Motivation modulates the P300 amplitude during brain–computer interface use

S.C. Kleih\textsuperscript{a,*}, F. Nijboer\textsuperscript{a}, S. Halder\textsuperscript{a}, A. Kübler\textsuperscript{a,b}

\textsuperscript{a}Institute of Medical Psychology and Behavioural Neurobiology, University of Tübingen, Tübingen, Germany
\textsuperscript{b}Department of Psychology I, Biological Psychology, Clinical Psychology and Psychotherapy, University of Würzburg, Würzburg, Germany

\textbf{ARTICLE INFO}

\textbf{Article history:}
Accepted 29 January 2010
Available online xxxx

\textbf{Keywords:}
EEG
ERP
Motivation
BCI
P300
Visual

\textbf{ABSTRACT}

\textbf{Objective:} This study examined the effect of motivation as a possible psychological influencing variable on P300 amplitude and performance in a brain–computer interface (BCI) controlled by event-related potentials (ERP).

\textbf{Methods:} Participants were instructed to copy spell a sentence by attending to cells of a randomly flashing $7 \times 7$ matrix. Motivation was manipulated by monetary reward. In two experimental groups participants received 25 ($N = 11$) or 50 ($N = 11$) Euro cent for each correctly selected character; the control group ($N = 11$) was not rewarded. BCI performance was defined as the overall percentage of correctly selected characters (correct response rate = CRR).

\textbf{Results:} Participants performed at an average of 99%. At electrode location Cz the P300 amplitude was positively correlated to self-rated motivation. The P300 amplitude of the most motivated participants was significantly higher than that of the least motivated participants. Highly motivated participants were able to communicate correctly faster with the ERP-BCI than less motivated participants.

\textbf{Conclusions:} Motivation modulates the P300 amplitude in an ERP-BCI.

\textbf{Significance:} Motivation may contribute to variance in BCI performance and should be monitored in BCI settings.

© 2010 International Federation of Clinical Neurophysiology. Published by Elsevier Ireland Ltd. All rights reserved.

1. Introduction

Brain–computer interfaces (BCI) allow users a motor activity independent interaction and communication with the environment on the basis of classifying and translating neurophysiological signals into commands for applications such as a spelling program, wheelchair or prosthesis (Kübler et al., 2001a; Birbaumer et al., 2008; Donoghue, 2008). Therefore, BCIs are a valuable tool for patients who are in the locked-in state, meaning they are conscious and can still control few muscles but are otherwise paralyzed (Laureys et al., 2005). Since amyotrophic lateral sclerosis (ALS) is a neurodegenerative disease which can lead to the locked-in state (Mitchell and Borasio, 2007), ALS patients constitute a potential BCI user population. It has been shown repeatedly that locked-in ALS patients are able to gain control over a BCI by regulating the amplitude of slow cortical potentials (SCPs) or sensorimotor rhythms (SMR) or by using event-related potentials (ERP) of the EEG (Birbaumer et al., 1999; Kübler et al., 2001a, 2005; Neuper et al., 2003; Piccione et al., 2006; Hoffmann et al., 2008; Nijboer et al., 2008a).

One advantage of ERP use for BCI control is that it does not require neurofeedback training because ERPs are elicited autonomously as a response of the brain to external stimulation (Pritchard, 1981; Farwell and Donchin, 1988; Sellers and Donchin, 2006; Nijboer et al., 2008a). The typical ERP-BCI setup uses an oddball-paradigm (Sutton et al., 1965; Fabiani et al., 1987) which consists of a sequence of standard irrelevant stimuli with interspersed rare deviant but relevant stimuli (the oddballs). An oddball is typically followed by a P300 event-related potential and consists of a positive deflection in the ERP occurring over the central-parietal region (Cz, Pz electrodes) within 200–700 ms after stimulus onset (Fabiani et al., 1987). The use of the P300 within a BCI for communication was first introduced by Farwell and Donchin (1988). The authors instructed healthy participants to spell words on a computer screen by selecting characters from a character matrix (see Fig. 1 for a $7 \times 7$ spelling matrix). The matrix contained 36 characters in 6 rows and 6 columns. Rows and columns of the matrix flashed randomly and the participant focused attention on the letter to be selected. In one sequence of 12 flashes (6 rows, 6 columns) the target character flashed only twice (1 column, 1 row) and thus, constituted an oddball and elicited a P300. To determine the target character stepwise discriminant analysis (SWDA) was
used and offline analysis revealed 95% accuracy with an average time of 26 s per selection.

