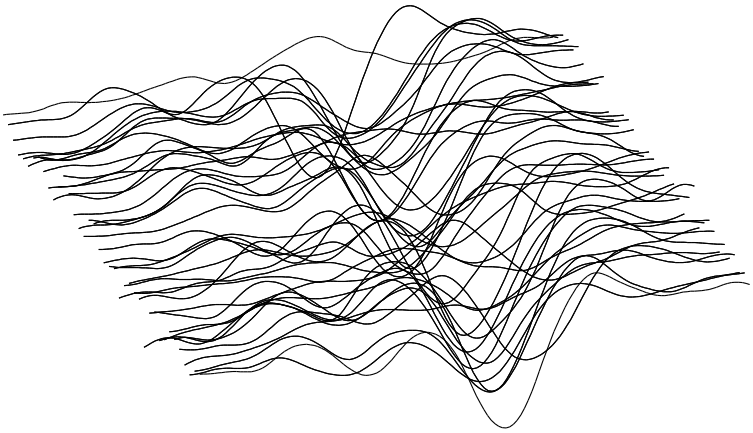
Variability of the event-related single sweep potentials

Anu Holm



Licentiate of Philosophy thesis
2004

University of Kuopio
Department of Applied Physics

UNIVERSITY OF KUOPIO, Faculty of Natural and Environmental Sciences,
Department of Applied Physics, and,
Finnish Institute of Occupational Health, Section of Clinical Neurosciences

Medical Physics

HOLM ANU S. J.: Variability of event-related single sweep potentials
Phil. Lic. thesis, 41 pages,

Supervisors:

Jari Kaipio, Ph.D., Professor
Pasi Karjalainen, Ph.D., Research director
Juhani Partanen, MD. Ph.D., Professor

August 2004

Keywords: Event related potentials, single trial, P300, sequential effect, habituation, amplitude variation, latency variation

The common approach in research on event-related brain potentials (ERP) is to assume that for repeated stimuli the responses are exactly the same. However, when measuring the ERPs, the responses are corrupted by ongoing brain activity (electroencephalography, EEG). Thus, to extract the ERP from the background electroencephalography, dozens of responses are averaged. However, the psychophysiological state of the subject may vary during the measurement session.

In this study, the variability of widely reported ERP, called P300, is studied at single-trial level with a method that is based on Bayesian estimation and regularization theory. The background of the P300 component, it's variability and commonly used analysis methods as well as the background of the used single trial method are presented briefly.

The variability of the P300 was studied in 12 clinical measurements. The results of the single trial method were compared with those of the conventional averaging method. In short measurement sessions (i.e., 10 min), the effect of sequential changes in the stimulus train is more pronounced than the effect of the time on the task both on the amplitude and the latency of the P300. The effect of the changes in the stimulus train is more clear in the latency than in the amplitude of the P300.

Application of the single-trial method to clinical data confirmed some previously reported results and revealed information about the dynamical changes of the P300 that are lost with the conventional averaging method. The evaluated single-trial method can be used to model changes in dynamical behavior of the P300 component and regarded as an alternative to other methods.

This thesis is available on http://it.uku.fi/biosignal

Acknowledgments

This study was carried out in the BrainWork Laboratory of the Finnish Institute of Occupational Health (FIOH) as part of project partially funded by the National Technology Agency (Tekes).

I want to express my gratitude to my supervisors, Research director Pasi Karjalainen, Ph.D., Professor Jari Kaipio, Ph.D., and Professor Juhani Partanen, M.D., Ph.D., for their guidance and advices during this work.

I thank the official reviewers Research director Pasi Karjalainen, Ph.D., and Medical Physicist Ari Pääkkönen, Ph.D., for reviewing the manuscript and giving constructive criticism.

I wish to express my deepest gratitude to Research Chief Medical Officer of the BrainWork laboratory in FIOH, Kiti Müller, M.D., Ph.D., for the opportunity to carry out this work in the laboratory. I also thank her for all long discussions which have helped me to clarify the focus of this thesis.

The comments and encouragement given by Mikael Sallinen, Ph.D., and Joel Hasan, M.D., Ph.D., are also greatly appreciated.

I thank Mika Tarvainen, Ph.D., and Perttu Ranta-aho, M.Sc., for their skillful collaboration in Tekes -project and for developing Biosignal analysis software. In addition, Perttu's help with the estimation method and LaTeX has been valuable.

The personnel of the Department of Clinical Neurophysiology, University Hospital of Kuopio, is greatly appreciated for introducing me the field of clinical neurophysiology.

I wish to thank the colleagues at the Brain and Work Research Units (FIOH) for the unforgettable moments during past five years and for stimulating atmosphere, especially in the 4th floor coffee room.

I warmly thank my friends for helping me get through the difficult times, and for all the emotional support, entertainment, and caring they provided.

My special thanks belong to my family, who have always given me the strength and encouragement to follow my dreams, and have never left me in doubt of their love for me.

Finally, I thank my husband Micke, who has taken care of me during this work and brought to my life happiness that only love can bring.

Espoo, August, 2004

Anu Holm

Contents

Chapter I

Introduction

Electroenchephalography (EEG) measures the spontaneuous electrical activity of the brain. EEG was first measured from human by Hans Berger in 1929. In 1947 Dawson showed that the certain sensory stimuli are followed by specific cortical changes in voltages. Sensory stimuli can be from different modalities: visual, auditory and even odor stimuli can be used. The event related potentials, ERPs, are time-locked changes in voltage in response to either external or internal stimuli. Event-related potentials can be divided into two categories: exogenous and endogenous evoked potentials. Exogenous potentials are determined by the physical characteristics of the stimulus whereas endogenous potentials are stimulus independent and generated solely as a function of the mental processing allocated to the stimulus. Sometimes it is difficult to separate purely sensory potentials from cognitive ones. Thus the term event related potential includes information receiving, transmitting, as well as handling.

Electrical activity of the brain is generated by a potential difference across the membranes of cells in the nervous system. The nervous impulses cause postsynaptical potentials, which reflect effects of the free electrolytes. These postsynaptical potentials (PSPs) are long lasting compared to action potentials and are thus able to summate both temporally and spatially. Postsynaptic potentials rather than action potentials are responsible for scalp recorded EEG and ERP waves. Volume conduction theory describes the extent to which event related potentials are recorded at a distance from their origin. The solid angle theorem of volume conduction states that the potential generated by a dipole layer in a volume conductor measured at any point in the conductor depends upon the generator's potential and the solid angle subtended by the dipole layer at the point of measurement. Thus the potentials recorded at the certain point can be a summation of potentials induced by several separate generators.

ERPs and EEG are usually recorded from scalp. Electrodes are placed according to the International 10-20 system [35]. Intracranial and depth electrodes are used when studying e.g. epilepsy.

Event-related potentials are usually labeled according to their polarity (e.g. positive or negative) and latency measured from the onset of the stimulus. Thus peak labeled N100 is a negative peak at a latency of 100 milliseconds from the stimuli onset and P300 is a positive peak at a latency of 300 ms.

The sensory evoked potentials are widely used as clinical tests. They provide information about the sensory systems that is not available with other methods. The presence of a normal sensory evoked potential is reasonable evidence for the functioning of the peripheral receptor. The absence of the potential indicates malfunction in the sensory end-organ, the generator mechanism of the evoked potential, or the connection between them. Cognitive event related potentials have helped to understand how the normal brain works. Cognitive event related potentials are also used in the studies of cognitive impairment in neurological disorders such as various dementias and mental disorders e.g. schizophrenia. It has been suggested that also cognitive event-related brain potentials (especially P300) could be used as a clinical assay. Intersubject variability in normative P300 amplitude and latency is comparable to standard clinical test procedures like cholesterol and serum rate [50]

Event-related potentials are very small changes in voltage (0.1-30 uV) whereas spontaneous EEG activity is 10-100 uV. So one of the most important problems in ERP studies is signal extraction, i.e., the extraction of the ERP evoked by sensory stimulus from the background EEG. As the signal to noise ratio of event-related potentials is quite poor, averaging dozens of responses is generally used to estimate responses. With averaging it is assumed that the ERP is a deterministic signal, that is, exactly the same for repeated stimuli produced during long lasting measurement sessions. However, this assumption has been shown to be a quite rough estimate and the idea of the trial-to-trial variation of ERP was presented early in ERP studies [2]. As successive single-trial ERPs are averaged, the signal often undergoes continual changes in both the amplitude and the latency of components. It is quite conceivable that such variation carries significant information about the underlying nervous system. This ERP variability has been studied on single-trial level.

Many single-trial methods are not very far from signal averaging. Selective averaging techniques are widely used. Single trials are divided to different classes, e.g. according to the performance and then averaged. Procedures with correlation techniques are also used in attempts to take the trial-to-trial variability into account [63]. In correlation techniques, one obtains a template by averaging a set of ERPs and performs a cross-correlation between a single trial ERP and the template. This is especially useful when studying latencies. Wiener filters and Kalman filters are also used. These filtering techniques are based on the assumption that the background EEG and the evoked potential are sample functions from stationary random processes. Other assumptions lead to

filters of different types such as the minimum mean square error (MMSE). The interdependence between members of a set of waveforms is studied with principal component analysis (PCA). If priori information about the evoked potential waveform is available it can be incorporated into the signal processor to further improve the estimation of single trial ERPs.

In this study, a widely reported event-related potential, called P300, has been chosed for further analysis. The P300 is generated whenever infrequently occurring sensory stimuli are detected or when unexpected and highly deviant stimuli are delivered to the subject. The P300 component has been showed to provide useful information about how the brain discriminates stimuli and evaluates probability. A new single trial method that uses prior information about the responses is used [31]. The method's most important property is its general structure that allows one to develop estimators with more realistic observation models. Two well known phenomena, i.e., the habituation of the P300 component and the effect of preceding stimulus train, are studied in healthy subjects with this new method and with the conventional averaging method. Results of the single trial analysis are compared with those of the conventional averaging.

Chapter II

Review of the literature

2.1 Variability of the P300 component

P300 is one of the most prominent positive peaks occurring around 300–600 ms after the infrequently presented target stimuli, to which the subject is requested to respond by a certain way such as button pressing.

The classic 'oddball task' is a paradigm in which stimuli that can be classified into two categories are presented sequentially. In the auditory oddball paradigm, one of the tones is rare and the other is frequent. Rare tones are designated as a 'target', and the subjects have to recognise these target tones, for example by counting them or by pressing a button. The probability of the target tone is typically 0.15-0.25. The inter-stimulus interval usually varies between one and three seconds and can be either fixed or variable.

A few studies have examined variability of the P300 component under more complex experimental conditions, such as in prediction task [46]. These experiments are more complicated than the oddball task and thus results are not comparable. For that reason results of these experiments are not discussed here.

There is evidence that the P300 phenomenon is elicited both by unpredictable stimulus changes when stimuli are not attended and by those that are being actively attended [58]. The earlier component, P3a (latency 220-280 millisecond), is elicited by infrequent unpredictable changes in either intensity or frequency of tone irrespective of the subject is ignoring or actively attending to the tones. The later component, P3b (latency 310-380 milliseconds), occurrs only when the subject is actively attending to the tones. Thus, P3a is related to the orienting response of the brain with frontocentral distribution and P3b is seen as stimulus evaluation that occurs later and has a centroparietal scalp distribution. For general reviews of the P300 component, see [59, 10, 51, 12, 17].

