



Effects of mental workload and fatigue on the P300, alpha and theta band power during operation of an ERP (P300) brain–computer interface

Ivo Käthner^{a,*}, Selina C. Wriessnegger^b, Gernot R. Müller-Putz^b, Andrea Kübler^a, Sebastian Halder^a

^a Institute of Psychology, University of Würzburg, Würzburg, Germany

^b Laboratory of Brain-Computer Interfaces, Institute for Knowledge Discovery, Graz University of Technology, Austria

ARTICLE INFO

Article history:

Received 10 February 2014

Accepted 19 July 2014

Available online 1 August 2014

Keywords:

Brain–computer interface

EEG

P300

Alpha

Mental workload

Fatigue

ABSTRACT

The study aimed at revealing electrophysiological indicators of mental workload and fatigue during prolonged usage of a P300 brain–computer interface (BCI). Mental workload was experimentally manipulated with dichotic listening tasks. Medium and high workload conditions alternated. Behavioral measures confirmed that the manipulation of mental workload was successful. Reduced P300 amplitude was found for the high workload condition. Along with lower performance and an increase in the subjective level of fatigue, an increase of power in the alpha band was found for the last as compared to the first run of both conditions. The study confirms that a combination of signals derived from the time and frequency domain of the electroencephalogram is promising for the online detection of workload and fatigue. It also demonstrates that satisfactory accuracies can be achieved by healthy participants with the P300 speller, despite constant distraction and when pursuing the task for a long time.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

A brain–computer interface (BCI) can serve as a muscle-independent communication channel (Birbaumer, 2006; Kübler, Kotchoubey, Kaiser, Wolpaw, & Birbaumer 2001; Wolpaw, Birbaumer, McFarland, Pfurtscheller, & Vaughan, 2002). Event-related potentials (ERPs) extracted from the electroencephalogram (EEG) are often used as a control signal for the BCI. Of these ERPs the P300 is often the most prominent, therefore the term P300 BCI was coined. Until today, no study explored the effects of experimentally manipulated mental workload and fatigue during usage of a P300 BCI, although the potential benefits of automatic mental workload/fatigue detection with the BCI would improve usability of BCIs. A system is desirable that adapts to the mental state of the user, e.g. pauses if mental overload or fatigue is detected or adapts the stimulation parameters according to the detected level of fatigue and/or workload. A mental workload detector could be implemented that is based on the signals that are acquired anyhow with the BCI. As a first step toward this goal, this study investigated

the effects of time and mental workload on performance and electrophysiological signals during prolonged usage. During spelling with the BCI, mental workload was systematically manipulated with dichotic listening tasks. Two levels of difficulty (medium and high workload) were tested.

As a primary task, participants used the well-established visual P300 speller based on the paradigm by Farwell and Donchin (1988; see Kleih et al., 2011 for a review). To select a particular symbol, participants had to focus attention on a defined character of a 6 × 6 matrix that was presented on a computer screen. Rows and columns of the matrix flashed randomly during the task. Hence, the attended cells can be classified, because rows and columns with the target character constitute an oddball that elicits a P300. The P300 ERP is a positive deflection in the EEG following a rare, task-relevant stimulus that is presented among frequent, non-target stimuli and is reliably elicited in an oddball paradigm (Polich, 2007).

As a consequence of the prolonged BCI usage, we expected participants to experience mental fatigue. In general, the term describes the feeling that may result from prolonged periods of cognitive activity. It is associated with tiredness or exhaustion and a decrease in task performance and commitment (Boksem, Meijman, & Lorist, 2005; Hockey, 1997).

The hope that specific mental activity and states of wakefulness would be reflected in the EEG was already expressed by Berger in 1929. Today, specific frequency bands of the EEG are widely used to

* Corresponding author at: Institute of Psychology, University of Würzburg, Marcusstr. 9–11, 97070 Würzburg, Germany. Tel.: +49 9313189620; fax: +49 9313182424.

E-mail address: ivo.kaethner@uni-wuerzburg.de (I. Käthner).

identify levels of alertness or arousal and both, the information contained in specific frequency bands of the EEG and in ERPs have been shown to discriminate between different levels of mental workload (Borbély, Baumann, Brandeis, Strauch, & Lehmann, 1981; Brouwer et al., 2012; Coull, 1998; Dement & Kleitman, 1957; Gevins, Yeager, Zeitlin, Ancoli, & Dedon, 1977; Van Erp, Veltman, & Grootjen, 2010).

In the following we will first focus on ERP findings concerning mental workload and fatigue and secondly on differences in the ongoing EEG that have been found in response to changes in the level of mental workload and fatigue. Since it was demonstrated that the P300 decreases or disappears if attention was diverted from the P300 eliciting stimuli (Wickens, Isreal, & Donchin, 1977), it has received much attention as a potential indicator of mental workload. Based on the observation that P300 amplitude was reduced in dual-task studies, it was suggested that the P300 amplitude represents the distribution of processing capacity between concurrent tasks (Isreal, Chesney, Wickens, & Donchin, 1980; see Kramer, 1991; Kok, 2001 for reviews). More recent studies demonstrated reduced P300 amplitude under high workload conditions in ecologically valid tasks. Allison and Polich (2007) and Miller, Rietschel, McDonald, and Hatfield (2011) used varying levels of computer game difficulty to manipulate workload and in the study by Ullsperger, Freude, and Erdmann (2001) subjects performed gauge monitoring and/or arithmetic tasks. Unlike P300 amplitude, which was found to be related to the depth of processing, latency of the P300 is often increased if categorization of the eliciting stimulus becomes more difficult, thus, representing timing of mental processing (Kutas, McCarthy, & Donchin, 1977; see review by Kok, 2001). The study by Trejo et al. (2005) is one of few that investigated the effects of fatigue on event-related potentials. No significant effect of fatigue on N100, P200 or P300 was found.

For the estimation of mental workload and fatigue most studies use information from the frequency domain of the EEG. While, to our knowledge, only Brouwer et al. (2012) applied measures of event-related potentials and EEG spectral power to estimate workload (with an N-back task), most studies use exclusively information from the ongoing EEG. For instance, Kohlmorgen et al. (2007) were able to detect high mental workload under real life driving conditions. The applied classifier computed the power of selected frequency bands to identify high workload. Several studies identified differences in the alpha band (~8–12 Hz) and theta band (~4–7 Hz) to be most sensitive for distinguishing between different levels of workload. Increases in task difficulty and mental workload were most often associated with a decrease of alpha and an increase of theta power (Klimesch, 1999; Scerbo, Freeman, & Milkulka, 2003; Smith, Gevins, Brown, Karnik, & Du, 2001). For instance, this pattern was found in two studies that used continuous interactive control tasks to manipulate workload (Brookings, Wilson, & Swain, 1996; Fournier, Wilson, & Swain, 1999) and a study that used an N-back task to vary working memory load (Gevins et al., 1998). With increasing memory load, workload or attention, changes in theta activity are especially apparent at frontal sites (Gevins et al., 1998; Gundel & Wilson, 1992; Holm, Lukander, Korpela, Sallinen, & Müller, 2009; Mecklinger, Kramer, & Strayer, 1992).

Changes in alpha and/or theta power seem to be not only the most reliable indicators of mental workload but also that of fatigue. During a reduced level of arousal, an increase in activity in the alpha and theta band can often be observed (Klimesch, 1999). Increases in power of the theta and alpha band during sleepiness (with eyes open) were repeatedly found in sleep deprived persons and/or during vigilance tasks (Boksem et al., 2005; Lal & Craig, 2001; Makeig & Jung, 1996; Paus et al., 1997; Stampi, Stone, & Michimori, 1995; Yamamoto & Matsuoka, 1990).

It was previously pointed out that in some studies the relationship between workload and alpha band was only found for participants that were classified as high alpha generators. For this

reason, there may exist strong individual differences in the sensitivity of alpha frequencies for changes in workload (Kramer, 1991; Pigeau, Hoffmann, Purcell, & Moffitt 1987). There are not only individual differences with regard to the sensitivity of indicators of mental workload in the EEG, but there appear to be also strong differences depending on the studied task and environmental factors (Humphrey & Kramer, 1994; Van Erp et al., 2010). Further, task dependency is also often observed in studies investigating the effects of fatigue and while a variety of approaches were taken to implement EEG based drowsiness detection algorithms, generalizability across tasks is rarely addressed (Johnson et al., 2011). Due to the task dependence of the effects of fatigue and workload on the EEG, it seems to date inevitable to study the task of interest.

So far, BCIs have mainly been tested in laboratory settings. Ultimately, they are intended to be used as an assistive technology in a home-environment and thus, efficiency, and usability in general are of utmost importance (Kleih et al., 2011; Zickler et al., 2011). In order to function in real-life settings, the P300 speller must also be robust during conditions of distraction. Robustness can either mean good performance despite a noisy environment or good online detection of mental overload and automatic standby mode of the BCI. One study (Ortner et al., 2011) showed that 50 participants were able to spell with an average accuracy of 86% correctly selected letters with a visual P300 BCI during ad-hoc testing at the CEBIT 2011 in Hannover, where other visitors were walking around and talking. As in most P300 BCI studies, subjects spelled only few letters (five) and a study that systematically manipulates the level of distraction, using this type of BCI, was missing.

Consequently, the current study aimed at identifying markers of mental workload and fatigue during prolonged usage of a visual P300 BCI. The effects on P300 and the EEG were investigated. Mental workload was experimentally manipulated with dichotic listening tasks. Thus, the effects of different levels of distraction on performance with a P300 BCI could also be revealed. The following hypotheses were derived from the studies mentioned above: In the high workload condition, we expected a significantly lower performance in the BCI task along with a subjectively higher level of workload compared to the medium workload condition. As for the electrophysiological signals, we hypothesized a decrease in P300 amplitude and alpha activity and an increase in theta band power with increasing workload. Over the course of the experiment, we expected performance to decrease and the subjective level of fatigue to increase. We hypothesized that higher mental fatigue would be reflected in an increase in activity in the alpha and theta bands and a reduced P300 amplitude in the last compared to the first run of each workload condition.

2. Methods

2.1. Participants

Twenty healthy subjects (12 female, $M = 24.9$ years, $SD \pm 5.1$) participated in the study. All participants were paid 8€ per hour. None of the subjects had previously participated in a P300 BCI study nor reported a history of psychiatric or neurological disorders. Further, no hearing impairments were reported. Before the start of the experiment, each participant gave informed consent to the study that was carried out in accordance with the Declaration of Helsinki and the guidelines of the Ethical Review Board of the Institute of Psychology, University of Würzburg.