The clinical relevance of the ERP-BCI was stressed repeatedly (Sellers and Donchin, 2006; Hoffmann et al., 2008; Nijboer et al., 2008a). Nijboer and colleagues (2008a) demonstrated the successful use of the ERP-BCI for communication in severely paralyzed ALS patients. P300 amplitude and latency remained stable over a time span of 40 weeks which suggests that there is no habituation effect for the P300 (Nijboer et al., 2008a). Thus, ERP-BCIs could offer a long-term communication solution for ALS patients.

Even though most, but not all healthy individuals react to stimulation with the 6 × 6 matrix oddball-paradigm with a P300, and are thus, able to control an ERP-BCI with high accuracy, paralyzed patients, present with a higher variability in correct response rates. For example, Nijboer and colleagues (2008a) investigated ERP-BCI use in 6 ALS patients. After 10 training sessions with the ERP-BCI the patients were able to spell between 1.5 and 4.1 characters per minute (4.8/19.2 bits/min, respectively). In the study by Hoffmann and colleagues (2008) the information transfer rates (ITRs) per minute for disabled people ranged between 10 and 25. Sellers and Donchin (2006) found an ITR of .43–1.80 bits per selection. Even though these ITRs cannot directly be compared due to different methods used, they still reveal a between-subjects variability in performance about which little is known.

As suggested by Kübler and colleagues (2001a) inter-individual differences in BCI performance might be ascribed to the influence of psychological variables. Johnson’s triarchic model of the P300 amplitude (1986) explains how psychological variables influence the P300 amplitude which might directly affect signal detection in an ERP-BCI. Johnson (1986) presented evidence that three influencing components subjective probability, stimulus meaning and information transmission determine the P300 amplitude. Within each of these three components there were sub-dimensions of components (for further details: Johnson, 1986) of which the attention as part of the information transmission and stimulus value as part of the stimulus meaning are especially interesting with respect to psychological impact factors on the P300 amplitude. Motivation being a psychological variable will, if manipulated with a reward for good performance, increase attention and the value of a stimulus. Therefore, we would expect a higher P300 amplitude if participants were motivated. Studies indicate that motivation increases the P300 (Hömberg et al., 1981; Begleiter et al., 1983; Carillo de la Pena and Cadaveira, 2000; Goldstein et al., 2006). Carillo de la Pena and Cadaveira (2000) for instance, manipulated motivation by informing participants that their results in an oddball-paradigm would be contrasted with the results of other peers; the control group was not challenged. The instructions lead to a higher P300 amplitude in the experimental group as compared to the control group. Goldstein and colleagues chose a Go/NoGo task to investigate motivation effects and used money for motivation manipulation. The participants were offered either 0 or 1 or 45 cents for correct responses, i.e. to press a button for the Go stimuli and to withhold presses for the NoGo stimuli. Additionally the participants had to rate their interest in and excitement about the task on a seven point Likert scale. The highest P300 amplitude was found in the 45 cents condition in which also the highest ratings of interest and excitement were reported.

But the quality of a monetary reward as an incentive is highly debated. Even though a negative effect of extrinsic rewards like money on intrinsic motivation (=motivation of doing something out of inherent interest, Heckhausen and Heckhausen, 2008) is suggested by Deci and colleagues (1999) there is also evidence that extrinsic rewards do have a positive effect on motivation (Salancik, 1975; Cameron and Pierce, 1994). Following the successful manipulation of the P300 with a monetary reward by Goldstein and colleagues (2006) we assumed that money can be used successfully as a motivator.

