Electroencephalography and clinical Neurophysiology, 1986, 65: 59–71
Elsevier Scientific Publishers Ireland, Ltd.

LATENCY VARIABILITY AND TEMPORAL INTERRELATIONSHIPS OF THE AUDITORY EVENT-RELATED POTENTIALS (N1, P2, N2, AND P3) IN NORMAL SUBJECTS [1]

H.J. MICHALEWSKI, D.K. PRASHER [2] and A. STARR

Department of Neurology, University of California, Irvine, CA 92717 (U.S.A.)

(Accepted for publication: July 15, 1985)

Summary Peak latency variation and the temporal interrelationships of the auditory event-related potential were investigated in 12 normal adults (ages 28–42). Measures of variation were based on both conventional averages and single trials. Estimates of N1, P2, N2 and P3 latencies were made on a trial-by-trial basis to target stimuli recorded from Fz, Cz and Pz scalp locations.

Results showed that single-trial latency variability of the auditory ERP differed both among the various components and between subjects. Larger standard deviations were measured for the later N2 and P3 components than the earlier N1 and P2 components. Regression analyses between various component latencies indicated a strong covarying relationship between N2 and P3, with N2 accounting for up to 61% of the variance of P3 latency at Pz. Earlier N1 and P2 components added little to the overall prediction of either P3 or N2. For the other components, P2 accounted for 9–16% of the variance of N2, while N1 accounted for approximately 1% of the variance of N2; N1 accounted for 8–10% of the latency variation of P2. The correlations between single-trial peak latencies and RTs were positive but of low magnitude. The highest correlations between peak latency and RT were found for N2 ($r = 0.33$) and P3 ($r = 0.24$).

The low correlations between the single-trial latencies of N1 and P3 suggest that the processes reflected by these components are independent and support a distinction between the earlier and the later components of the ERP. The close temporal coupling between N2 and P3 suggests that N2 may reflect cognitive properties in common to P3 in stimulus evaluation processes.

Keywords: *variability – ERP – auditory – latency – interrelationships*

The event-related potential (ERP) recorded from the scalp is comprised of a group of components presumed to be involved in human information processing, reflecting factors such as stimulus registration, attention and evaluation (see Picton and Hillyard 1974; Picton et al. 1974; Donchin et al. 1978; Hillyard et al. 1978; Donchin 1979; Pritchard 1981; Magliero et al. 1984). The timing and duration of the electrical events represented by the ERP may aid in comprehending the processes that take place between stimulus presentation and perception. In particular, an understanding of component variability and component interrelationships within the ERP wave form may provide insights into the transmission of signal information during these processing stages.

There is some difficulty in this analytic approach, since conventional signal averaging procedures obscure features of the ERP such as variations in latency, or temporal relations between peaks within the wave form. A possible solution to the limitations imposed by the averaging procedure is to determine peak latencies on a trial-by-trial basis using correlational procedures (e.g., Woody 1967; Weinberg and Cooper 1972). In the study described below we adapted a correlational technique to derive latencies for the sequence of N1, P2, N2 and P3 peaks of the auditory ERP based on single trials and to examine variations in ERP component latencies. A first objective of this study was to define latency variation for each of the successive peaks of the ERP in a group of normal individuals using both conventional averages and measures derived from single trials in order to provide a descriptive picture of component variation. A second objective of this study

[1] Supported by Grant No. 11876.

[2] MRC Neuro-Otology Unit, Institute of Neurology, National Hospital, Queen Square, London WC1N 3BG, U.K.

was to investigate the temporal interrelationships between components of the ERP. We are uncertain, for example, whether on a particular trial a delayed N1 component necessarily results in a prolonged P2, or whether the latency of N1 has any consequences for the latencies of the later N2 and P3 components. There already is evidence suggesting a covariation of the latencies of N2 and P3 (Renault and Lesèvre 1979; Renault et al. 1982; Renault 1983) but the relation of the earlier peaks to these later waves has not been directly investigated. Single-trial latencies allow us to test various component relationships by using the latencies of preceding peaks to predict the latencies of subsequent peaks. In this way we were able to evaluate the strength of temporal coupling between components and to determine whether latency variations in the later components were associated with variations in the earlier portions of the ERP.

Methods

Subjects

Twelve normal adult individuals (9 males, 3 females) ranging in age from 28 to 42 years (mean = 34.6) participated in the testing sessions.

Recording and test procedures

The scalp electroencephalogram (EEG) was recorded from midline electrode sites Fz, Cz and Pz referenced to linked ears. Eye movements were monitored from a pair of electrodes placed above and at the outer canthus of the right eye; an additional electrode was placed on the forehead as a ground. Skin impedance for the scalp sites measured below 3.0 kΩ. The EEG was amplified (2×10^5) with a bandpass of 0.1–100 Hz (3 dB down, 6 dB/octave slope).

Auditory event-related potentials were collected in a standard *oddball* paradigm. Subjects were instructed to detect an occasional (20%) high frequency target tone (640 Hz, 50 msec duration) interspersed among a series of frequent (80%) low frequency (440 Hz, 50 msec) tones. The order of target tones within the series of frequent tones was determined on a pseudorandom basis with the restriction that not more than 3 targets occurred in succession. A total of 300 tones comprised the stimulus series. The interstimulus interval between tones varied between 2 and 3 sec. Subjects were instructed during practice to press a reaction time (RT) key *promptly* whenever a target occurred. The tones were presented binaurally at an intensity of 72 dB SPL. Four channels of data were digitized (256 data points/channel) for a total sweep period of 1000 msec, beginning 200 msec (baseline period) before tone onset. Individual trials were stored to disk for subsequent processing by computer. Subjects were tested in a sound-attenuating chamber in a seated position with their eyes open and directed toward a fixation point.

Conventional averaging

Averages for each subject to correctly detected target tones were computed from stored single trials free of excessive eye movement artifacts. Peak latencies for N1, P2, N2 and P3 were determined from stimulus onset to the point of maximum voltage, or extrapolated from the intersection of ascending and descending limbs if the component was broad. Component amplitudes were computed as the difference between the maximum voltage of a selected peak and the average voltage of the 200 msec prestimulus baseline period.