Originally the P300-component has been seen as a measure of the amount of activity of general-purpose cortical processor, which is invoked on demand of data processing requirements [14]. In later studies a single general-purpose

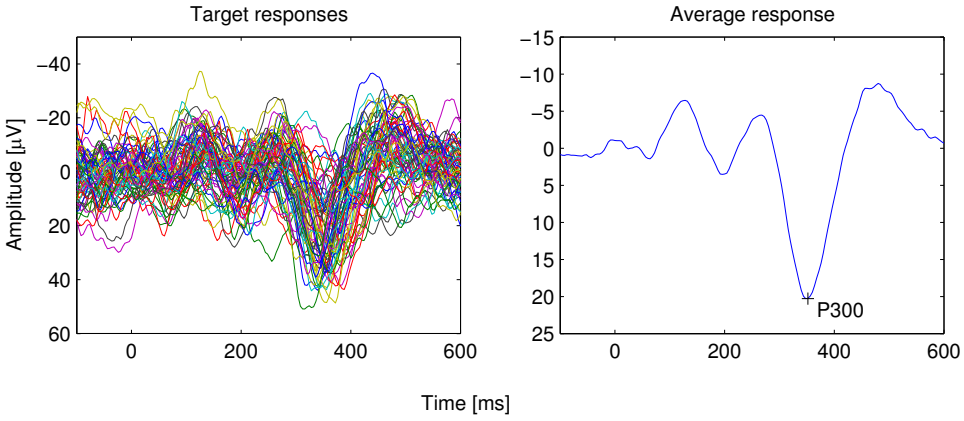


Figure 2.1: The P300 component.

processor has been displaced with mul tiple, specific-purpose cortical processors [27, 52]. It has been suggested, based on intra-cranial recordings, that the scalp recorded P300 has at least two different generators: Medial temporal lobe (MTL: hippocampus, parahippocampal gyrus and amygdala) and frontal lobe [27]. The P300 recorded at any given electrode site represents the volume conducted summation of activity from different neural processors. Each of these specific-purpose generators is related to the processing of a different type of information and are connected to e.g. the modality of a stimulus [52].

Psychological processes are considered as hierarchical steps and ERPs are manifestations of activities related to an elementary level in this hierarchy [11]. The highest level includes interactions between the organism and the environment, whisch, as behaviour, can be described in terms such as motivation, emotion, attention, and learning. These interactions may evoke a broad set of information-processing activities, requiring the activation of psychological processes such as choice, recall, decision or discrimination. Information-processing activities, in turn, require activation of more elementary functions, like mismatch detecting and encoding. Thus P300 is currently not seen as a manifestation of a certain cognitive function per se, rather it is a manifestation of a activation invoked by certain psychological process, like the updating of a cognitive model [11]. The context updating theory has its roots in Sokolov's theory of the orienting response. The orienting-response theory proposes a brain mechanism, which initiates the orienting-response release [55]. The theory proposes that a new stimuli elicits an orienting response which gets weaker whith stimulus repetition, and finally habituates completely. According the theory, habituation occurs because stimulus repetition leads to a gradual development of the neuronal model.

The template-restoration model for the P3 has been suggested by Gonsalvez in [26]. The model states that the P3 is related to the extent of restoration of a template that a stimulus evokes in working memory. The template of the target in memory is assumed to decrease with time in a systematic way. The major distinction between Sokolov's model and the template-restoration model is the importance between temporal versus sequential factors associated with the P300.

Latency of the P300 component is related to stimulus evaluation and categorisation time, transfer of information to 'consciousness' and memory systems, and stimulus saliency. An extensive review on the utility of P300 latency as an index of mental chronometry has been done by Verleger [60]. In this review, it is concluded that only in simple paradigms P300 is involved solely with stimulus processing. More complicated test designs involve P300 also in response selection and increase the P300 latency as a function of the complexity of the cognitive processing in the task. In the study of Kutas et al. [38] the latency of the P300 has been shown to reflect the time required for stimulus evaluation and categorisation and to be relatively independent of response selection and executive processes. The correlation between P300 latency and reaction time varies with the performance accuracy demanded from the subjects [38]. The P300 reflects both the stimulus evaluation and orientation capability of the subject and does not require an overt motor response for its generation [58].

The amplitude of the P300 is related to processing capacity and context updating. The P300 component (especially P3a) has seen as an index of a basic sensory mechanism which register any change in a background stimulus, perhaps by means of mismatching a specific neural "model" established by repetition of the background [56, 12]. In this context a neural model is defined as a schema that exists in the subjects brain about the situation and the P300 is an output of a mismatch detector. The larger the amplitude of the P300, the larger the change in the model. The context updating theorem states that P300 amplitude indexes brain activity that is required when the mental model of the stimulus environment (i.e. the context) is updated. This context updating theorem has its roots in Sokolov's model of the orienting response as it somehow satisfies properties related to the orientation reaction and is related to the processes involved in shifts of attention [12]. For a review on the P300 component as a manifestation of context updating, see [11]. The P300 amplitude is related to the allocation of perceptual resources and could be used for discriminating between levels of mental workload [29, 28]. The utility of P3 amplitude as a measure of processing capacity is reviewed in [36].

Time of day has captured minimal attention when recording event related potentials. Time of day, repeated tests and interblock interval also affect the P300 component [61]. Amplitudes are higher in the morning compared to afternoon

across six times repeated blocks suggesting a circadian variation in cognitive processes underlying P300.

Variability of the P300 component during a measurement session (latency and amplitude) has been widely studied. Main areas of research of intra-subject variability of the P300 component have been within 1) habituation of amplitude and changes in latencies when time on task increases [3, 48, 61, 49, 40, 22, 54] and 2) sequential effects of stimulus train on P300 [15, 56, 30, 39, 23].

2.1.1 Habituation of the P300 component

Habituation is defined as a decrease in a response resulting from repeated stimulation. It has been proven that habituation of vertex potential evoked by a train of auditory stimuli is more rapid and more pronounced the faster the stimuli are presented and the decrease for the stimulus repetition followed approximatelly a negative exponential function [21]. Several studies have shown that the human cortex meets the main criteria of habituation [61, 40, 54]. However, habituation of the P300 component has not been so obvious in studies [48, 49, 22].

Amplitudes of the P300 component showed habituation when comparing six blocks of repeated P300 measurement sessions to each others [61]. Same phenomenon was confirmed in study of Polich [49], where ten block of repeated measurement sessions were compared to each others. P300 amplitude seems to dishabituate with the presentation of a new target stimulus [49]. The habituation of auditory P300 is also seen when using a single-stimulus paradigm, where just one kind of of stimulus are presented and the subject has to respond when a stimulus occurs [47]. Even though the latencies are shorter and amplitudes smaller for single-stimulus elicited P300 than for the oddball task, the habituation effect is similar in both paradigms.

The response mode and stimulus modality can also affect the habituation of P300 amplitude and latency [40, 22]. Lew and Polich [40] speculated that if P300 is sensitive to the amount of attentional resources allocated to the processing of a target stimulus, P300 should habituate more slowly for a target trial counting task than for a "press a button when a target occurs" -task. When subjects pressed a button in response to the target tone, P300 amplitude declined rapidly and produced a highly reliable habituation effect whereas counting stimuli did not produce a reliable overall decline with trial block. Counting the number of targets increases memory load and requires more attentional resources which inhibits the reduction in P300 amplitude with repeated target presentation. On the other hand it has been speculated that the count mode requires more available processing resources and thus the overall P300 amplitude is diminished and therefore habituation may not be seen [13].

P300 has not been seen to habituate for repeated visual stimuli [22]. It has been speculated that visual stimulus processing demands more attentional

resources than auditory because of the dominance of visual input on attentional processes. On the other hand, slight habituation for visual stimulus was seen in [54]. Another study showed a fast response decrement at Fz electrode when visual stimuli were presented with random inter-stimulus-interval (ISI) between 16 to 29 seconds without any task requirements with respect to the stimulus [62]. No short time span habituation was seen in other electrode locations Cz, Pz and Oz.

The number of target trials needed for complete habituation of P300 component has also been studied. Single-trial responses were compared to the average of responses and the P300 component was detected visually from the single trials and wide inter-subject variation for the single-trial latencies and amplitudes was seen. Amplitude habituation was not seen during the 15 first stimuli in study [49]. The habituation of amplitudes was noticed when ten blocks of oddball paradigms were repeated until 20 successive target trials for each block were achieved and these block averages were compared. This comparison showed that P300 habituates when about 200 target stimuli are presented.

In conclusion, habituation in P300 paradigm has been seen in long recording sessions. A difference in results concerning different modalities e.g. habituation in auditory or visual stimuli has been described. The response mode also seems to effect the habituation of P300 response.

2.1.2 Sequential effects of the stimulus train

The effect of stimulus sequence on the P300 component was first introduced in 1976 [57]. Systematic variation in the P300 component as a function of the sequence of preceding stimuli train was found. The amplitude of the P300 invoked by a target stimulus increased with the number of standard stimuli that preceded it. This variation seemed to result mainly in changes in amplitude and not in large changes in the latencies of the P300s. Squires has also pointed out that sequential effects appear both in auditory and visual oddball series and no fundamental differences between the two modalities was found [56].

The effect of stimulus sequence depends also on the global probability of the target stimuli; the effect is more efficient with the lower global probability [15]. In some other studies this global probability related sequential effect was not found [20, 3]. Neither the amplitude nor the latency were affected by the sequential effect when probabilities of 0.083, 0.042 and 0.021 were used [20]. However, it was discussed that the probability range used may not be wide enough to invoke that effect.

The amplitude of P300 has been found to be inversely proportional to the subjective probability, that varies from trial to trial depending on the specific sequence of stimuli preceding each target event [30]. Thus the variability in P300 amplitude as a function of stimulus sequence is resulting from cognitive

processes rather than habituation or adaptation.

Several studies have investigated whether the P300 amplitude is related to the sequential probability (probability of the certain stimulus within a number of stimuli) or the temporal probability (the probability of the certain stimulus within a period of time). Influence of the inter-stimulus interval on sequential effects in P300 component was studied in [39], as ISI modulates the effect of the global stimulus probability. However, higher order effects were unmodulated by the ISI, indicating that P300 is determined by structural aspects of the stimulus sequence rather than time-dependent memory decay caused by prolonged inter-target interval. The finding that P300 amplitude does not vary with the amount of time passing between targets was confirmed in a study of Rasmusson & al. [53]. Inter-target interval was systematically varied between 5–20 seconds and no effects were found. However, in another study, in which both the inter-stimulus interval and the sequential probability of the target tone were systematically manipulated [18], the result indicated that the amplitude of P300 is related to temporal rather than sequential probability. Target-to-target interval was also varied in [4] and [26]. These studies also suggest that P300 amplitude and latency variability are affected by target to target interval rather than structure of stimulus train and target probability per se.

