Assessment of mental workload: A new electrophysiological method based on intra-block averaging of ERP amplitudes

Sibylle K. Horat [a], François R. Herrmann [c], Grégoire Favre [b], Jorge Terzis [d,e], Damien Debatisse [d,e], Marco C.G. Merlo [a], Pascal Missonnier [a,b,*]

[a] Unit of Psychiatric Neuroscience and Psychotherapy, Department of Medicine, Faculty of Science, University of Fribourg, Chemin du Musée 5, 1700 Fribourg, Switzerland
[b] Mental Health Network Fribourg (RFSM), Sector of Psychiatry and Psychotherapy for Adults, L'Hôpital 140, Case postale 90, 1633 Marsens, Switzerland
[c] Department of Rehabilitation and Geriatrics, University Hospitals of Geneva, 2 chemin du Petit-Bel-Air, Chêne-Bourg, Geneva 1225, Switzerland
[d] HELIOS Privatkliniken GmbH – Wuppertal-Universität/Barmen, Heusnerstrasse 40, 42283 Wuppertal, Germany
[e] Department of Neurosurgery, Universität Kliniken der Stadt Köln gGmbH, Krankenhaus Merheim, 51109 Köln, Germany

The present study contributes to the current debate about electrophysiological measurements of mental workload. Specifically, the allocation of attentional resources during different complexity levels of tasks and its changes over time are of great interest. Therefore, we investigated mental workload using tasks varying in difficulty during an auditory oddball target paradigm. For data analysis, we applied a novel method to compute event-related potentials (ERPs) by intra-block epoch averaging of P2, P3a and P3b amplitude components for the infrequent target stimuli. We obtained eight consecutive blocks of 5 epochs each, which allowed us to develop an electrophysiological parameter to measure mental workload. In both the easy and the more constraining tasks, the amplitude of P2 decreased beginning with the second block of the sequence. In contrast, the amplitudes of P3a and P3b components linearly decreased following the repetition of the target in the more constraining task, but not in the easy task. Statistical analysis revealed intra-block differences on amplitudes of ERPs of interest between the easy and the more constraining tasks, confirming this method as a measure to assess mental workload. Since a subject is his own control, the present method represents an electrophysiological parameter for individual measurement of mental workload and may therefore be applicable in clinical routine.

1. Introduction

Human mental workload is referred to as the time-limited working memory processing load. The required processes are the capacity to store information in short-term registers and simultaneously manipulate it 'on-line'. Functionally, mental workload is inversely related to the attentional reserve under the control of efficient allocation of neuronal resources, whose depletion results in performance decrement (Wickens et al., 1983). The efficient allocation of attentional resources over time is crucial when individuals perform mentally-demanding tasks. Thus, it is of major interest to assess if the time-dependent modulation of workload-related functional indices is likely to highlight the manner by which neuronal resources are allocated during the performance of a task (Korsnes et al., 2008; Miller et al., 2011;

Vingerhoets and Luppens, 2001).

Neurophysiological assessment of cognitive workload is performed using event-related potentials (ERPs), where the typical protocols of cognitive workload combine an oddball procedure while subjects perform a task with incrementally-varying difficulty (e.g. Donchin et al., 1986; Sirevaag et al., 1993). The auditory oddball task consists of detecting an infrequent deviant stimulus (i.e. target) amongst a series of frequent stimuli (i.e. standards). This procedure is suited to explore information processing ranging from stimulus detection to cognitive operations, including attentional and memory processing. The oddball processing elicits the P3 ERP component that has been suggested to indicate brain activities underlying revision of the mental representation induced by incoming stimuli (Donchin, 1981; Sutton et al., 1965). According to the context-updating hypothesis, a memory process is engaged in the oddball task after initial sensory processing, thereby evaluating whether the current stimulus is either the same as the previous one, or different. If the incoming stimulus is a deviant stimulus (i.e. target), attentional processes engage a change of the stimulus representation (Polich, 2007). This updating of the model

* Corresponding author at: Unit of Psychiatric Neuroscience and Psychotherapy, Department of Medicine, Faculty of Science, University of Fribourg, Chemin du Musée 5, 1700 Fribourg, Switzerland.
  E-mail address: pascal.missonnier@unifr.ch (P. Missonnier).

of the environment is concomitant with the P3 component that is preceded by sensory-evoked potentials (N1, P2, N2) (Donchin and Coles, 1988; Polich, 2007). However, early accounts of P3 revealed that this component originated from two functionally and topographically distinct components, P3a and P3b (Polich, 2007). It is now well established for cognitive P3 activity (for review, Polich, 2007) that an infrequent event distracts the subject and reorients attention by capturing attentional resources (P3a). Reflecting the attention directed to the stimulus (Chong et al., 2008; Daffner et al., 1998), the infrequent event facilitates subsequent memory processing updating or categorizing of an event (tagged by P3b) by engagement of appropriate resources. In this theoretical context, it clearly appears that the components of P3 represent a more reliable physiological index to study cognitive workload than the P3 component alone.

Globally, studies on mental workload have reported that both P3 amplitude and latency are inversely related to task difficulty (Allison and Polich, 2008; Miller et al., 2011; Ullsperger et al., 2001). This inverse relationship can be interpreted as a reflection of the amount of variation of attentional resources engaged. Likewise, using adapted workload paradigms, several studies (Allison and Polich, 2008; Lean et al., 2011; Tong et al., 2009) revealed that the amplitude of other ERP components was also inversely related to task difficulty. Among them, the exogenous P2 component is frequently cited. From a cognitive point of view, the P2 component could reflect pre-attentive alert mechanisms contributing to improve the perception of the stimuli. In this context, a reduction of P2 could diminish the stimulus information available to the later processes indexed by P3, thereby leading to a further reduction in amplitude of this component (Miller et al., 2011). Together, this supports the idea that attentional-dependent ERP components such as P2, P3a and P3b (Campanella et al., 2002; Goodin et al., 1978; Hansenne, 2000a, 2000b; Johnson, 1986) have a reliable sensitivity to workload changes (Miller et al., 2011).