2.2. Data acquisition

The electroencephalogram (EEG) was recorded with 31 active Ag/AgCl electrodes mounted in an elastic fabric cap (g.GAMMAcap, g.tec Austria) following the modified version of the 10–20 system of the American Electroencephalographic Society (Sharbrough et al., 1991). These were located at positions F3, Fz, F4, FC5, FC3, FCz, FC4, FC6, C5, C3, Cz, C4, C6, CP5, CP3, CPz, CP4, CP6, P3, P1, Pz, P2, P4, PO7, PO3, POz, PO4, PO8, O1, Oz and O2. Channels were referenced to the right earlobe and a ground electrode was positioned at AFz. The EEG was sampled at 256 Hz with two g.USBamp amplifiers (g.tec, Austria). A high pass filter of 0.1 Hz and a low pass of 30 Hz were applied and data was notch filtered at 50 Hz. Data collection and

stimulus presentation were controlled by the BCI2000 software package (Schalk, McFarland, Hinterberger, Birbaumer, & Wolpaw, 2004). Data processing, storage and online display were conducted on a laptop (Intel Core i5 2.5 GHz, 4 GB RAM, Microsoft Windows 7 Professional 64bit).

2.3. Procedure

During the experiment, participants simultaneously performed two tasks, a primary, P300 brain–computer interface (BCI) spelling task and a secondary, dichotic listening task (see description of the tasks below and Fig. 1a). The experiment was divided into ten blocks. In each block, participants had to spell 25 characters with the P300 BCI. In half of the blocks, difficulty of the additional dichotic listening task was intended to induce medium and in the other half, high workload. Blocks of medium and high workload alternated according to a particular scheme (see caption of Fig. 1B). Participants were seated about 1 m from a 22-in. monitor that displayed the P300 speller matrix. Stories for the dichotic listening tasks were presented using circumaural headphones to attenuate ambient noise (Sennheiser HD 280 Pro). Prior to recording periods, participants were asked to minimize eye movements and muscle contractions during the experiment.

2.4. P300 speller

To obtain data for classifier training, participants had to copy the words BRAIN and POWER during two screening runs, before the start of the experiment. The discriminant function was subsequently used for online spelling during the consecutive runs (see online signal classification). In each of the ten online spelling blocks, participants had to select 25 predefined letters, character by character, from a 6×6 matrix using the copy spelling mode of the BCI2000 software. All 25 characters were displayed in a line above the matrix and the current letter to be spelled was shown in parentheses next to it. Participants were instructed to focus on the target letter in the matrix and to notice/silently count whenever it was highlighted. Rows and columns of the matrix flashed randomly during spelling.

After a discriminant function is derived from the EEG data, the attended cell can be identified at the crossing of rows and columns containing the target character.

In the current study, flash duration was 62.5 ms and the inter-stimulus interval (ISI) was set to 125 ms. Each row and column was highlighted 5 times, before signals were classified and the selected letter was fed back to the participants. Therefore, participants had to focus on 10 target stimuli and ignore 50 non-target stimuli for the selection of one letter. In between trials, there was a break of 3 s to focus on the next letter. Hence, the time needed for one character selection was 14.25 s.

2.5. Dichotic listening tasks/mental workload manipulation

During the first 5 min of all online spelling runs (20 characters to be copied), two concurrent stories were presented over headphones. One story was presented over the left and another over the right headphone speaker. Both stories were from the Arabian Nights and one was told by a female, the other by a male professional narrator. The presentation side (left or right speaker male/female voice) was changed from block to block.

In half of the blocks, participants were instructed to ignore the stories (medium workload condition). In the other five, participants were instructed to pay attention to the context of both stories (high workload condition). The flashing of the rows and columns of the speller matrix started 10 s after the onset of the stories. After 5 min, the stories stopped. Participants were asked to spell the remaining five letters without additional mental load (Spelling Only condition).

After the first and last run of the medium and high workload conditions, the NASA-TLX questionnaire was administered as a measure of subjective workload. After every run, participants answered 6 questions about the content of the stories to ensure successful manipulation of workload.

2.6. Questionnaires

Subjective workload was assessed with an electronic version of the NASA Task Load Index (NASA-TLX; NASA Human Performance Research Group, 1987). It consists of six factors (mental, physical and temporal demands, effort, frustration and own performance). A total score (range 0–100) can be obtained, where a higher score indicates higher overall workload. First, the individual factors have to be rated by the participants on a 20 point scale with anchors such as high/low. Subsequently, a pair-wise comparison of all factors is performed to determine the degree each of them contributes to the overall workload. The NASA-TLX is a well validated instrument (Hart, 2006), has been introduced for evaluation of BCI-controlled applications (Zickler et al., 2011) and has been used to measure subjective workload in several BCI studies (Holz, Höhne, Staiger-Sälzer, Tangermann, & Kübler, 2013; Käthner et al., 2013; Riccio et al., 2011; Zickler et al., 2011; Zickler, Halder, Kleih, Herbert, & Kübler, 2013).

As manipulation check, we asked participants to fill in a multiple response questionnaire about the content of the stories after every block. For every block, it contained three questions about the content of each story. For each question, there were three possible answers, thus the chance level was .33. To verify that the difficulty level of all response options was similar, participants were asked to mark

the questions that they had guessed. We checked if the number of correctly guessed questions differed from chance.

Before and after the experiment, the level of fatigue was assessed with a visual analog scale (VAS). This was a 10 cm horizontal line with the anchor points “0 = not at all tired” and “10 = extremely tired”. Participants should mark the position on the line that best represented their level of fatigue. Visual analog scales were frequently used by sleep researchers to assess fatigue and the validity of the instrument was demonstrated repeatedly (Dinges et al., 1997; Gift, 1989; Smets, Garssen, Bonke, & Dehaes, 1995).

2.7. Online signal classification

For data classification, stepwise linear discriminant analysis (SWLDA) was used. It is an established method that was shown to surpass other classification algorithms with regard to performance and implementation characteristics in a P300 speller application (Krusienski et al., 2006). The applied algorithm determines a function that discriminates between epochs containing the target row/column and non-target rows/columns. It determines features from amplitude values from each of the 31 electrodes. To do so, the data was decimated and filtered with a moving average of 20 Hz. An epoch of 800 ms was used (starting at stimulus onset) that yielded 16 voltage values for each stimulus (flashing row or column). From this set, the feature that best predicted the target label using least square regression was added to the discriminant function (p -value ≤ 0.1). After adding another feature to the function, a backward stepwise discriminant analysis was performed to remove the features that were no longer significant ($p > .15$). This process was repeated until no additional features met the inclusion/exclusion criteria or 60 features were reached (Krusienski et al., 2006; Sellers, Vaughan, & Wolpaw, 2010). The parameters were determined with Matlab (The MathWorks, version R2012b) using the P300-GUI (part of the BCI2000 software). After the parameters had been determined for each participant, they were used for online classification during BCI operation.

2.8. Offline data analysis

The EEG was analyzed with Matlab (version R2012b) in combination with the EEGlab and ERPLab toolboxes (version 10.2.2.4b; Delorme & Makeig, 2004). No artifacts were removed or corrected with regard to online classification systems where this would also not be possible.

Auditory stimulation stopped during spelling of the 21st letter, therefore, data was analyzed for the first 20 letters of the medium and high workload conditions (first 295 s), and separately for the last 4 letters without the dichotic listening tasks. Spelling of these 4 letters was labeled as Spelling Only condition. This condition served as a control condition to reveal the effects of time on spelling accuracy, P300, alpha and theta power during spelling without auditory distraction.

For the ERP analysis, data was segmented into epochs of 800 ms starting at the onset of a stimulus and baseline corrected with a pre-stimulus interval of 200 ms. Averages were calculated for targets and non-targets. This yielded 200 target and 1000 non-target trials per block per condition (medium and high workload) per participant. After visual inspection of the waveforms for each participant, a post-stimulus interval between 300 and 550 ms was chosen for peak detection. The maximum positive peak in this interval was defined as P300.

Fast Fourier transform (FFT) was used to generate EEG spectra. The average power in the theta (4–7 Hz) and alpha (8–12 Hz) bands was log-transformed for normalization (Gasser, Bächer, & Möcks, 1982). Statistical analysis was performed with SPSS 19.

To assess the effect of workload, data of the five medium workload and the five high workload runs was averaged. Afterwards, repeated measures analyses of variance (ANOVAs) with the two levels of workload (medium/high) were calculated separately for the dependent variables spelling accuracy, correct answers to control questions, P300 amplitude and latency, alpha and theta power.

To assess differences between the first and last run of both workload conditions, 2×2 repeated measures analyses of variance (ANOVAs) with workload (high/medium) and time (first/last run) as within subject factors were calculated for the dependent variables spelling accuracy, NASA-TLX score, P300 amplitude and latency, alpha and theta power.

To assess the sole effect of time (over all ten runs) independent of workload, we calculated repeated measures ANOVAs for the Spelling Only condition for spelling accuracy, P300 amplitude and latency, and alpha and theta power. We checked for trends in the data when main effects were found. Post hoc t -tests (Bonferroni corrected) were used for pairwise comparisons.

