Methods for the quantification and statistical testing of ERP differences across conditions

J. HOORMANN, M. FALKENSTEIN, P. SCHWARZENAU, and J. HOHNSBEIN
Universität Dortmund, Dortmund, Germany

Several standard methods, as well as a new method for the quantification of event-related potential (ERP) differences across conditions, are described. The standard methods are (1) peak analysis, (2) the calculation of mean values, and (3) the calculation of difference waveshapes. The new method, called *window analysis*, was designed to quantify and statistically test in a very simple way any shape differences between two ERP curves in certain time intervals (windows) when clear peaks are lacking in one or all conditions. The window analysis is based on a conventional analysis of variance with sample time as an additional within-subjects factor. The significance of a shape difference between the curves for a factor of interest can then be determined with an *F* test for the interaction of this factor with the factor time. The usefulness of the window analysis is demonstrated in an example with real data.

STANDARD METHODS FOR ERP COMPONENT QUANTIFICATION AND THEIR PROBLEMS

The crucial task in event-related potential (ERP) research is the quantification, statistical testing, and interpretation of ERP differences across conditions. Ideally, the averaged ERP, as recorded from the scalp (*surface structure*), consists of a series of peaks and troughs, which can be labeled by their polarity (P for positive, N for negative) and sequence (N1, N2, P1, P2, etc.), typical latency (P300, N400), or mean peak latency (N134, P230, etc.). Such a simple structure allows a simple and straightforward quantification or characterization of the ERP by measuring the latency and amplitude of these extrema (*peaks*). The peaks are assumed to be generated by underlying components (*component structure*).

Peak Amplitude Measures

The peak analysis is usually conducted in two steps: First, for each peak a specific ERP segment has to be determined (*search window*) in which the peak is assumed to be located for each subject and each condition. One approach for determining search windows is to compute the grand means of the ERPs (averages for one condition across all subjects) and average the grand means across all conditions. This overall mean ERP contains all electroencephelographic (EEG) epochs and is an estimate for the general component structure of the ERP across conditions. In the case of large ERP differences across conditions (e.g., for different stimulus modalities), separate averages of grand means should be computed for subsamples of those conditions that yield a similar ERP structure. ERP segments of certain width are now centered on each of the peaks in the overall mean ERP. The width of these search windows should be chosen so as to include all individual peaks of the same component while excluding peaks of different, adjacent components of the same polarity. We propose to choose the search window width slightly larger than the width of the affiliated component in the overall mean ERP. The component width can be defined, for example, as the time distance of its zero crossings. Search windows can also be determined more analytically—for example, by using sample-by-sample tests of the zero deviation of a curve. However, in such multiple tests, the alpha level has to be corrected, using the Bonferroni method (see, e.g., Bortz, 1985; Brown, Michels, & Winer, 1991). As a consequence of correction, the use of many samples leads to an unacceptable loss of test power, so only a few tests should be performed. Hence, such an analytic approach could preferably be used for adjusting search window margins by testing only few samples around the margins predetermined by the described inspection of the overall mean ERPs.

The second step is to search, for each subject and each condition, for the peak within the affiliated search window. A peak can, for example, be defined as the largest local extremum within the search window, provided that the voltage differences to the adjacent peaks of opposite polarity exceed some predefined criterion (Falkenstein, Hohnsbein, & Hoormann, 1993; Hall, Rappaport, Hopkins, & Griffin, 1973). The search can be done by hand as well as by suitable programs (see, e.g., Daskalova, 1988). After determination of the peaks, their latency and amplitude can be measured for all conditions. Figure 1 illustrates the setting of a search window for an auditory P2 component. The two thick lines show the grand aver-
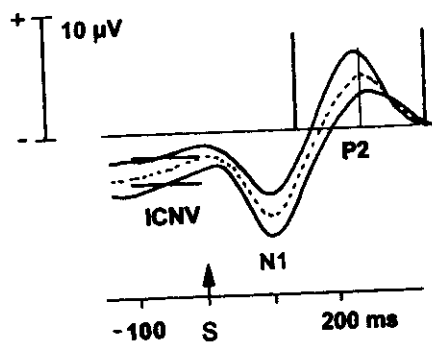
Figure 1. Example of the parametrization of an averaged ERP under two different conditions (thick straight curves). lCNV, late contingent negative variation, a tonic component before stimulus onset (S). It is quantified by the mean amplitude in an ERP segment before S (horizontal lines above lCNV). N1, P2, phasic components with distinct peaks. The peak is searched for in each subject and condition in a search window (margins, thick vertical lines) that is centered on the peak latency of the P2 (thin vertical line) in the mean ERP across conditions (dashed line).