The present study extends earlier efforts aimed at evaluating how mental workload affects the ERPs elicited by infrequent auditory target stimuli. We used the classical oddball paradigm which is comprised of tasks of varying difficulty requiring upgrading of cognitive resources, and is therefore particularly suited to study varying levels of attention allocation. The ERP components P2, P3a and P3b, elicited by the infrequent target tones are affected by attentional allocation and are probably interrelated. We therefore considered that it is crucial to determine whether the amplitude of the P3 components, not just the P3, declined according to the variation of attentional resources related to repetition-dependent stimuli. Moreover, we questioned the influence of early cognitive processing, as indexed by P2, on the amplitude of latter P3 components suspected to be more sensible to the variation of attentional resources.

The aims of the present study were thus to investigate: (a) whether ERP amplitudes displayed an intra-blocks decline in response to the repetition of the stimuli; (b) from which block this decline occurred; (c) how the decline of amplitude of these three components can be modulated by workload difficulty; and (d) whether mental workload-related changes in amplitude in the earlier P2 ERP component may modulate later P3 components response. Moreover, in order to exclude an effect of brain fatigue or habituation on the task difficulty, a control task was included.

Methodologically, a block-epochs averaging of P2, P3a and P3b amplitudes for infrequent target stimuli that consists of eight consecutive blocks of 5 epochs each was performed. This block-epochs approach allows studying the relationships which can exist between P2, P3a and P3b ERP components in a demanding workload environment. Previous investigations using oddball paradigms (Borchard et al., 2015; Lasaponara et al., 2015; MacDonald et al., 2015) reported that P2 and P3 components may be

observed and measured on the vertex. The components of interest were also distinguished over the vertex in our study. Therefore, we decided to measure ERPs over central electrode locations only. Additional source reconstruction was performed to ensure the distinct neuronal origin of both P3 components.

When successfully achieved, this averaging technique may provide a valuable method for an easy individual measurement of mental workload.

2. Materials and methods

2.1. Participants

Sixteen healthy, right-handed volunteers (9 women, 7 men; mean age: 27.5 ($\pm$ 4.3 SD) years; age range: 22–35) participated in the study. All participants showed a normal cognitive functioning when tested with the extensive neuropsychological CogState Battery (see www.cogstate.com for details) including (Table 1): the Groton Maze Learning Test and Set-Shifting Task for executive functions; Detection Task for psychomotor function; Identification Task for visual attention; Groton Maze Learning Test delayed recall for visual learning and memory; International Shopping List Task and Delayed Recall for verbal memory; One Back Task for working memory; as well as the Social-Emotional Cognition Task for social cognition. Participants all had normal or corrected-to-normal visual acuity, and neither reported a history of sustained head injury, nor neurological or psychiatric disorders. Moreover, none exhibited alcohol or drug abuse or the presence of a severe physical impairment. Participants with regular use of psychotropic drugs, stimulants and β-blockers were not included. Data were recorded in the Mental Health Network Fribourg (RFSM), Switzerland. Informed consent was obtained from all subjects. The study was approved by the Ethical Committee of the University of Fribourg, and was conducted in line with the Helsinki Declaration.

2.2. Experimental design

The subjects were seated in a comfortable chair while listening to the stimuli presented through loud speakers. Stimuli consisted of pure sine-wave tones including standard, or 'frequent' (80%), low-pitch tones (1000 Hz), and deviant or 'infrequent' (20%), high-pitch tones (2000 Hz) binaurally presented to each ear. Standard and deviant tones were presented in a random order. Sound intensity was adjusted for each subject at the beginning of the experiment, in order to obtain the same subjective loudness at both ears. In practice, these values ranged between 80 and 90 dB SPL. Each tone lasted 100 ms with a 10 ms rise and fall time. Inter-stimulus intervals randomly varied between 800 and 1200 ms. Subjects were instructed to remain quiet and to only move their right index finger in accordance with the nature of the task in order to minimize muscle artifacts.

Three different tasks were tested in the following order: (1) in a simple oddball detection task (used as control condition) (Kramer et al., 1995; Trejo et al., 1995), the participants were instructed to respond as quickly as possible by pressing a button with their right hand as soon as they detected a target (infrequent tones).

Table 1
Neuropsychological performances assessed with CogState ($n=16$).

Characteristics		Mean (SD)
Executive function		
Groton Maze learning test	ER tot	36.63 (11.77)
Set-Shifting Task	ER tot	16.25 (9.18)
Psychomotor function		
Detection task speed ($\log_{10}$ (ms))		2.51 (0.1)
Visual attention		
Identification task speed ($\log_{10}$ (ms))		2.67 (0.07)
Visual memory		
Groton Maze learning test	DRE	4.06 (2.67)
Verbal memory		
International shopping list		
	CR tot	29.06 (3.43)
	DRC	10.94 (1.06)
Working memory		
OneBack task accuracy (Acc)		1.29 (0.12)
Social cognition		
Social-emotional cognition task (Acc)		1.17 (0.05)

Notes: Data are presented as mean (SD). ER tot, total number of errors; DRE, delayed recall, total number of errors; CR tot, total number of correct responses, DRC, delayed recall, total number of correct responses, Acc, accuracy as the arcsine transformation of the square root of the proportion of correct responses.

For non-target trials (frequent tones), no motor response was required; (2) in the counting forward task, participants were told to press the button when a target was heard, and to further count the number of target occurrences in the sequence; (3) in the counting backward task, instead of counting the number of target occurrences, participants subtracted from the date of the experiment a day to the precedent (i.e. counting in reverse order) for each target detected. At the end of the session, subjects were asked to report the date calculated after backward counting. An interval of 3 min separating each condition and the order of presentation of the tasks both enabled us to exclude an effect of brain fatigue or habituation to the task difficulty, as well as to optimize the allocation of attentional resources involved by the protocol.

Thus, the workload increased from control (workload-free condition) to an easy task (counting forward; weak workload) and a more constraining task (counting backward; strong workload). Each task was tested in a unique session lasting 3.3 min composed of 200 sequential stimuli. Subjects were informed about the nature of the forthcoming task right before each sequence. For all conditions, subjects were asked to keep their eyes closed and to avoid blinking and eye movements. The total duration of the experiment was approximately 30 min. Reaction times (RT) were systematically recorded. Electrophysiological and neuropsychological assessments were performed in the morning.