Despite the evidence of the effect of psychological variables such as motivation on the P300 amplitude in an oddball-paradigm little effort has been made to quantify and verify such possible psychological effects in BCI use. Leeb and colleagues (2007) showed that healthy participants who were highly motivated could move around more accurately in a three-dimensional virtual reality environment than participants who were less motivated. However, motivation was only assessed qualitatively by asking the participants to verbally report their motivation. Nijboer and colleagues (2008b) quantitatively measured motivation and mood of healthy participants (N = 16) before each training session with a BCI requiring regulation of the SMR amplitude. Within three consecutive daily training sessions participants were presented with visual or auditory feedback (Nijboer et al., 2008b). Four motivational factors were distinguished according to the cognitive-motivational model introduced by Vollmeyer and Rheinberg (1998). The first factor was referred to as mastery confidence indicating the belief in successfully mastering a task. The second factor incompetence fear indicated different levels of anxiety about failing the task. Third, participants were possibly challenged in various degrees by the task, and finally, the task might or might not have evoked the interest of the participant. Higher ratings of mood and mastery confidence were related to better SMR self-regulation with visual feedback, whereas higher incompetence fear was related to worse self-regulation. On the contrary, with auditory feedback higher scores of incompetence fear were related to better self-regulation. The authors suggested that when performance is high at the beginning of the training (as found with visual feedback), the incompetence fear in further sessions or loss of interest may hamper performance and further learning. However, when performance is low at initial training (as found with auditory feedback) incompetence fear might boost performance (Nijboer et al., 2008b). As both, the study by Leeb and colleagues (2007) and Nijboer and colleagues (2008a,b) did not systematically manipulate motivation and used a BCI which requires regulation of one specific EEG component; no conclusion can be derived for an ERP-BCI. Thus, there is evidence that motivation influences the P300 amplitude in an oddball-paradigm (e.g., Carillo de la Pena and Cadaveira, 2000), but nothing is known about how motivation influences performance in an ERP-BCI.

The goal of this study was to quantitatively investigate the effect of motivation on ERP-BCI performance with a 7 × 7 spelling matrix comprising 7 rows and 7 columns in which the letters of the alphabet and several punctuation marks are arranged from left to right and top to bottom. Above the matrix are two lines. In the upper line the text to copy appears (here: BRAIN), in the lower line the feedback is presented to the user (here: the first three letters BRA have been copied).
matrix. We hypothesized firstly, that motivation to spell with the ERP-BCI can be manipulated by monetary reward; secondly, that the P300 amplitude in the ERP-BCI is related to motivation and finally, that communication speed and performance would increase with higher motivation.

2. Methods

2.1. Participants

Thirty-three healthy students from the University of Tübingen (N = 4 male, N = 29 female), all naïve to BCI training, participated in the study. None of the participants reported a history of psychiatric or neurologic symptoms. Mean age was 23.4 years (SD ± 4.24, range = 19–42 years). The participants received either 8 Euro or 1 research credit1 per hour. Participants gave informed consent to the study which had been reviewed and approved by the Ethical Review Board of the Medical Faculty, University of Tübingen.

2.2. Design

A one factorial between subjects design was chosen comprising three groups: two experimental groups and one control group to which the participants were randomly assigned. In the experimental groups participants received 25 (25ct), or 50 (50ct) Euro cent for every letter that was correctly spelled with the BCI; the control group (0ct) could not earn any extra money. The rewards for accurate spelling were given additionally to the 8 Euro or to 1 research credit. The dependent variables were motivation and mood (assessed by questionnaires described in Section 2.4; we assessed mood as it may influence executive function and attention) and the P300 amplitude measured at electrode positions Cz and Pz. Further dependent variables were the number of flashes needed for correct letter selection and correct response rate (CRR = percentage of correctly selected letters).

2.3. Data acquisition

Data collection and experimental design were controlled by the BCI2000 (Schalk et al., 2004). The EEG was measured using an electrode cap (electro-cap International) with 16 Ag/AgCl electrodes located at positions F3, Fz, F4, T7, C3, Cz, C4, T8, C3p, Cp4, P3, Pz, P4, PO7, PO8 and Oz following the international 10–20 standard system (Sharbrough et al., 1991) referenced to the right and grounded to the left mastoid. Data were filtered using a high pass of 0.1 Hz and a low pass of 30 Hz. The EEG signals were amplified using a g-tec USB amplifier. Impedances were kept below 5 kΩ and the sampling rate was 256 Hz. Data were processed and stored with BCI2000 using an IBM ThinkPad (Pentium 4 M 1.6 GHz, 512 MB RAM, Windows XP). After acquisition, the data were visually inspected for artifacts and baseline corrected.