All of the analyses described above were restricted to electrode positions Pz for P300 amplitude and latency and alpha power and to Fz for theta power. P300 amplitude was analyzed for the electrode at which the amplitude had the highest predictive value for spelling accuracy. We found significant positive correlations (Pearson) between P300 amplitude and spelling accuracy at parietal electrodes for the medium workload condition (P1 and Pz) and at parietal-occipital electrodes for the high workload condition (P1, P3, Pz, PO3 and POz). Predictive value of the P300 amplitude was higher at Pz than at P1 for both conditions, therefore, analysis of the EEG data was restricted to electrode Pz. P300 amplitude at Pz explained 23.9% of the variance in spelling accuracy in the medium workload

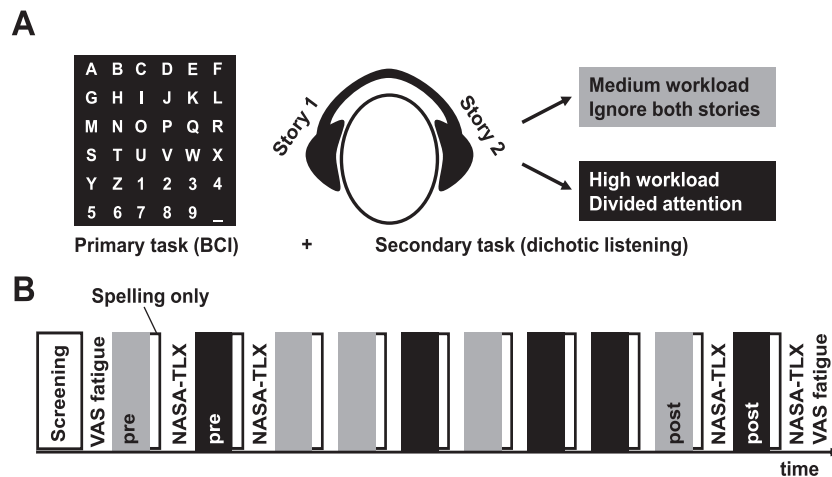


Fig. 1. Experimental tasks and design. Part A illustrates the tasks that had to be performed during the different conditions. During the medium and high workload conditions, participants had to spell predefined letters with the P300 BCI and two concurrent stories were presented over headphones. In the medium workload condition, they were instructed to ignore the stories and in the high workload condition they were supposed to pay attention to both stories. Part B depicts the block design of the experiment. After the screening, runs of medium and high workload alternated. The experiment always started with a medium workload condition followed by a high workload condition and ended that way. The six runs in between were pseudo randomized between subjects, with the constraint that no more than two runs of the same condition were presented consecutively. Each run ended with a “Spelling Only” condition in which subjects spelled 4 letters without auditory distraction. After every run, there was a short break during which subjects answered six questions about the content of the stories that were presented in that run. In addition, NASA-TLX and VAS were prompted at the indicated time points.

condition ($r = .49, p < .05$) and 25.2% in the high workload condition ($r = .50, p < .05$). Further, explorative analyses were performed that are described in the results section.

2.9. Information transfer rate

To measure the amount of data transmitted per time unit, we used the formula suggested by Wolpaw, Ramoser, McFarland, and Pfurtscheller (1998) and first described in Pierce (1980) to calculate bit rate. Despite its disadvantage of not taking into account unequal error distribution, it is the most widely used in BCI studies (Kronegg, Valoshynovskiy, & Pun, 2005).

$$B = \log_2 N + P \log_2 P + (1 - P) \log_2 \left(\frac{1 - P}{N - 1} \right)$$

N represents the number of possible targets (36 with the 6×6 matrix) and P represents the probability that these are correctly classified (average spelling accuracy). The Information transfer rate in bits/min was obtained by multiplying the bit rate B by the average number of selections per minute.

3. Results

3.1. Behavioral data

3.1.1. Effects of mental workload

The results of the experimental manipulation of mental workload on BCI spelling accuracy, subjective workload and number of control questions answered correctly are summarized in Fig. 2. The spelling accuracies for all experimental conditions are listed in Table 1.

A repeated measures ANOVA with the 2 levels medium workload and high workload revealed a main effect of workload on spelling accuracy, $F_{(1, 19)} = 45.79, p < .001$. BCI spelling accuracy was significantly higher for the medium workload condition (80% accuracy; 14.75 bits/min) as compared to the high workload condition (65% accuracy; 10.78 bits/min).

The number of correct answers to the multiple-choice questions served as manipulation check. Participants answered more questions correct after the high workload (4.28) than the medium workload conditions, in which participants were instructed to ignore the stories (3.01), $F_{(1, 19)} = 30.06, p < 0.001$. The total number of correctly answered questions (38%) of all the questions marked as guessed, did not differ significantly from chance (33%), $\chi^2 = 2.17,$

$df = 2, p = .338$, which demonstrates that correct answers were not merely guessed.

3.1.2. Time effects (pre/post differences)

On average, participants needed 101 min ($SD \pm 10$) to complete the experimental protocol after the classifier had been trained. Fig. 3 shows the subjective level of fatigue before and after the experiment and the BCI performance over the course of the experiment for the medium and high workload conditions. A repeated measures ANOVA with time (pre/post) as within subject factor revealed a main effect of time for the subjective ratings of fatigue (VAS scores). They were significantly higher at the end of the experiment (5.98) than at the beginning (2.37), $F_{(1, 19)} = 115.04, p < .001$.

The repeated measures 2×2 ANOVA revealed main effects of workload, $F_{(1, 19)} = 53.74, p < .001$, and time, $F_{(1, 19)} = 19.11, p < .001$, on spelling accuracy. Average spelling accuracies for all conditions are listed in Table 1. For both, the medium, $t(19) = 3.18, p < .01$, and high workload conditions, $t(19) = 2.13, p < .05$, spelling accuracy in the last run was significantly lower than in the first.

There were main effects of workload, $F_{(1, 19)} = 12.97, p < .01$, and time, $F_{(1, 19)} = 7.33, p < .05$, and a significant workload $\times$ time interaction, $F_{(1, 19)} = 6.47, p < .05$, for the subjective workload (NASA-TLX score). Post hoc t -tests revealed that only for the medium workload condition, ratings of workload were significantly higher after the last run (62.43) than after the first (50.51), $t(19) = -3.29, p < .01$. Only after the first run, subjective workload was significantly higher in the high workload condition (61.83) as compared to the medium workload condition (50.51), $t(19) = -4.78, p < .001$.

3.1.3. Spelling Only condition

For the Spelling Only condition, a repeated measures ANOVA (10) revealed a main effect of time, $F_{(9, 171)} = 2.51, p < .01$, on spelling accuracy. Post hoc t -tests revealed significantly lower spelling accuracy in run 8 compared to runs 2 and 3 ($p < .05$). Mean spelling accuracies over the course of the experiment are depicted in Fig. 4.

3.2. Electrophysiological data

Mean P300 amplitudes and latencies for all experimental conditions are listed in Table 1.

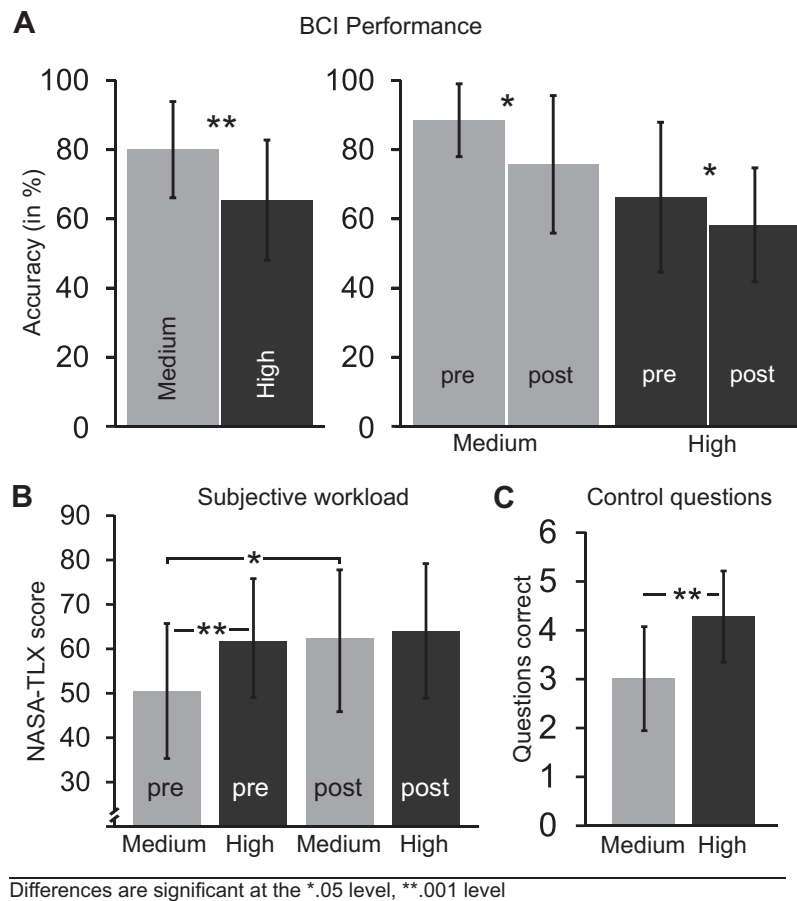


Fig. 2. Behavioral data for the different workload conditions. A shows the average BCI performance for the medium and high workload conditions and the differences between the accuracy in the first (pre) and last (post) run, respectively. B depicts the subjective level of workload and C the average number of correctly answered multiple response questions per run.

Table 1
Average P300 amplitude, latency, BCI performance and bitrate for all experimental conditions.

Condition	P300 amplitude at Pz in μV (SD)	P300 latency at Pz in ms (SD)	Accuracy in % correct	Bitrate (bits/min)
Medium (all)	3.45 (1.23)	427.98 (53.89)	80.00 (13.90)	14.41
High (all)	3.02 (1.38)	415.63 (63.83)	65.40 (17.41)	10.38
Medium pre	4.32 (1.23)	430.07 (66.85)	88.50 (10.52)	17.12
Medium post	3.59 (1.63)	421.93 (56.50)	75.75 (19.89)	13.17
High pre	3.55 (1.22)	427.51 (67.37)	66.25 (21.70)	10.60
High post	2.89 (1.57)	429.95 (55.00)	58.25 (16.49)	8.62
Spelling Only	3.34 (1.25)	427.27 (49.56)	78.40 (14.67)	13.82

3.2.1. Effects of mental workload

The grand average waveforms for the medium and high workload conditions are depicted in Fig. 5. The repeated measures ANOVA (2) revealed a main effect of workload, $F_{(1, 19)} = 5.70$, $p < .05$, on P300 amplitude. It was significantly smaller in the high as compared to the medium workload condition. For the latency, no effect for condition was found, $F_{(1, 19)} = 3.51$, $p = .08$.

The repeated measures ANOVA (2) revealed no difference in the alpha band (8–12 Hz) at Pz between the medium and high workload condition, $F_{(1, 19)} = .01$, $p = .938$. Further, no differences in the theta band between the two workload conditions were found, $F_{(1, 19)} = .33$, $p = .577$.

3.2.2. Explorative analyses

The effects of workload on the alpha and theta band might be too small to be visible in a comparison of the two workload conditions. However, they might be apparent compared to a baseline condition that does not include a secondary task. In our study design

we did not include such a condition, however, the data from the screening run and the data from the Spelling Only condition may serve as baselines, since they do not include auditory distraction. To reveal differences in the medium and high workload conditions as compared to conditions without distraction, we performed the following explorative analyses.