2.3. Electrophysiological recordings

In order to identify easily applicable electroencephalography (EEG) markers in routine clinical settings, continuous EEG (Advanced Neuro Technology Company-ANT, The Netherlands) was recorded using 32 surface electrodes placed over the scalp according to the 10–20 international electrode placement system (Homan et al., 1987). Linked mastoid right and left electrodes were used as a reference to respect equidistance between electrodes in order to achieve a central equipotency of the upper half sphere for later source reconstruction of ERPs. Skin impedance was kept below 5 kΩ. Physiological signals were sampled at 256 Hz, the lower cut-off was 0.03 Hz and the upper cut-off was 100 Hz (DC amplifiers, ANT). Right, left, supra- and infra-orbital electrodes monitored horizontal and vertical electro-oculograms (EOG). Simultaneously to stimulus onsets (standards or targets), TTL-pulses (Transistor–Transistor Logic) were sent to the EEG-recording system. These TTL-pulses were used off-line to segment the continuous EEG-data into epochs, synchronized with stimulus onset.

2.4. Data processing

Data were analyzed with the Advanced Source Analysis software (Advanced Source Analysis [ASA] 4.0 software, ANT, The Netherlands). After removal and off-line correction of ocular artifacts (threshold reduction algorithm), data from trials with correct answers were averaged according to task conditions (weak and strong mental workloads) and stimuli tones (frequent, target). EEG data were averaged over a window of 700 ms with a 200 ms pre-stimulus onset and band-pass filtered between 0.1 Hz and 30 Hz, 24 dB/octave for a low-pass filter. ERPs corresponding to correct answers in infrequent tones were analyzed for each condition.

The ERP components of interest were the exogenous N1 and P2 components and the P3a and P3b components. Component amplitude was assessed from the pre-stimulus baseline to the maximum peak within specified windows to each component. This automatic analysis was completed by a visual inspection by a trained neurophysiologist. The temporal limits for the measurement of these components were delimited in the time-range window between 70 and 120 ms for N1, 120 and 170 ms for P2, 220 and 260 ms for P3a and 270 and 330 ms for P3b, respectively. P2, P3a and P3b components were best distinguished at central (C3, Cz and C4 electrode combined) electrode locations in the individual average waveforms (implemented in ASA 4.0 software, ANT). Therefore, this site was selected and averaged for the amplitude measurements.

To explore workload-related ERPs, we performed an ERP waveform averaging in a sequence block of 5 stimuli restricted to P2, P3a and P3b. This procedure was realized for the first five epochs, then repeated for the five epochs following and so on, so that a total of 8 consecutive blocks were obtained in each task condition (control, weak workload, strong workload). Secondarily, for each component of interest we compared the amplitude of each block (i.e. block-2 to block-8) to the amplitude of the first block (i.e. block-1). Thus, the amplitude of components for each block was expressed as a percentage (i.e. magnitude) of variation, as compared to the first block. This procedure allowed the quantification of the attentional resources used during the test for each task and consequently represents a new electrophysiological parameter to measure mental workload.

2.5. Brain electric tomography

To ascertain that neuronal generators triggered two separate P3 components, topographical voltage maps of the scalp-recorded data were realized by plotting color-coded isopotentials derived by interpolating voltage values between scalp electrodes at specific latencies (Proverbio et al., 2008). An improved version of standardized Low Resolution Electromagnetic Tomography incorporating a singular value decomposition-based lead field weighting (swLORETA) was used to estimate intracranial sources of the scalp-recorded data (Nazari et al., 2010; Palmero-Soler et al., 2007; Pascual-Marqui et al., 1994). Intracranial source localizations were estimated and statistically processed on average P3a and P3b measured at the time of the peak maximum over central electrodes using the ASA4 software (ANT). The best fitting source model on the grand average potential distribution was computed at the 3 time points at which P3a and P3b were the largest in the easy and strong workload conditions, respectively. Source space properties were: grid spacing = 10 mm (using 5 mm did not change results; Tikhonov regularization: estimated SNR = 3 (see Nazari et al., 2010 for details).

2.6. Statistical analysis

To normalize the variance of the EEG data (i.e. magnitude of ERPs), a square power transformation was used. The normality of the data distribution was verified with the Shapiro-Francia test. Infrequent tones, task conditions (3 levels: control or without workload, weak workload, strong workload) and blocks (eight blocks) were included as independent variables in a repeated-measure regression model to analyze their respective influence on each of the dependent variables (i.e. behavioral responses and EEG measures). Blocks were considered as a unique categorical independent variable with varying reference values allowing a simultaneous comparison of all the blocks to a reference one, thus avoiding adjustment for multiple comparisons.

All statistical analyses were pre-planned in the protocol. Analyses were performed using the Stata software package, version 14.0. We used the Stata regress command with the "vce (cluster)" option which specifies that the standard errors allow for intragroup correlation. Observations are thereby independent. The statistical threshold for α was set at $p < 0.05$.

Corrected significance level was computed for each ERP component using the Benjamini–Hochberg procedure (Benjamini and Hochberg, 1995), implemented with a spreadsheet developed by Dr Manuel Weinkauf (http://www.marum.de/en/Leere_Seite_5.html#Section12094).

3. Results

3.1. Behavioral results

All participants had normal cognitive functions. During the mental workload testing, reaction times were not related to task difficulty (control vs. weak and strong, respectively: $p = 0.70$ and $p = 0.22$; weak vs. strong, $p = 0.42$). Similarly, no significant difference was observed for block-epoch measurements (Table 2).

3.2. Electrophysiological results

3.2.1. Event-related potentials (ERPs)

Amplitude of both grand averaged ERP waveforms and block-averaging ERP did not differ between the weak workload and control tasks. Fig. 1 (insets A2–B2) displays grand averaged ERP waveforms plotted over central (C3, Cz and C4 combined) regions for each task requiring workload and tone. As expected, all auditory stimuli elicited a distinct negative response at about 100 ms (N1 component) after stimulus onset. For infrequent stimuli (black line), analysis of grand averaged ERPs revealed a series of ERP

Table 2
Mean reaction times (in milliseconds) with standard errors (SE) for the control and the two workload conditions in participants ($n = 16$).