Single-trial analysis procedure

Identification of the major peaks of the evoked potential in single trials was performed by using a modified version of the Woody correlational-template procedure (Woody 1967). Separate templates, derived from the individual components (N1, P2, N2 and P3) of the averaged wave form, were used to define component shapes for each subject. Each template was made up of a different number of points: 32 points (or approximately 125 msec) for N1, 37 points (145 msec) for P2, 46 points (180 msec) for N2, and 64 points (250 msec) for P3. The number of points in each component template was sufficient to include the descending and ascending limbs for negative peaks, and ascending and descending limbs for the positive peaks. Each individual's average served as the source of the templates. The computational procedures for a component were as follows. Single

trials to correct target responses were first examined to determine whether the eye movement channel fell within acceptable voltage ranges. An accepted trial was then digitally smoothed to attenuate high frequency activity (bandpass was equivalent to 0–37.5 Hz, 3 dB down at 37.5). A series of Pearson product-moment correlations were performed between the template and a corresponding region of the single trial. The template was positioned on an individual basis before the expected component and was 'moved' along the single trial so that the pattern of correlations increased, reached a maximum, and then decreased as the template approached, reached, and passed the component. Peak latency was determined by knowing how far the template was moved to the point of maximum correlation. Latency adjusted averages for each component were formed by summing the points in the single trials that corresponded to the maximum correlation between the template and the points comprising the detected peak. Corrected peak amplitudes were measured as the difference between the maximum average peak voltage and the average voltage of the 200 msec baseline period. The computed template latencies had a temporal resolution of 3.9 msec. The temporal window scanned by the template in searching for a potential was different for each component: 104 msec for N1, 117 msec for P2, 203 msec for N2, and 250 msec for P3. The selection of each template window was made on the basis of estimating typical component ranges from the visual inspection of plotted single trials; each component window was set to minimize the possible misidentification of a preceding (or succeeding) peak of the same polarity, yet large enough to accommodate the expected peaks. The correlational computations were applied to each of the 3 midline electrode sites for one iteration or pass of the template along the single trial.

Data were analyzed by regression, t tests, and analysis of variance (ANOVA) procedures. Correlations were z-transformed before analysis. Post-hoc tests were conducted with the Tukey test (Keppel 1973). Significance levels were set at $P < 0.05$, or better.

Testing the single-trial procedure

Our estimates of peak latencies, and subsequently peak variations, depended in part on how well the template procedure was able to identify peaks in the single trials. The effectiveness of the template procedure can be compromised by unfavorable signal-to-background ratios (e.g., Ruchkin and Sutton 1979; McCarthy et al. 1984) and when background frequencies approach the dominant frequencies contained in the signal (Van der Tweel et al. 1980). We attempted to assess the accuracy of the template procedure in determining peak latencies by mixing a calibrated positive half sine wave (2.5 Hz), representing a 'P3' component, with a series of unsynchronized ongoing 10 μV sine waves of either 7, 10, or 13 Hz (the 'background' EEG). The amplitude of the model P3 was varied in steps from 5 to 25 μV and was added to the simulated trials of background activity at a fixed latency corresponding to 300 msec. A P3

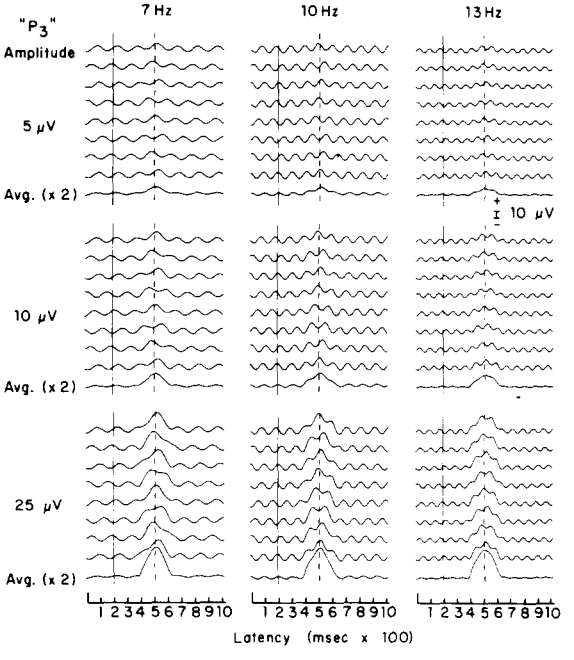

Fig. 1. Sample trials of unsynchronized sine wave background mixed with a simulated 'P3' component used in template program test. A conventional average to the sample trials is shown below each set of sine waves. The dashed vertical line references 300 msec after a simulated stimulus onset indicated by the solid vertical line at 200 msec.

TABLE I

Results of single-trial template analysis using a simulated 'P3' peaking at 300 msec mixed in an unsynchronized sine wave background. $\bar{X}$ = mean latency (msec); S.D. = standard deviation; r = average template correlation coefficient.

'P3' signal-to-background ratio	Frequency								
	7 Hz			10 Hz			13 Hz		
	$\bar{X}$	S.D.	r	$\bar{X}$	S.D.	r	$\bar{X}$	S.D.	r
2.50 : 1	295.8	1.2	0.93	299.7	4.2	0.90	303.8	3.3	0.95
1.75 : 1	304.2	2.5	0.87	298.6	9.8	0.83	295.9	5.1	0.91
1.00 : 1	304.1	6.7	0.71	294.4	21.9	0.71	295.6	12.0	0.81
0.50 : 1	302.1	32.4	0.54	294.1	26.4	0.57	294.2	17.7	0.65

template was derived by conventional averaging (n = 60 trials) for each of the 'P3' signals in each of the background frequencies. The template program swept a 250 msec window of the 1 sec epochs beginning before the fixed latency 'P3' and searched for the signal in the background. This method allowed the testing of the template program in detecting a model P3-like potential embedded in ongoing-like EEG at several signal-to-background ratios and several background frequencies.

The findings illustrated in Fig. 1 and summarized in Table I show that the means computed by the single-trial procedure were, on the average, within 10 msec of the fixed value of 300 msec for all the combinations of conditions tested. Variability increased, however, as signal amplitude decreased and a standard deviation as large as ±32 msec was measured with a signal-to-background ratio of 0.5 : 1 in a 7 Hz background. Correlation values of 0.7 and above reflected favorable signal-to-background ratios for determining peak latencies and were within the range of correlation values (0.65–0.75) obtained for the P3 single-trial analysis of normal subjects.

Results

Conventional averages

Before proceeding to the single-trial results it may be useful to provide a descriptive picture of component variation derived from conventional measures. Fig. 2 shows the conventional ERP averages for all of the subjects to target tones. The P3 peak, indicated by the slanting arrow in each of the subject averages for the Pz electrode, occurred at a latency from 277 to 363 msec. The superimposed records of the same conventional averages portray component variation between individuals by showing the variability for the sequence of peaks in the ERP. In the superimposed records there is a relative increase in component dispersion from the earlier N1 and P2 peaks to the later N2 and P3 peaks. The grand averages to the targets computed from the averages of normal subjects are shown at the bottom of Fig. 2.

Peak latencies determined from these conventionally averaged wave forms can be used to compute measures of peak variation between subjects. Table II presents mean component latencies and standard deviations obtained from the conventional averages.

Single-trial measures

Component variability within an individual can be appreciated in the sample of single trials to target tones shown in Fig. 3 for one representative

TABLE II

Mean latencies (msec) and standard deviations (in parentheses) computed from peak latencies derived from conventional averages.

