

P300 amplitude age reductions are not caused by latency jitter

KRISTINE B. WALHOVD, HANNAH ROSQUIST, AND ANDERS M. FJELL

Center for the Study of Human Cognition, Institute of Psychology, University of Oslo, Oslo, Norway

Abstract

Studies of the event-related potentials (ERPs) P3a/P3b have given insights into age effects on cognitive processes in the brain, and it has been established that latency increases and amplitude decreases with age. However, if latency jitter, that is, variation in single trial latencies, is larger in elderly than in younger participants, this may create an artificial age–amplitude correlation. The aim of this article is to test whether correction for latency jitter affects the P3a/P3b age correlations. One hundred thirty-three healthy adults (20–88 years old) went through a 3-stimuli visual oddball paradigm. Latency jitter was corrected by use of a Maximum Likelihood Estimation method. The results showed that corrections for latency jitter did not significantly affect the correlations between P3a/P3b and age. It is concluded that previous reports of amplitude reduction as a function of age seem to be valid regardless of whether latency jitter correction has been applied.

Descriptors: Aging, ERP, P3a, P3b, Maximum likelihood estimation, Latency jitter

Event-related potentials (ERPs) have provided important insights about age-related changes in the human neurocognitive system. ERPs provide the opportunity to study cognitive processes in the brain directly and independently of behavioral measures. Because increasing age is associated with a range of physical alterations, both in the central and peripheral nervous system as well as in the motor system (Woodruff-Pak, 1997), ERPs may yield information that is useful to complement behavioral findings. The ERP components P3a/P3b are especially often used in aging research. Starting with Goodin, Squires, Henderson, and Starr (1978), several studies have demonstrated prolongation of P3 latency and reduction of P3 amplitude with increasing age (e.g., Pfefferbaum, Ford, Roth, & Kopell, 1980; Pfefferbaum, Ford, Wenegrat, Roth, & Kopell, 1984; Picton, Stuss, Champagne, & Nelson, 1984). In a meta-analysis of 32 studies of P3 age-effects, Polich (1996) concluded that a robust relationship exists between age and P3 latency. The latency prolongation can be interpreted as indexing slowing of information processing (Salthouse, 2000a, 2000b) as a result of generally less efficient neurocognitive functions, whereas amplitude reduction may reflect that less cognitive resources are available for the given task. Recently, Daffner et al. (2006) found that cognitively high performing old adults generated larger P3s than cognitively

average performing ones, suggesting that they managed the task successfully by appropriating more resources and that higher P3 amplitude represents a beneficial compensatory mechanism.

The P3 can be separated into P3a and P3b, which both reflect attentional processes in the brain. In a three-stimulus oddball task, P3a is elicited to distractor stimuli that the participant is instructed to ignore, whereas P3b is elicited to target stimuli that the participant is instructed to discriminate from standard stimuli and attend to by button press (see Methods section below for details). The two components are distinguished primarily by topography and have different neural generators (Baudena, Halgren, Heit, & Clarke, 1995; Bledowski et al., 2004; Halgren, Marinkovic, & Chauvel, 1998). In young adults, P3a has a fronto-central maximum, whereas P3b has a parietal maximum. In addition, P3a habituates more rapidly than P3b (see Polich, 2003, and Linden, 2005, for details on the two components).

Most ERP aging studies have focused on the P3b and have quite consistently reported latency prolongations and amplitude reductions in aging (e.g., Barrett, Neshige, & Shibasaki, 1987; Brown, Marsh, & LaRue, 1983; Fein & Turetsky, 1989; Fjell & Walhovd, 2001, 2003a, 2003b, 2003c; Goodin et al., 1978; Iragui, Kutas, Mitchiner, & Hillyard, 1993; Picton et al., 1984; Polich, 1996; Walhovd & Fjell, 2001, 2002). More recent studies have confirmed age effects also for P3a (Fjell & Walhovd, 2004, 2005; Stige, Fjell, Smith, Lindgren, & Walhovd, 2007; Walhovd & Fjell, 2001; Wang, Friedman, Ritter, Bersick, & Latif, 2006; but see Daffner et al., 2005). In fact, Fjell and Walhovd (2004) reported stronger age effects for visual P3a than visual P3b. One may speculate that a stronger age effect for P3a may be related to

This research was supported by the Norwegian Research Council (grant to Walhovd, no. 177404/W50) and the Institute of Psychology, University of Oslo (grant to Fjell).

Address reprint requests to: Kristine B. Walhovd, University of Oslo, Institute of Psychology, POB 1094 Blindern, 0317 Oslo, Norway. E-mail: k.b.walhovd@psykologi.uio.no

its possibly more frontally oriented neural generators being more prone to neural age changes. However, because equal age effects have been found for auditory P3a and P3b (Walhovd & Fjell, 2001), such an interpretation may be premature.

There appears to be a linear amplitude reduction and latency prolongation of P3a from childhood to old age (Stige et al., 2007), probably reflecting maturational processes early on, with a shift toward nonbeneficial changes to cognition in older age. For P3b, there may be a more complex age relationship. Although there appears to be a linear amplitude reduction, P3b latency shortens in adolescence, but later increases with age (Stige et al., 2007). This may be due to normal aging processes, including neuroanatomical volumetric reductions (e.g., Walhovd et al., 2005a, 2005b), and can be associated with a moderate decrease in attention capacity characteristic of normal aging (Kramer & Kray, 2006). Recently, Chapman et al. (2007) also reported that ERPs including P3s to task-relevant and -irrelevant stimuli successfully discriminated early Alzheimer's Disease from normal aging. The implications of P3a and P3b amplitude reduction and latency prolongation in electrophysiological attention components in aging may be decline of automatic orienting as well as disordered effortful attention. Understanding such age changes is pivotal to understanding age changes in a broad span of cognitive functions. This is so because automatic and effortful attention serve as gatekeepers of the mind, filtering information for further processing. Any limits set here will ultimately have consequences for what information is acted upon, reasoned about, and remembered.

P3a and P3b amplitude reductions may indicate that cognitive decline in aging is found at a basic neural level. However, one important methodological problem complicates the interpretation of the P3a/P3b amplitude–age correlations. This is due to the averaging procedure used to obtain the ERPs. The averaging procedure is based on the assumption that the ERP signal does not vary in latency from trial to trial. Variation in latency from trial to trial is called latency jitter. If the ERP signal is stochastic, latency jitter will deflate ERP components based on the average of multiple single trials, thereby increasing the error in the amplitude estimation. This may constitute a problem if the amount of latency jitter correlates with the variable of interest. It has been suggested that latency jitter increases with age (McDowell, Kerick, Maria, & Hatfield, 2003). If this is true, P3 amplitude reduction with age may be caused by the averaging method and not reflect real amplitude reductions. Differences in amount of jitter between age groups may also have an effect on the age–latency correlation. If the amplitude of the ERP signal is always the same even though it may be shifted in time, latency jitter will only affect the component average amplitude. However, if the amplitude of the single trial component varies, this may also have an effect on the latency of the average component.