	Control	Weak workload	Strong workload
Mean	306.23 (16.44)	309.70 (12.64)	320.38 (14.76)
Block 1	304.65 (17.12)	294.62 (13.64)	304.95 (13.72)
Block 2	306.02 (19.61)	301.75 (15.36)	304.41 (10.69)
Block 3	308.45 (17.56)	310.91 (14.77)	320.08 (16.20)
Block 4	311.18 (19.04)	303.38 (12.42)	324.74 (16.49)
Block 5	310.05 (18.48)	326.01 (19.18)	326.46 (17.71)
Block 6	311.37 (16.93)	322.20 (10.92)	348.40 (25.75)
Block 7	309.34 (16.36)	313.69 (14.82)	313.07 (17.30)
Block 8	306.80 (17.26)	303.39 (12.33)	311.21 (17.10)

Note: There was no statistically significant difference in reaction times between conditions. See text for details.
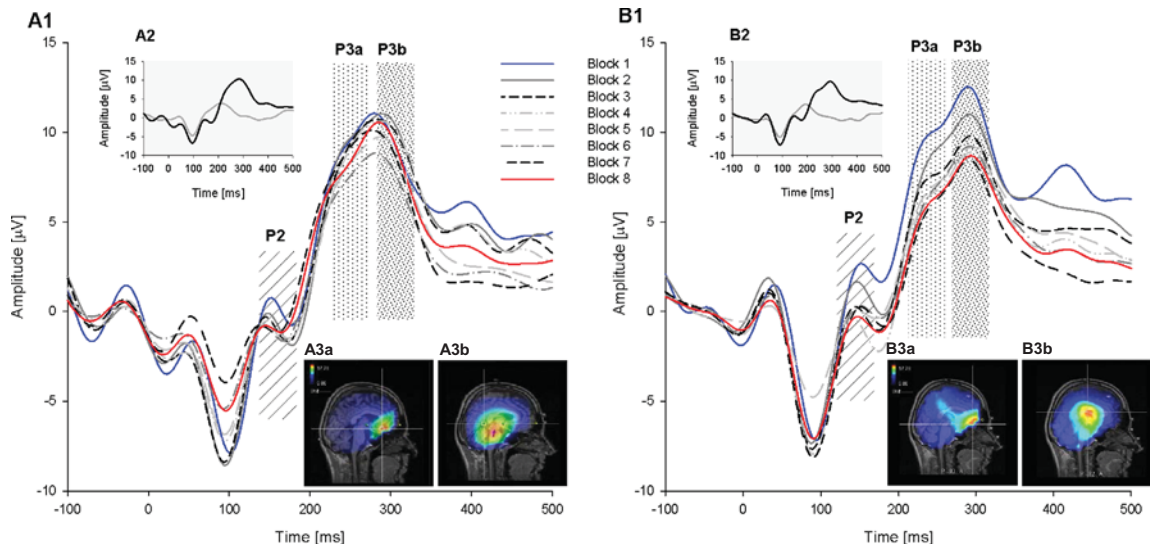
Fig. 1. Data show the averages of cumulative of ERPs in 8 blocks of 5 epochs at Central (C3, Cz and C4) electrode sites following infrequent target stimuli during the easy (A1) and more constraining (B1) task conditions. Insets: Grand average ERP waveforms at Central (C3, Cz and C4) electrode sites following frequent (gray line) and infrequent target (black line) stimuli during the easy (A2) and strong constraining (B2) task conditions. Note the longer latency of the P3 component for the strong constraining, i.e. more demanding task. Illustration of source localization of the P300 components on sagittal section (corresponding topography on the horizontal axis) using the swLORETA inverse solution performed for the easy (A3) and the strong constraining task condition (B3). Source localization revealed distinct activated regions for the P3a (A3a and B3a) and P3b (A3b and B3b) components, respectively.

peaks (P2 and N2) followed by the P3 component, which culminated over the central region and could be distinguished into two components (P3a and P3b). The amplitude of these ERPs did not differ between task conditions. Similarly, P2 latency was not modulated by the task (control vs. weak and strong: $p=0.303$ and $p=0.766$, respectively; weak vs. strong: $p=0.639$). In contrast, P3b latency was longer in the higher demanding task compared to both the control ($p < 0.001$) and the easier demanding task ($p < 0.01$). Yet, this effect was not observed for the P3a component (control vs. weak and strong: $p=0.070$ and $p=0.082$, respectively; weak vs. strong: $p=0.696$). No significant difference was observed between control and easy demanding tasks in both P3a ($p=0.242$) and P3b ($p=0.242$) components. The adjusted p threshold computed according to the Benjamini–Hochberg procedure (Benjamini and Hochberg, 1995) remained unchanged.

Examples of the five-epoch averages of target stimuli for constraining conditions are presented in Fig. 1 (A1–B1). Similar numbers of artifact-free epochs per block were obtained for both workload conditions. The average numbers for the easier task were 4.88 ± 0.34; 4.94 ± 0.25; 4.88 ± 0.50; 4.88 ± 0.50;

4.88 ± 0.50; 4.81 ± 0.75; 4.00 ± 1.37 and 2.50 ± 1.32, respectively from block-1 to block-8. For the more constraining task, average numbers were 4.63 ± 1.26; 4.69 ± 1.25; 4.69 ± 1.25; 4.69 ± 1.25; 4.69 ± 1.25; 4.69 ± 1.25; 4.00 ± 1.37 and 2.31 ± 1.25, respectively from block-1 to block-8. Finally, for the control task, average numbers were 5.00 ± 0.00; 5.00 ± 0.00; 4.96 ± 0.20; 4.85 ± 0.46; 4.92 ± 0.27; 4.96 ± 0.20; 4.31 ± 1.05 and 2.38 ± 1.39, respectively from block-1 to block-8.

There was no significant task effect (control vs. weak and strong: $p=0.857$ and $p=0.108$, respectively; weak vs. strong: $p=0.201$) on the number of accepted trials. No other main effect (i.e. blocks) or interaction between factors was observed.