Electrode	Components			
	N1	P2	N2	P3
Fz	97.9 (13.8)	180.0 (11.5)	238.4 (25.6)	335.9 (21.6)
Cz	96.7 (10.3)	172.3 (19.7)	231.3 (26.6)	336.3 (25.1)
Pz	92.3 (8.2)	171.0 (14.7)	227.1 (25.7)	337.3 (24.5)

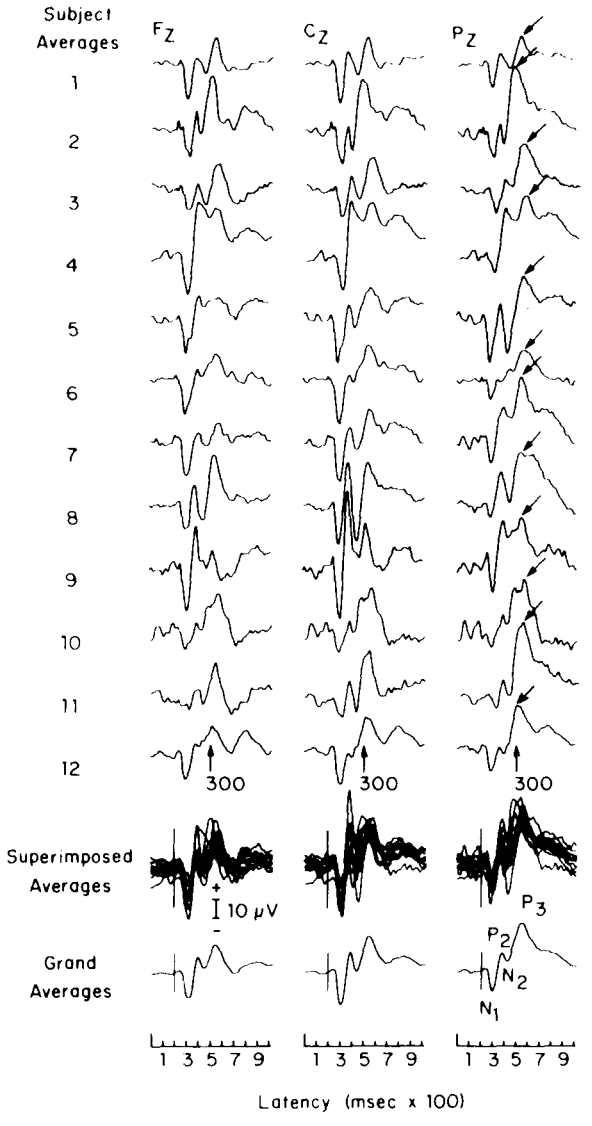

Fig. 2. Conventional ERP averages for the 12 subjects to target tones. The P3 component for Pz is identified by a slanting arrow for each subject. The grand average (n = 12) is shown below the overlayed averages. Stimulus onset is indicated in the superimposed and grand averages by a short vertical line 200 msec after the baseline period. The vertical arrows below the individual averages reference 300 msec after stimulus onset.

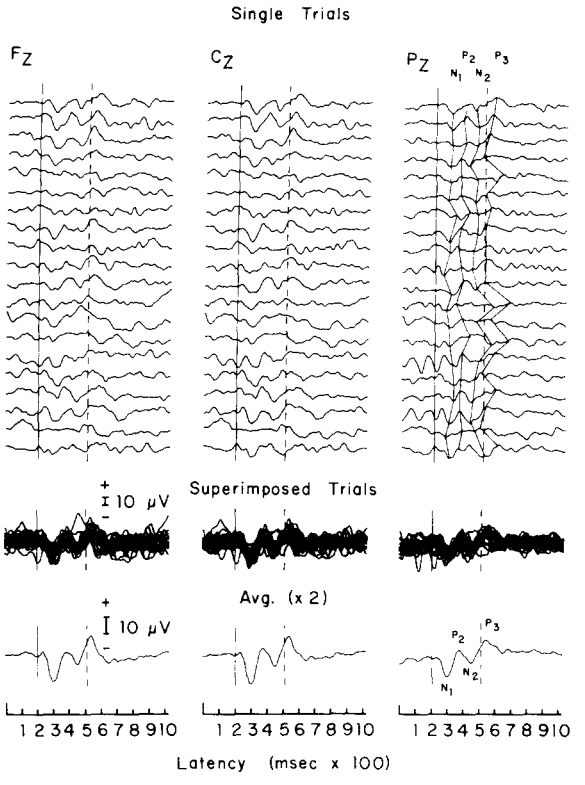

Fig. 3. Sample of single trials for one subject to target tones. Peaks identified by the template procedure are connected by vertical lines between trials for Pz. Conventional averages to the sample single trials are shown below the superimposed sample trials. Stimulus onset and the end of the baseline period are indicated by the solid vertical line at 200 msec; the dashed vertical line indicates 300 msec after stimulus onset.

subject (no. 1 of Fig. 2). For each trial the peaks identified as N1, P2, N2 and P3 by the template procedure are connected by vertical lines drawn between the trials. Latency variation of the components is indicated by the relative size of the displacement of the connected lines. When the same sample of single trials was superimposed as shown in the middle tracings of Fig. 3, the separate components were evident and showed an increase in component dispersion from N1 and P2 to the later N2 and P3 peaks. Conventional averages to the sample of single trials shown are displayed at the bottom of Fig. 3.

Peak latencies derived from single trials were used to compute measures of component variation. Table III presents grand mean peak latencies and standard deviations for each component. Grand means for each component were computed by averaging the means of the single-trial latencies for each subject. Standard deviations for each compo-

TABLE III

Grand mean latencies (msec) and standard deviations (in parentheses) computed from peak latencies derived from single trials.

Electrode	Components			
	N1	P2	N2	P3
Fz	89.6 (17.0)	174.9 (22.0)	248.5 (46.3)	335.5 (52.1)
Cz	88.7 (15.6)	167.2 (21.0)	247.5 (48.1)	345.3 (50.2)
Pz	90.8 (19.0)	171.1 (23.3)	243.8 (45.4)	327.4 (46.3)

nent were derived by averaging the single-trial standard deviations of each subject. The component means computed from single trials (Table III) were, in general, comparable to the component means determined from conventional ERP averages (Table II). Average standard deviations computed from single trials (Table III) were only slightly larger for N1 and P2 but were approximately twice as large for N2 and P3 compared to the respective standard deviations computed from conventional averages (Table II). Some sense of the range of component dispersion within and between subjects can be gained from the graphic representation shown in Fig. 4. For each subject, horizontal lines representing one standard deviation were drawn about the single-trial mean latencies of each component. Inspection of this figure suggests a pattern of increased component variation from N1 and P2 to the later N2 and P3 components of the ERP.