Rationale

The purpose of the present article is to test whether corrections for latency jitter will affect the correlations between age and P3a/P3b amplitude and latency. Data from overlapping samples of that presented here have previously been reported with standard averaging procedures without correction for latency jitter (Fjell & Walhovd, 2004, 2005). Readers more broadly interested in P3a/P3b age effects may consult those papers, because the present analyses and discussion will primarily center on the effects of P3a/P3b latency jitter and Maximum Likelihood Estimation

(MLE) correction in aging. If latency jitter corrections have only minor impact on the age correlations, this will imply that the averaging method is valid in aging studies, at least for P3a/P3b. Further, it will imply that previously reported age–P3a/P3b correlations probably reflect real changes in amplitude and latency, not side effects of the averaging procedure. If, however, the age correlations change substantially when corrected for latency jitter, this would imply that such procedures should be employed in aging studies. Further, it would imply that previous studies reporting age correlations that are not corrected for latency jitter must be interpreted with care.

To correct for latency jitter, the single sweeps need to be shifted in time, which requires a reliable estimate for the single latency shifts. Pham, Möcks, Köhler, & Gasser (1987) and Möcks, Köhler, Gasser, & Pham (1988) suggested a MLE method that allowed latency variability and demonstrated that the procedure performed better than the often-used Woody algorithm (Woody, 1967). Jaskowski and Verleger (1999, 2000) extended this method to also include amplitude variability. In the present study, the basic procedure from Jaskowski and Verleger (2000) was used, with some modifications, allowing the latency and amplitude to vary, the only assumption being that the shape of the P3 curve is constant. In the present project, the procedure was first tested on synthetic data and then applied to a data set with 133 participants ranging from 20 to 88 years of age.

Methods

Sample

Sample characteristics are presented in Table 1. The original sample consisted of 137 healthy volunteers (70 women) evenly distributed from 20 to 89 years of age (mean 51 years, *SD* 21 years) and screened for diseases and traumas known to affect central nervous system functioning by a set of health-related questions. All received a small monetary reward for their participation. One participant was excluded on the basis of IQ criteria, and 3 more were excluded due to low performance on the ERP task (see below), reducing the sample to 133. All participants were examined with the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999) and excluded if $IQ < 85$. WASI consists of two verbal (vocabulary, similarities) and two performance (matrix reasoning, block design) tests, which were used to calculate an age-adjusted IQ score (sample mean 114, *SD* = 10, range 85–136). One participant did not complete the matrix reasoning test, and the IQ score for this person was calculated based on the other three tests. Age was not correlated with IQ ($r = .05$, $p = .56$) or education ($r = -.14$, $p = .11$). The Beck Depression Inventory (BDI; Beck & Steer, 1987) and the Mini Mental Status Exam (MMS; Folstein, Folstein, & McHugh, 1975) were included in the study at a later stage, and we thus have data from only 73 participants for the BDI and 84 for the MMS. All participants scored below 15 on the BDI and above 25 on the MMS. Due to bad MLEs (see below), the P3a data from 2 and the P3b data from 14 participants were excluded from analyses (see section on MLE estimation).

ERP Task and Recording Procedures

A three-stimuli visual oddball task with a total of 210 stimuli, .10 target and .10 distractor probability, was used. The task is a variation of one used by Comerchero and Polich (1999), shown to elicit both P3a and P3b (Polich, 2003). In the present task, the standard stimuli were blue elliptic shapes with a height of 15 cm

Table 1. Sample Characteristics

	Young (20–33 years) <i>n</i> = 44	Middle-aged (34–63 years) <i>n</i> = 44	Elderly (65–88 years) <i>n</i> = 45
Sex	26 f/18 m	19 f/25m	25 f/20 m
Age	27.7 (3.4)	51.3 (8.0)	75.6 (6.4)
Education	15.0 (2.3)	16.0 (2.4)	14.3 (3.0)
IQ	113.7 (8.0)	115.1 (9.7)	114.3 (12.5)
BDI ^a	3.0 (3.5)	2.4 (2.3)	6.5 (3.5)
MMS ^b	29.2 (0.8)	29.2 (0.7)	28.5 (1.2)

^aBeck Depression Inventory.

^bMini Mental Status Exam.

and a width of 12.5 cm. The targets, to which the participant was told to press a button, were blue elliptic shapes with height and width of 17.5 and 14.5 cm, respectively. The distractor stimuli, which the participant was told to ignore, were blue rectangles of 21 × 17 cm. Viewing distance was 100 cm. The stimuli were presented on a 21-in. computer screen with a black background, with a visual field of about 9° × 7°, 10° × 8°, and 12° × 10° for the standard, target, and distractor stimuli, respectively. The small difference between targets and standards and the large difference between targets and distractors were chosen to maximize the P3a (Comerchero & Polich, 1999). Presentation time was 0.5 s and interstimulus interval ISI was 1.5 s. Cutoff criteria for task performance were set to 20% target misses, 20% responses to standards, or 25% responses to distractors. Before recording, an example task with eight standard and three target stimuli was presented to prime the participants for the task and to ascertain that all could discriminate targets from standards.

The task was administered while the participants sat in a reclining chair within a sound-attenuating recording chamber. The electrodes were placed in accordance with the international 10–20 system. A total of 20 electrodes (Ag/AgCl) were used for recording: Fp1, Fp2, F7, F3, Fz, F8, F4, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, Oz, and O2, referred to the left mastoide. In the present article, statistical analyses are confined to Fz, Cz, and Pz (see below). A VEOG channel was obtained by placing one electrode above and one below the left eye, and ground was placed anteriorly on the right side. Inter-electrode impedance was generally measured to be less than 10 kΩ. For the recording of EEG activity, A/D rate was 500 Hz, and filter setting was 0.10 Hz (high pass) and 70 Hz (low pass). In addition, a 50-Hz notch filter was applied. The signals were amplified by a SynAmp DC amplifier (Neuroscan Inc.). Epochs were rejected from averaging if amplitude exceeded ± 110 μV, and eyeblinks were corrected for statistically in accordance with recommendations from Semlitsch, Anderer, Schuster, and Presslich (1986).

Analyses were performed for targets and distractors separately. After the above-mentioned cutoff criteria were employed, analyses were performed for all accepted target and distractor trials, without exclusion of the remaining few error trials for targets (see below). This was done in light of a low error rate and because it was deemed a possible bias to exclude trials for targets while not being able to perform similar procedures for distractors, to which no response was required. EEG was segmented in epochs of 900 ms duration (–100 ms to 800 ms relative to stimulus onset), and was digitally filtered (3 Hz low pass) and baseline corrected. The harsh filter setting was chosen because single sweep data are more contaminated by high frequency noise than average data. This is close to what Smulders, Kenemans,

and Kok (1994) recommend as the optimal filter setting for peak picking in terms of sensitivity and reliability and to what Jaskowski and Verleger (2000) used in their paper on maximum-likelihood estimation of single sweep ERP data. Neuroscan software was used to present stimuli and record data. A combination of Neuroscan software, Matlab code generously provided by Piotr Jaskowski and Rolf Verleger (see Jaskowski and Verleger, 2000), and custom-made procedures implemented in Matlab were used to preprocess the data and calculate peak values, average values, and corrections of latency jitter. P300 peaks were determined algorithmically, in accordance with Pfefferbaum, Ford, and Kraemer's (1990) recommendations, defined as the most positive point constituting a peak within 250 and 650 ms poststimulus. This was done using both standard peak picking and with maximum-likelihood estimation (see below). Statistical analyses were mainly confined to Cz for distractor stimuli (P3a) and Pz for target stimuli (P3b), because the P300 components in the present paradigm generally are most pronounced at these recording sites (Fjell & Walhovd, 2004). Still, because it has been demonstrated in several previous studies that the midline topography of both the P3a and the P3b changes with age (Fabiani, Friedman, & Cheng, 1998; Fjell & Walhovd, 2004; Fjell, Walhovd, & Reinvang, 2005; Friedman & Fabiani, 1995; Nielsen-Bohlman & Knight, 1995), correlation analyses will be performed also for Fz.