The sequential effect did not seem to be age-related [19]. The data from younger and older adults behave in a similar manner, although P300 latencies are longer for the elder group. The sequential effect is more pronounced in children than in adults [33], reflecting more rapid loss of the neural model of the target stimulus during a long inter-target interval in children than in adults.

The sequential structure effect behave differently in some diseases compared to normals subjects. In a study of an Alzheimer group it was also found that P300 latency changes with sequence position in a linear manner compared to controls but with a diminished amplitude [23]. Schizophrenic patients have reduced P3 amplitudes compared to healthy subjects and this amplitude reduction is specific to certain sequential structure conditions [25].

Compared to normal children, shorter P300 latencies and more frontal P300 amplitude distribution to the target stimuli are seen in easily distractible children when target-to-target interval increases [33]. The authors speculate that in distractible children, the neural model of the P300 decays faster and thus a long inter-target interval results in activation of the brain's orienting network that are known to generate shorter latency brain responses (P3a like potential).

In conclusion, sequential structure of the stimulus train is shown to affect to both P300 latency and amplitude data. It is still unclear whether the effect is due to the structure of the stimulus cue per se (local probability and the number of preceding standard stimuli) or temporal factors (inter-target interval). Thus, in study conditions that use fixed ISI, sequential and temporal factor effects can

not be distinguished.

2.2 Analysis methods in ERP-analysis

Recorded ERP waveforms z are often modelled with an additive model:

$$z_k(t) = s_k(t) + n_k(t), \tag{2.1}$$

where t is time, k corresponds to response number, s_k represents the signal (i.e. the desired event related potential) and n_k the noise (i.e. unwanted activity, not systematically related to the stimulus). The signal is frequently modelled as a deterministic waveform, in which it is assumed that it is repeated exactly same for each stimulus application. The noise process is modelled as a stationary random process. Thus the average value of the noise is assumed to be zero. As seen in previous sections, it is well known that at least the cognitive ERPs may show considerable trial-to-trial variability during measurement session. More realistic assumptions allow randomness in the signal s and non-stationary behaviour in the EEG. Need for more sophisticated signal analysis methods was noticed early in ERP studies. Currently several methods have been developed for different purposes as the post processing of the measurements has become possible.

In this section some of the ERP-analysis methods are briefly reviewed and their benefits and drawbacks are discussed. Some of these methods are presented more widely in the next chapter. For review of different analysis methods, see also [1, 41, 43]. Comparisons of different methods are presented in [44, 45]

2.2.1 Averaging

The most widely used method of ERP analysis is the so-called ensemble averaging. The method is easily implemented, simple to use and the interpretation of the results is quite simple and does not require an extensive background in signal analysis.

The measurement corresponding to the k'th stimulus can be written

$$z_k(t) = s(t) + n_k(t) \tag{2.2}$$

The stimuli are repeated for a sufficient number of times K and the stimulus-related average is computed in order to obtain the estimate of the mean value of the measured data points.

$$\hat{s}(t) = \frac{1}{K}\sum_{k=1}^{K} z_k(t) = \frac{1}{K}\sum_{k=1}^{K} s(t) + \frac{1}{K}\sum_{k=1}^{K} n_k(t)\ . \tag{2.3}$$

If the background is zero mean and in a wide sense stationary, we have

$$\frac{1}{K}\sum_{k=1}^{K} n_k(t) \overset{K\to\infty}{\longrightarrow} 0\ . \tag{2.4}$$

With these assumptions the average $\hat{s}(t)$ is the best estimator in the mean-square sense for the underlying evoked potential $s(t)$.

2.2.2 Cross-correlation method, Woody-averaging

Cross-correlation methods are based on analysis of similarities between the ERP waveform and a template. Choice of template should be made in order to have a similar waveshape between the template and the component of interest. The model assumes that the ERP waveform is the same for all stimuli but the latency is not. Thus the additional noise model can be written as

$$z_k(t) = s(t - \tau_k) + n_k(t), \tag{2.5}$$

where τ_k represents a random time variable.

Woody proposed the use of an average waveform as the first template [63]. Before applying cross-correlation, each ERP epoch is normalised. This normalisation implies that the amplitude of P300 is not a defining characteristic, but the shape of it is. With normalisation the information about amplitude variation is lost. The ERP waveforms are shifted temporaly to aim maximum correlation with the template. The shifted waveforms are then re-averaged and the new average is used as a template for a new iteration. This iteration procedure is repeated until the template does not show any change from the previous iteration. The Woody method has been useful for the estimation of single trial latencies. If, however different components within each ERP have latency shifts that are independent of each other, improvement is not achieved.

2.2.3 Latency-corrected averaging

With the assumption that each component of ERP has latency changes independent of each other, it is reasonable to detect each component separately. After detecting components a histogram of their latencies is generated. By suitable statistical test, the peaks corresponding to signal components are separated from peaks corresponding to background EEG. Then peaks corresponding to the same ERP component are aligned and averaged. This method is called latency corrected averaging, LCA [42]. LCA is not a continuous waveform, as short time segments of ERPs are averaged independently of each other. This method allows the latencies of components in the same waveform to vary independently. By using LCA a number of additional parameters of ERP waveforms are achieved; the average latency and amplitude of each peak, the variance of the latency and amplitude of each peak and the number of peaks detected at each latency. The procedure for converting the disjoint segments of the LCA into a continuous waveform is presented for example in [44].

2.2.4 Wiener-filtering

The technique of a posteriori Wiener filtering is optimal for minimizing the mean squared error of the estimate when the following three assumptions hold: the signal and noise are additive, statistically independent and each is a sample function from a wide-sense stationary random process [7]. Under these assumptions the transfer function of the Wiener filter for a signal embedded in additive noise can be written in the frequency domain as

$$H(\omega) = \frac{S_s(\omega)}{S_s(\omega) + S_n(\omega)} \tag{2.6}$$

where $S_s(\omega)$ is the spectral density of the signal and $S_N(\omega)$ is the spectral density of the noise.

For the average evoked potential, the optimal Wiener filter transfer function is

$$H(\omega) = \frac{S_s(\omega)}{S_s(\omega) + \frac{1}{K}S_n(\omega)} \tag{2.7}$$

where K is the number of waveforms used to obtain the average. In practise the power density spectra of the signal and the noise are unknown, but can be estimated from the ensemble. Spectras can be estimated as the squared magnitude of the Fourier transform, as the Fourier transform of the autocorrelation function after appropriate windowing or as the smoothed periodogram. In practise the theoretical least-squared error is not reached, since the optimal Wiener filtrer transfer function is in itself an estimate.

Wiener filtering is a noncausal procedure, thus the entire waveform has to be available for processing. This is not a problem in ERP studies; usually the whole data is recorded at the same time and processed afterwards.

As ERP is a transient time-varying response, it is unlikely that time-invariant systems will produce optimal results. In Wiener filtering this means that the power density spectra which appear in equation 2.7 could be replaced with their time-varying counterparts. After this procedure the optimal time-varying system function is given by

$$\hat{G}(t,\omega) = \frac{S_s(t,\omega)}{S_s(t,\omega) + \frac{1}{K}S_n(t,\omega)} \tag{2.8}$$

where t refers to the time-lag relative to stimulus onset.

$\hat{G}(t,\omega)$ is no longer a transfer function in the classical sense, rather it is a weighing function. Time-varying filtering is a time-varying generalisation of a posteriori Wiener filtering. The benefit of this time-varying version is that it tries to cope with non-stationarities of the noise and the ERP signal. However, the estimation of a time-varying power spectrum is not very obvious. If the signal and the noise overlap simultaneously in time and frequency, both the

noise and signal are attenuated. The most apparent benefit is achieved when the waveform can be easily recognized in the background. If this is not a case, Wiener filter should not be used. The method does not subtitute averaging, but it can be used to improve signal to noise ratio whenever further improvement by other methods is unlikely. The obvious solution to obtain the time-varying power spectra is the use of a bank of proportional bandwith filters [5, 6].

2.2.5 Principal component analysis

The method of principal component analysis (PCA) allows to compare large sets of ERP waveforms by producing a set of 'principal components', or waves that are common to all of the potentials under consideration, the so-called basic waveforms [13]. The set of measurements is presented as a weighted sum of orthogonal basis waveforms, i.e. basis vectors. The weights (eigenvalues) and the basis vectors (eigenvectors) are obtained from the eigendecomposition of either the covariance or correlation matrix of the set of measurement vectors. The principal component analysis gives information on the second order statistics of the set of waveforms. By analysing the contributions of the basis vectors needed to represent each waveform in the set with a given degree of accuracy, one can qualitatively determine the degree of interdependence between the set of waveforms. The property of orthogonality has sometimes lead to misinterpretation of the results obtained with this method: the eigenvectors are thought to correspond to independent physiological generators. This assumption is erroneous since the eigenvectors are necessarily orthogonal due to the symmetry of the correlation (covariance) matrix.

2.2.6 Wavelet analysis

The wavelet analysis has been used to decompose the ERP signal onto a space with basis functions. With this technique the ERP is assumed to be the result of the superimposition of wave packets in various frequencies with varying degrees of frequency stabilization, enhancement and time locking within conventional frequency bands of the ongoing EEG activity such as delta, theta, alpha and gamma ranges. The wavelet analysis treats ERP responses in time frequency plane and has yielded to new knowledge about ERP components [9, 8]. However, the selection of relevant frequency band and the interpretation of the results in frequency domain is a challenging task.

2.2.7 Other single trial methods

Use of a deconvolution method, called enhanced averaging, is presented in [44]. The recorded waveform is convolved with a composite filter that relates the true signal to the estimate. The method is easy to use and does not require complex

programming. It reveals some high frequency components present in the average but provides much less resolution of the smaller peaks than the LCA.

In [37] Krieger et al. propose the use of an iterative Kernel estimator to estimate cognition related signals varying from trial to trial. They also pointed out the problem of parametrisation of single-trial estimates. Typically the parametrisation is carried out by manual or semi-automatic determination of the maximum peak within a predetermined latency range. Neither of these determination is an objective method, thus more automatic methods should be used.

A method that combines techniques from non-linear time series analysis with the wavelet transform was used in [16]. A method of non-linear de-noising was applied to both intra- and extracranially recorded ERPs in order to obtain additional information on dynamic neural processes. It was shown that with this method more accurate information was available, for example identification of possible correlations between various ERP components was possible.

2.3 Signal analysis

This chapter presents the background of the estimation theory that is essential for understanding the single trial analysis method used in this thesis. The method itself is introduced in chapter 6.

2.3.1 Introduction

The problem in estimation theory of evoked potentials is to model reality from some measured signal that may contain errors. With single-trial estimation our goal is to get an estimate for every single event-related potential s based on measured data z. A single-trial estimator thus filters out the unwanted effect of the background EEG v from the measurements z. The additive noise model, which is mostly used to model ERP measurements, can be written in the form

$$z = s + v = H\theta + v \tag{2.9}$$

where H is an observation matrix that does not contain parameters to be estimated whereas the parameters that are to be estimated are denoted with θ. The event related potential is thus modelled as a linear combination of some basis vectors, namely the columns of the matrix H. The equation 2.9 is called the linear observation model. It connects sampled measurements z with the parameters θ and the measurement errors v.