The distributions of the earlier N1 and P2 components were slightly skewed in a positive direction; for the later N2 and P3 components, the distributions were moderately skewed in a positive direction. Latency histograms and related summary statistics are shown in Fig. 5 for comparison purposes. The median is sometimes the preferred measure of central tendency when skewed distributions are suspected. The similarity of the median and mean measures for N1 and P2 reflect the approximately symmetrical shape of the distributions. The medians for the moderately asymmetri-

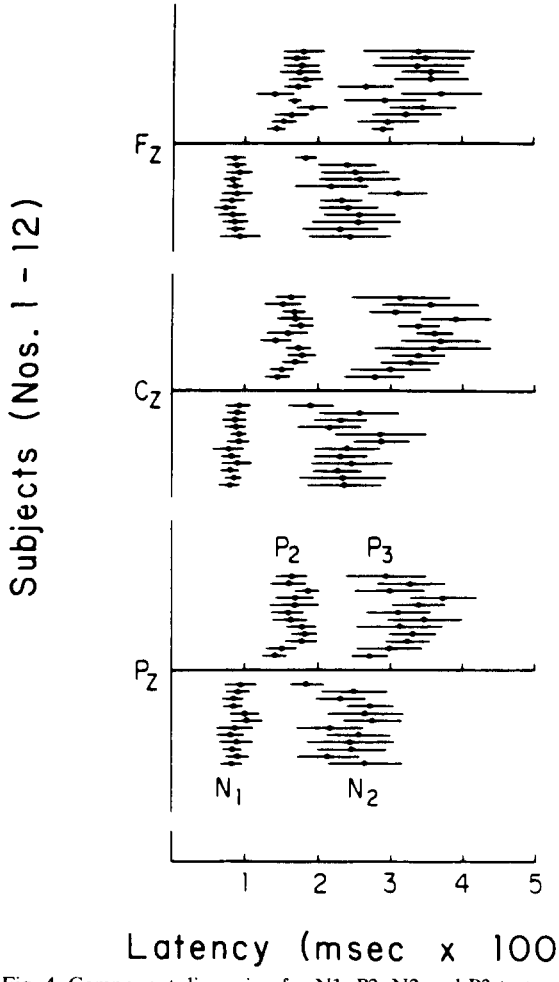

Fig. 4. Component dispersion for N1, P2, N2 and P3 to target tones for each subject tested. The ±1.0 S.D. lines were drawn around the mean single-trial latency of each subject.

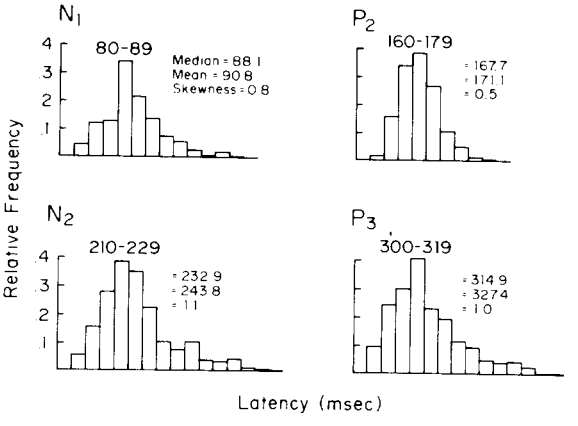

Fig. 5. Latency histograms of single-trial latencies to target tones for N1, P2, N2 and P3 at the Pz site for the group of 12 subjects. The most frequently occurring latencies are represented by the tallest bar in each histogram, and the median and mean latency are shown for each component histogram.

TABLE IV

Component amplitudes (μV) and standard deviations before and after latency adjustment.

Component	Electrode					
	Fz		Cz		Pz	
	Before	After	Before	After	Before	After
N1	−9.2 (2.7)	−10.2 (2.3)	−11.6 (3.0)	−12.8 (2.5)	−6.9 (2.6)	−8.2 (2.4)
P2	6.0 (5.2)	8.8 (4.5)	7.6 (8.2)	11.7 (7.4)	6.8 (5.4)	10.6 (4.8)
N2	0.1 (4.8)	−3.8 (3.6)	−2.0 (5.6)	−5.8 (5.0)	1.5 (5.9)	−1.8 (5.9)
P3	9.7 (5.8)	13.2 (5.1)	12.7 (4.3)	15.4 (4.4)	15.7 (5.5)	18.6 (5.5)

cal distributions of N2 and P3 latencies were shorter by approximately 10 msec than the corresponding means obtained from single trials.

Component amplitudes measured from averages before latency adjustment and after latency adjustment are summarized in Table IV. Analyses of before and after amplitudes for each component indicated significant overall increases after adjustment across the scalp. The average increase in amplitude for N1 was approximately 1.2 μV, while the increases for P2, N2 and P3 were approximately 3.0 μV.

Magnitude differences in component variation

Inspection of the relative magnitude of the standard deviations in Table III suggested that the variances of the components may differ. A series of t tests for related variances compared average variances derived from single-trial latencies for all component combinations. The results indicated that the variances of N1 and P2 were not different from each other; the variances of N2 and P3 were also not different from each other. However, the N2 and P3 variances were separately larger than the variances of either N1 or P2 ($P < 0.01$). This pattern of variance differences among components was the same for all 3 recording locations, and paralleled the impression of greater component dispersion for N2 and P3 than for N1 and P2 suggested by the overlayed ERP averages (Fig. 2) and the sample of single trials (Fig. 3).

Temporal relationships among components

The relationships between the latency of P3 and the latency of the other peaks were evaluated on a trial-by-trial basis in a series of regression analyses. Predictors for P3 included N1, P2 and N2 considered separately or in combination. A similar analysis strategy using P2 and N1 to predict N2, and N1 to predict P2 latencies was also performed. The multiple correlations (R) and the proportion of variance (R^2) for the various predictors and predictor combinations are summarized in Table V. Significant coefficients within a particular regression are also indicated. From 55% to 61% of the variance of P3 was accounted for by N2 across the scalp (line 3 of Table V). The addition of N1 or P2, or N1 and P2 combined, contributed little to the overall prediction of P3 latency. For the prediction of N2 latencies, P2 accounted for 9–16% of the N2 variance, and N1 accounted for only approximately 1% of the variance of N2. For the prediction of P2 latencies, N1 accounted for 8–12% of the variance in P2 latency. Differences in the correlation values among electrode sites did not reach significant levels. Among the P3 predictor combinations tested across subjects, the correlations containing N2 were all significantly larger than other correlation combinations. The correlations between N1 and P3 were smaller than for any of the other correlation combinations. For N2, the correlations between N1 and N2 were smaller than any of the other correlations. The bivariate scatterplots shown in Fig. 6 for Pz pictorially present the strong relationship between N2 and P3 latencies, along with the lesser associations of P2 latencies and P3 latencies, and N1 latencies and P3 latencies found in the regression analyses. The scatterplots for N2 and P2 latencies, and for P2 and N1 are also shown in Fig. 6.