ERP Analyses—Maximum Likelihood Estimation

Illustrations of individual MLEs are presented in Figure 1. In the present study, the basic MLE procedure as described by Jaskowski and Verleger (2000) was used, with some modifications, allowing the latency and amplitude to vary, the only assumption being that the shape of the P300 curve is constant. In the model employed, the output signal can be described as

$$r_j(t) = a_j s(t + \tau_j) + e_j(t), \quad (j = 1, \dots, N, t) \in [0, T], \quad (1)$$

where $r_j(t)$ is the EEG output, $s(t)$ is the stimulus-related ERP signal, a_j is the normalization factor for the amplitude, τ_j denotes the latency shift, and $e_j(t)$ is noise. N is the number of sweeps and j indexes the sweeps. The noise was assumed to be zero mean stationary Gaussian. Equation (1) was Fourier transformed, and an MLE approach applied, which can be seen as a linear regression method (Pham et al., 1987). The Fourier transform of Equation (1) is

$$R_j(\lambda) = a_j S(\lambda) e^{-i\lambda\tau_j} + E_j(\lambda), \quad \lambda = 0, 2\pi/N, 4\pi/N, \dots, N\pi/N, \quad (2)$$

where it is assumed that $S(t + \tau_j)$ is periodic so that the Fourier transform of s becomes

$$S(\lambda) e^{-i\lambda\tau_j}.$$

In the frequency domain, the E_j 's can be treated as independent complex variates with mean zero and variance $f(\lambda)$. Further, $s(t)$ has negligible power for higher frequencies; thus the calculations can be restricted to, for example, the 10 lowest λ (Jaskowski & Verleger, 2000).

The log likelihood, L , given by Brillinger (1981) is

$$L = - \sum_j \sum_{\lambda} \left[\log\{f\} + \frac{1}{f} |R_j - a_j S e^{-i\lambda\tau_j}|^2 \right]. \quad (3)$$

R_j is the Fourier transform of the EEG output. Thus the unknown parameters are τ_j , a_j , S , and f . It is described in the Ap-

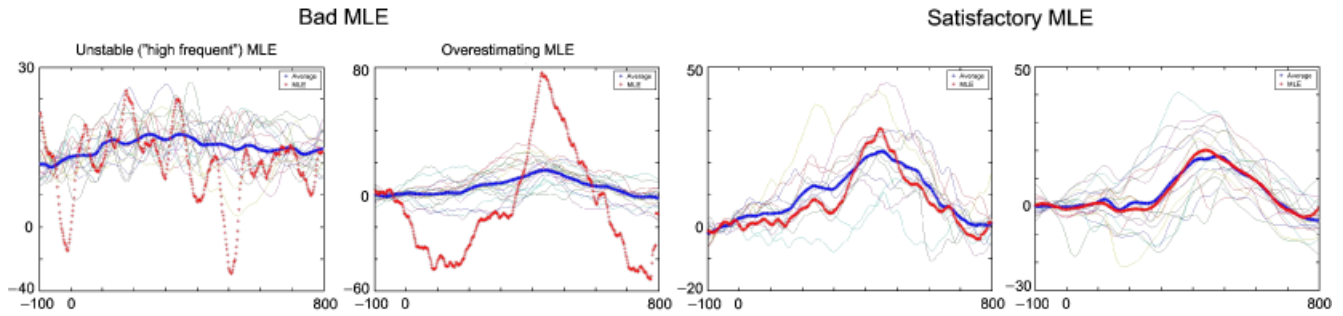


Figure 1. Examples of individual MLEs with corresponding average curves are presented. The two plots to the left are examples of bad MLEs. These were discarded from statistical analyses. Bad MLEs were mainly due to one of two reasons: The MLE signal was unstable and did not show a clearly defined peak (plot number one from the left) or the MLE overestimated the amplitude due to many trials near 0 μV (plot number two from left). The two plots to the right are examples of satisfactory MLEs, where the MLE amplitude is larger than the average amplitude, due to correction for latency jitter.

pendix how these parameters can be calculated, through log likelihood and Fisher scoring.

To aid the decision of which parameters were optimal to use in the present study, a simulated data set was constructed (see Figure 2) with 20 sweeps, standard deviation (SD) of latency = 150 ms, of amplitude = 0.5 (scaling factor), and with

different amounts of noise. Signal-to-noise ratio (SNR) was calculated as described by Handy (2004), where the maximum amplitude from the averaged signal was divided by the SD of the noise (EEG signal minus averaged signal). Because the signal has negligible power above a certain frequency in the frequency domain, the analysis can be restricted to few frequency components.

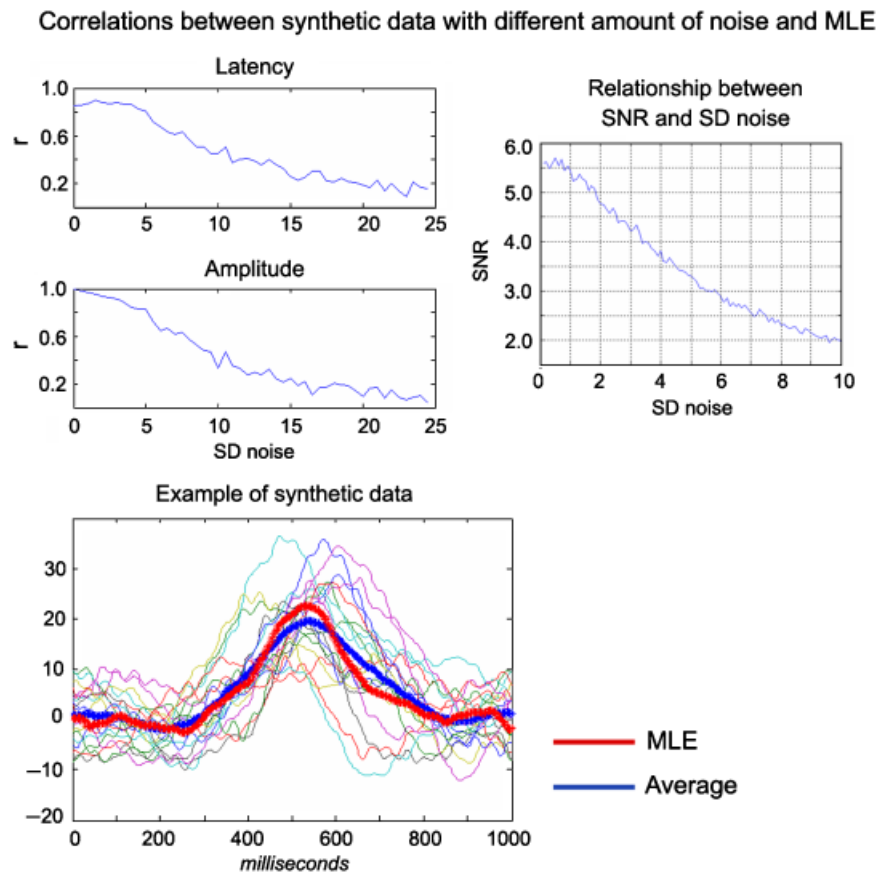


Figure 2. Main results from the simulated data set used to test the Maximum Likelihood Estimation (MLE) approach. The lower left panel shows an example of synthetic data, with the average ERP (blue curve) and the MLE signal (red curve) shown in thicker lines. The upper left panels show the correlation between the synthetic data and the MLE approach for different signal-to-noise (SNR) ratios. For an SNR between 3 and 5, the calculations from the simulations correlated >0.8 . However, for $SNR < 2$ the correlation was <0.4 . Larger variations in latency or amplitude had little influence on the correlation. The right panel shows the relationship between the SNR and the SD of noise (two ways of expressing noise levels in the synthetic data).