ages of the ERPs for two different conditions; the dashed line is the mean of the grand averages across the two conditions. The search window is centered at the peak of the P2 in the mean of the grand averages, and its boundaries are set to ±100 msec from the center, which is slightly larger than the width of the P2.

An important issue in amplitude measuring is the question, in relation to which reference point or reference line should the amplitudes be measured? A common method is to measure the amplitude A of a certain component relative to the mean amplitude B of a (presumably neutral) baseline. The basic idea behind this approach is that all components are superposed (*riding*) on the same common baseline. The problem with this approach is twofold: First, the assumption that the underlying baseline is the same for all components is questionable; second, it is often difficult to find such a neutral baseline. In most reports, the mean value of a short ERP segment before stimulus onset is chosen as the baseline. However, this baseline is often influenced by the late CNV (Walter, Cooper, Aldridge, McCallum, & Winter, 1964), which reflects the preparation for the present trial. In such a case, the variation of A is confounded with a possibly independent variation of B (Rösler, 1979). Hence, to avoid this influence, A and B should at first be measured relative to technical zero, which is truly neutral. After this, correlations should be computed between A and B for all conditions and subjects. Only in the case of a significant positive correlation of A with B can the affiliated ERP component or peak be indeed assumed to ride on the baseline. In this case, A should be measured relative to B by transforming A to $A' = A - B$. Such significant positive correlations are often found for earlier components (e.g., for the N100). If A and B show no significant correlation, which is usually found for

later ERP components, A and B vary independently, so A should be measured relative to technical zero (Rösler, 1979). However, there are also limitations to this approach. A correlation may be observed not because of superposition, but because of a common influence—such as arousal—on baseline, component amplitude, latency, or all of these. For example, a high arousal could cause a large late CNV (negative) and a large P300. We could show (Falkenstein, Hohnsbein, & Hoormann, 1994b; Hohnsbein, Falkenstein, & Hoormann, in press) that a large late CNV appears to be associated with a short latency of the late P3 subcomponent (P-CR), while the early P3 subcomponent (P-SR) remains stable. This leads to a larger superposition of both subcomponents and to a seemingly larger P3 complex. In this case, the correlation did not meet the sign criterion (the correlation was negative instead of positive), but, for a negative component such as the N200, there would have been a positive correlation, but one not caused by superposition.

Problems With Peak Analysis

A general problem with peak analysis in average waveforms is that the components in the single epochs underlying the average are usually affected by considerable latency jitter. Such jitter is caused by a variety of influences, such as trial-by-trial fluctuations in arousal and effort. The latency jitter is usually larger in later components, such as the P3 complex. Latency jitter in itself smears the average component and makes peak detection difficult, so eventually noise peaks can be mistaken for true peaks. Moreover, differences in latency jitter across conditions smear the average components differentially, seemingly indicating attenuation of the peak amplitude of the component in the condition with the larger latency jitter.