The event related potential s is modelled with a linear model $s = H\theta$. The single event related potentials are modeled as linear combinations of some basis vectors ψ_i that are the columns of the matrix $H = (\psi_1, \ldots, \psi_p)$. The task is then

to form an estimate $\hat{\theta}$ for the parameters θ corresponding to each measurement z. The single-trial estimate of the event related potential is then of form

$$\hat{s} = H\hat{\theta}. \tag{2.10}$$

2.3.2 Bayes cost method

If we assume that θ is a random vector with a known linear joint density $p(\theta, z)$ with the measurement z, we have made the so-called Bayesian assumption. For solving the estimator $\hat{\theta}(z)$ this assumption leads to the Bayes cost method. The Bayes estimation criterion states: For a given cost function $C(\theta, \hat{\theta})$, the Bayesian estimator $\hat{\theta}_B$ is selected so that

$$B(\hat{\theta}_B) \leq B(\hat{\theta}) \tag{2.11}$$

for all $\hat{\theta}$. So the Bayesian estimator is the one that minimises the Bayes cost.

Different choices of the cost function lead to different estimators. Estimators that are essential for the single trial method used in this thesis are presented.

2.3.3 Linear minimum mean square estimator

In this thesis the form of the estimator is restricted to be a linear function of data

$$\hat{\theta} = Kz. \tag{2.12}$$

Let θ and z be random vectors with zero means and known covariances. The task now is to derive the estimator that is of the form 2.12. In minimum mean square estimation the squared norm of the estimation error $\tilde{\theta} = \theta - \hat{\theta}$ is minimized.

$$\begin{aligned} C_{MS}(\theta, \hat{\theta}) &= \| \theta - \hat{\theta} \|^2 && (2.13) \\ &= \tilde{\theta}^T \tilde{\theta} && (2.14) \end{aligned}$$

Noted that

$$\begin{aligned} B_{MS}(\hat{\theta}) &= E\{\tilde{\theta}^T \tilde{\theta}\} && (2.15) \\ &= E\{(\theta - \hat{\theta})^T (\theta - \hat{\theta})\} && (2.16) \\ &= \text{trace} E\{(\theta - \hat{\theta})(\theta - \hat{\theta})^T\} && (2.17) \end{aligned}$$

where trace(A) is defined to be the sum of the diagonals of a square matrix A.

Thus the estimation error covariance is

$$\begin{aligned} C_{\tilde{\theta}} &= E\{(\theta - \hat{\theta})(\theta - \hat{\theta})^T\} && (2.18) \\ &= E\{(\theta - Kz)(\theta - Kz)^T\} && (2.19) \\ &= \theta\theta^T - \theta(Kz)^T - Kz\theta^T + Kz(Kz)^T && (2.20) \\ &= C_\theta - KC_{z\theta} - C_{\theta z}K^T + KC_z K^T && (2.21) \\ &= C_\theta + (K - C_{\theta z}C_z^{-1})C_z(K - C_{\theta z}C_z^{-1})^T - C_{\theta z}C_z^{-1}C_{z\theta} && (2.22) \end{aligned}$$

where it is assumed that C_z is invertible. Only the second term on the right hand side depends on the matrix K. The Bayes cost is thus minimised when

$$K = C_{\theta z} C_z^{-1} \tag{2.23}$$

so the linear minimun mean square estimate is

$$\hat{\theta}_{LMMS} = C_{\theta z} C_z^{-1} z \tag{2.24}$$

and the estimation error covariance is form

$$C_{\tilde{\theta}_{LMMS}} = C_\theta - C_{\theta z} C_z^{-1} C_{z\theta} \tag{2.25}$$

The estimator 2.24 is optimum when its structure is restricted to be linear. No assumptions have been made about the densities of the measurements and the parameters.

2.3.4 Minimum mean square estimation

If θ and z are jointly Gaussian, then the linear minimum mean square estimator is not only the optimal linear estimator but the overall optimal estimator for θ [31]. So the minumum mean square estimate is

$$\hat{\theta}_{MS} = C_{\theta z} C_z^{-1} z \tag{2.26}$$

and the estimation error covariance is in the form

$$C_{\tilde{\theta}_{MS}} = C_\theta - C_{\theta z} C_z^{-1} C_{z\theta} \tag{2.27}$$

that is exactly the linear minimum mean square estimator.

Next a model for the dependence between observations and parameters is examined. The additive noise model 2.9 where v and θ are random is used. Let v and θ have zero means and known covariances. The measurements z then have a zero mean and a covariance

$$\begin{aligned} C_z &= zz^T && (2.28)\\ &= E\{(H\theta + v)(H\theta + v)^T\} && (2.29)\\ &= H\theta(H\theta)^T + H\theta v^T + v(H\theta)^T + vv^T && (2.30)\\ &= H\theta\theta^T H^T + H\theta v^T + v\theta^T H^T + vv^T && (2.31)\\ &= HC_\theta H^T + HC_{\theta v} + C_{v\theta} H^T + C_v && (2.32) \end{aligned}$$

and the cross covariances are

$$\begin{aligned} C_{\theta z} &= E\{\theta(H\theta + v)^T\} && (2.33)\\ &= \theta(H\theta)^T + \theta v^T && (2.34)\\ &= C_\theta H^T + C_{\theta v} && (2.35)\\ & && (2.36) \end{aligned}$$

and

$$\begin{aligned} C_{z\theta} &= E\{(H\theta + v)\theta^T\} && (2.37)\\ &= \theta(H\theta)^T + \theta v^T && (2.38)\\ &= HC_\theta + C_{v\theta} && (2.39) \end{aligned}$$

With these the linear mean square estimate can be written as

$$\hat{\theta}_{LMMS} = C_{\theta z}C_z^{-1}z = (C_\theta H^T + C_{\theta v})(HC_\theta H^T + HC_{\theta v} + C_{v\theta}H^T + C_v)^{-1}z \quad (2.40)$$

with error covariance matrix

$$C_{\tilde{\theta}_{LMMS}} = C_\theta - (C_\theta H^T + C_{\theta v})(HC_\theta H^T + HC_{\theta v} + C_{v\theta}H^T + C_v)^{-1}(HC_\theta + C_{v\theta}) \quad (2.41)$$

A special case of this model is when θ and v are uncorrelated, i.e. $C_{\theta v} = C_{v\theta} = 0$. Then the equations for the estimate and the error covariance reduce to

$$\begin{aligned} \hat{\theta}_{LMMS} &= C_\theta H^T(HC_\theta H^T + C_v)^{-1}z && (2.42)\\ C_{\tilde{\theta}_{LMMS}} &= C_\theta - C_\theta H^T(HC_\theta H^T + C_v)^{-1}HC_\theta && (2.43)\\ & && (2.44) \end{aligned}$$

Applying the so-called matrix inversion lemma [24]

$$\begin{pmatrix} C_\theta & C_{\theta z} \\ C_{z\theta} & C_z \end{pmatrix}^{-1} = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} \quad (2.45)$$

where

$$\begin{aligned} C_{11} &= (C_\theta - C_{\theta z}C_z)^{-1} = C_\theta^{-1} + C_\theta^{-1}C_{\theta z}C_{22}C_{z\theta}C_\theta^{-1} && (2.46)\\ C_{22} &= (C_z - C_{z\theta}C_\theta^{-1}C_{\theta z})^{-1}C_z^{-1} + C_z^{-1}C_{z\theta}C_{11}C_{\theta z}C_z^{-1} && (2.47)\\ C_{12} &= C_{21} = -C_{11}C_{\theta z}C_z^{-1} = -C_\theta^{-1}C_{\theta z}C_{22} && (2.48) \end{aligned}$$

the model for linear mean square estimate and error covariance matrix is obtained

$$\begin{aligned} \hat{\theta}_{LMMS} &= (C_\theta^{-1} + H^T C_v^{-1}H)^{-1}H^T C_v^{-1}z && (2.49)\\ &= C_{\tilde{\theta}_{LMMS}}H^T C_v^{-1}z && (2.50)\\ C_{\tilde{\theta}_{LMMS}} &= (C_\theta^{-1} + H^T C_v^{-1}H)^{-1} && (2.51)\\ & && (2.52) \end{aligned}$$

2.3.5 Gauss Markov methods

Gauss-Markov theorem states that when estimating parameters in a linear model, the linear least squares estimator is the most efficient of all unbiased estimators which can be reduced to linear functions of the data. There are cases where other estimators are more efficient, but they are not linear functions of the data.

Next the situation where the parameters θ are unknown but non-random is to be considered. The observation model is, again, linear, that is

$$z = H\theta + v \tag{2.53}$$

where v is random and has a zero mean and covariance C_v. The linear unbiased estimator $\hat{\theta}$ minimises the criterion

$$\begin{aligned} E\{\| \theta - \hat{\theta} \|^2\} &= E\{(\theta - \hat{\theta})^T(\theta - \hat{\theta})\} && (2.54)\\ &= \text{trace}E\{(\theta - \hat{\theta})(\theta - \hat{\theta})^T\} && (2.55)\\ &= \text{trace}C_{\tilde{\theta}} && (2.56) \end{aligned}$$

Let

$$\hat{\theta} = Kz + k \tag{2.57}$$

with the requirement

$$E\{\hat{\theta}\} = \theta \tag{2.58}$$

Then

$$\begin{aligned} E\{\hat{\theta}\} &= E\{Kz + k\} && (2.59)\\ &= KE\{z\} + k && (2.60)\\ &= KE\{H\theta + v\} + k && (2.61)\\ &= KH\theta + k && (2.62)\\ & && (2.63) \end{aligned}$$

Taking in account the requirement 2.58 K and k must satisfy

$$KH = I, K = 0 \tag{2.64}$$

For the error covariance C_v it can be written

$$\begin{aligned} C_{\tilde{\theta}} &= E\{(\theta - \hat{\theta})(\theta - \hat{\theta})^T\} && (2.65)\\ &= E\{(\theta - Kz)(\theta - Kz)^T\} && (2.66)\\ &= E\{(\theta - KH\theta - Kv)(\theta - KH\theta - Kv)^T\} && (2.67)\\ &= E\{Kvv^TKT\} && (2.68)\\ &= KC_vK^T && (2.69) \end{aligned}$$

The Gauss-Markov Theorem states that the linear minimum variance unbiased estimate of θ is given by

$$\hat{\theta}_{GM} = (H^T C_v^{-1} H)^{-1} H^T C^{-1} z \tag{2.70}$$

if each element of z is unbiased i.e. $E\{\hat{\theta}\} = \theta$. The error covariance of the solution is

$$C_{\tilde{\theta}_{GM}} = (H^T C_v^{-1} H)^{-1} \tag{2.71}$$

It is worth to notice that the ordinary Gauss-Markov estimate $\hat{\theta}_{GM}$ can be obtained from $\hat{\theta}_{LMMS}$ formally by setting $C_\theta^{-1} = 0$. This means that no *a priori* information is available about θ that is, nothing is assumed on the properties of the parameters $\hat{\theta}$. The Gauss-Markov estimate $\hat{\theta}_{GM}$ also minimises the variance of the estimator and is thus also called the linear minimum variance estimator $\hat{\theta}_{LMV}$

2.3.6 LS-estimation

In the least squares solution neither the parameters θ or the error v in observations is interpreted as random. The observation model is

$$z = h(\theta) + v \tag{2.72}$$

where θ and v are unknown but non-random. The least squares estimator $\hat{\theta}_{LS}$ minimises the 2-norm of residual $r = z - \hat{z}$ and is form

$$\hat{\theta}_{LS} = (H^T H)^{-1} H^T z \tag{2.73}$$

The least squares and the Gauss-Markov estimates can be seen as Bayesian estimates with no prior information about the parameters θ.