A repeated measures ANOVAs with condition (4 levels: screening run, Spelling Only, medium workload and high workload condition) as factor yielded a main effect of condition on power in the alpha band at Pz, $F_{(3, 57)} = 6.85$, $p < .01$. Post hoc tests revealed significantly less activity in the alpha band during the screening runs as compared to the medium ($p < .01$) and high workload condition ($p < .001$) and compared to the Spelling Only condition ($p < .05$). Likewise, a main effect of condition was found for power in the theta range at Fz, $F_{(1.88, 35.69)} = 5.59$ Greenhouse–Geisser corrected, $p < .01$. Post hoc tests revealed higher theta power in the medium and high workload condition as compared to the screening run ($p < .05$). In addition, theta power was significantly higher in the

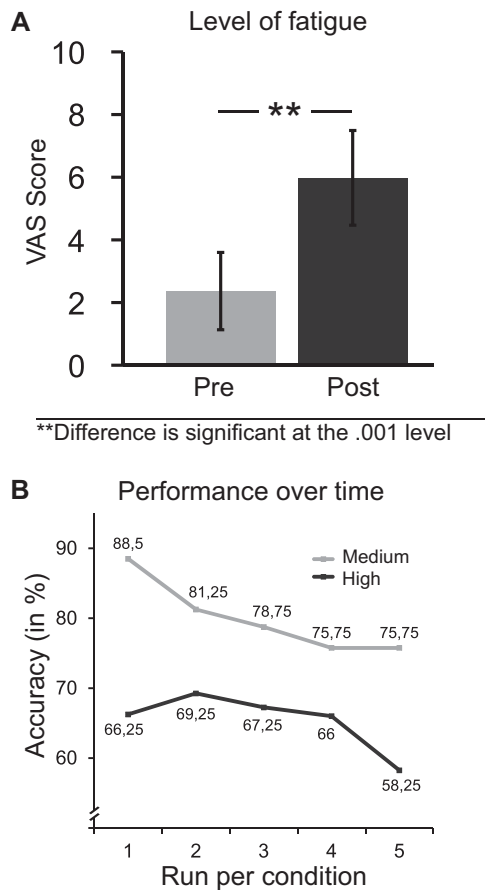


Fig. 3. Effects of time on performance and level of fatigue. A depicts the differences in the subjective level of fatigue as assessed with the visual analog scale before (pre) and after (post) the BCI tasks. B shows the course of performance as a function of time for the medium and high workload conditions.

high workload condition as compared to the Spelling Only condition ($p < .05$). The amount of data that was included in the analysis differed for the four conditions. It consisted of the averaged data of 200 target and 1000 non-target trials for the screening run, 400 target and 2000 non-target trials for the Spelling Only condition and 1000 target and 5000 non-target trials for each of the workload conditions.

The primary analysis was restricted to Pz (for alpha power) and Fz (for theta power). To further explore differences at all 31 electrode positions for the medium and high workload condition, we compared alpha and theta power in these conditions to the screening run. Multiple t -tests were conducted and

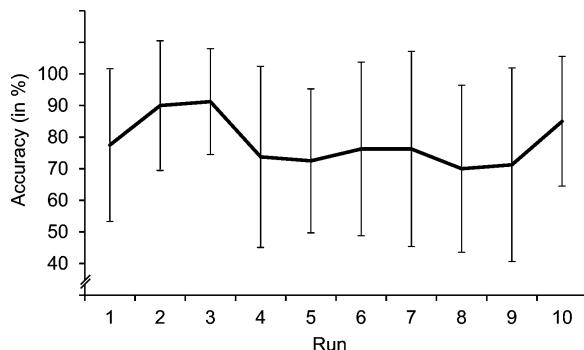


Fig. 4. Spelling accuracy over the course of the 10 Spelling Only runs.

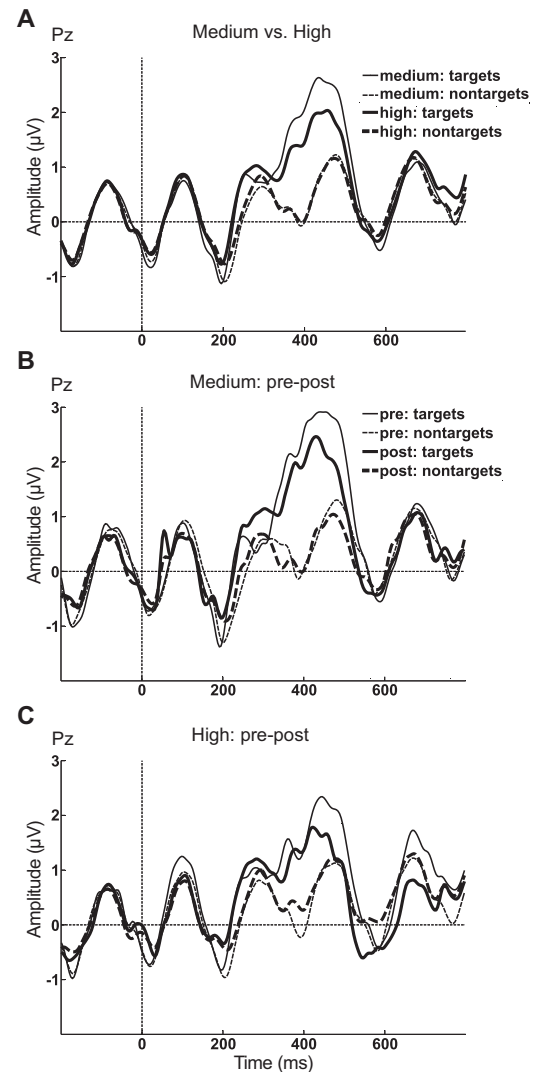


Fig. 5. Grand average waveforms at Pz. A shows the average waveforms for the medium and high workload conditions. The averaged waveforms for the first and last run of the medium workload condition are depicted in B and for the high workload condition in C.

differences that were significant at the uncorrected levels ($p < .05$ and $p < .01$) are reported along with the differences that remained significant after Bonferroni correction ($0.05/31 = p < .0016$ and $0.01/31 = p < 3.2 \times 10^{-4}$).

Significant differences are depicted in Fig. 6. Alpha power was significantly higher at widespread electrode positions for both workload conditions compared to the screening run. Significance values were highest at parietal and parieto-occipital sites. For theta power significant differences were only found at frontal and fronto-central electrode sites in the high workload condition as compared to the screening run using adjusted p -values.

3.3. Time effects (pre/post differences)

Main effects of time, $F_{(1, 19)} = 19.36$, $p < .001$, and workload, $F_{(1, 19)} = 6.40$, $p < .05$, and a significant workload $\times$ time interaction, $F_{(1, 19)} = 21.83$, $p < .001$, were found for P300 amplitude at Pz. P300 amplitude was significantly smaller during the last run ($3.59 \mu V$) of the medium workload condition as compared to the first run (4.32 , $p < .05$). The P300 was also smaller during the last run of the high workload condition ($2.89 \mu V$) as compared to the first run, but the difference was not significant ($3.55 \mu V$, $p = .084$). Differences

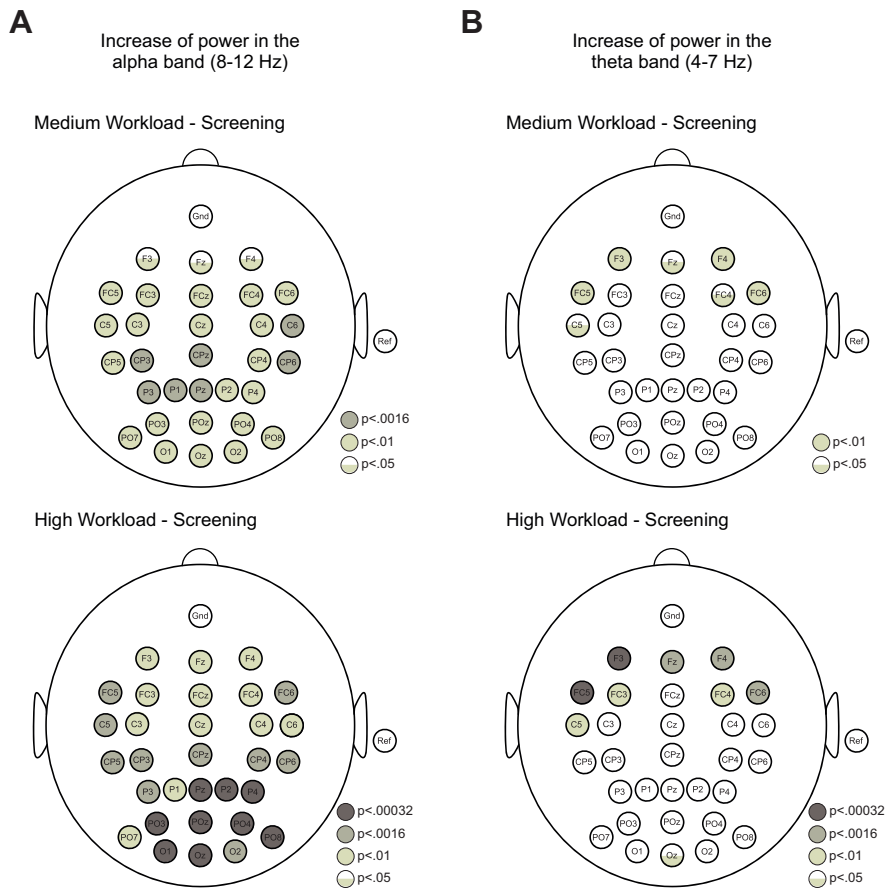


Fig. 6. Differences of power in the alpha (A) and theta band (B) at individual electrode sites. Significance values of the power differences are shown for the medium and high workload condition compared to the screening run.

between the conditions were significant for the first run ($p < .05$), but not for the last run. There was no significant effect of workload or time on P300 latency at Pz.

A 2×2 repeated measures ANOVA revealed a main effect of time for activity in the alpha band, $F_{(1, 19)} = 19.62$, $p < .001$, with an increase in activity in the last as compared to the first run but no significant effect of workload, $F_{(1, 19)} = .24$, n.s. No significant effects of workload or time were found for activity in the theta range at Fz.