A common method for reducing latency jitter and preserving latency information is the Woody filter (Woody, 1967). This procedure assumes a constant form of the component(s) across single epochs. A segment of each single epoch is crosscorrelated with a template (usually a sine half wave or the respective segment of the average ERP), which results in different lags for maximum correlation for each epoch. The single epochs are then timeshifted by the respective lags and averaged only after this correction. This is aimed at compensating for the time jitter. The procedure can be repeated by using the new average as the new template. For complex structures, such as the P3 complex (see Figure 2), the mean amplitude computation is suboptimal, since any differential latency information about the underlying components is lost in the mean values. Also, the Woody filter may be misleading, since the form of the complex is likely to change in each single epoch because of differential overlap of the subcomponents. Hence, in this case, not only the latency estimation by crosscorrelation is impaired, but, more severely, the subcomponents are likely to be smeared together to one single peak. So the Woody fil-
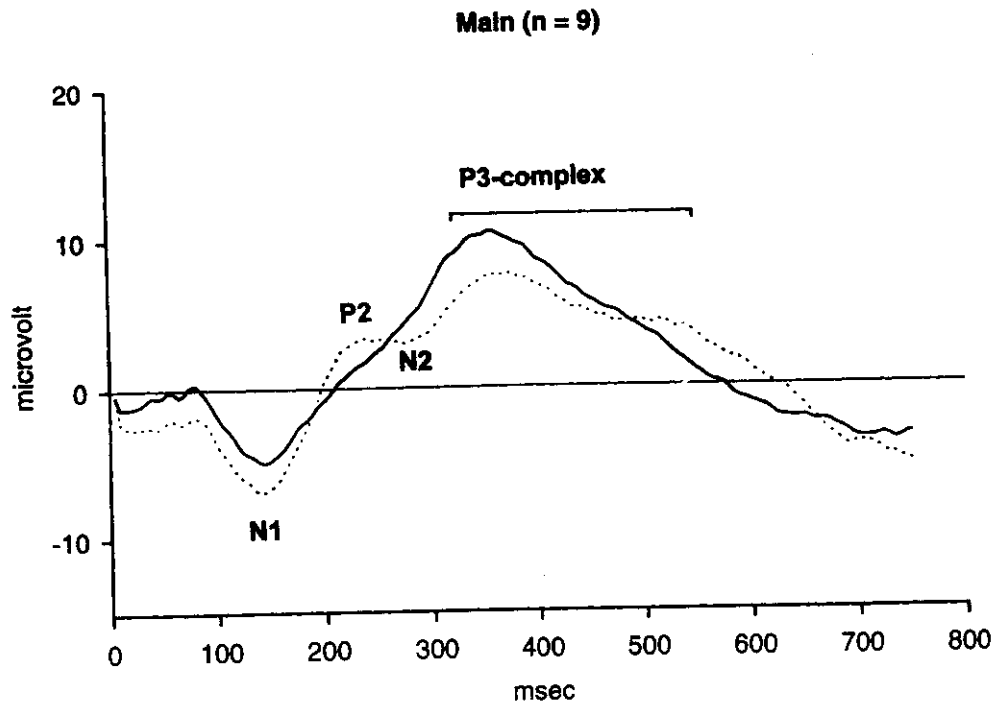
Main (n = 9)

Figure 2. Grand means (averaged across all subjects) of the ERPs in two attention conditions (focused atten-. tion [FA], solid line; divided attention [DA], dotted line). The N1 is followed by P2 and N2 components in DA, which are not seen in FA. These early components are followed by a large P3 complex, which also differs between FA and DA.

ter should be applied only if the segment of interest definitely contains no overlapping components. This precondition, however, is rarely met.

Two more simple methods to reduce latency jitter are (1) to train the subjects thoroughly and (2) to impose some degree of time pressure (Falkenstein et al., 1993; Falkenstein, Hohnsbein, & Hoormann, 1994a, 1994b; Hohnsbein, Falkenstein, Hoormann, & Blanke, 1991). Both measures reduce variance in performance and. thereby, the latency jitter of late components that are (at least partially) time-locked to the response, such as the lateralized readiness potential (LRP; cf. the contributions of Eimer, 1998, and of Schwarzenau, Falkenstein, Hoormann, & Hohnsbein, 1998) and the late P3 subcomponent (P-CR; Falkenstein et al., 1994a).