2.4. Questionnaires

To measure mood and motivation three questionnaires were used. These were the Questionnaire for Current Motivation (QCM) for BCI (Nijboer et al., 2008b) which is a slightly changed version of the original QCM (Rheinberg et al., 2001) to account for using a BCI instead of making predictions about performance and success in cognitive learning tasks. The QCM-BCI comprises 18 items divided into four subscales incompetence fear, mastery confidence, interest and challenge which have to be rated on a seven point Likert scale. To assess mood we used a subscale of the questionnaire “Skalen zur Erfassung der Lebensqualität” (SEL, engl.: “Scales for the assessment of quality of life”, Averbeck et al., 1997) which comprises a total of 10 items to be answered on a five point Likert scale. Finally, we assessed motivation and mood on a visual analogue scale (VAS) which ranged from 1 to 10 on a 10 cm long horizontal line (1 = extremely unmotivated/ extremely bad mood and 10 = extremely motivated/ extremely good mood). Participants indicated their motivation and mood with regards to the BCI task by marking one position on the line that best represented their subjective motivation and mood.

2.5. Procedure and stimuli

First participants completed the QCM-BCI and the SEL (see Section 2.4). For the ERP-BCI session each participant sat 1 m from a 43 cm video screen (60 Hz refresh rate) and viewed the matrix display. The 49 cells of the 7 × 7 matrix (7 rows and 7 columns, see Fig. 1) contained the German alphabet, the numbers 1–9 and several punctuation marks. The participants completed 9 runs of copy-spelling (Kübler et al., 2001b), each run consisting of one word with up to 6 characters. The word-to-copy appeared above the matrix. Directly next to the word-to-copy the first letter of the word-to-copy (target character) appeared in parenthesis. The participant’s task was to pay attention to the target character and to count silently the number of times it intensified. For each target character 15 sequences of flashes (1 trial) were presented. Each sequence comprised 14 randomly distributed flashes (each row and column flashed once per sequence). Thus, each target cell flashed 30 times per trial which equalled a probability for target flashes of 14.3% (210 flashes total and 30 target flashes). Each flash lasted 62.5 ms followed by an inter stimulus interval of 125 ms. Thus, one sequence lasted for 2.63 s and 39.38 s were needed to select one character. After 1 trial of 15 sequences the matrix stopped flashing for 2.5 s in which the selected character appeared in the line below the word-to-copy (feedback) and the participant had time to locate the next target character to be spelled in the matrix. This was repeated until the entire word was copy-spelled. Participants completed two calibration runs in which the words “BRAIN” and “POWER” had to be spelled and no feedback was provided. The calibration runs were needed to derive classification coefficients for the following 7 experimental copy-spelling runs. Before copy-spelling with the ERP-BCI participants were informed to which of the three groups (0ct, 25ct or 50ct) they were randomly assigned. The experimental design was explained to the subjects to enhance the motivational effect. The sentence used for copy-spelling was “XAVER SAH 34 MÜTZENQUALLEN IM ZOO!” (Engl. “Xaver saw 34 cap jellyfish in the zoo!”). This sentence included at least one character of each column and row of the matrix. The information about the amount of money earned and the CRR appeared on the screen after every run. At the end of the BCI session, participants were presented with the VAS to assess mood and motivation.

2.6. Classification method

Stepwise linear discriminant analysis (SWLDA) was applied for online and offline analysis due to its reliability for ERP classification (Donchin et al., 2000; Sellers and Donchin, 2006; Krusienski et al., 2008; Furdea et al., 2009) and its low computational requirements (Lotte et al., 2007). SWLDA separates the data into two classes (target and non target signals) both obtaining equal covariance matrices. It calculates a linear equation which depends on the spatiotemporal features of the signals and separates the data classes as a hyperplane. This maximizes the distance between the means of the two classes while also minimizing the variance within one
class (Lotte et al., 2007). Input features that predict the target label statistically significant are added to the linear equation to explain the largest amount of unique variance. New features are added if they account for a significant amount which results in an improvement of the model. Through backward stepwise regression features are removed that no longer meet the criterion to remain in the model. In this study we predefined the number of features included to 60 which allowed us an exact classification of the data and was still very effective. Details of the SWLDA are described in Krusienski and colleagues (2006), Sellers and Donchin (2006) and Sellers and colleagues (2006).