3.2.2. Block repetition analysis

In order to apply our epoch evaluation method, averaging of ERP waveforms of five consecutive stimuli was performed with a total of eight consecutive blocks in each task condition. ERP values of interest for each block were compared to those of the first block (Fig. 2). For the easy condition, the magnitude of the P2 component was significantly decreased from the fourth block (block-1 vs.
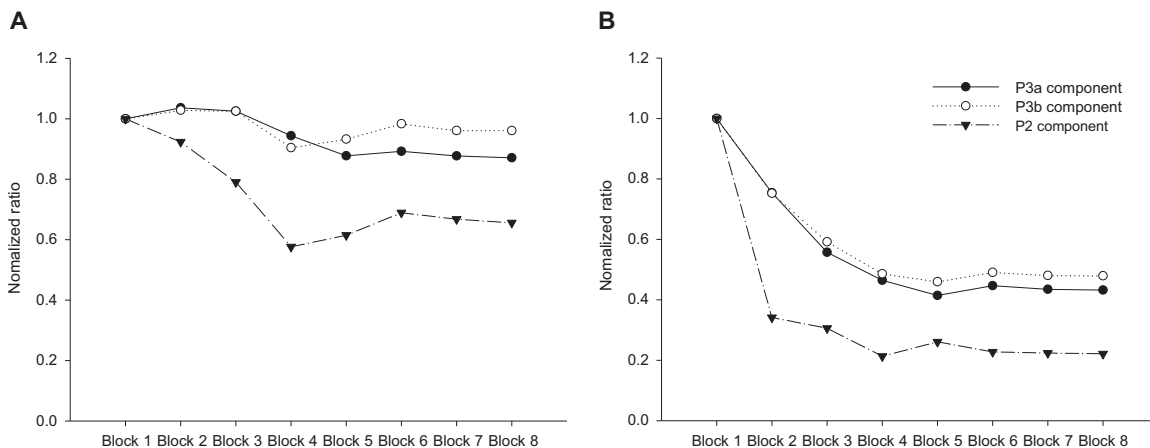

Fig. 2. Averaging of ERP waveforms in 8 blocks of 5 epochs for P2 (triangles), P3a (black circles) and P3b (white circles) components in infrequent target auditory stimuli. Normalized ratios from each block were compared to those of the first block in the two easy (A) and strong constraining (B) task conditions (intra-block comparison). Note the greater decrease of magnitude for the P2 ERP component with block-repetition in the two tasks.

blocks 4 to 8: $p < 0.03$). In contrast, no difference between blocks was observed on the normalized magnitude of both P3a and P3b components. For the more constraining condition, statistical analysis revealed block-related effects for the three ERP components with a significant lower magnitude starting from the second block (block 1 *vs.* blocks 2 to 8, P2 component: $p < 0.001$; P3b component: $p < 0.02$; block 1 vs. blocks 3 to 8, P3b components: $p < 0.01$). Moreover, an additional decrease was noted from the third block compared to the second block for both P3a and P3b components (block 2 *vs.* blocks 3 to 8, P3a components: $p < 0.02$; P3b components: $p < 0.02$). A higher magnitude was also observed for block 3 compared to block 5 ($p < 0.04$) for the P3b component.

Concerning the control task, a significant decreased was present from the fourth block (block-1 *vs.* blocks 4 to 6 and 8: $p < 0.05$) for the P2 component (blocks 1 to 8: 1.00 ± 0.00; 0.92 ± 0.12; 0.86 ± 0.20; 0.59 ± 0.12; 0.59 ± 0.17; 0.69 ± 0.16; 0.65 ± 0.15; 0.66 ± 0.14). No difference between blocks was observed on the normalized magnitude of both P3a (blocks 1 to 8: 1.00 ± 0.00; 1.02 ± 0.12; 0.85 ± 0.11; 0.84 ± 0.15; 0.84 ± 0.16; 0.84 ± 0.15; 0.84 ± 0.16; 0.86 ± 0.15) and P3b (blocks 1 to 8: 1.00 ± 0.00; 1.04 ± 0.12; 1.01 ± 0.13; 0.89 ± 0.11; 0.89 ± 0.11; 0.94 ± 0.16; 0.95 ± 0.16; 0.92 ± 0.14) components. The adjusted p threshold computed according to the Benjamini–Hochberg procedure (Benjamini and Hochberg, 1995) remained unchanged.

3.2.3. The "electrophysiological parameter to measure mental workload"

Fig. 3 shows normalized magnitude values for measurement of mental workload. Regarding the three ERP components, the normalized magnitude was significantly higher in the easier condition compared to the more difficult condition (P2 component: $p < 0.01$; P3a and P3b components: $p < 0.05$). Differences between blocks analysis yielded significant task effects on the magnitude of P2 ($p < 0.002$) and P3 components (P3a and P3b: $p < 0.04$), thus revealing a lower normalized magnitude for the highly demanding task from the second block for the P2 component, as well as from the fourth and third blocks for the P3a and P3b components, respectively. The adjusted p threshold computed according to the Benjamini–Hochberg procedure (Benjamini and Hochberg, 1995) remained unchanged.

As expected, no significant task effects were observed on the P2 component when comparing the control to the easier condition ($p = 0.916$), neither on both P3a ($p = 0.819$) and P3b ($p = 0.937$) components.

3.2.4. Electrical tomography analysis

Current density maps of the swLORETA for the P3 components are shown in Fig. 1(A3–B3). In both constraining conditions, the P3a elicited a positive activity in the anterior cingulate region (Fig. 1 A3a–B3a), whilst P3b was associated with a positive activity in centro-posterior medial regions (Fig. 1 A3b–B3b). These findings demonstrate a separate activation of the electrical sources of P3a and P3b that correspond to neuronal generators of auditory P3 components reported in literature (Wronka et al., 2012).