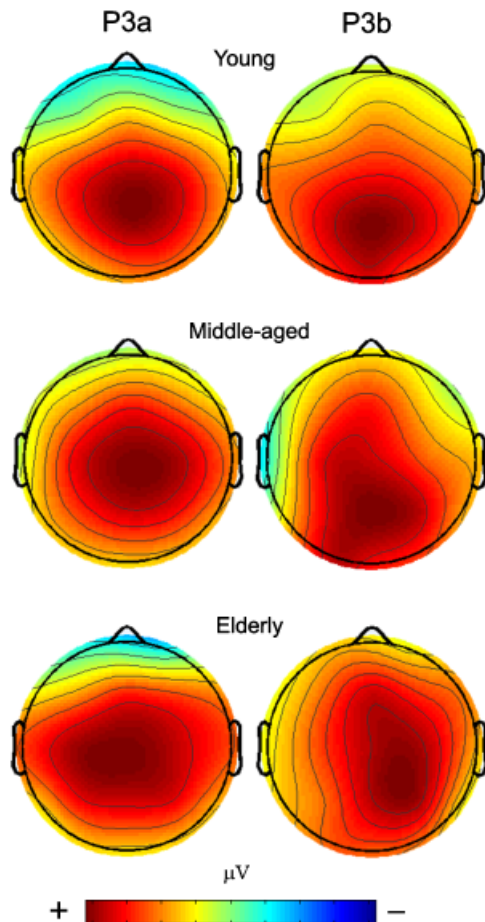


Figure 3. Voltage maps split by age group and component (P3a, P3b), showing the distribution of activity at the point in the P300 time window where the topographical variance was the largest for each age group and component. The age groups consist of equally sized groups of young (20–33 years), middle-aged (34–64 years), and elderly (65–88 years) participants. The voltage scale is optimized for each age group and thus cannot be compared across plots. As can be seen, the distribution of activity is central for the P3a and more parietally distributed for the P3b, especially in the group of young and middle-aged participants.

In accordance with Jaskowski and Verleger (1999), the 10 lowest frequencies acquired after the Fourier transform were chosen. The simulations showed that, for high SNR, the choice of number of frequencies was not crucial. However, for lower SNR and more than 15 frequency components, the noise contributed much, and 10 frequencies were thus chosen as the optimal solution. Simulations further showed that the number of frequency components influenced the latency correlation more than the amplitude correlation.

For the most extreme signals, τ_j often oscillates. To cope with this, τ_j and a_j were calculated as averages of the fourth and the fifth iterations. This improved the correlations slightly. For an SNR between 3 and 5, the calculations from the simulations correlated >0.8 (see Figure 2). However, for $\text{SNR} < 2$ the correlation was less than 0.4. Larger variations in latency or amplitude had little influence on the correlation.

Null trials (trials where no ERP signal can be found) constitute a huge problem for single sweep analyses (Jaskowski & Verleger, 2000). In case of MLE, null signals may enhance the

amplitude too much. We tested three different criteria for removing null trials: minimum correlation coefficients, minimum covariance coefficients, and mean amplitude not exceeding $0 \mu\text{V}$ in the time window around the peak. After close inspection of the data, it was decided that removal of epochs where the mean amplitude was below $0 \mu\text{V}$ in the time window ± 100 ms from the peak value yielded the best results both in terms of sensitivity and specificity. Thus, this approach was used to exclude null trials from further analyses. A null trial was then defined as a trial in which the mean amplitude in the time window ± 100 ms around the peak did not exceed $0 \mu\text{V}$, and all such trials were removed from the analyses. The removal of null trials naturally lowered the number of accepted trials to be analyzed. The mean number of accepted trials was 17.2 for distractors, which did not correlate with age ($r = -.15, p = .088$), and 14.4 for targets, which did correlate with age ($r = -.25, p < .01$). In addition, the final MLE results were carefully inspected, and estimations judged to be inaccurate were discarded. Criteria for exclusion of bad MLEs were based primarily on two rules: MLEs were discarded if the curve had higher amplitude than most of the single sweep curves or if the MLE curve deviated much from the average in terms of shape, typically with several positive and negative peaks. Illustrations of bad MLEs are presented in Figure 1. This led to the exclusion of 2 participants for P3a Cz MLE analyses, 7 for P3a Fz MLE analyses, and 14 for both P3b Pz and P3b Fz MLE analyses.

Statistical Analyses

Statistical analyses were done with SPSS 12.0.1, and EEG-lab 4.515 (Delorme & Makeig, 2004) was used to create voltage maps. First, average and MLE amplitude was compared by t tests. If the MLE procedure successfully corrects for latency jitter, the mean MLE amplitude should be significantly larger than the mean value obtained by the traditional averaging procedure. Both traditional peak picking and MLE results were correlated with age, and the results were compared with t tests for dependent correlations. All statistical analyses were performed with statistical outliers excluded (deleted studentized residual value ≥ 2.0). The number of outliers in each analysis was small, typically < 5 . The number of participants will vary slightly between analyses, and the n or df is always presented along with the statistical results.

Results

Behavioral Data

Mean reaction time (RT) in the sample was 520 ms ($SD = 72$ ms). Mean target hit rate was .96, and rates of false alarms to distractor and standard were .03 and .01, respectively. Age correlated with mean RT ($r = .20, p < .05$), number of target hits ($r = -.21, p < .05$), number of false alarms to the distractors ($r = -.24, p < .01$), and number of false alarms to the standards ($r = -.25, p < .01$).