3.3.1. Spelling Only condition

For the Spelling Only condition, a repeated measures ANOVA with runs (10) as factor revealed a main effect of time on alpha power at Pz, $F_{(4.14, 78.69)} = 4.99$, Greenhouse–Geisser corrected, $p < .01$. Post hoc tests revealed that alpha power was significantly higher in run 8 compared to run 1 ($p < .01$) and higher in run 6 compared to 2, and in run 4 compared to 2 ($p < .05$). A linear trend could be observed and was confirmed by a linear regression analysis. Time (run number) significantly predicted alpha power at Pz, $b = .115$, $t(199) = 2.18$, $p < .05$. Time explained a significant proportion of variance in alpha power at Pz, $R^2 = .024$, $F_{(1, 199)} = 4.77$, $p < .05$.

Repeated measures ANOVAs (10) revealed no main effects of time on P300 amplitude, $F_{(9, 171)} = 1.05$, $p = .399$, P300 latency, $F_{(9, 171)} = 1.28$, $p = .278$, or theta power at Fz, $F_{(4.52, 85.93)} = 2.08$, Greenhouse–Geisser corrected, $p = .082$.

3.3.2. Explorative analyses

To explore differences in alpha and theta power between the first and last runs of the medium and high workload condition

at individual electrodes, multiple t -tests were conducted. Fig. 7 displays the level of significance for individual electrodes. The three electrodes with the lowest p -values in the high workload condition were CPz ($p = 1.8 \times 10^{-5}$), Pz (2.6×10^{-5}) and P1 (3.4×10^{-5}). For the medium workload condition a general increase in activity in the alpha band could be observed, but it did not reach significance at the adjusted level. The three electrodes with the lowest p -values were Pz ($p = .004$), P2 and CPz ($p = .005$).

Differences in the theta band (4–7 Hz) did not reach significance at the adjusted levels of significance. For the high workload condition an increase in activity for the last as compared to the first run was found at Fz, FC5, FC6, C5, CP5, P3, O1, O2 and PO8 ($p < .05$, not adjusted for multiple comparisons). A significant increase in activity in the last run compared to the first run of the medium workload condition was only found at electrode C5 ($p < .05$).

4. Discussion

We succeeded in the experimental manipulation of mental workload to induce changes in behavioral and electrophysiological indicators of performance during prolonged operation of a P300 BCI. As an indicator of high mental workload, a reduced P300 amplitude was found. The effects of the prolonged operation were evident in the time and frequency domain of the EEG. An increased activity in the alpha band was observed for the last run as compared to the first run for both, the medium and high workload condition.

The amplitude of the P300 was significantly smaller for the last run of the medium workload condition as compared to the first.

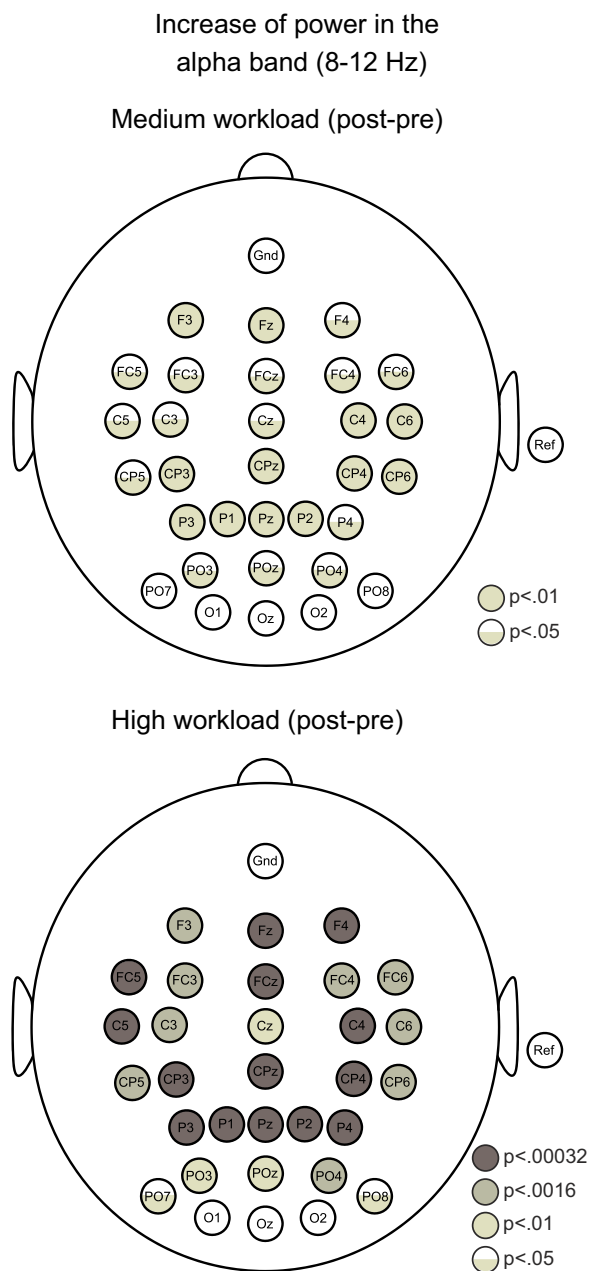


Fig. 7. Differences of power in the alpha band (8–12 Hz) at individual electrode positions. Significance values of the power differences are shown for the last run compared to the first run of the high and medium workload conditions.

The study also demonstrated that users were able to maintain a satisfactory level of performance with the P300 BCI despite long operation and constant auditory distraction.

4.1. Effects of the manipulation of mental workload

The behavioral data suggests that the experimental manipulation succeeded since it significantly affected performance with the BCI. Participants spelled better in the medium workload condition than in the high workload condition.

An accuracy of 70% correct selections was defined as the threshold value necessary to achieve satisfactory communication with a BCI (Kübler, Neumann, Kaiser, Kotchoubey, Hinterberger, & Birbaumer, 2001). The average accuracy of 80% achieved in the medium workload condition is above this level,

despite constant distraction through background noise. In the high workload condition, accuracy dropped below this value (65%). Thus, mental workload was too high for the participants to perform at a satisfactory level. Accordingly, subjective workload as rated with the NASA-TLX was significantly higher than in the medium workload condition. Both ratings were substantially higher than ratings in a previous study in which participants used solely the visual speller without an additional task (Käthner et al., 2013). The multiple choice questionnaire, which was administered after every run, served as an index of compliance and task performance. Participants answered more questions correct in the high as compared to the medium workload condition, suggesting that task compliance was high and the manipulation successful. In addition, the fact that the number of correctly guessed answers did not differ significantly from chance level, indicates suitability of this instrument to assess (secondary) task performance.

As expected, P300 amplitude was significantly reduced in the high workload condition as compared to the medium workload condition and no significant differences in latency were found. This result confirms previous studies that reported also a decreased P300 amplitude with higher workload (Allison & Polich, 2007; Miller et al., 2011; Ullsperger et al., 2001; for a review see Kok, 2001) and again indicates that cognitive demands during task processing influence P300 amplitude. Although we analyzed P300 separately from changes in the frequency bands, there is generally a strong connection between the event-related potentials and oscillatory activity in specific frequency bands. P300 consists primarily of phase-locked delta and theta-range synchronized oscillations (Başar-Eroglu, Başar, Demiralp, & Schürmann, 1992; Spencer & Polich, 1999). Başar-Eroglu et al. (1992) observed diffuse delta increase and prolonged activation in the theta band in response to rare tones in an oddball paradigm. Further, Yordanova, Kolev, and Polich (2001) demonstrated strong relationships between event-related alpha power suppression (event-related desynchronization, ERD) and P300 latency and amplitude in auditory oddball tasks. ERD onset was negatively associated with P300 amplitude and positively with P300 latency. The authors also found that P300 preceded ERD indicating that the underlying cognitive processes are distinct, but that the internal processes underlying P300 influence (guide or modify) those underlying ERD.

Unlike hypothesized, there were no significant differences in activity in the alpha and theta bands between the medium and high workload conditions. However, the explorative analysis revealed that alpha power at Pz and theta power at Fz were increased in both workload conditions compared to the screening run, in which no additional workload was induced. The increase in alpha power with increasing task demands was unexpected, but had previously been observed in the study by Kohlmoorgen et al. (2007), who induced workload by means of a mental calculation and an auditory task during a real-life driving condition. The authors explained their findings with the nature of their workload inducing tasks, which was not visual in contrast to the secondary tasks used in many studies. An increase of alpha activity with increasing task demands, especially during auditory stimulation had previously been reported (Galín, Johnstone, & Herron, 1978; Legewie, Simonova, & Creutzfeldt, 1969; Markand, 1990). These findings are descriptive and do not allow conclusions on the functionality of alpha oscillations. A possible functional interpretation is the alpha inhibition hypothesis (Klimesch, Doppelmayr, Schwaiger, Auinger, & Winkler, 1999; Klimesch, Sauseng, & Hanslmayr, 2007 for a review). According to this theory, low alpha activity reflects active neuronal processing, whereas high alpha activity indicates the inhibition of task-irrelevant brain regions. In our study, the secondary, workload inducing task was either to ignore two simultaneously presented stories or listen to both stories to recall

their content later on. In line with the inhibition theory, top-down inhibition of the brain regions for auditory processing in the ignore condition could have led to the increase in alpha power. In the high work load condition that required divided attention, participants had to retain information in working memory. Klimesch et al. (1999) found that alpha activity was increased during encoding and retention of a working memory task and argued that the increase in alpha power reflects inhibitory top-down control to block retrieval of information. This could also be the case for the high workload condition of our study. However, Palva and Palva (2007) suggest that the working memory related alpha oscillations represent active processing of the frontoparietal network involved in sustaining the neuronal representations of memorized items. A claim that is supported by the positive correlation of alpha activity and working memory load and task difficulty (Jensen, Gelfand, Kounios, & Lisman, 2002; Sauseng et al., 2005). The mixed results indicate that the empirical data on the functional role of alpha band oscillations is not yet conclusive.

The explorative analysis of our study revealed differences in alpha power at multiple, widespread electrode sites for the workload conditions compared to the screening run, while for theta power differences were only found at frontal and frontocentral electrode sites. These topographic findings are in line with previous findings that show an increase of frontal theta with increasing workload (Gevins et al., 1998; Gundel & Wilson, 1992; Holm et al., 2009; Mecklinger et al., 1992). Further, Jensen and Tesche (2002) found an increase of frontal theta with increasing amount of encoded information and Sarnthein, Petsche, Rappelsberger, Shaw, and von Stein (1998) reported prefrontal to left temporo-parietal theta coupling especially for retention of verbal information. Depending on the task, the cognitive processes involved in working memory vary, but usually entail several processes such as encoding, storage and retrieval of information, attentional control, interfacing with long-term memory, multi modal integration and others (Sauseng, Griesmayr, Freunberger, & Klimesch, 2010). Nevertheless, an increase of theta is consistently reported. These observations led Sauseng et al. (2010) to suggest that theta might be an integrative brain mechanism to coordinate the processes in several brain regions involved in working memory. In the workload conditions of our study working memory related processes such as attentional control, encoding and storage of information were involved and might be indicated by the increase in frontal theta.