A second problem with peak analysis, which has already been mentioned above, is that often the individual ERP components are not aligned in a strict sequential manner, but overlap to different degrees, often forming one broad complex. This can cause plateau-like ERP segments, which cannot be easily quantified by peak analysis. Moreover, the form of the ERP segment may change across conditions. In particular, certain components may be only visible for one, but not for the other condition. An example for a different component structure of the ERP is illustrated in Figure 2, which shows ERPs from two-alternative choice reaction tasks for two

attention conditions (the paradigm is described in more detail below). In one condition (solid line), no P2 is visible, and the P3 complex (the positive complex beginning at about 300 msec) consists of a broad positivity with flat falling slope. In the other condition (dotted line), a clear P2 is visible, and the P3 complex is divided into two parts, the first part being attenuated compared to the first condition. So it is a problem to define ERP parameters common to both conditions because of the profound shape change of the ERP across conditions. A method for quantifying such form differences will be presented below.

One possibility for disentangling subcomponents is given by the principal components analysis (PCA), which is presented in detail by van Boxtel (1998). This method, though, has some caveats and problems. One of the main problems with PCA is that it has difficulties in dealing with latency variations of components. For example, Möcks (1986) could demonstrate that a strong latency variation of one single component C resulted in a PCA solution with a basic component related to C and a fictitious secondary component related roughly to the first derivative of C.

Mean Amplitude Measures

The preceding section has shown considerable problems with peak analysis in the case of component over-

lap. Moreover, there is a variety of very slow ERP phenomena, such as the readiness potential (Kornhuber & Deecke, 1965) or the above mentioned CNV, which exhibit no peak, so that peak analysis is not possible.

A standard method for dealing with plateau-like peaks, form changes, and very slow ERP components is the computation of the mean amplitude of the ERP within the time segment of interest. Figure 1 illustrates how this can be done for the late CNV by computing, for example, the mean amplitude (offset of the two horizontal lines) of the ERP segment of 100 msec before stimulus onset for the two different conditions. Mean amplitude measures are certainly the suitable method for tonic phenomena such as the late CNV, which seems not to be composed of subcomponents. A recent example of the application of the mean amplitude analysis can be found in Gratton, Corballis, and Jain (1997). The computation of mean amplitudes also reduces the influence of latency jitter and, consequently, peak amplitude fluctuations across components, because the area and also the mean amplitude across a fixed time segment remain rather constant despite differential latency jitter. However, for phasic components, the peak latency is lost in the mean amplitudes.

Difference Amplitude Measures

A straightforward and very simple method for highlighting subcomponents is the calculation of difference waveshapes (DIFs). DIFs suppress ERP activity that is constant across two conditions, electrodes, or both. DIFs between conditions have been used extensively by attention researchers (e.g., Hansen & Hillyard, 1980) to highlight attention-related components. The above mentioned LRP is also calculated as the difference between the activity at two lateralized electrodes, whereby activity common to both electrodes (mainly activity that is not movement related) is suppressed. Falkenstein et al. (1994b) could not only demonstrate that the P3 complex consists of two subcomponents, but also, by using DIFs, that the later of these subcomponents (P-CR) is preceded by a negative wave (N-CR), which is only seen indirectly in the *raw* ERPs. One assumption, however, that has to be met when calculating DIFs is the invariance of the common activity across conditions, which is difficult to prove. Also, DIFs can sometimes produce artificial peaks, which are merely due to time shifts of the underlying components across conditions. We used the basic idea of the DIF approach—namely, to highlight differences across conditions—as a simple method by which to quantify and statistically test form differences between ERP curves in certain time intervals (windows) when clear peaks are lacking in one or all conditions, as shown in Figure 2. The new approach (called *window analysis*) does not necessarily assume that components underlie the form differences. It only attempts to verify statistically whether ERP curves differ significantly across conditions in certain time windows.