Since some positive skewing of the distributions for N2 and P3 latencies was indicated (Fig. 5), a

TABLE V

Results of a series of regression analyses testing the relationship between P3 latency on a trial-by-trial basis and the latencies of the other peaks. Predictors for P3 included N1, P2, and N2 considered separately or in combination. Similar analyses using preceding peaks to predict N2 and P2 latencies were also computed. The multiple correlation (R) and proportion of variance (R^2) accounted for by the various predictors are summarized in the table.

Predictor(s)						Electrode					
						Fz		Cz		Pz	
						R	R^2	R	R^2	R	R^2
$N1^{3}$	+	$P2^{12}$	+	$N2^{123}$	→ P3	0.75 *	0.56	0.77 *	0.60	0.78 *	0.61
		$P2^{12}$	+	$N2^{123}$	→ P3	0.75 *	0.56	0.77 *	0.59	0.78 *	0.61
				$N2^{123}$	→ P3	0.74 *	0.55	0.77 *	0.59	0.78 *	0.61
		$P2^{123}$			→ P3	0.34 *	0.11	0.30 *	0.09	0.33 *	0.11
N1					→ P3	0.06	0.00	0.05	0.00	0.04	0.00
N1	+	$P2^{12}$			→ P3	0.33 *	0.11	0.31 *	0.09	0.34 *	0.11
N1	+			$N2^{1}$	→ P3	0.74 *	0.55	0.77 *	0.59	0.78 *	0.61
N1	+	$P2^{123}$			→ N2	0.33 *	0.11	0.30 *	0.09	0.39 *	0.16
$N1^{23}$					→ N2	0.11	0.01	0.12	0.01	0.12	0.01
		$P2^{123}$			→ N2	0.33 *	0.11	0.30 *	0.09	0.39 *	0.16
$N1^{123}$					→ P2	0.29 *	0.08	0.34 *	0.12	0.28 *	0.08

* $P < 0.01$.

[123]Significant coefficients at $P < 0.05$, or better; [1] = Fz, [2] = Cz, [3] = Pz.

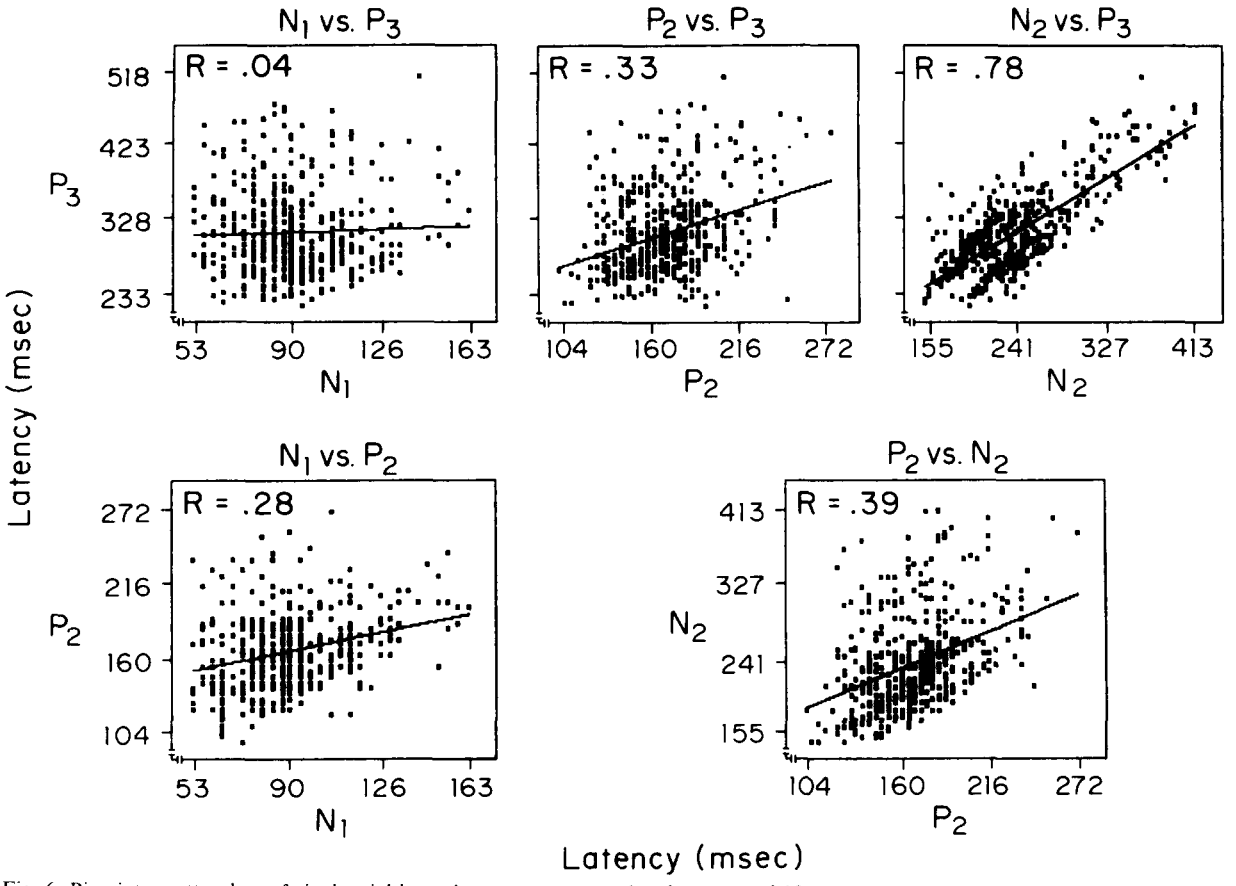

Fig. 6. Bivariate scatterplots of single-trial latencies to target tones for the group of 12 subjects at the Pz site. Note that the largest correlation occurred between N2 and P3.

separate set of regression analyses using transformed data to normalize the component distributions was performed in order to determine whether the skewness of the distributions affected the correlations obtained. There were no differences in the results of the transformed data compared to the untransformed data for any of the component combinations.

Some consequences of the lack of temporal coupling between components are depicted in Fig. 7. Event-related potentials to targets averaged on the basis of particular component latencies are shown superimposed for a single individual for the 3 midline electrode locations. On the left, the averages were computed on a group of 'fast' N1 latencies (light trace, trials with N1 latencies greater than 0.5 S.D. below the mean) and compared to a group of 'slow' N1 latencies (bold trace, trials with N1 latencies greater than 0.5 S.D. above the mean) at Pz. While a difference of approximately 39 msec was measured between N1 peaks for the fast and slow averages, the latencies of both N2 and P3 for each average were comparable (within 8 msec). Conversely, the averages shown on the right of Fig. 7 were computed from 'fast' (light trace) and 'slow' P3 trials (bold trace). A P3 latency difference of approximately 82 msec between these two averages showed only negligible differences in the latencies of N1 and P2. These re-computed averages reflect the low association found between N1 and P3 latencies defined in the regression analyses.