4. Discussion

The present study contributes to the current debate about electrophysiological measurements of mental workload. In respect to recent assessment protocols, mental workload was investigated using tasks varying in difficulty during an auditory oddball target paradigm. Applying this experimental approach and a novel method to analyze ERPs by block-epochs averaging, we were able to provide a new reliable physiological method for individual measurement of mental workload.
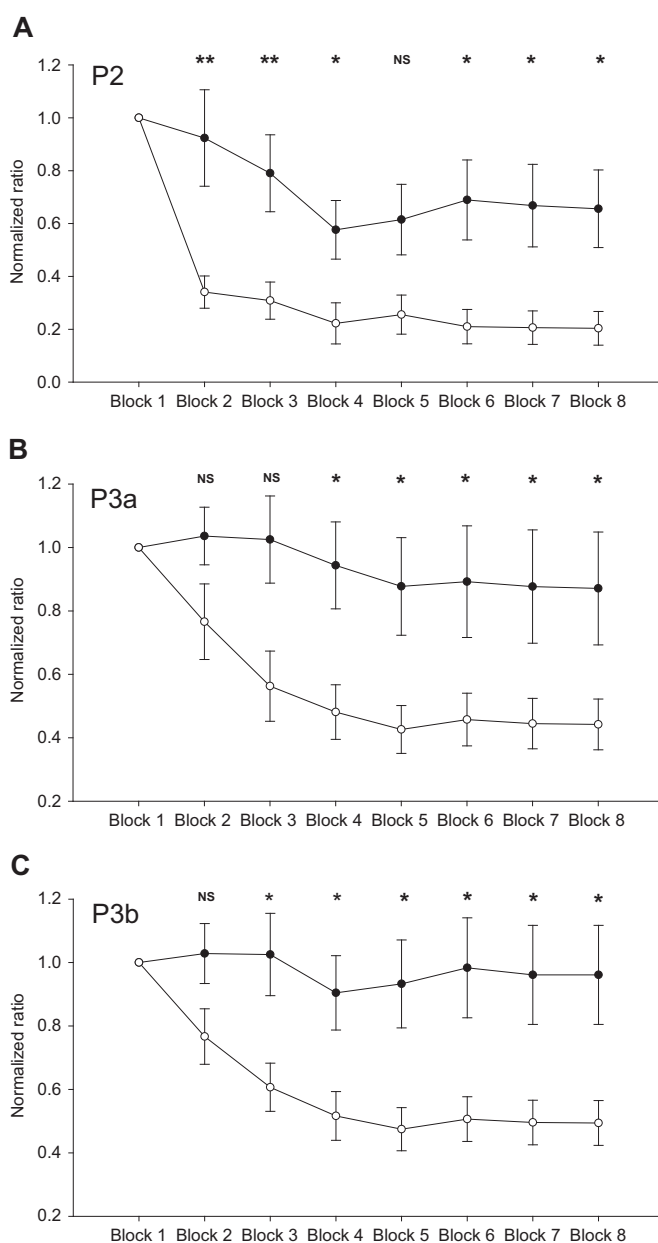

Fig. 3. For each individual block, comparison between easy (black line) and strong (gray line) constraining task conditions on normalized ratios and error bars (SD) for P2 (A), P3a (B) and P3b (C) components in infrequent target auditory stimuli (inter-block comparison). Note that task effects on normalized magnitude ratios according to block differed between the three components. * $p < 0.05$; ** $p < 0.01$; strong compared to easy task. The adjusted p threshold remained unchanged after the Benjamini–Hochberg procedure.

Sensory-evoked potentials were recorded in both unchanged (frequent) and changed (infrequent) context stimulus environments. As expected, infrequent target stimuli evidenced the P3 component in both task conditions (Fig. 1). Therefore, we opted to perform analyses on infrequent target stimuli. The modulation of grand average ERPs related to task difficulty for target stimuli usually reported in the literature (Allison and Polich, 2008; Brouwer et al., 2012; Kramer et al., 1985) was not observed in our findings. Several methodological differences may explain this discrepancy. First, we chose to analyze ERP components plotted on combined electrodes over the central region, although each evoked potential is usually measured over an electrode corresponding to the maximum peak amplitude for the considered component. Second, our observations concerned two separate

endogenous ERP components, notably the P3a and P3b components as supported by source reconstruction analysis, which is clearly not the case in studies restricted to the analysis of P3. The present results further contrast with previous studies reporting a smaller P2 amplitude as workload increased (Allison and Polich, 2008; Miller et al., 2011). However, in these studies, subjects performed a visuo-motor task under levels of incrementally-varying difficulty, while the second task of interest was presented in a verbal sensory modality.

Finally, the grand average of ERPs did not seem sensitive enough to reflect subtle effects related to task difficulty in the present design. Accordingly, the literature reports that single epoch-analysis or single block of stimuli subsequently divided into sub-blocks constitute a more effective method for workload assessment (Allison and Polich, 2008; Carrillo-de-la-Pena and Garcia-Larrea, 1999; Ivey and Schmidt, 1993; Lindin et al., 2004; Murphy and Segalowitz, 2004; Woestenburg et al., 1983). Restricting the analysis to the P3 component using averaging of stimulus repetition, the amplitude measurement may not be the best option to measure individual changes related to task difficulty. To address this issue, we thus performed an intra-block epoch averaging of P2, P3a and P3b amplitudes for infrequent target stimuli that consists of eight consecutive blocks of 5 epochs each. Since these components reflect the neuronal generators activity in a specific period of the task processing, this method allows us to elaborate a task-related measurement of mental workload. Our intra-block averaging method showed that the easy and the more constraining memory tasks presented differences in specific ERP patterns, thereby supporting the present method to assess mental workload. Most importantly, block comparisons to the first one allowed the use of the subject as his own control. Finally, in absence of any difference between the control task and the weaker constraining task in both ERPs and RTs, neither fatigue nor habituation may be accountable for task-related effects.

In line with early EEG studies that emphasized the task difficulty as a major variation factor of the P3 amplitude (Miller et al., 2011; Polich, 2007), in the present study the amplitudes of its components (i.e. P3a and P3b) were also inversely related to task difficulty (Fig. 1). Functionally, a smaller P3 amplitude reflects a reduction of the efficiency of task-related information processing. It has been suggested that a high workload increases demands on processing resources for operations linked to the task (Pinal et al., 2014). Furthermore, reduced processing resources available for these processes are reflected by the P3 (McEvoy et al., 1998; Morgan et al., 2008; Pinal et al., 2014), thus providing fewer attentional resources to target stimuli. Supporting the strong relationship between the amplitude of the P3 components and the level of engagement of attentional resources, there was no block repetition effect on the amplitude of both P3 components for the lower demanding task (i.e. weaker workload), while it was clearly the case in the more constraining one (Fig. 2). From a cognitive point of view, this strongly suggests a link between the speed of depletion of attentional resources and the difficulty of cognitive processes. It is well established in literature (for review, Kok, 2001; Polich, 2007) that P3a reflects mechanisms to evaluate incoming stimuli, whilst P3b indexes more difficult processes such as updating operations for subsequent working memory. In this context, the lower amplitude with incremental task difficulty observed earlier for the P3b compared to the P3a component (i.e. from the third vs. from the fourth block, respectively) may reflect a more rapid disengagement in the allocation of attentional resources to higher processing levels (Fig. 3). As a consequence, attentional resources may be reallocated to perform the supplementary operations engaged by the more constraining task and no differences in reaction times per block in either of the tasks may be observed. Accordingly, reaction times were very similar among blocks independently of the nature of the difficulty, as well as between the tasks (Table 2).