WINDOW ANALYSIS

Assuming an ERP interval of fixed length $(1, \ldots, k)$ denote the equidistant sample points within the interval under two different conditions $C(1,2)$ for $S(1, \ldots, l)$ subjects. Let $y(t, c, s)$ denote the ERP amplitude at sample point t, in condition c, for subject s. Can $y(t, c, s)$ be distributed normally? The question of interest is whether the two ERP intervals differ statistically between the two conditions. The straightforward solution to this question yields an analysis of variance (ANOVA) with the within-subjects factor condition (C) and an additional within-subjects factor time (T). A significant main effect of T would simply reflect the fact that the time course of the mean ERP deviates significantly from zero. A significant main effect of C would reflect a difference in the mean amplitude in the window across conditions—that is, the two curves were significantly shifted against each other on average. One could get this result also by a mean values analysis, as described above. More interestingly, a significant interaction $C \times T$ would indicate that the two curves have a different time course, independent of a difference of the mean amplitude in the window. In the case of a significant interaction $C \times T$, simple effects may be used to test the significance of the curve difference at prespecified sample points. Such points of maximum difference can be chosen a priori by inspection of the grand means. As already mentioned above, the alpha levels have to be Bonferroni-corrected in multiple comparisons, which leads to a strong decrease of the test power. Hence, the test of many sample points is not advisable. The location and size of the window is usually determined after visual inspection of the grand means. It is also possible to determine the exact window size that is most sensitive for the detection of the form difference with the aid of the ANOVA. To that end, the length and the location of the interval can be varied to determine the particular interval that yields the highest significance for the interaction. To reduce computation time, the distance between the sample points should be in the range of 10 to 20 msec. Since the factor T contains multiple levels, and the mutual correlation across these levels of T differs strongly (adjacent sample points correlate higher than do distant points), one of the essential preconditions of ANOVA, the circularity assumption (see, e.g., Bortz, 1985; Brown et al., 1991), is clearly violated. Hence it is absolutely necessary to correct the degrees of freedom for the ANOVA by using the conservative Greenhouse-Geisser procedure (Geisser & Greenhouse, 1958), which is, for example, implemented in the 4V program of BMDP.

A caveat for the method is that interactions are sometimes misleading and difficult to interpret. This is due to the fact that the ANOVA is based on an additive model, whereas the components are multiplicative in nature (McCarthy & Wood, 1985). For example, a doubling of strength of a component generator leads to a doubling of amplitudes at each sample point of that component. This

results not only in a main effect, but usually also in an interaction, because the additive enhancement of the component is larger at the peak than at the flanks. In such cases, it would be misleading to interpret the significant interaction as a difference in waveshape, because the real reason for this difference is obviously an enhancement in generator strength. The ordinary way to compensate for such multiplicativity effects is the normalization of the data for the condition in question—that is, for each level of the condition, each sample point is divided by the mean value for this level within the window. This sets the main effect of the condition to zero and thereby eliminates the significance of form differences that are due solely to multiplicative effects. In any case, a window analysis would not be used, when a single component is seen with clear peaks in two conditions; instead, a peak or mean value analysis, as described above, should be appropriate.

A related (but different) technique to deal with multiplicativity has recently been described by Tucker, Liotti, Potts, Russell, & Posner (1994).

EXAMPLE WITH REAL DATA

In the following example, an application of the method is shown. Visual and auditory letter stimuli (F or J) were presented in a train (ISI about 1,750 msec); the occurrence of Js and Fs was equiprobable. In different blocks, the stimuli were presented either visually (focused atten-

tion [FA], visual) or auditorily (focused attention [FA], auditory), or the stimulus modality was randomized within the block (divided attention [DA]; Hohnsbein et al., 1991). Nine subjects performed speeded binary choice reactions with the right and the left index fingers to one letter each (J and F, respectively). The EEG was measured at four midline electrodes—Fz, Cz, Pz, Oz—and sampled with a rate of 200 Hz. The ERPs of the correct trials were averaged with letter onset as trigger. The number of trials was virtually the same for both attention conditions. Figure 2 exhibits the grand mean (across all subjects) of the ERPs after auditory stimuli for the two conditions, FA and DA, at the vertex electrode Cz.