2.3.7 Selection of the basis vectors

The linear least squares problem can be seen as an interpretation of the measurements z as a linear combination of some basis vectors, in other words the basis vectors span a subspace that determines the model for the evoked potentials. A special case is obtained when the basis vectors ψ_i are mutually orthonormal. This means that $H^T H = I$ and the least squares solution is

$$\hat{\theta} = (H^T H)^{-1} H^T z = H^T z. \tag{2.74}$$

The estimate for the observations z using this base is then of form

$$\hat{z} = H\hat{\theta} = H H^T z = \sum_{i=1}^{p} \psi_i c_i \tag{2.75}$$

where $c_i = \psi_i^T z$, that is, the inner product of the data with the basis vector.

An infinite number of different sets of basis vectors span the same subspace and thus it is important to choose a suitable basis for the estimation. If we have some information about the characteristic of the measurements we can use this information in basis vectors selection. If we do not have any kind of information about measurements, some generic vectros that we assume can model the data can be used. For example sampled polynomials, trigonometric functions or sets of Gaussian shaped vectors are possible choises. Any mixture of these is also possible.

2.3.8 Principal component analysis and regression

When measurements z are random, the coefficients c_i are also random parametres. Now we require that the chosen basis can express the evoked potentials with a minimum number of basis vectors. Thus if the statistical properties of the measurements are known, the optimal set of basis vectors are formed. To achieve this goal our requirement for coefficients c_i is that they are uncorrelated and we can write

$$\begin{aligned} E\{cc^T\} &= E\{H^T z z^T H\} && (2.76)\\ &= H^T R_z H && (2.77)\\ &= \mathrm{diag}(\sigma_1^2, \ldots, \sigma_p^2) && (2.78) \end{aligned}$$

This is an eigenproblem. The desired basis vectors are obtained as the eigenvectors of the data correlation matrix R_z.

With the selection of the eigenvector decomposition of the data matrix as the basis in least-square estimation the model can be written in the form

$$z = H_S\theta + v \tag{2.79}$$

where H_S contain the p first eigenvectors of the data correlation matrix R_z. S stands for the subspace spanned by these vectors. The Gauss-Markov estimate is then

$$\begin{aligned} \hat{\theta} &= (H_S C_v^{-1} H_S)^{-1} H_S^T C_v^{-1} z && (2.80)\\ \hat{s} &= H_S\hat{\theta} && (2.81) \end{aligned}$$

The correlation matrix of the evoked potential is usually not known. In some cases the first eigenvectors of the data correlation matrix $R_z = E\{zz^T\}$ span nearly the same subspace than the first eigenvectors of the evoked potentials correlation matrix R_s. Then we can approximate the evoked potentials subspace by the first eigenvectors of R_z. This approach resembles the use of the principal component regression where evoked potentials are forced to lie in the principal subspace spanned by the comlumns of H_S.

2.3.9 Subspace regularization

A generalised least squares solution for the parameters θ can be written as

$$\hat{\theta}_\alpha = \arg\min\{\| L_1(z - H\theta) \|^2 + \alpha^2 \| L_2(\theta - \theta^*) \|^2\} \tag{2.82}$$

where $L_1^T L_1 = W_1$ and $L_2^T L_2 = W_2$ are positive definite weighting matrices. The solution $\hat{\theta}_\alpha$ is called the Tikhonov regularised solution and it is a modification of the ordinary weighted least squares solution $\hat{\theta}_{LS} = \arg\min_\theta\{\| L_1(z - H\theta) \|^2\}$ to the direction in which *the side constraint* (the norm $\{\| L_2(\theta - \theta^*) \|\}$) gets smaller. The regularisation parameter α controls the weight given to minimisation of the side constraint $\| L_2(\theta - \theta^*) \|^2$ relative to minimisation of the weighted least squares index $l_{LS} = \| L_1(H\theta - z) \|^2$. The term θ^* is the initial (prior) guess for the solution. The regularised solution can be written in the form

$$\hat{\theta}_\alpha = (H^T W_1 H + \alpha^2 W_2)^{-1}(H^T W_1 z + \alpha^2 W_2 \theta^*) \tag{2.83}$$

Generic vectors are robust and easy to generate. On the other hand the eigenvector basis is optimal for a set measurements. These two basis can be combined and use a generic basis as a model for the evoked potentials and the eigenvector basis to represent prior information about the measurement.

We assume that the columns of the matrix H_S contain an orthonormal basis of the subspace S and the evoked potential is close to this subspace. The projection of evoked potential $s = H\theta$ onto subspace S is $(H_S H_S^T)H\theta$ and distance of s from S is

$$\Omega(\theta) = \| H_S H_S^T \theta - \theta \| = \| (H_S H_S^T - I)\theta \| . \tag{2.84}$$

This quantity can be used as a side constraint in the regularised least squares solution and we obtain

$$\hat{\theta} = \arg\min\{\| L_1(z - H\theta) \|^2 + \alpha^2 \| (I - H_S H_S^T)\theta \|^2\}. \tag{2.85}$$

Since $L_2^T L_2 = H^T(I - H_S H_S^T)^T(I - H_S H_S^T)H = H^T(I - H_S H_S^T)H$, the desired solution for the parameters θ can be written

$$\hat{\theta}_s = (H^T L_1^T L_1 H + \alpha^2 H^T(I - H_S H_S^T)H)^{-1} H^T z. \tag{2.86}$$

If we select $\alpha = 0$, the solution $\hat{\theta}_S$ is equivalent to the gauss-Markov solution.

For event related potentials one can select $L_1^T L_1 = C_v^{-1}$ if the covariance of the background EEG can be estimated. The estimate for the event related potential is then

$$\hat{s}_S = H\hat{\theta}_S. \tag{2.87}$$

Chapter III

Purpose of the study

Aim of this study was to examine the hypothesis that single-trial estimates give more information about shifts in attention and alertness of subjects than the average of the event related responses. It is assumed that by using the single trial method it is possible to gain more precise information about the characteristics of the P300 component and the analysis of the single trial responses extracts features of the P300 not visible in the averaged waveforms.

The well known potential, P300, is known to vary both in latency and amplitude during the measurement session. Two phenomena have been reported to cause variability in the P300. These phenomena are: time on the task and the local changes in the oddball stimulus sequence. Thus this potential was chosen for further analysis.

In this study a single-trial estimation method introduced in [32] is used to analyse more accurately the latency and amplitude variation of the P300 component during the oddball paradigm. The method uses the most available a priori information in defining the P300 component.

The foci of this thesis are

1. The evaluation of the single-trial method introduced in [32] by comparing the analysis results of a set of clinical ERP measurements between the single trial and the conventional averaging method.

2. The examination of the single-trial variability of the P300 component during the auditive oddball test.

Chapter IV

Material and methods

4.1 Task

The ERPs were recorded using an auditory oddball-paradigm in an otherwise silent room. The subject was asked to press a button as soon as he/she heard the target stimulus, and was also instructed to keep the eyes open and to look at a fixation point straight ahead.

4.2 Subjects

A total of 12 normal females aged 30 $\pm$ 4.5 years (aged 25-39 years) with no known history of medication affecting the central nervous system participated this study. The subjects had at least a high-school diploma as their educational level. All subjects reported normal hearing. All subjects participated voluntarily in the experiment and the study was accepted by the local ethical committee.

4.3 Stimuli and Responses

The ERPs were recorded using an auditory oddball paradigm with 85 % standard (800 Hz) and 15 % of target (560 Hz) tones. The duration of each stimuli was 84 ms including 7 ms rise and fall times. Interstimulus interval (ISI) was 1 s and the total number of stimuli was 600. The stimuli sequence was generated as Bernoulli trials, so also two targets could occur in succession. The same stimuli sequence was used with all subjects. The experimenter tested the individual hearing level at the beginning of the recording session. The stimuli were delivered by earphones to the right ear of the subject at 60 dB above the hearing level. The experimenter observed each subject's performance. The subjects performed the task mostly without errors.

4.4 Electrodes

The ERPs were recorded using Ag/AgCl electrodes placed on the scalp according to the International 10-20 System. The vertical eye movements were recorded by electrodes placed at above and below of the right eye and horizontal eye movements by electrodes placed at the right and left canthus. All other electrodes were referred to the right mastoid. Electrodes were affixed with adhesive paste and the interelectrode impedance was less than 5 $k\Omega$.

4.5 Amplification and Analog-to-Digital (A/D) conversion

Signals were amplified and filtered by a Neuroscan Synamps amplifier with a bandpass 0.3-50 Hz and sampled continuously at 250 Hz. Trigger pulses from a Neuroscan Stim Audio System P/N 1105 controlling stimuli were stored together with the electrophysiological data in all subjects.

4.6 Data analysis

The midline electrodes Fz, Cz and Pz were chosen for further analysis of the P300 component. The continuous data were transformed off-line to epochs of 0 to 500 ms relative to the onset of each stimulus. Epochs containing eye movements artifacts were excluded from further analysis, using both automatic (trials exceeding $\pm$70 mV) and visual inspection of the data. After artifact rejection at least 50 artifact-free target responses remained for further analysis. Both conventional averaging and a single trial method were used to estimate P300 responses.

4.6.1 Analysis of the conventional average P300-response

For habituation comparisons the stimulus sequency was divided into two blocks, the first half of measurement and the second half of measurement. Both blocks contained 40 target stimuli and after artifact rejection there was 25 to 30 target responses in each block.

In order to study the sequential effects of the stimulus train, two averages were calculated: the average for responses to target tones that were preceded by five or fewer consecutive standard tones and the average for responses to target tones that were preceded by six or more consecutive standard tones.

The P300 peak was defined as a maximum peak within 250-500 ms. The amplitude of the P300 was measured relative to the 100 ms pre-stimulus baseline.

4.6.2 Calculation and analysis of the single sweep estimates

Single sweep estimates for epoched and artifact free data were calculated using the method described in Chapter 4. The steps for single-trial estimation were as follows:

1. The estimate for the background covariance C_v was chosen to be an identity matrix as it has been speculated that the background EEG may have changes time-locked to target stimulus.

2. The matrix H was formed as a set of Gaussian shaped vectors.

3. The correlation matrix of the measurement R_z was used and the ordinary eigendecomposition was solved.

4. The matrix H_s was formed with four eigenvectors that were associated with the four largest eigenvalues.

5. The regularization parameter α was fixed to 0.05.

6. Single trial estimates $\hat{s}_S$ for the event-related potentials were calculated with equation

$$\hat{\theta}_s = (H^T C_v^{-1} H + \alpha^2 H^T (I - H_S H_S^T) H)^{-1} H^T z. \tag{4.1}$$

The amplitude and latency of the P300 component were defined as a maximum of a single trial estimate in an individually defined time window. The window width was individually assessed so that it covered the whole P300 component in the average P300 waveform.