The current model of information processing assumes that attentional functions are engaged by sensory processes. Among the neuronal responses reflecting these processes a reliable effect of attention on P2 amplitude was previously reported (Missonnier et al., 2003). Given that mental workload is inversely related to attentional reserve, Miller et al. (2011) suggested that the P2 amplitude decrease may represent a reduction in the allocation of attention to the target stimuli. This would diminish the stimulus information available to the higher order processes indexed by the later P3 response. Our results confirm this hypothesis. First, the present P2 amplitude decreased with incremental task difficulty. Second and most importantly, a task difficulty effect on P2 amplitude was observed from the second block. Interestingly, this effect occurred earlier than for the amplitude of the P3 components. These differences between blocks on the amplitude of P2 strongly support the idea for a reduction of attention after sensory analysis of the stimuli for higher task difficulty. This would in turn lead the subject to allocate reduced attentional resources to the infrequent target stimuli. Therefore, the subsequent memory operations are weakened, this being reflected by the decrease of the amplitude of the P3 components. This interpretation is in line with the model of the P3 which stipulates that attentional resource allocation and P3 outcomes are modulated by the level of arousal and the amount of attention available in on-going tasks (Kahneman, 1973; Polich, 2007). Altogether, our data clearly indicate that in a sequence of repetitive stimuli, the first few ones are treated with much more attention because of their psychological importance. Yet, with stimulus repetition less attentional resources are engaged and can therefore be reallocated to supplementary processes required for mental effort enhancement.

In conclusion, using an adapted procedure of ERP measurement we report a novel measure of mental workload. This individual-related electrophysiological parameter may prove to be particularly valuable to evaluate the effect of psychiatric disorders on mental workload. Since mental workload is affected early in some neuropsychiatric disorders like prodromal Alzheimer's disease (Missonnier et al., 2007) and psychotic disorders (Missonnier et al., 2012), early detection of an alteration of our parameter could provide a powerful clinical tool to identify subclinical abnormalities in the early stages of illness and evaluate the performances of psychiatric patients with pre- and post-drug or neuropsychological treatments. As the subject is his own control, this method is particularly well suited to evaluation in the field of ergonomics or cognition in clinical settings.

Disclosure statement

This study has not been submitted elsewhere for publication, in whole or in part, and all the authors listed have approved the manuscript. The authors declare that there are no actual or potential conflicts of interest.

Contributors

P. Missonnier, D. Debatisse, J. Terzis and M.C.G. Merlo designed the study. P. Missonnier, G. Favre and S. Horat acquired the data. P. Missonnier, S. Horat and M.C.G. Merlo analyzed the data. P. Missonnier, F.R. Herrmann, S. Horat and M.C.G. Merlo wrote the article.

Acknowledgments

References

Allison, B.Z., Polich, J., 2008. Workload assessment of computer gaming using a single-stimulus event-related potential paradigm. Biol. Psychol. 77, 277–283.

Benjamini, Y., Hochberg, Y., 1995. Controlling the False Discovery Rate: A practical and powerful approach to multiple testing. J. R. Stat. Soc. B 57 (1), 289–300.

Borchard, J.P., Barry, R.J., De Blasio, F.M., 2015. Sequential processing in an auditory equiprobable Go/NoGo task with variable interstimulus interval. Int. J. Psychophysiol. 97, 145–152.

Brouwer, A.M., Hogervorst, M.A., van Erp, J.B., Heffelaar, T., Zimmerman, P.H., Oostenveld, R., 2012. Estimating workload using EEG spectral power and ERPs in the n-back task. J. Neural Eng. 9, 045008.

Campanella, S., Gaspard, C., Debatisse, D., Bruyer, R., Crommelinck, M., Guerit, J.M., 2002. Discrimination of emotional facial expressions in a visual oddball task: an ERP study. Biol. Psychol. 59, 171–186.

Carrillo-de-la-Pena, M.T., Garcia-Larrea, L., 1999. On the validity of interblock averaging of P300 in clinical settings. Int. J. Psychophysiol. 34, 103–112.

Chong, H., Riis, J.L., McGinnis, S.M., Williams, D.M., Holcomb, P.J., Daffner, K.R., 2008. To ignore or explore: top-down modulation of novelty processing. J. Cogn. Neurosci. 20, 120–134.

Daffner, K.R., Mesulam, M.M., Scinto, L.F., Cohen, L.G., Kennedy, B.P., West, W.C., Holcomb, P.J., 1998. Regulation of attention to novel stimuli by frontal lobes: an event-related potential study. Neuroreport 9, 787–791.

Donchin, E., 1981. Presidential address, 1980. Surprise!…Surprise? Psychophysiology 18, 493–513.

Donchin, E., Coles, M.G.H., 1988. Is the P300 component a manifestation of context updating. Behav. Brain Sci. 11, 357–374.

Donchin, E., Kramer, A.F., Wickens, C.D., 1986. Applications of brain event-related potentials to problems in engineering psychology. In: Coles, M.G.H., Donchin, E., Porges, S. (Eds.), Psychophysiology: Systems, Processes, and Applications. Guilford Press, New York, pp. 702–718.

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

Hansenne, M., 2000a. Le potentiel évoqué cognitif P300 (I): aspects théorique et psychobiologique. Neurophysiol. Clin. 30, 191–210.

Hansenne, M., 2000b. Le potentiel évoqué cognitif P300 (II): variabilité inter-individuelle et application clinique en psychopathologie. Neurophysiol. Clin. 30, 211–231.

Homan, R.W., Herman, J., Purdy, P., 1987. Cerebral location of international 10–20 system electrode placement. Electroencephalogr. Clin. Neurophysiol. 66, 376–382.

Ivey, R.G., Schmidt, H.B., 1993. P300 response: habituation. J. Am. Acad. Audiol. 4, 182–188.

Johnson Jr., R., 1986. A triarchic model of P300 amplitude. Psychophysiology 23, 367–384.

Kahneman, D., 1973. Attention and Effort. Prentice-Hall, Englewood Cliffs, N.J.

Kok, A., 2001. On the utility of P3 amplitude as a measure of processing capacity. Psychophysiology 38, 557–577.