In the context of BCIs, it was recently shown in Mak et al. (2012) that an increase in theta power was associated with a decrease in P300 BCI performance, which is also the case for our study since performance decreased with higher mental load and theta increased.

Apart from changes in the frequency bands reported in this paper, changes could also occur in other frequency bands, or in particular sub bands (e.g. lower or upper alpha) and should be taken into account for workload detection algorithms. Therefore, we provide information for frequency bands not described in this paper in supplementary Fig. 1.

Supplementary Fig. 1 related to this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.biopsycho.2014.07.014>.

In this section, we discussed those processes modulating alpha and theta band activity with regard to mental workload that are under volitional control. It has to be noted, however, that tonic changes that occur, for instance as a consequence of fatigue, also play a role. This holds especially true for the explorative analyses, because the averaged effects from several runs were compared to the screening run that was performed in the beginning.

4.2. Effects of prolonged BCI operation

In accordance with the hypothesized effects, performance decreased over the course of the experiment. During the last run of the medium and high workload conditions, spelling accuracy was significantly lower than in the first. Further, subjective ratings of fatigue were significantly higher at the end of the experiment as compared to the beginning. Likewise, the P300 amplitude was reduced during the last as compared to the first run of the medium workload condition and the increase of power in the alpha band was significant in the high and medium workload conditions. An explorative analysis revealed a general increase of alpha power at multiple frontal, central and parietal electrode positions with strongest effects around parietal and central-parietal electrode sites. While for the first four runs of the high workload condition spelling accuracy was just below the critical value of 70%, it dropped below 60% in the last run. Considering the behavioral data, it is likely that the increase of activity in the alpha band indicates an increase of fatigue and a drop in the level of arousal and subsequent attention deficits. This interpretation is in line with the following findings from previous studies. If a person gets sleepier the activity in the alpha band typically decreases and eventually disappears (Santamaria & Chiappa, 1987). However, this holds only true for an eyes closed condition. If eyes are open, activity in the alpha band increases globally with increasing sleepiness (Åkerstedt, Torsvall, & Gillberg, 1985; Daniel, 1967; Stampi et al., 1995; Torsvall & Åkerstedt, 1987). Further, O'Connell et al. (2009) found that increasing activity in the alpha band was the strongest electrophysiological predictor of lapsing attention during the vigilance task used in their study. However, in contrast, Braboszcz and Delorme (2010) found a decrease in alpha power during periods of self-reported mind wandering, a state that constitutes low-alertness. Klimesch (1999) pointed out that the increase of alpha activity with increasing sleepiness is especially apparent in the lower alpha band. Interestingly, this is only the case if participants are not allowed to fall asleep and thus, the increase in lower alpha might indicate the increased effort of study participants to stay alert and awake. On the other hand, changes in upper alpha are probably task-induced and associated to storage processes during retention (Schack, Klimesch, & Sauseng, 2005).

Along with the changes in the alpha band, widespread increase of theta activity is seen in sleep deprived persons and during sustained wakefulness (Borbély et al., 1981; Cajochen, Brunner, Krauchi, Graw, & WirzJustice, 1995). The explorative analysis of the present study showed that theta was also increased at multiple widespread scalp positions (note that this was only observed at a level of significance not adjusted for multiple comparisons). Thus it is likely that the increase of theta, observed in the present study, represents an increase of fatigue.

As for mental workload, effects of time/fatigue could be apparent in frequency bands not reported in this paper, therefore, we provide information for other bands in supplementary Fig. 2.

Supplementary Fig. II related to this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.biopsycho.2014.07.014>.

The Spelling Only condition allowed to identify the changes in spelling accuracy, P300 amplitude and latency, alpha and theta power that occurred over the ten runs without the direct influence of a secondary task. Since the Spelling Only condition always followed the dual task conditions, spillover effects cannot be excluded. A main effect of time was revealed for spelling accuracy and alpha power at Pz. The performance with the BCI without auditory distraction does not follow a linear trend. It is generally high, but substantially varies around the mean accuracy of 78% (range 70–91%). The high variance within each run can be explained by the low number of letters to be spelled (4). Spelling accuracy is

significantly higher in the beginning (runs 2 and 3: $\geq 90\%$) than toward the end (run 8: 70%), but is again high in the last run. Participants probably increased their mental effort, when they realized that the experiment was almost finished.

4.3. Toward a BCI that adapts to the mental state of the user

Ultimately, BCIs are intended as assistive technology to be used more or less independently by severely paralyzed patients. In this study, a high average performance of healthy users even during constant distraction was demonstrated. This is encouraging for BCI use in a home environment, where distraction is unavoidable.

However, classification accuracies are usually lower for patients than healthy users and in some the required pattern of EEG activity cannot be classified by the BCI (Kübler & Birbaumer, 2008; Mak et al., 2011; Nijboer et al., 2008). Countermeasures are necessary to overcome the so-called BCI inefficiency (Kübler, Blankertz, Müller, & Neuper, 2011). Different solutions were proposed to solve this problem. For instance, Kaufmann et al. (2013) have shown that changes in the stimulus material can boost performance for individual patients with neurodegenerative disease. Another approach may be to use a BCI that adapts to the mental state of the user. Thus, a BCI is desirable that automatically detects high workload and fatigue and either adapts the stimulation parameters or pauses the system, if voluntary selections are no longer possible. The current study demonstrates that ERPs and changes in the alpha and theta frequency bands are suitable candidates for the automatic detection of workload during operation of a P300 BCI.

Further, the automatic detection of mental states with EEG is of interest also for other user groups. A major advantage of the EEG over other functional neuroimaging methods is its portability. Recently it was demonstrated that the acquisition of good EEG signals is possible with a low cost mobile EEG, even when users are walking outdoors (Debener, Minow, Emkes, Gandras, & De Vos, 2012). Such developments increase the likelihood that the detection of mental states with the EEG will be possible during real life situations – a goal that has been pursued for decades (Brookings et al., 1996; Gevins et al., 1995; Kramer, 1991; Lindholm, Cheatham, Koriath, & Longridge, 1984; Sirevaag et al., 1993).

4.4. Limitations of the study

The study reported electrophysiological findings on the group level for fixed frequency bands. Previous studies demonstrated that alpha frequency not only varies with task demands, but is also strongly influenced by individual characteristics such as age and memory performance (Klimesch, Schimke, & Pfurtscheller, 1993; Niedermeyer, 1993), albeit the results require validation on the single subject level.

While the study succeeded in the manipulation of mental workload, the study design limits comparisons of the high and medium workload with baseline conditions (spelling without secondary task). Runs without auditory tasks, which were equal in length to the workload conditions, were not included to keep the total duration of the experiment at an acceptable level. Nevertheless, as an explorative analysis, the workload conditions were compared to the screening run and the Spelling Only condition. When interpreting the results of this analysis, one has to be aware of the limited comparability of these conditions. In particular, it is important to note that the Spelling Only condition always started after the spelling of 20 letters, thus, it cannot be ruled out that concentration had dropped. On the other hand, concentration was probably high during the screening run, which was performed before all online spelling runs.

5. Conclusions

The study demonstrates that satisfactory accuracies can be achieved by healthy participants with the P300 speller despite constant distraction and prolonged operation. This is promising for the use of BCIs in a home environment where distraction is inevitable. To improve usability for the target user group of severely paralyzed patients, it was suggested to implement a BCI that adapts to the mental state of the user. As a further step in this direction, the current study identified behavioral and electrophysiological changes that could serve as indicators of high workload or fatigue. Our results confirm that both, the frequency and time domains of the EEG contain valuable information that could be used for the automatic detection of workload and/or fatigue during operation of a P300 BCI.

Acknowledgements

The study was funded by the European ICT Program Project FP7-288566 and supported by Deutsche Forschungsgemeinschaft (DFG) RTG 1252/1. This paper only reflects the authors' views and funding agencies are not liable for any use that may be made of the information contained herein.

References

- Åkerstedt, T., Torsvall, L., & Gillberg, M. (1985). *Sleepiness in laboratory and field studies*. In W. P. Koella, E. Ruther, & H. Schulz (Eds.), *Sleep '84*. New York: Gustav Fischer.
- Allison, B. Z., & Polich, J. (2008). Workload assessment of computer gaming using a single-stimulus event-related potential paradigm. *Biological Psychology*, 77(3), 277–283.
- Başar-Eroglu, C., Başar, E., Demiralp, T., & Schürmann, M. (1992). P300-response: Possible psychophysiological correlates in delta and theta frequency channels. A review. *International Journal of Psychophysiology*, 13(2), 161–179. [http://dx.doi.org/10.1016/0167-8760\(92\)90055-G](http://dx.doi.org/10.1016/0167-8760(92)90055-G)
- Berger, H. (1929). Über das Elektroencephalogramm des Menschen. *Archiv für Psychiatrie und Nervenkrankheiten*, 87(1), 527–570. <http://dx.doi.org/10.1007/BF01797193>
- Birbaumer, N. (2006). Breaking the silence: Brain–computer interfaces (BCI) for communication and motor control. *Psychophysiology*, 43(6), 517–532.
- Boksem, M. A. S., Meijman, T. F., & Lorist, M. M. (2005). Effects of mental fatigue on attention: An ERP study. *Cognitive Brain Research*, 25(1), 107–116.
- Borbély, A. A., Baumann, F., Brandeis, D., Strauch, I., & Lehmann, D. (1981). Sleep deprivation: Effect on sleep stages and EEG power density in man. *Electroencephalography and Clinical Neurophysiology*, 51(5), 483–493.
- Braboszcz, C., & Delorme, A. (2011). Lost in thoughts: Neural markers of low alertness during mind wandering. *NeuroImage*, 54(4), 3040–3047. <http://dx.doi.org/10.1016/j.neuroimage.2010.10.008>
- Brookings, J. B., Wilson, G. F., & Swain, C. R. (1996). Psychophysiological responses to changes in workload during simulated air traffic control. *Biological Psychology*, 42(3), 361–377.
- Brouwer, A. -M., Hogervorst, M. A., van Erp, J. B. F., Heffelaar, T., Zimmerman, P. H., & Oostenveld, R. (2012). Estimating workload using EEG spectral power and ERPs in the n-back task. *Journal of Neural Engineering*, 9(4). <http://dx.doi.org/10.1088/1741-2560/9/4/045008>
- Cajochen, C., Brunner, D. P., Krauchi, K., Graw, P., & WirzJustice, A. (1995). Power density in theta/alpha frequencies of the Waking EEG progressively increases during sustained wakefulness. *Sleep*, 18(10), 890–894.
- Coull, J. (1998). Neural correlates of attention and arousal: Insights from electrophysiology, functional neuroimaging and psychopharmacology. *Progress in Neurobiology*, 55(4), 343–361.
- Daniel, R. S. (1967). Alpha and theta EEG in vigilance. *Perceptual and Motor Skills*, 25(3), 697–703.
- Debener, S., Minow, F., Emkes, R., Gandras, K., & De Vos, M. (2012). How about taking a low-cost, small, and wireless EEG for a walk? *Psychophysiology*, 49(11), 1617–1621.
- 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(1), 9–21.
- Dement, W., & Kleitman, N. (1957). Cyclic variations in EEG during sleep and their relation to eye movements, body motility, and dreaming. *Electroencephalography and Clinical Neurophysiology*, 9(4), 673–690.
- Dinges, D. F., Pack, F., Williams, K., Gillen, K. A., Powell, J. W., Ott, G. E., et al. (1997). Cumulative sleepiness, mood disturbance, and psychomotor vigilance performance decrements during a week of sleep restricted to 4–5 h per night. *Sleep*, 20(4), 267–277.