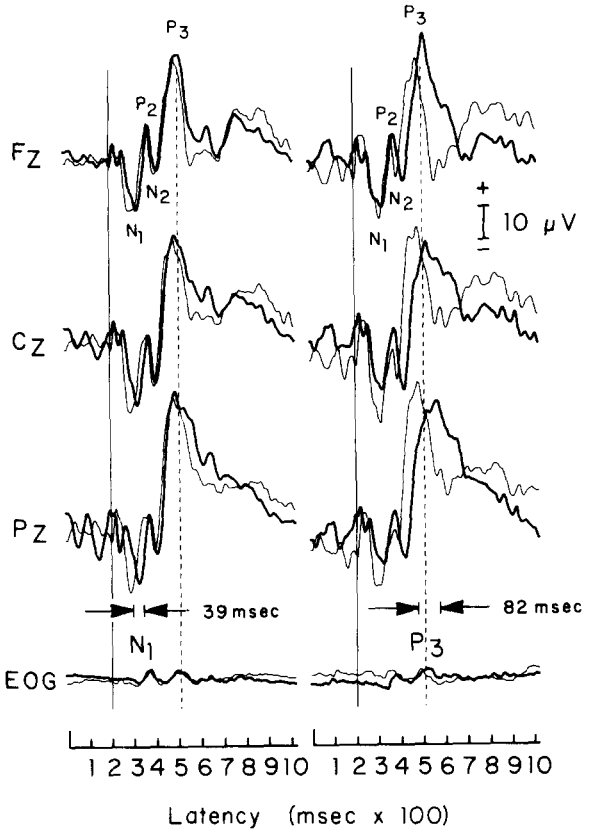

Fig. 7. ERPs to target tones for one subject re-averaged (conventional) on the basis of single-trial latencies at Pz. Left panel averages were derived from fast (light trace) and slow (bold trace) N1 latencies. Right panel averages were derived from fast (light trace) and slow (bold trace) P3 latencies. Note that the latencies of P3 in the left panel were approximately the same between the averages while N1 latencies differed by 39 msec; conversely, P3 latencies in the right panel differed by 82 msec while the differences in N1 latencies were negligible. Twenty-two trials make up each of the averages in the left panel and 15 trials make up the averages in the right panel.

Reaction time measures

Reaction times (RTs) to correctly detected targets for the entire group ranged from 207 msec to 400 msec with a mean of 299.6 and a standard deviation of 48.3. The distribution of RTs for the group of 12 subjects was moderately skewed in a positive direction. The average error rate for detecting target tones for the group was less than 5%.

Correlations between component latencies determined from single trials and corresponding RTs for the group were significant but of low magnitude. For example, at Pz the correlation (r) between N1 and RT was 0.12 ($P < 0.03$), between P2 and RT was 0.14 ($P < 0.01$), between N2 and RT was 0.33 ($P < 0.001$), and between P3 and RT was 0.24 ($P < 0.001$). Tests of the correlations between component latencies and RTs were not different except for the correlations between N2 and RT which were larger than between N1 and RT. Regression analyses indicated that multiple correlations based on combinations of component latencies to predict RTs were low, never exceeding 0.35 at Pz for instance (N1 + P2 + N2 + P3 → RT), with N2 latencies contributing significantly more to the correlations than the other components. Correlations computed between peak latencies determined from conventional averages and mean

RTs were low and non-significant (the correlations were 0.25, 0.48, 0.39, and 0.02 for N1, P2, N2, and P3, respectively).

Discussion

The results of this study show that the single-trial latency variability of auditory ERPs differ both among the various components and between subjects. Component variation based on single trials yielded larger standard deviations than standard deviations from conventional averages. This difference reflects the measures on which the estimates of component variation were made. The peak selected from the conventional average is itself a kind of summary statistic built upon the accumulation of potentials representing both the more frequently occurring latencies and perhaps larger responses. Intersubject variance measures computed from such conventional ERP averages yield smaller deviations around the mean than the deviations around the mean computed from single trials, which include *all* peak latencies identified within the range of the template window and therefore make the estimate of variation greater.

The mean peak latencies computed from conventional average wave forms and the means computed from single trials were, in general, comparable. However, some allowance for discrepancies between the means calculated from conventional averages and single trials should be made, since the shape of the distribution of single trials will affect the values computed.

The larger component variance measured for the N2 and P3 components than the earlier N1 and P2 components is consonant with the more labile nature of the so-called 'endogenous' peaks presumed to accompany decision-making compared to the 'stimulus-bound' N1 and P2 components. There was also a marked positive skewing of N2 and P3 in the direction of longer latencies. A similar positive skewing of P3 latencies based on an analysis of single trials was shown by Pfefferbaum et al. (1983) during task regimes that emphasized either speed or accuracy of responding. Our single-trial estimates of P3 variation were generally comparable to the variation measured by other investigators also using single trials (e.g., Goodin and Aminoff 1984 reported an S.D. of 23 msec; Pfefferbaum et al. 1984a reported an S.E. of 54 msec).

Alternatively, the differences in variance among components may have been influenced by the procedures used to analyze the single trials. Because unequal-sized temporal windows were used to scan the EEG for each component, the smaller variances measured for N1 and P2 may simply reflect the shorter intervals analyzed compared to N2 and P3. Component windows were selected to prevent the misidentification of a peak of the same polarity within a trial. Extending the template window of N1 to equal that of P3 (250 msec), for example, increases the overall computation of N1 variability primarily due to the inclusion of N2 peaks or negative waves in the EEG. Even extending the temporal window of P3, say to 500 msec, increases measures of P3 variation as other later positive peaks or other features resembling P3 in the EEG may be selected. If only one positive component is known to be present in the single trial, as in the test of the correlational procedure shown in Fig. 1, the window searching for P3 can extend throughout the sweep without affecting estimates of variability. The choice of temporal windows and the placement of the template starting point may also influence the shape of the component distributions as wider windows tend to include peaks at longer latencies and extend the tail regions of the curve. However, examination of component dispersion, as represented in the single trials shown in Fig. 3, suggests that the restrictions placed on the temporal windows were probably reasonable as a practical resolution of estimating component variability. The problem of pattern recognition techniques in general was aptly stated by Weinberg (1978): 'The limitation of all methods is the inability of these methods to identify patterns within the time domain without having *a priori* knowledge of exactly when these patterns occurred' (p. 600). The use of unequal template windows is a limitation of the procedure dictated by the constraints of the correlational technique and the properties of the ERP wave form.

The temporal interrelationships found among the sequence of components in the ERP support a

distinction between the earlier N1 and P2, and the later N2 and P3 component complexes. For example, the generally low correlations between earlier and later components (e.g., N1 vs. P3) suggest independent or even perhaps parallel processing phases. The moderate correlations found between N1 and P2 may similarly reflect distinct processing stages within the sensory-perceptual complex.