The grand means first show an N1 peaking around 140 msec. Subsequently, a P2, peaking around 220 msec, and a later N2, peaking around 280 msec, are visible for the DA condition, whereas the FA condition shows no discernible peaks. After the N2, a late positive complex (P3 complex) emerges, which has a quite different shape in the two conditions. Hence, there are clear form differences in the regions of the P2/N2 and the P3 complexes across conditions in the grand means. In Figure 3, the individual ERPs are given for both conditions. The form differences as seen in Figure 2 are present for most, but not for all subjects.

The grand mean structure suggests the choice of three different windows, which are (arbitrarily) centered on the regions of interest: Window 1 (0–180 msec) contains the N1, Window 2 (200–320 msec) is centered on the inter-
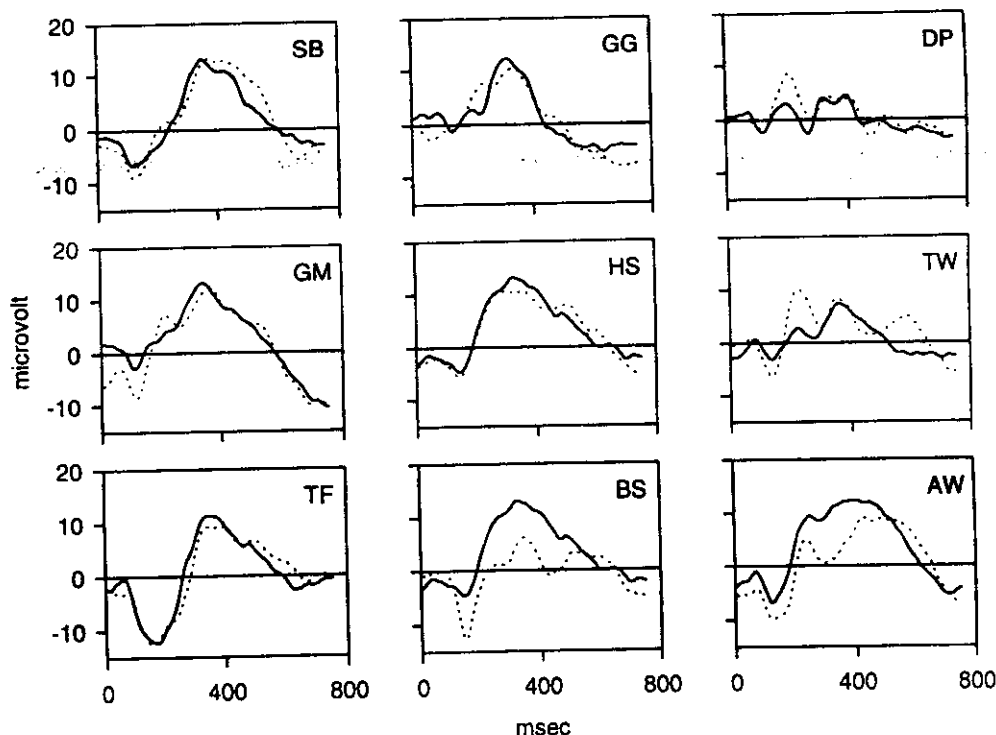

Figure 3. Individual ERPs for all 9 subjects in focused attention (solid lines) and divided attention (dotted lines) conditions.

Table 1
Results of the Two-Way Repeated Measures ANOVA With the Factors C(FA, DA) and T

	p value		
Effect	Window 1	Window 2	Window 3
T	.0167	.0181	.0000
C	.0176	.9552	.09
C × T	.2822	.0000	.0117

Note—The degrees of freedom have been corrected after Geisser and Greenhouse (1958).

section of the curves of the FA and DA conditions and contains the P2 and the N2, and Window 3 (320–750 msec) contains the P3 complex.