In order to compare the single sweep method to conventional averaging, the single sweep amplitude and latency data were sorted in a same manner to habituation groups (the first and the second half of stimulus sequence) and to sequential groups (1-5 or 6-20 preceding standard tones) as the conventional average responses.

Single-sweep estimates also give the oppurtunity to make some additional analysis. The amplitude and latency data were arranged according the number of preceding standard tones. The straight line was fitted to this rearranged data. The degree of sequential effect was defined as the slopes of the linear functions fitted to both the amplitude and latency data.

The amplitude and latency data were then arranged according to their appearance time, measured from the beginning of the stimulus sequence. A straight line for both the amplitude and latency data was fitted in a least square sense. The degree of habituation was defined as a slope of both amplitude and latency data.

4.7 Statistical analysis

All statistical analyses were computed with the SPSS for Windows 11.5 statistical program. Repeated measures analysis of variance (ANOVA) for two methods (single trial or convantional averaging), two situation sorting (either two stimuli block (the first and the second half of stimulus sequence) or two stimulus sequences (1-5 or 6-20 preceding standard stimuli) and three electrode sites (Fz, Cz, and Pz) was performed. Further analysis was calculated with repeated measures ANOVA separately for each electrode (method X situation) and for both situations (method X electrode). For all analyses, Greenhouse Geisser degrees of freedom were used to correct the violations of the spherical assumption when appropriate.

Chapter V

Results

Single trial estimates of one representative subject of this study are shown in Figure 5.1. It can be seen that both latency and amplitude are varying during an oddball paradigm.

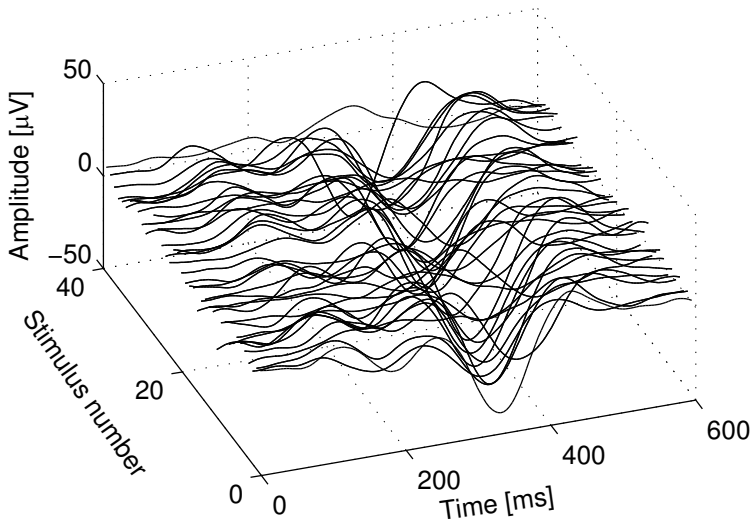


Figure 5.1: An example of single trial estimates from one representative subject

5.1 Habituation

The grand-average waveforms for the habituation study are shown in Figure 5.2.

The mean (+ 1 SE) P300 amplitude and latency data for both the conventional averaging method and for the single trial method are presented as a

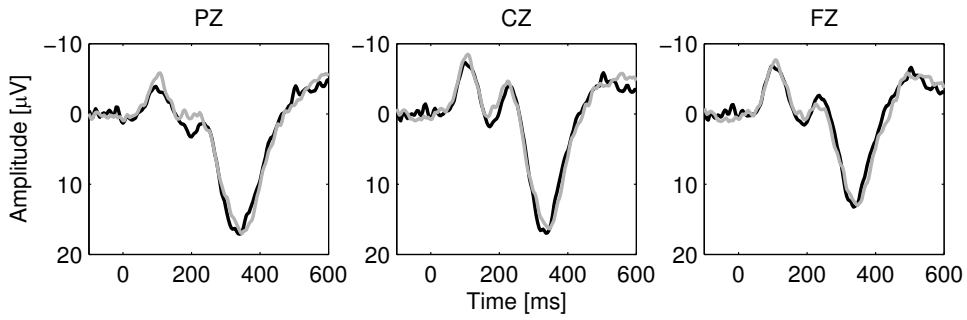


Figure 5.2: The grand average waveforms of the event-related brain potentials to target stimuli at the first half (black) and at the second half (grey) of an auditory oddball task

function of electrode site in Figure 5.3

In general, the mean amplitude was larger and the mean latency was longer with the single trial than with the conventional averaging method. This difference yielded a significant main effect for the used method for both the amplitude $F(1, 11) = 82, 8, p < 0.001$ and the latency $F(1, 11) = 19.5, p < 0.001$ data.

The amplitude of the P300 was affected by electrode site $F(1.26, 13.87) = 5.10, p = 0.034$. The amplitude was smaller at electrode Fz than at elctrode Cz ($p = 0.011$). The amplitude of the P300 component was not affected by situation, i.e., amplitudes were not larger at the beginning of the task compared to those at the end of the task. The time on task did not have an effect on the latency of the P300, nor did the electrode site.

5.2 Sequential effects

The grand-average waveforms for the sequential effect study are shown in Figure 5.4.

As in the habituation situation, the mean amplitude was larger and the mean latency was longer with the single-trial than with the conventional averaging method. This difference yielded a significant main effect for the used method for both the amplitude $F(1, 11) = 123.45, p < 0.001$ and the latency $F(1, 11) = 9.06, p = 0.012$ data.

The mean (+ 1 SE) P300 amplitude and latency data for both the conventional averaging method and for the single trial method are presented as a function of electrode site in Figure 5.5

The P300 amplitude was smaller when the number of preceding standard tones was smaller $F(1, 11) = 8.13, p = 0.016$. Also electrode site had an

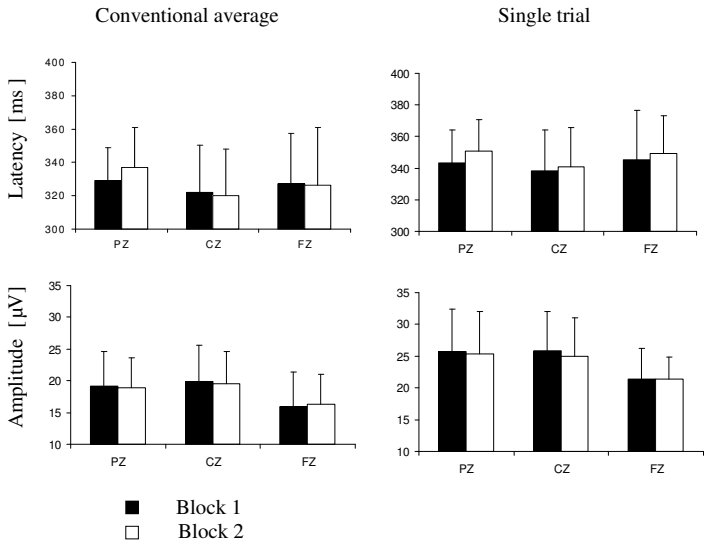


Figure 5.3: Mean(+ 1 SE) P300 amplitude (in microvolts) and latency (in milliseconds) data as a function of electrode site measured from the conventional average and with the single trial method). The responses to target tone are sorted in two blocks: Block 1) First half of measurement and Block 2) Second half of the measurement.

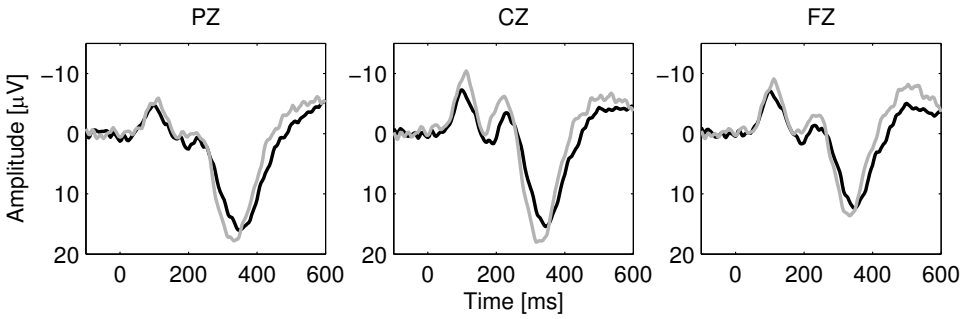


Figure 5.4: The grand average waveforms of the event-related brain potentials to target stimuli preceded by 1-5 standard stimuli (black) and preceded by 6-20 standard stimuli(grey).

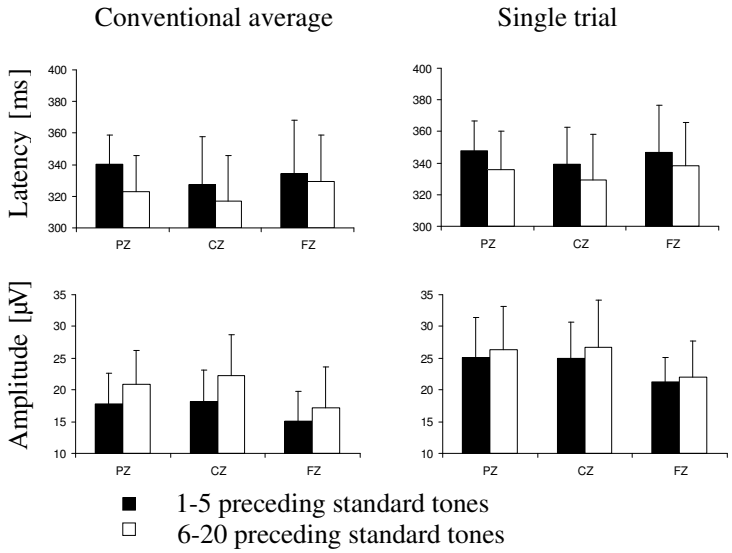


Figure 5.5: Mean(+ 1 SE) P300 amplitude and latency data as a function of electrode site measured from the conventional average and with the single trial method. Responses to target tones are sorted according to the number of standard tones preceding the target tone.

effect on the P300 amplitude $F(1.3, 14.3) = 6.06, p = 0.021$. The amplitude at Fz was smaller than at Cz $F(1, 11) = 11.38, p = 0.006$ or at Pz $F(1, 11) = 4.65, p = 0.054$. There was an interaction between the used method and the number of preceding standard tones $F(1, 11) = 13.44, p = 0.004$. Closer analysis showed that there was no effect to response amplitude when the single trial method was used $F(1, 11) = 2.48, p = 0.14$, whereas the conventional averaging yielded to significant decrease in amplitude when there was a smaller number of preceding standard tones compared to large number of standard tones $F(1, 11) = 14.27, p = 0.0003$. The latency of the P300 component was longer when number of preceding standard tones was smaller $F(1, 11) = 6.94, p = 0.023$.