Topography of P3a versus P3b

Uncorrected voltage maps for each age group and component are shown in Figure 3, and mean amplitude for each age group and electrode are presented in Figure 4. As can be seen, the P3a has a more central distribution than the P3b in all age groups. A 3 (electrodes: Fz, Cz, Pz) \times 2 (components with uncorrected amplitudes: P3a, P3b) \times 3 (age groups: young, middle-aged, elderly) ANOVA confirmed this impression by yielding a significant

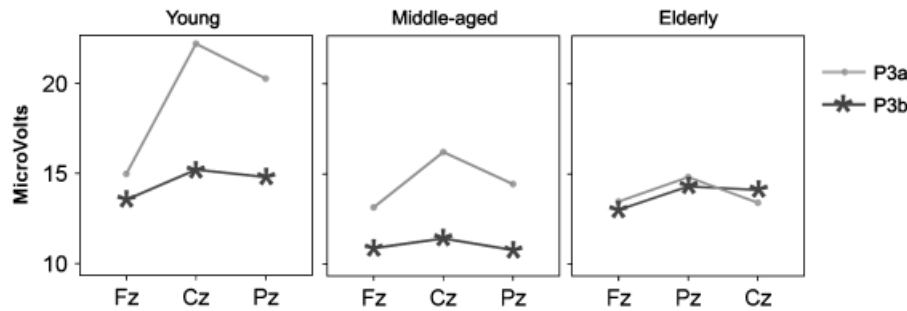


Figure 4. Mean amplitude for P3a and P3b across the midline electrodes (Fz, Cz, Pz) for each of the three age groups (young: 20–34 years, middle-aged: 34–64 years, elderly: 65–88 years). As can be seen, the P3a peaks at Cz, whereas the P3b is more parietally distributed. This was confirmed by a significant Component \times Electrode ANOVA. It can also be seen that the activity is more evenly distributed with increasing age.

Electrode \times Component interaction effect, $F(2,254) = 16.36$, $p < 10^{-6}$, Greenhouse Geisser epsilon = .957. This indicates that the two components were topographically different. There was also a main effect of stimuli, $F(1,126) = 31.23$, $p < 10^{-6}$. As can be seen from Figure 4, these effects were caused by a central

maximum for P3a and a parietal maximum for P3b, and generally larger amplitudes for P3a. These results indicate that a valid P3a versus P3b distinction was obtained. Further, a main effect of age was found, $F(2,126) = 6.53$, $p < .005$, as well as an Age \times Electrode interaction effect, $F(4,252) = 4.44$, $p < .01$, and an Age \times Electrode \times Stimuli interaction effect, $F(4,252) = 7.18$, $p < 10^{-4}$. Inspections of Figure 4 revealed that these effects were due to relatively more frontal activation with increasing age (i.e., the frontal shift), and that this was more prominent for P3a than for P3b.

MLE versus Average Amplitude

MLE and average amplitude was contrasted by use of paired samples t tests. P3a amplitude was analyzed at Cz and P3b amplitude at Pz. As expected, MLE amplitude was significantly larger than average amplitude (P3a: 19.06 μ V vs. 17.64 μ V, for MLE and average, respectively, $t[129] = 7.126$, $p < 10^{-10}$; P3b: 14.80 μ V vs. 12.79 μ V, for MLE and average, respectively, $t[115] = 7.867$, $p < 10^{-11}$).

P3a/P3b–Age Correlations

Grand average P3a and P3b ERP curves as well as MLE grand average curves for the youngest, middle, and oldest parts of the sample are presented in Figure 5, and scatterplots showing individual data points are presented in Figure 6. Age effects across groups are visible for both components, both for the uncorrected curves and the MLE curves. P3a latency measures and age correlated significantly in the case of average latency ($r = .59$, $p < 10^{-12}$, $n = 129$) and MLE latency ($r = .58$, $p < 10^{-12}$, $n = 129$), and the correlations were not significantly different from each other ($t = 0.439$, n.s.). For P3a amplitude measures, age correlated with both average amplitude ($r = -.46$, $p < 10^{-6}$, $n = 123$) and MLE amplitude ($r = -.47$, $p < 10^{-7}$, $n = 124$), and the difference between the coefficients was not significant ($t = 0.350$, n.s.). The analyses were repeated for P3a at Fz, where age did not correlate with either MLE amplitude ($r = -.02$, n.s., $n = 118$) or average amplitude ($r = -.07$, n.s., $n = 118$). The same was the case for P3a latency at Fz (average: $r = .00$, n.s., $n = 118$; MLE: $r = -.03$, n.s., $n = 118$).

P3b latency at Pz and age correlated significantly in the case of both average latency ($r = .40$, $p < 10^{-4}$, $n = 110$) and MLE latency ($r = .32$, $p < .001$, $n = 111$), but not significantly different ($t = 1.208$, n.s.). For P3b Pz amplitude, age correlated with both average amplitude ($r = -.36$, $p < 10^{-4}$, $n = 114$) and MLE amplitude ($r = -.40$, $p < 10^{-4}$, $n = 112$), also not significantly different ($t = 1.329$, n.s.). The analyses were repeated for Fz, where age

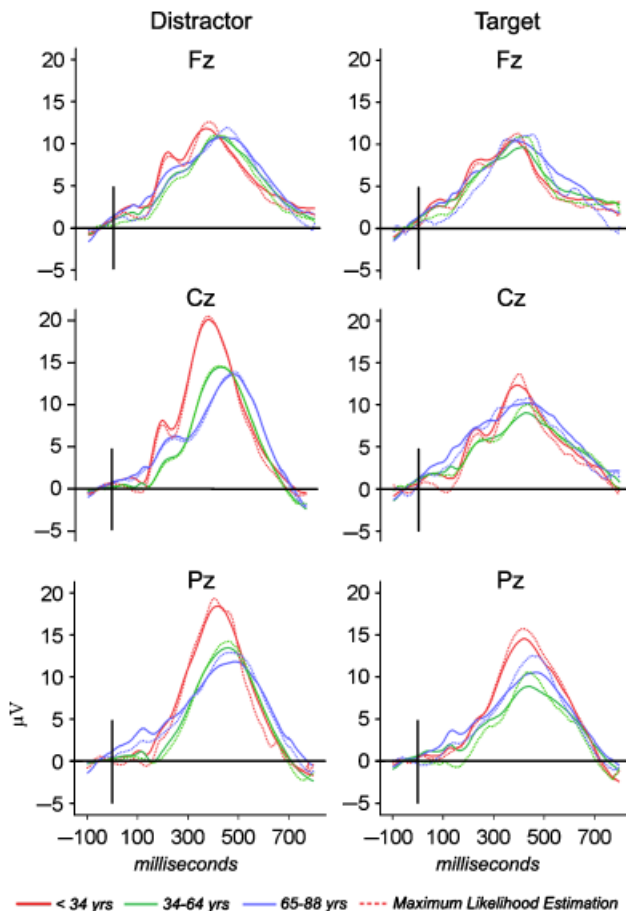


Figure 5. The ERP curves elicited to the distractor (P3a) and the target (P3b), split by age group. The age groups consist of equally sized groups of young (20–33 years), middle-aged (34–64 years), and elderly (65–88 years) participants. The dotted lines are the Maximum Likelihood Estimation (MLE) curves, and the colored solid lines are the averages. As can be seen, the MLE curves do not differ much from the averages in case of P3a, whereas the differences for P3b are larger. For both P3a and P3b, clear age effects on P3 amplitude and latency can be seen.