To reduce the levels of the factor T, only every second sample was used, which results in a sampling rate of 100 Hz.

Table 1 shows the results of an ANOVA (BMDP4V; Dixon, 1990) with the two within-subjects factors C(FA, DA) and $T(1, \ldots, N)$, with $N = 19$ for Window 1, $N = 13$ for Window 2, and $N = 44$ for Window 3. The significant main effect C in Window 1 shows that the curves are shifted against each other by a certain voltage. (Further analysis—that is, a correlation with the baseline—revealed that this difference was due to an amplitude shift of the late CNV across conditions, while the N1 was constant.) Significant interactions $C \times T$ are found in Windows 2 and 3. This shows that the form differences in the grand means in these windows are significant. There was also a slight trend for a main effect C in Window 3, whereas in Window 2 there was no significant main effect at all. When the data in Window 3 were normalized in order to remove the trend of the main effect, the $C \times T$ interaction remained significant. So it appears that the form differences of the curves in Windows 2 and 3 were not caused by multiplicative effects—for example, by enhancement of one component within the windows.

The result of the window analysis is dependent on the location and the size of the window. Table 2 shows how the level of significance decreases with a stepwise narrowing of Window 2. However, even with a 20-msec window around 260 msec a significant interaction is obtained, showing that the focus of the curve difference is centered on 260 msec, where the ERPs intersect in the grand means.

Table 2
Change of the Significance Level of the C × T Interaction in Window 2 When the Window Size is Changed

Beginning of the Window (msec)	End of the Window (msec)	p value of C × T Interaction	
		Raw	Bonferroni
200	320	.0000	.0001
210	310	.0000	.0001
220	300	.0001	.0006
230	290	.0005	.0030
240	280	.0025	.0149
250	270	.0068	.0401

Note—The p values have also been corrected after Bonferroni ($n = 6$) for multiple comparisons (last column).

Furthermore, two simple effects were computed in order to test the curve differences at the two sample points that correspond to the P2 and N2 peaks in the grand means (Figure 2)—namely, at 220 and 290 msec. At both sample points, the simple effects showed only tendencies (220 msec, .0703; 290 msec, .0602), which vanished after Bonferroni correction. This shows that the window analysis is superior to single sample analyses for determining curve differences.

SUMMARY AND CONCLUSION

One of the main tasks in ERP research is the quantification, statistical testing, and interpretation of ERP differences across conditions. If clear peaks are discernible in the averaged ERPs, a simple quantification can be conducted by determining the peaks and measuring their latency and amplitude. However, in cases of tonic components or component overlap, this method fails. The amplitude (but not the latency) of tonic components can be assessed by calculating the mean amplitude in an ERP segment of interest (window). In the case of overlapping components, difference waveshapes between conditions can be used to suppress ERP contributions that are common to both conditions. The idea of the difference approach was used to design a very simple method for statistically testing ERP differences of any kind between conditions—especially form differences, which are often not detectable by the analysis of mean values. This approach, called *window analysis*, simply uses sample time (T) as an additional within-subjects factor in the ANOVA. Hence, the approach requires no sophisticated data transformations, but simply uses all sampled ERP data for the ANOVA. The significance of form differences is tested by the interaction of T with the condition factor. The application of the window analysis with real data showed a high degree of sensitivity of the method for the detection of curve differences, even in small ERP segments. The window analysis can also be used to specify points of maximum difference across conditions. In summary, the window analysis may be helpful to test whether ERP segments without consistent peaks are statistically different across conditions.

REFERENCES

BORTZ, J. (1985). *Lehrbuch der Statistik*. Berlin: Springer-Verlag.

BROWN, D. B., MICHELS, K. M., & WINER, B. J. (1991). *Statistical principles in experimental design* (3rd ed.). New York: McGraw-Hill.

DASKALOVA, M. I. (1988). Wave analysis of the electroencephalogram. *Medical & Biological Engineering & Computing*, 26, 425-428.