- Farwell, L. A., & Donchin, E. (1988). Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and Clinical Neurophysiology*, 70(6), 510–523.
- Fournier, L. R., Wilson, G. F., & Swain, C. R. (1999). Electrophysiological, behavioral, and subjective indexes of workload when performing multiple tasks: Manipulations of task difficulty and training. *International Journal of Psychophysiology*, 31(2), 129–145.
- Galín, D., Johnstone, J., & Herron, J. (1978). Effects of task difficulty on EEG measures of cerebral engagement. *Neuropsychologia*, 16(4), 461–472.
- Gasser, T., Bächer, P., & Möcks, J. (1982). Transformations towards the normal distribution of broad band spectral parameters of the EEG. *Electroencephalography and Clinical Neurophysiology*, 53(1), 119–124.
- Gevens, A., Yeager, C., Zeitlin, G., Ancoli, S., & Dedon, M. (1977). On-line computer rejection of EEG artifact. *Electroencephalography and Clinical Neurophysiology*, 42(2), 267–274.
- Gevens, A., Leong, H., Du, R., Smith, M. E., Le, J., DuRousseau, D., et al. (1995). Towards measurement of brain function in operational environments. *Biological Psychology*, 40(1–2), 169–186.
- Gevens, A., Smith, M. E., Leong, H., McEvoy, L., Whitfield, S., Du, R., et al. (1998). Monitoring working memory load during computer-based tasks with EEG pattern recognition methods. *Human Factors*, 40(1), 79–91.
- Gift, A. (1989). Visual analog scales – Measurement of subjective phenomena. *Nursing Research*, 38(5), 286–288.
- Gundel, A., & Wilson, G. F. (1992). Topographical changes in the ongoing EEG related to the difficulty of mental tasks. *Brain Topography*, 5(1), 17–25.
- Hart, S. G. (2006). *NASA-Task Load Index (NASA-TLX) 20 years later*. Santa Monica, CA, USA: Paper presented at the human factors and ergonomics society 50th annual meeting 2006.
- Hockey, G. R. J. (1997). Compensatory control in the regulation of human performance under stress and high workload: A cognitive-energetical framework. *Biological Psychology*, 45(13), 73–93.
- Holm, A., Lukander, K., Korpela, J., Sallinen, M., & Müller, K. M. I. (2009). Estimating brain load from the EEG. *Scientific World Journal*, 9, 639–651.
- Holz, E. M., Höhne, J., Staiger-Sälzer, P., Tangermann, M., & Kübler, A. (2013). Brain–computer interface controlled gaming: Evaluation of usability by severely motor restricted end-users. *Artificial Intelligence in Medicine*, 59(2), 111–120.
- Humphrey, D. G., & Kramer, A. F. (1994). Toward a psychophysiological assessment of dynamic changes in mental workload. *Human Factors*, 36(1), 3–26.
- Isreal, J. B., Chesney, G. L., Wickens, C. D., & Donchin, E. (1980). P300 and tracking difficulty: Evidence for multiple resources in dual-task performance. *Psychophysiology*, 17(3), 259–273.
- Jensen, O., & Tesche, C. D. (2002). Frontal theta activity in humans increases with memory load in a working memory task. *European Journal of Neuroscience*, 15(8), 1395–1399. <http://dx.doi.org/10.1046/j.1460-9568.2002.01975.x>
- Jensen, O., Gelfand, J., Kounios, J., & Lisan, J. E. (2002). Oscillations in the alpha band (9–12 Hz) increase with memory load during retention in a short-term memory task. *Cerebral Cortex*, 12(8), 877–882. <http://dx.doi.org/10.1093/cercor/12.8.877>
- Johnson, R. R., Popovic, D. P., Olmstead, R. E., Stikic, M., Levendowski, D. J., & Berka, C. (2011). Drowsiness/alertness algorithm development and validation using synchronized EEG and cognitive performance to individualize a generalized model. *Biological Psychology*, 87(2), 241–250.
- Käthner, I., Ruf, C. A., Pasqualotto, E., Braun, C., Birbaumer, N., & Halder, S. (2013). A portable auditory P300 brain–computer interface with directional cues. *Clinical Neurophysiology*, 124(2), 327–338.
- Kaufmann, T., Schulz, S. M., Köblitz, A., Renner, G., Wessig, C., & Kübler, A. (2013). Face stimuli effectively prevent brain–computer interface inefficiency in patients with neurodegenerative disease. *Clinical Neurophysiology*, 124(5), 893–900.
- Kleih, S. C., Kaufmann, T., Zickler, C., Halder, S., Leotta, F., Cincotti, F., et al. (2011). Out of the frying pan into the fire – The P300-based BCI faces real-world challenges. *Progress in Brain Research*, 194, 27–46.
- Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis. *Brain Research Reviews*, 29(2–3), 169–195.
- Klimesch, W., Schimke, H., & Pfurtscheller, G. (1993). Alpha frequency, cognitive load and memory performance. *Brain Topography*, 5(3), 241–251.
- Klimesch, W., Doppelmayr, M., Schwaiger, J., Auinger, P., & Winkler, T. (1999). 'Paradoxical' alpha synchronization in a memory task. *Cognitive Brain Research*, 7(4), 493–501. [http://dx.doi.org/10.1016/S0926-6410\(98\)00056-1](http://dx.doi.org/10.1016/S0926-6410(98)00056-1)
- Klimesch, W., Sauseng, P., & Hanslmayr, S. (2007). EEG alpha oscillations: The inhibition–timing hypothesis. *Brain Research Reviews*, 53(1), 63–88. <http://dx.doi.org/10.1016/j.brainresrev.2006.06.003>
- Kohlmorgen, J., Dornhege, G., Braun, M., Blankertz, B., Müller, K. R., Curio, G., et al. (2007). Improving human performance in a real operating environment through real-time mental workload detection. In G. Dornhege, J. del, R. Millan, T. Hinterberger, D. McFarland, & K. R. Müller (Eds.), *Toward brain–computer interfacing*. Cambridge, MA: MIT Press.
- Kok, A. (2001). On the utility of P3 amplitude as a measure of processing capacity. *Psychophysiology*, 38(3), 557–577.
- Kramer, A. F. (1991). Physiological metrics of mental workload: A review of recent progress. In D. L. Damos (Ed.), *Multiple-task performance*. London: Taylor & Francis.
- Kronegg, J., Voloshynovskiy, S., & Pun, T. (2005). *Analysis of bit-rate definitions for brain–computer interfaces*. Las Vegas, NV, USA: Paper presented at the 2005 International Conference on Human–Computer Interaction.
- Krusinski, D. J., Sellers, E. W., Cabestaing, F., Bayoudh, S., McFarland, D. J., Vaughan, T. M., et al. (2006). A comparison of classification techniques for the P300 Speller. *Journal of Neural Engineering*, 3(4), 299–305.
- Kübler, A., & Birbaumer, N. (2008). Brain–computer interfaces and communication in paralysis: Extinction of goal directed thinking in completely paralysed patients? *Clinical Neurophysiology*, 119, 2658–2666.
- Kübler, A., Kotchoubey, B., Kaiser, J., Wolpaw, J. R., & Birbaumer, N. (2001). Brain–computer communication: Unlocking the locked in. *Psychological Bulletin*, 127(3), 358–375.
- Kübler, A., Neumann, N., Kaiser, J., Kotchoubey, B., Hinterberger, T., & Birbaumer, N. (2001). Brain–computer communication: Self-regulation of slow cortical potentials for verbal communication. *Archives of Physical Medicine and Rehabilitation*, 82, 1533–1539.
- Kübler, A., Blankertz, B., Müller, K. R., & Neuper, C. (2011). *A model of BCI-control*. Graz, Austria: Paper presented at the fifth international BCI conference.
- Kutas, M., McCarthy, G., & Donchin, E. (1977). Augmenting mental chronometry: The P300 as a measure of stimulus evaluation time. *Science*, 197, 792–795.
- Lal, S. K. L., & Craig, A. (2001). A critical review of the psychophysiology of driver fatigue. *Biological Psychology*, 55(3), 173–194.
- Legewie, H., Simonova, O., & Creutzfeldt, O. (1969). EEG changes during performance of various tasks under open- and closed-eyed conditions. *Electroencephalography and Clinical Neurophysiology*, 27(5), 470–479.
- Lindholm, E., Cheatham, C., Koriath, J., & Longridge, T. (1984). *Physiological assessment of aircraft pilot workload in simulated landing and simulated hostile threat environments* (Technical Report AFHRL-TR-83-49). Williams Air Force Base, AZ: Air Force Systems Command.
- Mak, J. N., Arbel, Y., Minett, J. W., McCane, L. M., Yuksel, B., Ryan, D., et al. (2011). Optimizing the P300-based brain–computer interface: Current status, limitations and future directions. *Journal of Neural Engineering*, 8(2).
- Mak, J. N., McFarland, D. J., Vaughan, T. M., McCane, L. M., Tsui, P. Z., Zeitlin, D. J., et al. (2012). EEG correlates of P300-based brain–computer interface (BCI) performance in people with amyotrophic lateral sclerosis. *Journal of Neural Engineering*, 9(2), 026014. <http://dx.doi.org/10.1088/1741-2560/9/2/026014>
- Makeig, S., & Jung, T. P. (1996). Tonic, phasic, and transient EEG correlates of auditory awareness in drowsiness. *Cognitive Brain Research*, 4(1), 15–25.
- Markand, O. (1990). Alpha-rhythms. *Journal of Clinical Neurophysiology*, 7(2), 163–189.
- Mecklinger, A., Kramer, A. F., & Strayer, D. L. (1992). Event related potentials and EEG components in a semantic memory search task. *Psychophysiology*, 29(1), 104–119.
- Miller, M. W., Rietschel, J. C., McDonald, C. G., & Hatfield, B. D. (2011). A novel approach to the physiological measurement of mental workload. *International Journal of Psychophysiology*, 80(1), 75–78.
- NASA Human Performance Research Group. (1987). *Task Load Index (NASA-TLX)*. Retrieved from: <http://humansystems.arc.nasa.gov/groups/TLX>
- Niedermeyer, E. (1993). Maturation of the EEG: Development of waking and sleep patterns. In E. Niedermeyer, & F. H. Lopes da Silva (Eds.), *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Baltimore, MD: Lippincott Williams & Wilkins.
- Nijboer, F., Furdea, A., Gunst, I., Mellinger, J., McFarland, D. J., Birbaumer, N., et al. (2008). An auditory brain–computer interface (BCI). *Journal of Neuroscience Methods*, 167(1), 43–50.
- O'Connell, R. G., Dockree, P. M., Robertson, I. H., Bellgrove, M. A., Foxe, J. J., & Kelly, S. P. (2009). Uncovering the neural signature of lapsing attention: Electrophysiological signals predict errors up to 20 s before they occur. *Journal of Neuroscience*, 29(26), 8604–8611.
- Ortner, R., Prueckl, R., Putz, V., Scharinger, J., Bruckner, M., Schnuerer, A., et al. (2011). *Accuracy of a P300 speller for different conditions: A comparison*. Graz, Austria: Paper presented at the 5th international brain–computer interface conference 2011.
- Palva, S., & Palva, J. M. (2007). New vistas for alpha-frequency band oscillations. *Trends in Neurosciences*, 30(4), 150–158. <http://dx.doi.org/10.1016/j.tins.2007.02.001>
- Paus, T., Zatorre, R. J., Hofle, N., Caramanos, Z., Gotman, J., Petrides, M., et al. (1997). Time-related changes in neural systems underlying attention and arousal during the performance of an auditory vigilance task. *Journal of Cognitive Neuroscience*, 9(3), 392–408.
- Pierce, J. (1980). *An introduction to information theory: Symbols, signals & noise* (2nd ed.). New York: Dover Publications.
- Pigeau, R., Hoffman, R., Purcell, S., & Moffitt, A. (1987). The effect of endogenous alpha on hemispheric asymmetries and the relationship of frontal theta to sustained attention. In K. Jensen (Ed.), *Electric and magnetic activity of the central nervous system: Research and clinical applications in aerospace medicine*. France: NATO Advisory Group for Aerospace Research and Development.
- Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, 118(10), 2128–2148.
- Riccio, A., Leotta, F., Bianchi, F., Aloise, F., Zickler, C., Hoogerwerf, E. J., et al. (2011). Workload measurement in a communication application operated through a P300-based BCI. *Journal of Neural Engineering*, 8(2), 025028. <http://dx.doi.org/10.1088/1741-2560/8/2/025028>
- Santamaria, J., & Chiappa, K. (1987). The EEG of drowsiness in normal adults. *Journal of Clinical Neurophysiology*, 4(4), 327–382.
- Sarnthein, J., Petsche, H., Rappelsberger, P., Shaw, G. L., & von Stein, A. (1998). Synchronization between prefrontal and posterior association cortex during human working memory. *Proceedings of the National Academy of Sciences of the United States of America*, 95(12), 7092–7096. <http://dx.doi.org/10.1073/pnas.95.12.7092>
- Sauseng, P., Klimesch, W., Doppelmayr, M., Pecherstorfer, T., Freunberger, R., & Hanslmayr, S. (2005). EEG alpha synchronization and functional coupling during