There is other evidence based on conventional averages for a distinction between N1 and P3 which can be drawn upon from the literature in support of our results. For example, the lack of a relationship between N1 and P3 has been noted in studies of aging in which P3 latencies are delayed with advancing age while the latencies of N1 change very little (in the auditory modality, e.g., Goodin et al. 1978; Pfefferbaum et al. 1980; Picton et al. 1984; in the visual modality, e.g., Beck et al. 1980). Furthermore, the latency of P3 is relatively insensitive to changes in stimulus intensity, whereas significant latency shifts with intensity changes can occur for N1 (reported in Squires et al. 1980). The stimulus omission paradigm (Picton and Hillyard 1974) elicits a late N2-P3 complex to an expected but missing stimulus. This emitted response lacks earlier components since it occurs independently of a specific evoking stimulus. Additional evidence to support the distinction between N1 and P3 has been provided by other investigators (Hillyard et al. 1978).

The strong covarying relationship found between N2 and P3 indicated that up to 61% of the variance in P3 latency could be accounted for by N2 at Pz. Similar single-trial relationships between N2 and P3 in a visual emitted potential paradigm were reported by Renault (1983). These findings suggest that many of the psychological attributes that are used to characterize P3 (e.g., stimulus evaluation, categorization, task complexity) may apply to N2 as well. In fact, Ritter and his colleagues (Ritter 1978; Ritter et al. 1979) defined a covariation of N2 and RT. These investigators found that the correlations between N2 and RT in many instances were as large or larger than between P3 and RT. Our single-trial correlations between N2 and RT were also slightly larger than the correlations between P3 and RT. Our correlations between component latencies and RTs, based either on single trials or conventional averages, were generally low or non-significant. In relatively simple task situations that emphasize speed, it is not unusual for RTs to be comparable to P3 latencies or shorter and would account for the low correlations between RT and P3. In more demanding task conditions, the relation of the later component latencies and RT appears to increase (McCarthy and Donchin 1983). Since N2 precedes P3 temporally in the sequence of components, the role of N2 in stimulus selection processes takes on added significance. Like P3, regression analyses indicated that N1 and P2 latencies accounted for only a small proportion of the N2 variance. Although highly speculative at this time, other peaks may exert an influence on the prediction of N2 latency such as the P165 peak (Goodin et al. 1983; Goodin and Aminoff 1984). Thus, with the possible exception of P165, the preceding peaks in the auditory ERP are relatively poor predictors of the latencies of N2 and P3.

Our own tests of the template procedure using the model 'P3' indicated that estimated peak latencies were more variable and that template correlations declined as signal amplitudes approached background levels of activity. The standard deviations for the model 'P3' were estimated to range between 6.7 and 21.9 msec for correlation values near 0.70 (Table I). This suggests that our standard deviations for P3 based on single-trial correlations may be in error by an amount within this range. For the peaks analyzed here in normals, the correlations between the template and components in the single trials averaged between 0.65 and 0.75. These relatively high correlations probably reflect favorable signal-to-background ratios overall but are also likely to include correlations below the average value. Instances of low correlations may indicate a poor signal-to-background ratio in a particular trial. It is possible that distribution differences for a peak, such as the more parietal distribution of P3, may affect signal-to-background ratios and result in different measures of variability for the same peak across the scalp. Our estimates of peak latencies were based on the maximum correlation obtained within a trial regardless of the actual magnitude of the correlation value. In future studies, it may be useful to esti-

mate signal-to-background ratios (e.g., Coppola et al. 1978) so that unfavorable trials are recognized and eliminated, or to adopt a cut-off criterion so that trials with low measures of association between the template are not included (see Pfefferbaum et al. 1984b).

Considerations of signal-to-background ratios apply equally to the other components of the ERP, especially since they are likely to be of smaller amplitude than P3. It could be argued that the low correlations between N1 and P3, for example, may have resulted because of a poor sensitivity of the template procedure in accurately determining the smaller N1 components from the background EEG. We do not think that this was the case. First, we regularly checked the peaks identified by the template program with single trial plots to verify that random points were not being selected (Fig. 3). Second, sorting of the single trials on the basis of peak latency and then averaging by conventional methods as was illustrated in Fig. 7 for fast and slow N1 peaks provides another check on the template method of defining peak latencies. Third, the average correlation between N1 and the single trials was 0.70 which indicated adequate signal-to-noise background ratios for peak detection. Nevertheless, it is important to recognize that the template technique can complement but not replace other methods of peak detection or signal extraction for examining properties of the ERP.

It is likely that component variations and interrelationships will be affected by stimulus variables, modality, and task demands. A knowledge of peak latencies from trial-to-trial is useful in understanding latency variations and clarifying component interrelationships during information processing. Such results demonstrate dynamic changes in ERP components not evident with standard averages.

Résumé

Variabilité des latences et relations temporelles dans les potentiels auditifs liés à l'événement (N1, P2, N2, et P3) chez des sujets normaux

La variation du pic de latence et les relations temporelles pour le potentiel auditif lié à l'événement ont été étudiées chez 12 adultes normaux âgés de 28 à 42 ans. Les mesures des variations étaient basées à la fois sur des moyennages conventionnels et sur des essais isolés. Les estimations des latences de N1, P2, N2 et P3 ont été effectuées sur la base d'essais isolés en réponse à des stimulus cible à partir d'enregistrements effectués sur le scalp en Fz, Cz et Pz.

Les résultats ont montré que la variabilité des latences du PLE en essais isolés différait tout à la fois entre composantes et entre sujets. Les écart-types étaient plus grands pour les composantes tardives N2 et P3 que pour les composantes précoces N1 et P2. Les analyses de régression entre les latences de diverses composantes ont indiqué une forte relation de covariance entre N2 et P3, avec N2 rendant compte jusqu'à 61% de la variance de la latence de P3 à Pz. Les composantes précoces, N1 et P2 n'ont que peu ajouté à la prévision totale concernant P3 et N2. Si l'on considère la variance de N2, P2 rendait compte de 9 à 16% de cette variance alors que N1 n'intervenait que pour 1%; si l'on considère celle de P2, N1 intervenait pour 8 à 10%. Les corrélations étaient positives entre les latences de pic en essais isolés et RT, mais de faible amplitude. Les plus hautes corrélations entre pic de latence et RT étaient trouvées pour N2 ($r = 0{,}33$) et P3 ($r = 0{,}24$).

Les faibles corrélations entre les latences de N1 et P3 pour des essais isolés suggèrent que les traitements traduits par ces composantes sont indépendants, et confortent la distinction entre latences les plus précoces et les plus tardives du PLE. L'étroite liaison temporelle entre N2 et P3 suggère que N2 pourrait refléter les propriétés cognitives communes avec P3 dans les processus d'évaluation du stimulus.