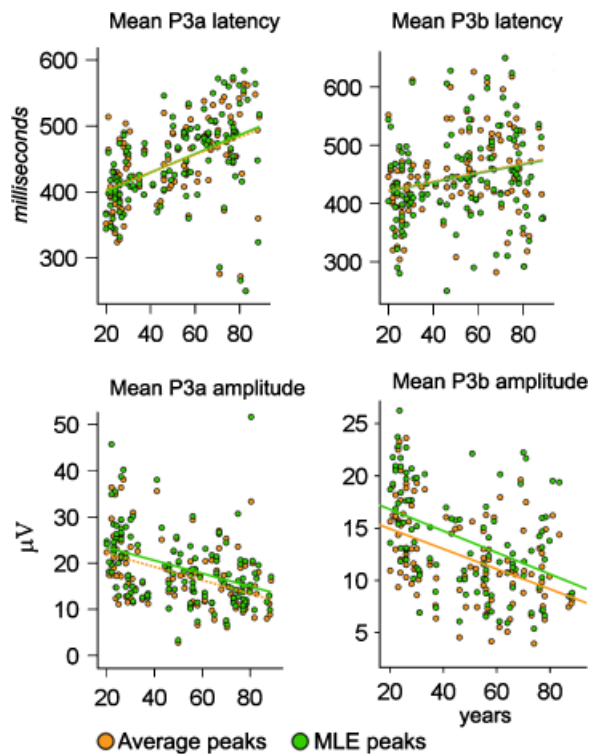


Figure 6. Relationships between age (in years) and P3a/P3b amplitude and latency. Both the results of the MLE method (green dots and lines) and the traditional averaging method (orange dots and lines) are shown. Outliers are excluded from the statistical analyses, but not from the scatterplots. The regression equations are as follows: P3a average latency = $367.87 + 1.58x$, $df = 1,127$, $p < 10^{-12}$, P3a MLE latency = $364.66 + 1.63x$, $df = 1,127$, $p < 10^{-12}$, P3a average amplitude = $23.06 - 0.12x$, $df = 1,121$, $p < 10^{-6}$, P3a MLE amplitude = $25.06 - 0.14x$, $df = 1,122$, $p < 10^{-7}$, P3b average latency = $393.44 + 1.01x$, $df = 1,108$, $p < 10^{-4}$, P3b MLE latency = $396.76 + 0.91x$, $df = 1,109$, $p < .001$, P3b average amplitude = $16.25 - 0.08x$, $df = 1,112$, $p < 10^{-4}$, P3b MLE amplitude = $18.61 - 0.10x$, $df = 1,110$, $p < 10^{-4}$, where $x = \text{age}$ (in years).

correlated significantly with MLE amplitude ($r = -.28$, $p < .01$, $n = 117$) but not with average amplitude ($r = -.09$, n.s., $n = 113$), a significant difference ($t = 3.916$, $p < .01$). Finally, age did not

correlate with average ($r = -.05$, n.s., $n = 110$) or MLE ($r = -.07$, n.s., $n = 109$) P3b latency at Fz.

Discussion

As expected, P3a latency correlated highly with age, and correction for latency jitter did not affect the correlation notably ($r = .59$ before vs. $r = .58$ after latency correction). The latency increase with advancing age can be interpreted as reflecting slower processing speed, resulting from less efficient neurocognitive processing, and has been demonstrated previously with visual three-stimuli tasks (e.g., Fjell & Walhovd, 2004, 2005, with overlapping samples). Also for P3a amplitude, the correlation with age was almost identical for the latency-corrected and the uncorrected component ($r = -.47$ and $-.48$, respectively). Amplitude reductions and changes have been interpreted as reflecting that less cognitive resources are available for the given task. However, it could be possible that such amplitude effects are due to different degrees of latency jitter in groups of participants, not genuine amplitude deflations. The present data indicate that amplitude reductions in aging probably are valid and that they cannot be attributed to effects of latency jitter.

For P3b, the age correlations were a bit smaller, but the general tendency from the P3a analyses was confirmed. Both uncorrected ($r = .40$) and corrected ($r = .32$) latency correlated with age, as did uncorrected ($r = -.36$) and corrected ($r = -.40$) amplitude. Thus, the conclusion from these analyses is that latency jitter correction has a minute effect on the age correlations for P3a/P3b latency and amplitude, and that previous reports of amplitude reductions as a function of age (e.g., Polich, 1996) seem to be valid regardless of whether latency jitter correction has been applied.

This suggests a real age reduction in both automatic orienting (P3a) and more controlled attention (P3a) at an early neural level in aging. Thus, even though correction for latency jitter may ensure that artificial amplitude effects are not observed, the need for this in P3 age studies seems to be limited. However, MLE has the potential to yield information not inherent in average ERP measures, that is, the quantification of single sweep variation, which in itself may be an important feature of ERPs (Fjell, Rosquist, & Walhovd, in press).

REFERENCES

- Barrett, G., Neshige, R., & Shibasaki, H. (1987). Human auditory event related potentials: Effects of response condition and age. *Electroencephalography and Clinical Neurophysiology*, *66*, 409–419.
- Baudena, P., Halgren, E., Heit, G., & Clarke, J. M. (1995). Intracerebral potentials to rare target and distractor auditory and visual stimuli. III. Frontal cortex. *Electroencephalography and Clinical Neurophysiology*, *94*, 251–264.
- Beck, A. T., & Steer, R. (1987). *Beck Depression Inventory Scoring Manual*. New York: The Psychological Corporation.
- Bledowski, C., Prvulovic, D., Hoechstetter, K., Scherg, M., Wibral, M., Goebel, R., et al. (2004). Localizing P300 generators in visual target and distractor processing: A combined event-related potential and functional magnetic resonance imaging study. *Journal of Neuroscience*, *24*, 9353–9360.
- Brillinger, D. R. (1981). Some aspects of the analysis of evoked response experiments. In H. Csörgő, D. A. Dawson, J. N. K. Rao, & A. J. M. E. Saleh (Eds.), *Statistics and related topics* (pp. 155–168). Amsterdam: North-Holland.
- Brown, W. S., Marsh, J. T., & LaRue, A. (1983). Exponential electrophysiological aging: P3 latency. *Electroencephalography and Clinical Neurophysiology*, *55*, 277–285.
- Chapman, R. M., Nowlis, G. H., McCrary, J. W., Chapman, J. A., Sandoval, T. C., Guillily, M. D., et al. (2007). Brain event-related potentials: Diagnosing early-stage Alzheimer's disease. *Neurobiology of Aging*, *28*, 194–201.
- Comerchero, M. D., & Polich, J. (1999). P3a and P3b from typical auditory and visual stimuli. *Clinical Neurophysiology*, *110*, 24–30.
- Daffner, K. R., Ryan, K. K., Williams, D. M., Budson, A. E., Rentz, D. M., Scinto, L. F. M., et al. (2005). Age-related differences in novelty and target processing among cognitively healthy high performing adults. *Neurobiology of Aging*, *26*, 1283–95.
- Daffner, K. R., Ryan, K. K., Williams, D. M., Budson, A. E., Rentz, D. M., Wolk, D. A., et al. (2006). Increased responsiveness to novelty is associated with successful cognitive aging. *Journal of Cognitive Neuroscience*, *18*, 1759–1773.

- Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*, 9–21.
- Fabiani, M., Friedman, D., & Cheng, J. C. (1998). Individual differences in P3 scalp distribution in older adults, and their relationship to frontal lobe function. *Psychophysiology*, *35*, 698–708.
- Fein, G., & Turetsky, B. (1989). P300 latency variability in normal elderly: Effects of paradigm and measurement technique. *Electroencephalography and Clinical Neurophysiology*, *72*, 384–394.
- Fjell, A. M., Rosquist, H., & Walhovd, K. B. (in press). Instability in the latency of P3a/P3b brain potentials and cognitive function in aging. *Neurobiology of Aging*.
- Fjell, A. M., & Walhovd, K. B. (2001). P300 and neuropsychological tests as measures of aging: Scalp topography and cognitive changes. *Brain Topography*, *14*, 25–40.
- Fjell, A. M., & Walhovd, K. B. (2003a). Effects of auditory stimulus intensity and hearing threshold on the relationship among P300, age, and cognitive function. *Clinical Neurophysiology*, *114*, 799–807.
- Fjell, A. M., & Walhovd, K. B. (2003b). On the topography of P3a and P3b—A factor-analytic study with orthogonal Procrustes rotation to maximize congruence. *Brain Topography*, *15*, 153–164.
- Fjell, A. M., & Walhovd, K. B. (2003c). P3a and neuropsychological ‘frontal’ tests in aging. *Aging, Neuropsychology, and Cognition*, *10*, 169–181.
- Fjell, A. M., & Walhovd, K. B. (2004). Life-span changes in P3a. *Psychophysiology*, *41*, 575–583.
- Fjell, A. M., & Walhovd, K. B. (2005). Age-sensitivity of P3 in high-functioning adults. *Neurobiology of Aging*, *26*, 1297–1299.
- Fjell, A. M., Walhovd, K. B., & Reinvang, I. (2005). Age-dependent changes in distribution of P3a/P3b amplitude and thickness of the cerebral cortex. *NeuroReport*, *16*, 1451–1454.
- Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-mental state. *Journal of Psychiatric Research*, *12*, 189–198.
- Friedman, D., & Fabiani, M. (1995). Memory and aging: An event-related brain potential perspective. In P. Allen & T. Bashore (Eds.), *Aging differences in word and language processing* (pp. 345–389). Amsterdam: Elsevier.
- Goodin, D. S., Squires, K. C., Henderson, B., & Starr, A. (1978). Age-related variation in evoked potentials to auditory stimuli in normal human subjects. *Electroencephalography and Clinical Neurophysiology*, *44*, 447–458.
- Halgren, E., Marinkovic, K., & Chauvel, P. (1998). Generators of late cognitive potentials in auditory and visual oddball tasks. *Electroencephalography and Clinical Neurophysiology*, *106*, 156–164.
- Handy, T. C. (2004). *Event-related potentials: A methods handbook*. Cambridge, MA: The MIT Press.
- Iragui, V. J., Kutas, M., Mitchiner, M. R., & Hillyard, S. A. (1993). Effects of aging on event-related brain potentials and reaction times in an auditory oddball task. *Psychophysiology*, *30*, 10–22.
- Jaskowski, P., & Verleger, R. (1999). Amplitudes and latencies of single-trial ERP estimated by a maximum likelihood method. *IEEE Transactions on Biomedical Engineering*, *46*, 987–993.
- Jaskowski, P., & Verleger, R. (2000). An evaluation of methods for single-trial estimation of P3 latency. *Psychophysiology*, *37*, 153–162.
- Kramer, A., & Kray, J. (2006). Aging and attention. In E. Bialystok & F. I. M. Craik (Eds.), *Lifespan cognition. Mechanisms of change* (pp. 57–69). New York: Oxford University Press.
- Linden, D. E. J. (2005). The P300: Where in the brain is it produced and what does it tell us? *The Neuroscientist*, *11*, 563–576.
- McDowell, K., Kerick, S. E., Maria, D. L. S., & Hatfield, B. D. (2003). Aging, physical activity, and cognitive processing: An examination of P300. *Neurobiology of Aging*, *24*, 597–606.
- Möcks, J., Köhler, W., Gasser, T., & Pham, D. T. (1988). Novel approaches to the problem of latency jitter. *Psychophysiology*, *25*, 217–226.
- Nielsen-Bohlman, L., & Knight, R. T. (1995). Prefrontal alterations during memory processing in aging. *Cerebral Cortex*, *5*, 541–549.
- Pfefferbaum, A., Ford, J. M., & Kraemer, H. C. (1990). Clinical utility of long latency ‘cognitive’ event-related potentials (P3): The cons. *Electroencephalography and Clinical Neurophysiology*, *6*, 6–12.
- Pfefferbaum, A., Ford, J. M., Roth, W. T., & Kopell, B. S. (1980). Age-related changes in auditory event-related potentials. *Electroencephalography and Clinical Neurophysiology*, *49*, 266–276.
- Pfefferbaum, A., Ford, J. M., Wenegrat, B. G., Roth, W. T., & Kopell, B. S. (1984). Clinical application of the P3 component of event-related potentials. I. Normal aging. *Electroencephalography and Clinical Neurophysiology*, *59*, 85–103.
- Pham, D. T., Möcks, J., Köhler, W., & Gasser, T. (1987). Variable latencies of noisy signals: Estimation and testing in brain potential data. *Biometrika*, *74*, 525–533.
- Picton, T. W., Stuss, D. T., Champagne, S. C., & Nelson, R. F. (1984). The effects of age on human event-related potentials. *Psychophysiology*, *21*, 312–325.
- Polich, J. (1996). Meta-analysis of P300 normative aging studies. *Psychophysiology*, *33*, 334–353.
- Polich, J. (2003). Theoretical overview of P3a and P3b. In J. Polich (Ed.), *Detection of change: Event-related potential and fMRI findings (83–98)*. Boston: Kluwer Academic Press.
- Salthouse, T. A. (2000a). Aging and measures of processing speed. *Biological Psychology*, *54*, 35–54.
- Salthouse, T. A. (2000b). Pressing issues in cognitive aging. In D. Park & N. Schwarz (Eds.), *Cognitive aging: A primer*. Philadelphia: Psychology Press.
- Semlitsch, H. V., Anderer, P., Schuster, P., & Presslich, O. (1986). A solution for reliable and valid reduction of ocular artifacts applied to the P300 ERP. *Psychophysiology*, *23*, 695–703.
- Smulders, F. T. Y., Kenemans, J. L., & Kok, A. (1994). A comparison of different methods for estimating single-trial latencies. *Electroencephalography and Clinical Neurophysiology*, *92*, 107–114.
- Stige, S., Fjell, A. M., Smith, L., Lindgren, M., & Walhovd, K. B. (2007). The development of visual P3a and P3b. *Developmental Neuropsychology*, *32*, 563–584.
- Walhovd, K. B., & Fjell, A. M. (2001). Two- and three-stimuli auditory oddball ERP tasks and neuropsychological measures in aging. *NeuroReport*, *12*, 3149–3153.
- Walhovd, K. B., & Fjell, A. M. (2002). The relationship between P3 and neuropsychological function in an adult life span sample. *Biological Psychology*, *62*, 65–87.
- Walhovd, K. B., Fjell, A. M., Reinvang, I., Lundervold, A., Dale, A. M., Eilertsen, D. E., et al. (2005a). Effects of age on volumes of cortex, white matter and subcortical structures. *Neurobiology of Aging*, *26*, 1261–1270.
- Walhovd, K. B., Fjell, A. M., Reinvang, I., Lundervold, A., Dale, A. M., Quinn, B. T., et al. (2005b). Neuroanatomical aging: Universal but not uniform. *Neurobiology of Aging*, *26*, 1279–1282.
- Wang, J. T., Friedman, D., Ritter, W., Bersick, M., & Latif, L. (2006). Aging effects of involuntary attentional capture in speech sound analysis. *Neurobiology of Aging*, *27*, 1164–1179.
- Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. San Antonio, TX: The Psychological Corporation.
- Woodruff-Pak, D. S. (1997). *The neuropsychology of aging*. Malden, MA: Blackwell Publishers.
- Woody, C. D. (1967). Characterization of an adaptive filter for the analysis of variable latency neuroelectric signals. *Medical and Biological Engineering*, *5*, 539–553.