- top-down processing in a working memory task. *Human Brain Mapping*, 26(2), 148–155. <http://dx.doi.org/10.1002/hbm.20150>
- Sauseng, P., Griesmayr, B., Freunberger, R., & Klimesch, W. (2010). Control mechanisms in working memory: A possible function of EEG theta oscillations. *Neuroscience and Biobehavioral Reviews*, 34(7), 1015–1022. <http://dx.doi.org/10.1016/j.neubiorev.2009.12.006>
- Scerbo, M. W., Freeman, F. G., & Mikulka, P. J. (2003). A brain-based system for adaptive automation. *Theoretical Issues in Ergonomics Science*, 4(1–2), 200–219.
- Schack, B., Klimesch, W., & Sauseng, P. (2005). Phase synchronization between theta and upper alpha oscillations in a working memory task. *International Journal of Psychophysiology*, 57(2), 105–114. <http://dx.doi.org/10.1016/j.ijpsycho.2005.03.016>
- Shalk, G., McFarland, D. J., Hinterberger, T., Birbaumer, N., & Wolpaw, J. R. (2004). BCI2000: A general-purpose, brain-computer interface (BCI) system. *IEEE Transactions on Biomedical Engineering*, 51(6), 1034–1043.
- Sellers, E. W., Vaughan, T. M., & Wolpaw, J. R. (2010). A brain-computer interface for long-term independent home use. *Amyotrophic Lateral Sclerosis*, 11(5), 449–455.
- Sharbrough, F. W., Chatrian, G. -E., Lesser, R. P., Luders, H., Nuwer, M., & Picton, T. W. (1991). American electroencephalographic society guidelines for standard electrode position nomenclature. *Journal of Clinical Neurophysiology*, 8, 200–202.
- Sirevaag, E., Kramer, A., Wickens, C., Reisweber, M., Strayer, D., & Grenell, J. (1993). Assessment of pilot performance and mental workload in rotary wing aircraft. *Ergonomics*, 36(9), 1121–1140.
- Smets, E., Garssen, B., Bonke, B., & Dehaes, J. (1995). The Multidimensional Fatigue Inventory (mfi) psychometric qualities of an instrument to assess fatigue. *Journal of Psychosomatic Research*, 39(3), 315–325.
- Smith, M. E., Gevins, A., Brown, H., Karnik, A., & Du, R. (2001). Monitoring task loading with multivariate EEG measures during complex forms of human-computer interaction. *Human Factors*, 43(3), 366–380. <http://dx.doi.org/10.1518/001872001775898287>
- Spencer, K. M., & Polich, J. (1999). Poststimulus EEG spectral analysis and P300: Attention, task, and probability. *Psychophysiology*, 36(2), 220–232. <http://dx.doi.org/10.1111/1469-8986.3620220>
- Stampi, C., Stone, P., & Michimori, A. (1995). A new quantitative method for assessing sleepiness: The alpha attenuation test. *Work & Stress*, 9(2–3), 368–376.
- Torsvall, L., & Åkerstedt, T. (1987). Sleepiness on the job: Continuously measured EEG changes in train drivers. *Electroencephalography and Clinical Neurophysiology*, 66(6), 502–511. [http://dx.doi.org/10.1016/0013-4694\(87\)90096-4](http://dx.doi.org/10.1016/0013-4694(87)90096-4)
- Trejo, L. J., Kochavi, R., Kubitz, K., Montgomery, L. D., Rosipal, R., & Matthews, B. (2005). J. A. Caldwell, & N. J. Wesensten (Eds.), *Biomonitoring for physiological and cognitive performance during military operations. Proceedings of the SPIE* (vol. 5797) (pp. 105–115). <http://dx.doi.org/10.1117/12.604286>
- Ullsperger, P., Freude, G., & Erdmann, U. (2001). Auditory probe sensitivity to mental workload changes – An event-related potential study. *International Journal of Psychophysiology*, 40(3), 201–209.
- Van Erp, J. B. F., Veltman, H. J. A., & Grootjen, M. (2010). Brain-Based Indices for User System Symbiosis. In D. S. Tan, & A. Nijholt (Eds.), *Brain-computer interfaces*. London: Springer.
- Wickens, C. D., Isreal, J. B., & Donchin, E. (1977). The event-related cortical potential as an index of task workload. In A. S. Neal, & R. F. Palasek (Eds.), *Proceedings of the human factors society 21st annual meeting*. Santa Monica, CA: Human Factors Society.
- Wolpaw, J. R., Ramoser, H., McFarland, D. J., & Pfurtscheller, G. (1998). EEG-based communication: Improved accuracy by response verification. *IEEE Transactions on Rehabilitation Engineering: A Publication of the IEEE Engineering in Medicine and Biology Society*, 6(3), 326–333. <http://dx.doi.org/10.1109/86.712231>
- Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002). Brain-computer interfaces for communication and control. *Clinical Neurophysiology*, 113(6), 767–791.
- Yamamoto, S., & Matsuoka, S. (1990). Topographic EEG study of visual display terminal (VDT) performance with special reference to frontal midline theta waves. *Brain Topography*, 2(4), 257–267.
- Yordanova, J., Kolev, V., & Polich, J. (2001). P300 and alpha event-related desynchronization (ERD). *Psychophysiology*, 38(1), 143–152. <http://dx.doi.org/10.1017/S0048577201990079>
- Zickler, C., Riccio, A., Leotta, F., Hillian-Tress, S., Halder, S., Holz, E., et al. (2011). A brain-computer interface as input channel for a standard assistive technology software. *Clinical EEG and Neuroscience*, 42(4), 236–244.
- Zickler, C., Halder, S., Kleih, S. C., Herbert, C., & Kübler, A. (2013). Brain Painting: Usability testing according to the user-centered design in end users with severe motor paralysis. *Artificial Intelligence in Medicine*, 59(2), 99–110.