The authors would like to thank Dr. Julie V. Patterson for her helpful comments.

References

Beck, E.C., Swanson, C. and Dustman, R. Long latency components of the visually evoked potential in man: effects of aging. Exp. Aging Res., 1980, 6: 523–545.

Coppola, R., Tabor, R. and Buchsbaum, M.S. Signal to noise ratio and response variability measurements in single trial

evoked potentials. Electroenceph. clin. Neurophysiol., 1978, 44: 214–222.

Donchin, E. Event-related brain potentials: a tool in the study of information processing. In: H. Begleiter (Ed.), Evoked Brain Potentials and Behavior. Plenum Press, New York, 1979: 13–88.

Donchin, E., Ritter, W. and McCallum, W.C. Cognitive psychophysiology: the endogenous components of the ERP. In: E. Callaway, P. Tueting and S.H. Koslow (Eds.), Event-Related Brain Potentials in Man. Academic Press, New York, 1978: 349–411.

Goodin, D.S. and Aminoff, M.J. The relationship between the evoked potential and brain events in sensory discrimination and motor response. Brain, 1984, 107: 241–251.

Goodin, D.S., Squires, K.C., Henderson, B.H. and Starr, A. Age-related variations in evoked potentials to auditory stimuli in normal human subjects. Electroenceph. clin. Neurophysiol., 1978, 44: 447–458.

Goodin, D.S., Squires, K.C. and Starr, A. Variations in early and late event-related components of the auditory evoked potential with task difficulty. Electroenceph. clin. Neurophysiol., 1983, 55: 680–686.

Hillyard, S.A., Picton, T.W. and Regan, D. Sensation, perception and attention: analysis using ERPs. In: E. Callaway, P. Tueting and S.H. Koslow (Eds.), Event-Related Brain Potentials in Man. Academic Press, New York, 1978: 223–321.

Keppel, G. Design and Analysis: A Researcher's Handbook. Prentice-Hall, Englewood Cliffs, NJ, 1973.

Magliero, A., Bashore, T.R., Coles, M.G.H. and Donchin, E. On the dependence of P300 latency on stimulus evaluation processes. Psychophysiology, 1984, 21: 171–186.

McCarthy, G. and Donchin, E. Chronometric analysis of information processing. In: A.W.K. Gaillard and W. Ritter (Eds.), Tutorials in ERP Research: Endogenous Components. Elsevier/North-Holland, Amsterdam, 1983: 251–268.

McCarthy, G., Squires, K.C. and Schvaneveldt, R. Stimulus evaluation time and P300 latency. In: E. Donchin (Ed.), Cognitive Psychophysiology: Event-Related Potentials and the Study of Cognition. Lawrence Erlbaum, Hillsdale, NJ, 1984: 249–301.

Pfefferbaum, A., Ford, J.M., Roth, W.T. and Kopell, B.S. Age-related changes in auditory event-related potentials. Electroenceph. clin. Neurophysiol., 1980, 49: 266–276.

Pfefferbaum, A., Ford, J., Johnson, R., Wenegrat, B.G. and Kopell, B.S. Manipulation of P3 latency: speed vs. accuracy instructions. Electroenceph. clin. Neurophysiol., 1983, 55: 188–197.

Pfefferbaum, A., Ford, J.M., Wenegrat, B.G., Roth, W.T. and Kopell, B.S. Clinical applications of the P3 component of event-related potentials. I. Normal aging. Electroenceph. clin. Neurophysiol., 1984a, 59: 85–103.

Pfefferbaum, A., Wenegrat, B.G., Ford, J.M., Roth, W.T. and Kopell, B.S. Clinical application of the P3 component of event-related potentials. II. Dementia, depression and schizophrenia. Electroenceph. clin. Neurophysiol., 1984b, 59: 104–124.

Picton, T.W. and Hillyard, S.A. Human auditory evoked potentials. II. Effects of attention. Electroenceph. clin. Neurophysiol., 1974, 36: 191–199.

Picton, T.W., Hillyard, S.A., Krausz, H.I. and Galambos, R. Human auditory evoked potentials. I. Evaluations of components. Electroenceph. clin. Neurophysiol., 1974, 36: 179–190.

Picton, T.W., Stuss, D.T., Champagne, S.C. and Nelson, R.F. The effects of age on human event-related potentials. Psychophysiology, 1984, 21: 312–325.

Pritchard, W.S. Psychophysiology of P300. Psychol. Bull., 1981, 90: 506–540.

Renault, B. The visual emitted potentials: clues for information processing. In: A.W.K. Gaillard and W. Ritter (Eds.), Tutorials in ERP Research: Endogenous Components. Elsevier/North-Holland, Amsterdam, 1983: 159–175.

Renault, B. and Lesèvre, N. A trial by trial study of the visual omission response in reaction time situations. In: D. Lehmann and E. Callaway (Eds.), Human Evoked Potentials. Plenum, New York, 1979: 317–329.

Renault, B., Ragot, R., Lesèvre, N. and Rémond, A. Onset and offset of brain events as indices of mental chronometry. Science, 1982, 215: 1413–1415.

Ritter, W. Latency of event-related potentials and reaction time. In: D.A. Otto (Ed.), Multidisciplinary Perspectives in Event-Related Brain Potential Research. U.S. Environmental Protection Agency, Washington, DC, 1978: 173–174.

Ritter, W., Simson, R., Vaughan, H.G. and Friedman, D. A brain event related to the making of a sensory discrimination. Science, 1979, 203: 1358–1361.

Ruchkin, D.S. and Sutton, S. Latency characteristics and trial-by-trial variations of emitted cerebral potentials. In: J.E. Desmedt (Ed.), Cognitive Components in Cerebral Event-Related Potentials and Selective Attention. Progress in Clinical Neurophysiology, Vol. 6. Karger, Basel, 1979: 106–118.

Squires, K.C., Chippendale, T.J., Wrege, K.S., Goodin, D.S. and Starr, A. Electrophysiological assessment of mental function in aging and dementia. In: L.W. Poon (Ed.), Aging in the 1980s. American Psychological Association, Washington, DC, 1980: 125–134.

Van der Tweel, L.H., Estevez, O. and Strackee, J. Measurement of evoked potentials. In: C. Barber (Ed.), Evoked Potentials. University Park Press, Baltimore, MD, 1980: 19–41.

Weinberg, H. Comments on methods of signal analysis and signal detection. In: D.A. Otto (Ed.), Mutidisciplinary Perspectives in Event-Related Brain Potential Research. U.S. Environmental Protection Agency, Washington, DC, 1978: 593–600.

Weinberg, H. and Cooper, R. The Recognition Index: a pattern recognition technique for noisy signals. Electroenceph. clin. Neurophysiol., 1972, 33: 608–613.

Woody, C.D. Characterization of an adaptive filter for the analysis of variable latency neuroelectric signals. Med. biol. Engng, 1967, 5: 539–553.