Korsnes, M.S., Wright, A.A., Gabrieli, J.D., 2008. An fMRI analysis of object priming and workload in the precuneus complex. Neuropsychologia 46, 1454–1462.

Kramer, A.F., Trejo, L.J., Humphrey, D., 1995. Assessment of mental workload with task-irrelevant auditory probes. Biol. Psychol. 40, 83–100.

Kramer, A.F., Wickens, C.D., Donchin, E., 1985. Processing of stimulus properties: evidence for dual-task integrality. J. Exp. Psychol. Hum. Percept. Perform. 11, 393–408.

Lasaponara, S., Dragone, A., Lecce, F., Di Russo, F., Doricchi, F., 2015. The "serendipitous brain": low expectancy and timing uncertainty of conscious events improve awareness of unconscious ones (evidence from the Attentional Blink). Cortex 71, 15–33.

Lean, Y., Shan, F., Xuemei, Q., S., X., 2011. Effects of mental workload on long-latency auditory-evoked-potential, salivary cortisol, and immunoglobulin A. Neurosci. Lett. 491, 31–34.

Lindin, M., Zurron, M., Diaz, F., 2004. Changes in P300 amplitude during an active standard auditory oddball task. Biol. Psychol. 66, 153–167.

MacDonald, B., Barry, R.J., Bonfield, R.C., 2015. Trials and intensity effects in single-trial ERP components and autonomic responses in a dishabituation paradigm with very long ISIs. Int. J. Psychophysiol. 98, 394–412.

McEvoy, L.K., Smith, M.E., Gevins, A., 1998. Dynamic cortical networks of verbal and spatial working memory: effects of memory load and task practice. Cereb. Cortex 8, 563–574.

Miller, M.W., Rietschel, J.C., McDonald, C.G., Hatfield, B.D., 2011. A novel approach to the physiological measurement of mental workload. Int. J. Psychophysiol. 80, 75–78.

Missonnier, P., Deiber, M.P., Gold, G., Herrmann, F.R., Millet, P., Michon, A., Fazio-Costa, L., Ibanez, V., Giannakopoulos, P., 2007. Working memory load-related electroencephalographic parameters can differentiate progressive from stable mild cognitive impairment. Neuroscience 150, 346–356.

Missonnier, P., Herrmann, F.R., Zanello, A., Badan, Ba, M., Curtis, L., Canovas, D., Chantraine, F., Richiardi, J., Giannakopoulos, P., Merlo, M.C., 2012. Event-related potentials and changes of brain rhythm oscillations during working memory activation in patients with first-episode psychosis. J. Psychiatry Neurosci. 37, 95–105.

Missonnier, P., Leonards, U., Gold, G., Palix, J., Ibanez, V., Giannakopoulos, P., 2003. A new electrophysiological index for working memory load in humans. Neuroreport 14, 1451–1455.

Morgan, H.M., Klein, C., Boehm, S.G., Shapiro, K.L., Linden, D.E., 2008. Working memory load for faces modulates P300, N170, and N250r. J. Cogn. Neurosci. 20, 989–1002.

Murphy, T.I., Segalowitz, S.J., 2004. Eliminating the P300 rebound in short oddball paradigms. Int. J. Psychophysiol. 53, 233–238.

Nazari, M.A., Berquin, P., Missonnier, P., Aarabi, A., Debatisse, D., De Broca, A., Wallois, F., 2010. Visual sensory processing deficit in the occipital region in children with attention-deficit/hyperactivity disorder as revealed by event-related potentials during cued continuous performance test. Neurophysiol. Clin. 40, 137–149.

Palmero-Soler, E., Dolan, K., Hadamschek, V., Tass, P.A., 2007. swLORETA: a novel approach to robust source localization and synchronization tomography. Phys. Med. Biol. 52, 1783–1800.

Pascual-Marqui, R.D., Michel, C.M., Lehmann, D., 1994. Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. Int. J. Psychophysiol. 18, 49–65.

Pinal, D., Zurron, M., Diaz, F., 2014. Effects of load and maintenance duration on the time course of information encoding and retrieval in working memory: from perceptual analysis to post-categorization processes. Front. Hum. Neurosci. 8, 165.

Polich, J., 2007. Updating P300: an integrative theory of P3a and P3b. Clin. Neurophysiol. 118, 2128–2148.

Proverbio, A.M., Zani, A., Adorni, R., 2008. Neural markers of a greater female responsiveness to social stimuli. BMC Neurosci. 9, 56.

Sirevaag, E.J., Kramer, A.F., Wickens, C.D., Reisweber, M., Strayer, D.L., Grenell, J.F., 1993. Assessment of pilot performance and mental workload in rotary wing aircraft. Ergonomics 36, 1121–1140.

Sutton, S., Braren, M., Zubin, J., John, E.R., 1965. Evoked-potential correlates of stimulus uncertainty. Science 150, 1187–1188.

Tong, Y., Melara, R.D., Rao, A., 2009. P2 enhancement from auditory discrimination training is associated with improved reaction times. Brain Res. 1297, 80–88.

Trejo, L.J., Kramer, A.F., Arnold, J.A., 1995. Event-related potentials as indices of display-monitoring performance. Biol. Psychol. 40, 33–71.

Ullsperger, P., Freude, G., Erdmann, U., 2001. Auditory probe sensitivity to mental workload changes-an event-related potential study. Int. J. Psychophysiol. 40, 201–209.

Vingerhoets, G., Luppens, E., 2001. Cerebral blood flow velocity changes during dichotic listening with directed or divided attention: a transcranial Doppler ultrasonography study. Neuropsychologia 39, 1105–1111.

Wickens, C., Kramer, A., Vanasse, L., Donchin, E., 1983. Performance of concurrent tasks: a psychophysiological analysis of the reciprocity of information-processing resources. Science 221, 1080–1082.

Woestenburg, J.C., Verbaten, M.N., Van Hees, H.H., Slangen, J.L., 1983. Single trial ERP estimation in the frequency domain using orthogonal polynomial trend analysis (OPTA): estimation of individual habituation. Biol. Psychol. 17, 173–191.

Wronka, E., Kaiser, J., Coenen, A.M., 2012. Neural generators of the auditory evoked potential components P3a and P3b. Acta Neurobiol. Exp. (Wars.) 72, 51–64.