavioral Research*, **25**, 1–28.

Vokey, J. R. (1989). Matching factor structures using orthogonal rotation to congruence. *Behavior Research Methods, Instruments & Computers*, **21**, 84–85.

Walter, W. G., Cooper, R., Aldridge, V. J., McCallum, W. C., & Winter, A. L. (1964). Contingent negative variation: an eletrical sign of sensorimotor association and expectancy in the human brain. *Nature London*, **203**, 380–384.

Wood, C. C., & McCarthy, G. (1984). Principal Component Analysis of event-related potentials: Simulation studies demonstrate misallocation of variance across components. *Electroenceph. and Clin. Neurophysiol.*, **59**, 249–260.

Wood, C. C., & McCarthy, G. (1985). Principal Component Analysis of ERPs: further simulation studies. *Electroenceph. and Clin. Neurophysiol.*, **61**, S238.

Woody, C. D. (1967). Characterization of an adaptive filter for the analysis of variable latency neuroelectric signals. *Medical & Biol. Engn. & Computing*, **5**, 539–553.

Zwick, W. R., & Velicer, W. F. (1982). Factors influencing four rules for determining number of components to retain. *Multivariate Behavioral Research*, **17**, 253–269.

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin*, **99**, 432–442.