(RECEIVED April 18, 2007; ACCEPTED October 29, 2007)

APPENDIX

We describe how to estimate the four unknown parameters τ_j , a_j , S , and f . To find the unknown parameters the log-likelihood function was derived with respect to each of the parameters. For τ_j and a_j , the Fisher scoring was applied by iterating over p :

$$\tau_j^{p+1} = \tau_j^p - \frac{L'(\tau_j)}{\langle L''(\tau_j) \rangle},$$

where $\langle L''(\tau_j) \rangle$ is the expectation value of the second derivative. To make the solutions uniquely defined, the conditions $\sum_j \tau_j = 0$ and $\sum_j a_j = 1$ were defined.

For estimating τ_j , the first and second derivatives of the log likelihood with respect to τ_j need to be calculated:

$$\begin{aligned} L'(\tau_j) &= -\operatorname{Re} \left(\sum_{\lambda} \frac{1}{f} 2(R_j - a_j S e^{-i\lambda\tau_j}) (i\lambda a_j S^* e^{i\lambda\tau_j}) \right) \\ &= -\operatorname{Im} \left(2 \sum_{\lambda} \frac{1}{f} \lambda (a_j R_j S^* e^{i\lambda\tau_j} - a_j^2 |S|^2) \right), \end{aligned}$$

where $a_j^2 |S|^2$ is real and can be deleted. Thus

$$\begin{aligned} L'(\tau_j) &= -\operatorname{Im} \left(2 \sum_{\lambda} \frac{1}{f} \lambda a_j R_j S^* e^{i\lambda\tau_j} \right) \\ &= \operatorname{Im} \left(2 \sum_{\lambda} \frac{1}{f} \lambda a_j R_j^* S e^{-i\lambda\tau_j} \right). \end{aligned}$$

The second derivative is calculated as

$$\begin{aligned} L''(\tau_j) &= \frac{\partial}{\partial \tau_j} \operatorname{Im} \left(2 \sum_{\lambda} \frac{1}{f} \lambda a_j R_j^* S e^{-i\lambda\tau_j} \right) \\ &= \operatorname{Im} \left(2 \sum_{\lambda} \frac{1}{f} (-i\lambda) \lambda a_j R_j^* S e^{-i\lambda\tau_j} \right) \\ &= -\operatorname{Re} \left(2 \sum_{\lambda} \frac{1}{f} \lambda^2 a_j R_j S^* e^{i\lambda\tau_j} \right). \end{aligned}$$

The expected value of the second derivative becomes

$$\begin{aligned} \langle L''(\tau_j) \rangle &= \left\langle -\operatorname{Re} \left(2 \sum_{\lambda} \frac{1}{f} \lambda^2 a_j R_j S^* e^{i\lambda\tau_j} \right) \right\rangle \\ &= -\operatorname{Re} \left(2 \sum_{\lambda} \frac{1}{f} \lambda^2 a_j \langle R_j \rangle \langle S^* e^{i\lambda\tau_j} \rangle \right) \\ &= [[R_j = a_j S e^{-i\lambda\tau_j}]] = -2 \sum_{\lambda} \frac{1}{f} \lambda^2 a_j^2 |\hat{S}|^2. \end{aligned}$$

Thus

$$\tau_j^{p+1} = \tau_j^p - \frac{L'(\tau_j)}{\langle L''(\tau_j) \rangle} = \tau_j^p + \frac{\sum_{\lambda} \frac{1}{f} \lambda a_j \operatorname{Im} \left(R_j^* \hat{S} e^{-i\lambda\tau_j} \right)}{\sum_{\lambda} \frac{1}{f} \lambda^2 a_j^2 |\hat{S}|^2}.$$

For estimating a_j ,

$$\begin{aligned} L'(a_j) &= \frac{\partial L}{\partial a_j} = -\operatorname{Re} \left(\sum_{\lambda} \frac{1}{f} 2(R_j - a_j S e^{-i\lambda\tau_j}) (S^* e^{i\lambda\tau_j}) \right) \\ &= -2\operatorname{Re} \left(\sum_{\lambda} \frac{1}{f} (R_j S^* e^{i\lambda\tau_j} - a_j |S|^2) \right) \end{aligned}$$

$$\begin{aligned} L''(a_j) &= \frac{\partial^2 L}{\partial a_j^2} = \frac{\partial L}{\partial a_j} \left\{ -2\operatorname{Re} \left(\sum_{\lambda} \frac{1}{f} (R_j S^* e^{i\lambda\tau_j} - a_j |S|^2) \right) \right\} \\ &= 2 \sum_{\lambda} \frac{1}{f} |S|^2. \end{aligned}$$

Thus

$$a_j^{p+1} = a_j^p - \frac{L'(a_j)}{\langle L''(a_j) \rangle} = a_j^p + \frac{\operatorname{Re} \left(\sum_{\lambda} \frac{1}{f} (R_j S^* e^{i\lambda\tau_j} - a_j |S|^2) \right)}{\sum_{\lambda} \frac{1}{f} |\hat{S}|^2}.$$

For estimating S ,

$$\begin{aligned} \frac{\partial L}{\partial S} &= -\operatorname{Re} \left(\sum_j \sum_{\lambda} \frac{1}{f} 2(R_j - a_j S e^{-i\lambda\tau_j}) (-a_j e^{i\lambda\tau_j}) \right) \\ &= \operatorname{Re} \left(\sum_j \sum_{\lambda} \frac{1}{f} 2(a_j R_j e^{i\lambda\tau_j} - a_j^2 S) \right) \end{aligned}$$

$$\frac{\partial L}{\partial S} = 0 \Leftrightarrow a_j R_j e^{i\lambda\tau_j} - a_j^2 S = 0 \Leftrightarrow S = \frac{R_j e^{i\lambda\tau_j}}{a_j}$$

$$\hat{S}(\lambda) = \frac{1}{N} \sum_j \frac{R_j e^{i\lambda\tau_j}}{a_j}.$$

For estimating f ,

$$\frac{\partial L}{\partial f} = -\operatorname{Re} \left(\sum_j \sum_{\lambda} \frac{1}{f} - \frac{1}{f^2} |R_j - S e^{-i\lambda\tau_j}|^2 \right)$$

$$\frac{\partial L}{\partial f} = 0 \Leftrightarrow \frac{1}{f} = \frac{1}{f^2} |R_j - a_j S e^{-i\lambda\tau_j}|^2 \Leftrightarrow f = |R_j - a_j S e^{-i\lambda\tau_j}|^2$$

$$\hat{f}(\lambda) = \frac{1}{N} \sum_j |R_j - a_j S e^{-i\lambda\tau_j}|^2$$