5.3 Summary of habituation and sequential effect studies

In general the mean amplitude was higher and the mean latency was longer with the single trial than with the conventional averaging method. Conclusions for method comparison results are shown in Table 5.3.

Both methods showed significant increase in PZ latency and no changes in P300 amplitude when time on task increased. With the single-trial method

decreased latency with no amplitude effect was found with increasing number of preceding standard tones. The conventional averaging showed only a tendency for latency shortening and a significant amplitude increase when the number of preceding standard tones increased.

Table 5.1: The summary table of method comparison results. Statistically significant results are marked with + and non-significant results with -.

	Latency		Amplitude	
	Habituation	Stimulus history	Habituation	Stimulus history
Conventional averaging	-	+	-	+
Single trial	-	+	-	-

5.4 Single trial estimation

Single sweep estimates provide an opportunity to make some additional analysis.

To study the habituation effect further, the amplitude and latency data were arranged according to their appearance time, measured from the beginning of the stimulus sequence. A straight line for both the amplitude and the latency data was fitted in a least-square sense. The degree of habituation was defined as a slope of both amplitude and latency data. To study the sequential effects further, the amplitude and latency data were rearranged according to the number of preceding standard tones. The straight line was fitted to this rearranged data. The degree of sequential effect was defined as the slopes of the linear functions fitted to both the amplitude and the latency data. Figure 5.6 shows fitted lines for the latency data and figure 5.7 for the amplitude data for one representative subject.

Sequential changes in the stimulus train had more effect than time on task to the latency data, a significant main effect $F(1,11)=43.6$, $p<0.001$. Additionally, the stimulus history had more effect than the time on task on the P300 amplitude data, but this effect was not statistically significant. There was no interaction with the channel and the situation.

Table 5.4 shows that the slopes are nearly zeros for the amplitude and latency data of habituation. This means that the amplitude or the latency of the P300 did not change due habituation during measurement in any channels. When the amplitudes and latencies were rearranged according to the number of preceding stimulus sequences, there was a clear effect in the latency data slopes. Slopes for latency in sequential effect data were negative, meaning a decrease in the latency with an increasing number of the preceding standard tones. Slopes for the amplitude in stimulus history data are positive meaning that the amplitude increases with an increasing number of preceding standard tones.

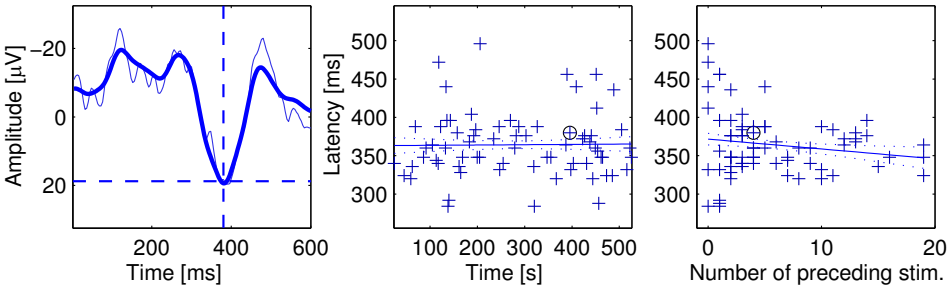


Figure 5.6: An example of a single trial estimate test and the fitted lines to the latency data for one representative subject in both the habituation (middle) and the stimulus history situation (right). The location of the example single-trial estimate (left) is marked with a circle in both situations.

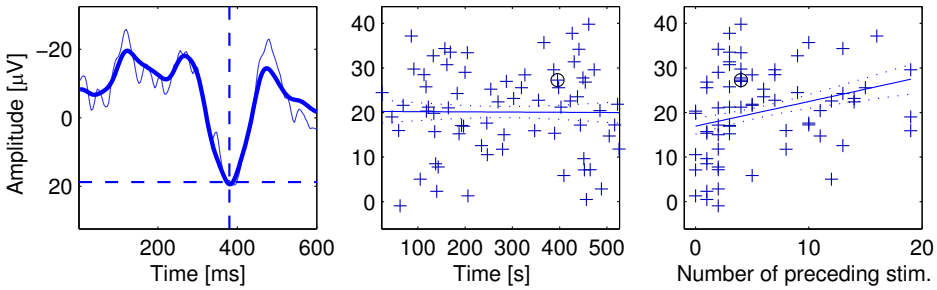


Figure 5.7: An example of single-trial estimate test and fitted lines to amplitude data for one representative subject in both the habituation (middle) and the stimulus history situation (right). The location of the example single-trial estimate (left) is marked with a circle in both situations.

Table 5.2: The average of the slopes (± SE). Slopes for both amplitude and latency data were estimated in a least-square sense for each individual subject.

	Amplitude		Latency	
	Habituation (μV/time on task)	Stimulus history (μV/no. of std.)	Habituation (ms/time on task)	Stimulus history (ms/no. of std.)
PZ	-0.002 ± 0.002	0.186 ± 0.093	0.029 ± 0.012	-1.877 ± 0.313
CZ	-0.005 ± 0.003	0.301 ± 0.127	0.011 ± 0.010	-1.816 ± 0.267
FZ	-0.002 ± 0.002	0.154 ± 0.124	0.018 ± 0.014	-1.646 ± 0.412

Chapter VI
Discussion

The present study examined habituation and stimulus sequence effects on the P300 component. P300 data from 12 normal subjects were analysed both with single-trial and conventional averaging methods. Results with the single-trial method were compared with results achieved with conventional averaging. Results of some previous reports were confirmed with the single-trial method and new insights into the variability of P300 component were revealed.

With many single-trial methods subjective decisions on data handling must be made before analysis, for example the choice of a suitable model in the Woody averaging and the selection of the independent components in the LCA-method. The determination of the estimates is also semi-automatic rather than automatic. In the presented single-trial method there is no need for subjective decisions before analysis and the method is very automatic. The estimates are easy to interpret when compared to the interpretation of the PCA components or to the analysis of frequency bands in wavelet methods.

In general, it was shown that the amplitudes were higher and the latencies were longer with the single trial than with the conventional averaging method. The most likely explanation for this amplitude decrease with the conventional averaging is a high variability in the P300 latency. The jitter in the P300 peak causes amplitude decrease which may lead to erroneous conclusions about the amplitudes.

In the habituation study no time on task effect was found in amplitude or latency with either of the methods. Amplitude decrease in stimulus repetition has been found in studies where recording sessions have been longer than in our study [61, 49, 40, 54]. Latency increase with successive trial blocks has been found in some studies [47] but not in all [54]. It might be that our recording session was too short to reveal the habituation in the P300 amplitude. Even the number of target responses was quite similar in our study than in other studies, but the inter-stimulus interval was relatively short compared to the other studies resulting shorter recording time that may explain our results.

In the present study the P300 latency was significantly shorter for target tones that were preceded by a large number of standard tones compared to the responses to target tones that were preceded by a small number of standard tones. The amplitude was affected, e.g., the amplitude increased, with a larger number of preceding standard tones, only when analysed with conventional averaging. The inter-stimulus interval was fixed in our study and thus either number of standard stimulus or the passed time between target tones (ITI) could explain this finding. Rasmusson and Allen [53] did not find an amplitude increase with increasing time between targets with the conventional averaging method. Latency differences were not studied. Their inter-target interval was 5, 10, 15, or 20 s and thus it is possible that the range they used was not well-suited to elicit amplitude differences. Studies that have found latency and/or amplitude changes due to sequential changes in the stimulus history have used a wider range of ITI, <5 s compared to >10 s in [33], 2-3 s vs. 4-9 s vs. 10-19 s in [34], 2.6-13.0 s in [25], 2-4 s vs. 4-6 s vs. 6-8s in [26] and 1-10 s in [4].

Further analysis of the single trial responses showed that even if the intersubject variability was large, the P300 latency was affected by the sequential effects in every subject in the channels CZ and PZ. The amplitude was affected in most of the subjects in the same channels too. Amplitude results are consistent with the study of Gonsalvez et al. [26]. Increases in the target to target interval resulted in significant increases of P3 amplitude in channels FZ, CZ and PZ. It was also found that the target to target interval affected more significantly at parietal sites than it did at frontal sites. With the single trial analysis the latency seemed to be affected more systematically than the amplitude. With the conventional averaging method this phenomena may be disturbed as the jitter in the P300 component latency causes an amplitude decrease. Thus the amplitudes are smaller when the variation in P300 component is larger.

This study suggests that the sequential changes in the stimulus train have more effect on P300 amplitude and latency than on habituation during a short (i.e. 10 min) measurement session. Many of the previously reported misclassifications of single-trial ERPs may have been made due to variations in cognitive expectancy produced by sequential effects. The same conclusion is made in [30], where the data provide evidence that even though the objective prior probability remains constant over a series of trials, the subjective probabilities vary from trial to trial, depending the specific sequence of stimuli preceding each event. In summary they conclude that variability in P300 amplitude as a function of stimulus sequence is the result of cognitive process rather than the adaptation or the habituation. Thus it is reasonable to expect that the variation in the P300 includes information about subjects cognitive state and this variation should not to be depressed with conventional averaging.

6.1 Conclusions

The method presented in this study needs a sample of responses to estimate single responses accurately. Reliable estimates of single or very few responses cannot be formed. Because of large intra-individual variation, the need for a representative sample of single-trial responses of each individual is obvious. However, this study shows that in addition to the information that can be obtained with the conventional averaging, the single-trial method provided information about the dynamical changes of the P300 component during oddball paradigm. The method was evaluated with 12 subjects and some previous, well-known, results were confirmed and new information about the dynamics of P300 was achieved. In short measurement sessions, the effect of sequential changes in the stimulus train is more pronounced than the effect of the time on the task. The effect of the changes in the stimulus train is more clear in the latency than in the amplitude of the P300.

The single-trial method could be used to model changes in dynamical behaviour of the P300 component during maturation and aging. It is also possible that this kind of dynamical behaviour is disturbed in certain diseases and thus single trial analysis may provide valuable information about abnormal brain functioning. Future applications in which the presented single-trial method may prove useful include studies about sleep deprivation, stress and effects of task load.

Bibliography

[1] J.I. Aunon, C.D. McGillem, and D.G. Childers. Signal processing in evoked potential research: averaging and modeling. *CRC Critical Rewievs in Biomedical Engineering*, 4:323–367, 1981.

[2] M.A.B. Brazier. Evoked responses recorded from the depths of the human brain. *Annals of New York Academy of Sciences*, 112:33–59, 1964.

[3] E. Courchesne. Neurophysiological correlates of cognitive development: changes in long-latency event-related potentials from childhood to adulthood. *Electroencephalography and Clinical Neurophysiology*, 45:468–482, 1978.

[4] R.J. Croft, C.J. Gonsalvez, C. Gabriel, and R.J. Barry. Target-to-target interval versus probability effects on P300 in one- and two-tone tasks. *Psychophysiology*, 40:322–328, 2003.

[5] J.P.C. de Weerd. A posteriori Time-varying filtering of averaged evoked potentials. I. Introduction and conseptual basis. *Biological Cybernetics*, 41:211–222, 1981.