DIXON, W. J. (ED.) (1990). *BMDP Statistical Software: Program 4V*. Berkeley.

EIMER, M. (1998). The lateralized readiness potential as an on-line measure of central response activation processes. *Behavior Research Methods, Instruments, & Computers*, 30, 146-156.

FALKENSTEIN, M., HOHNSBEIN, J., & HOORMANN, J. (1993). Late visual and auditory ERP components and choice reaction time. *Biological Psychology*, 35, 201-224.

FALKENSTEIN, M., HOHNSBEIN, J., & HOORMANN, J. (1994a). Effects of

choice complexity on different subcomponents of the late positive complex of the event-related potential. *Electroencephalography & Clinical Neurophysiology, 92,* 148-160.

FALKENSTEIN, M., HOHNSBEIN, J., & HOORMANN, J. (1994b). Time pressure effects on late components of the event-related potential (ERP). *Journal of Psychophysiology,* 8, 22-30.

GEISSER, S., & GREENHOUSE, S. W. (1958). An extension of Box's results on the use of the *F* distribution in multivariate analysis. *Annals of Mathematics & Statistics, 29,* 885-891.

GRATTON, G., CORBALLIS, P. M., & JAIN, S. (1997). Hemispheric organization of visual memories. *Journal of Cognitive Neuroscience, 9,* 92-104.

HALL, R. A., RAPPAPORT, M., HOPKINS, H. K., & GRIFFIN, R. B. (1973). Peak identification in visual evoked potentials. *Psychophysiology,* 10, 52-60.

HANSEN, J. C., & HILLYARD, S. A. (1980). Endogenous brain potentials associated with selective auditory attention. *Electroencephalography & Clinical Neurophysiology, 49,* 277-290.

HOHNSBEIN, J., FALKENSTEIN, M., HOORMANN, J., & BLANKE, L. (1991). Effects of crossmodal divided attention on late ERP components: I. Simple and choice reaction tasks. *Electroencephalography & Clinical Neurophysiology, 78,* 438-446.

HOHNSBEIN, J., FALKENSTEIN, M., & HOORMANN, J. (in press). Performance differences in reaction tasks are reflected in event-related brain potentials (ERPs). *Ergonomics.*

KORNHUBER, H. H., & DEECKE, L. (1965). Hirnpotentialänderungen bei Willkürbewegungen und passiven Bewegungen des Menschen:

Bereitschaftspotential und reafferente Potentiale. *Pflügers Archiv,* 284, 1-17.

MCCARTHY, G., & WOOD, C. C. (1985). Scalp distributions of event-related potentials: An ambiguity associated with analysis of variance models. *Electroencephalography & Clinical Neurophysiology,* 62, 203-208.

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

RÖSLER, F. (1979). Zur psychologischen Bedeutung evozierter Hirnrindenpotentiale: Methodische Probleme. *Psychologische Beiträge,* 21, 1-21.

SCHWARZENAU, P., FALKENSTEIN, M., HOORMANN, J., & HOHNSBEIN, J. (1998). A new method for the estimation of the onset of the lateralized readiness potential (LRP). *Behavior Research Methods, Instruments, & Computers, 30,* 110-117.

TUCKER, D. M., LIOTTI, M., POTTS, G. F., RUSSELL, G. S., & POSNER, M. I. (1994). Spatiotemporal analysis of brain electric fields. *Human Brain Mapping,* 1, 134-152.

VAN BOXTEL, G. J. M. (1998). Computational and statistical methods for analyzing event-related potential data. *Behavior Research Methods, Instruments, & Computers, 30,* 87-102.

WALTER, W. G., COOPER, R., ALDRIDGE, V. J., MCCALLUM, W. C., & WINTER, A. L. (1964). Contingent negative variation: An electric sign of sensorimotor association and expectancy. *Nature, 203,* 380-384.

WOODY, C. D. (1967). Characterization of an adaptive filter for the analysis of variable latency neuroelectric signals. *Medical & Biological Engineering,* 5, 539-553.