[6] J.P.C. de Weerd and J.I. Kap. A posteriori Time-varying filtering of averaged evoked potentials.II. Mathematical and computational aspects. *Biological Cybernetics*, 41:223–234, 1981.

[7] J.P.C. de Weerd and W.J.L. Martens. Theory and practise of a posteriori Wiener filtering of averaged evoked potentials. *Biological Cybernetics*, 30:81, 1978.

[8] T. Demiralp, A. Ademoglu, Y. Istefanopulos, C. Basar-Eroglu, and E. Basar. Wavelet analysis of oddball P300. *International Journal of Psychophysiology*, 39:221–227, 2001.

[9] T. Demiralp, A. Ademoglu, M. Schurmann, C. Basar-Eroglu, and E. Basar. Detection of P300 waves in single trials by the Wavelet transform (WT). *Brain and language*, 66:108–128, 1999.

[10] E. Donchin. Surprise !... surprise ? *Psychophysiology*, 18:493–513, 1981.

[11] E. Donchin and M.G.H. Coles. Is the P300 component a manifestation of context updating ? *Behavioral and Brain Sciences*, 11:357–374, 1988.

[12] E. Donchin, E. Heffley, S.A. Hillyard, N. Loveless, I. Maltzman, A. Öhman, F. Rösler, D. Ruchkin, and D. Siddle. Cognition and event-related potentials II. The orienting reflex and P300. *Annals New York Academy of Sciences*, (1):39–57, Oct 1984.

[13] E. Donchin and E.F. Heffley. Multivariate analysis of event-related potential data: A tutorial review. In D. Otto, editor, *Multidisciplinary perspectives in event-related brain potential research*, pages 555–572. U.S. Governement Printing Office, Washington DC, 1978.

[14] E. Donchin, M. Kubovy, R. Johnson JR., and R.I. Herning. Graded changes in evoked response P300 amplitude as a function of cognitive activity. *Perception & Psychophysics*, 14(2):319–324, 1973.

[15] C.C. Duncan-Johnson and E. Donchin. On quantifying suprise: The variation in event-related potentials with subjective probability. *Psychophysiology*, 14:456–467, 1977.

[16] A. Effern, K. Lehnertz, G. Fernandez, T. Grunwald, P. David, and C.E. Elger. Single trial analysis of event related potentials: non-linear de-noising with wavelets. *Clinical Neurophysiology*, 111:2255–2263, 2000.

[17] M. Fabiani, G. Gratton, D. Karis, and E. Donchin. Definition, identification and reliability of measurement of the P300 component of the event-related brain potential. In P. Ackles, J.R. Jennings, and M.G.H. Coles, editors, *Advances in psychophysiology, Vol. 2*, pages 1–78. JAI Press, Greenwich, 1987.

[18] P.G. Fitzgerald and T.W. Picton. Temporal and sequential probability in evoked potential studies. *Canadian Journal of Psychology*, 35:188–200, 1981.

[19] J.M. Ford, C.C. Duncan-Johnson, A. Pfefferbaum, and B.S. Kopell. Expectancy for events in old age: stimulus sequence effects on P300 and reaction time. *Journal of Gerontology*, 37:696–704, 1982.

[20] J.M. Ford, W.T. Roth, and B.S. Kopell. Auditory evoked potentials to unpredictable shifts in pitch. *Psychophysiology*, 13:32–39, 1976.

[21] H. Fruhstorfer, P. Soveri, and T. Järvilehto. Short-term habituation of the auditory evoked response in man. *Electroencephalography and Clinical Neurophysiolgy*, 28:153–161, 1970.

[22] M.W. Geisler and J. Polich. P300 habituation from visual stimuli. *Physiology & Behavior*, 56(3):511–516, 1994.

[23] E.J. Golob and A. Starr. Effects of stimulus sequence on event-related potentials and reaction time during target detection in alzheimer's disease. *Clinical Neurophysiology*, 111:1438–1449, 2000.

[24] G.H. Golub and C.F. van Loan. *Matrix Computation*. The Johns Hopkins University Press, 1989.

[25] G.J. Gonsalvez, E. Gordon, J. Anderson, G. Pettigrew, R.J. Barry, C. Rennie, and R. Meares. Numbers of preceding nontargets differentially affect responses to targets in normal volunteers and patients with schitzophrenia: A study of event-related potentials. *Psychiatry Research*, 58:69–75, 1995.

[26] G.J. Gonsalvez, E. Gordon, S. Grayson, R.J. Barry, I. Lazzaro, and H. Bahramali. Is the target-to-target interval a critical determinant of P3 amplitude. *Psychophysiology*, 36:643–654, 1999.

[27] E. Halgren, J.M. Stapelton, M. Smith, and I. Altafullah. Generators of the human scalp P3(s). In *Evoked potentials*, pages 269–284. Alan R. Liss, Inc, 1986.

[28] D.G. Humphrey and A.F. Kramer. Toward a psychophysiological assessment of dynamic changes in mental workload. *Human Factors*, 36(1):3–26, 1994.

[29] J.B. Isreal, C.D. Wickens, and G.L. Chesney. The event-related brain potential as an index of display-monitoring workload. *Human Factors*, 22(2):211–224, 1980.

[30] R. Johnson and E. Donchin. P300 and stimulus categorization: two plus two in not so different from one plus one. *Psychophysiology*, 17:167–178, 1980.

[31] P.A. Karjalainen. *Regularization and Bayesian methods for evoked potential esimation*. PhD thesis, University of Kuopio, Department of Applied Physics, 1997.

[32] P.A. Karjalainen, J.P. Kaipio, A.S. Koistinen, and M. Vauhkonen. Subspace regularization method for the single trial estimation of evoked potentials. *IEEE Transactions on Biomedical Engineering*, 46(7):849–860, July 1999.

[33] R. Kilpeläinen, A. Koistinen, M. Könönen, E. Herrgård, J. Partanen, and JR. Karhu. P300 sequence effects differ between children and adults for auditory stimuli. *Psychophysiology*, 36:343–50, 1999.

[34] R. Kilpeläinen, J. Partanen, and JR. Karhu. What does the p300 brain response measure in children? New insight from stimulus sequence studies. *NeuroReport*, 10:2625–2630, 1999.

[35] G.H. Klem, H.O. Luders, H.H. Jasper, and C. Elger. The ten-twenty electrode system of the international federation. The international federation of clinical neurophysiology. *Electroencephalogr Clin Neurophysiol Suppl*, 52:3–6, 1999.

[36] A. Kok. On the utility of P3 amplitude as a index of processing capacity. *Psychophysiology*, 38:557–577, 2001.

[37] S. Krieger, J. Timmer, S. Lis, and H.M. Olbrich. Some considerations on estimating event-related brain signals. *Journal of Neural Transm Gen Sect*, 99(1-3):103–129, 1995.

[38] M. Kutas, G. McGarthy, and E. Donchin. Augmenting mental chronometry: The P300 as a measure of stimulus evaluation time. *Science*, 197:792–795, 1977.

[39] H. Leuthold and W. Sommer. Stimulus presentation rate dissociates sequential effects in event-related potentials and reaction times. *Psychophysiology*, 30:510–517, 1993.

[40] G.S. Lew and J. Polich. P300, habituation, and response model. *Physiol Behav*, 53:111–117, 1993.

[41] C. D. McGillem, J. I. Aunon, and D. G. Childers. Signal processing in evoked potential research: Applications of filtering and pattern recogniton. *CRC Critical Rewievs in Biomedical Engineering*, 14:225–265, 1981.

[42] C.D. McGillem and J.I. Aunon. Measurements of signal components in single visually evoked brain potentials. *IEEE Transactions on Biomedical Engineering*, 24:232–241, 1977.

[43] C.D. McGillem and J.I. Aunon. *Analysis of event-related potentials*, chapter 5, pages 131–169. Elsevier Science Publisher B. V, 1987.

[44] C.D. McGillem, J.I. Aunon, and C.A. Pomalaza. Improved waveform estimation procedures for event-related potentials. *IEEE Transactions on Biomedical Engineering*, BME-32, 1985.

[45] J. Möcks, T.G. Gasser, W. Kohler, and J.P.C. De Weerd. Does filtering and smoothing of average evoked potentials really pay ? A statistical comparison. *Electroencephalography and Clinical Neurophysiology*, 64:469–480, 1986.

[46] R. Munson, D.S. Ruchkin, W. RItter, S. Sutton, and N.K. Squires. The relation of P3b to prior events and future behavior. *Biological Psychology*, 19:1–29, 1984.

[47] J. Pan, T. Takeshita, and K. Morimoto. P300 habituation from auditory single-stimulus and oddball paradigms. *Int J Psychophysiol*, 37(2):149–53, Aug 2000.

[48] J. Polich. Normal variation of P300 from auditory stimuli. *Electroencephalography and Clinical Neurophysiology*, 65:236–240, 1986.

[49] J. Polich. Habituation of P300 from auditory stimuli. *Psychobiology*, 17:19–28, 1989.

[50] J. Polich and K.L. Herbst. P300 as a clinical assay: rationale, evaluation, and findings. *Int J Psychophysiol*, 38(1):3–19, Oct 2000.

[51] W.S. Pritchard. Psychophysiology of P300. *Psychol Bull*, 90:506–540, 1990.

[52] Jr R. Johnson. On the neural generators of the P300 component of the event-related potential. *Psychophysiology*, 30:90–97, 1993.

[53] D.X. Rasmusson and J.D. Allen. Intertarget interval does not affect P300 within oddball series. *Psychophysiology*, 17:57–63, 1994.

[54] R. Romero and J. Polich. P300 habituation from auditory and visual stimuli. *Physiol Behav*, 59:517–522, 1996.

[55] E.N. Sokolov. Higher nervous functions: The orienting reflex. *Annual review of physiology*, 25:545–580, 1963.

[56] K. Squires, S. Petuchowski, C. Wickens, and E. Donchin. The effects of stimulus sequence on event related potentials: A comparison of visual and auditory sequences. *Perception & Psychophysics*, 22:31–40, 1977.

[57] K.C. Squires, C. Wickens, N.K. Squires, and E. Donchin. The effect of stimulus sequence on the waveform of the cortical event-related potential. *Science*, 193:1142–1146, 1976.

[58] N.K. Squires, K.C. Squires, and S.A. Hillyard. Two varietes of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalography and Clinical Neurophysiology*, 38:387–401, 1975.

[59] S. Sutton. P300 – thirteen years later. In *Proc. Conf. Evoked Brain Potentials and Behavior*, pages 107–126, 1997.

[60] R. Verleger. On the utility of P3 latency as an index of mental chronometry. *Psychophysiology*, 34:131–156, 1997.

[61] N.J. Wesensten, P. Badia, and J. Harsh. Time of day, repeated testing, and interblock interval effects on P300 amplitude. *Physiology & Behavior*, 47:653–658, 1988.

[62] J. Woestenburg, M. Verbaten, and J. Slangen. The removal of the eye-movement artifact from EEG by regression analysis in the frequency domain. *Biological Psychology*, 16:127–147, 1983.

[63] C.D. Woody. Characterization of an adaptive filter for the analysis of variable latency neuroelectric signals. *Medical and Biological Engineering and Computing*, 5:539–553, 1967.