2.7. Communication speed

The most common measure of communication speed in contemporary BCI publications is based on a formula by Pierce (1980) and was originally suggested by Wolpaw and colleagues (2000, see formula (1)). The bits per trial (B) were calculated by (1) with $N$ representing the number of possible selections in the matrix and $P$ the accuracy of the participant.

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

Formula (1) is a valid measure for BCIs in which all targets are selected with the same accuracy and all errors are made with the same probability. In particular for P300 speller BCIs this is not the case, because cells around the target cell, and in particular of the same row or column, are incorrectly selected with higher probabilities than other cells (see Fig. 2). To account for unequal probability of erroneous selections we followed the suggestion of Schlögl and colleagues (2007) and calculated the ITR based on the formula for mutual information (MI) derived in Nykopp (2001) for $M$ classes:

$$I(X; Y) = \sum_{i=1}^{M} \sum_{j=1}^{M} p(x_i) \cdot p(y_j|x_i) \cdot \log_2\left(\frac{p(y_j|x_i)}{p(y_j)}\right)$$

$$- \sum_{j=1}^{M} p(y_j) \cdot \log_2\left(p(y_j)\right)$$

with $p(x_i)$ being the a priori probability of the $i$th class defined as $p(x_i)/N$ and $p(y_j|x_i)$ being the conditional probability of selecting class $y_j$ when class $x_i$ is selected.

The final result is also given in bits/min and therefore divided by the time needed for one selection (see Section 3.5). Even though in this study a bias concerning the ITR might have occurred due to not using every field of the matrix when spelling, we believe that the equal distribution of characters throughout the matrix prevented such a bias.

We were interested in the ITR to reach 100%, 90% and 70% correctness in spelling. We chose 70% as the minimally acceptable correctness following earlier recommendations (Kübler et al., 2001a) and 90% and 100% because we wanted to investigate how much time (flashes) could possibly be saved by lowering the CRR.

2.8. Data analysis

Statistical analysis was performed with SPSS 11.0. To test the first hypothesis that motivation can be manipulated by extrinsic rewards we calculated a Multivariate Analysis of Variance (MANOVA) with reward as independent and the four subscales of the QCM as dependent variables. We considered appropriate a MANOVA
because the four subscales of the QCM measure different aspects of the same psychological construct “motivation” and thus, are mutually dependent. To test the effect of motivation on the P300 amplitude data were pre-processed by Brain Vision Analyzer 1.05* Software. The EEG data were visually inspected for artifacts and all trials were included. Because N = 26 participants showed the highest P300 amplitude under either Cz or Pz (the other 7 participants had a maximum P300 amplitude at either Cp3, C4 or Oz but showed very high P300 amplitudes also at Pz or Cz) we decided to restrict our analysis on electrodes Cz and Pz. The P300 was defined as the maximum positive peak occurring between 200 and 700 ms after stimulus onset. To investigate the effect of motivation on the P300 amplitude, a MANOVA was calculated with group as fixed variable and P300 amplitude on locations Pz and Cz as dependent variables. We chose MANOVA because a change in the P300 amplitude at Cz could affect the P300 amplitude at Pz and vice versa. The third hypothesis stated that by increasing motivation, spelling speed could be enhanced. Communication speed was defined here as the number of flashes needed in offline analysis to achieve 70%, 90% or 100% CRR. This was calculated by deriving the individual weights on the calibration runs and applying these weights to the copy-spelling runs. The hypothesis was tested using a MANOVA with group (0, 25 and 50ct) as independent and accuracy (70%, 90% and 100%) as dependent variables (offline analysis).

The level of significance was set to α = .05. All data were checked for Gaussian distribution and homogeneity of variances. For all data not meeting the requirements for parametric analysis and testing non-parametric procedures were used (Kruskal–Wallis H, Mann–Whitney U-test, and Spearman’s ρ).

3. Results

3.1. General findings

We did not find any differences between the three groups according to demographic variables (age: \( F(2,30) = .98, p = .39 \); gender: \( F(2,30) = .39, p = .68 \), MANOVA) or the basic allowance method (money vs. research credits: \( F(10,23) = .38, p = .93 \), MANOVA). The VAS for motivation revealed a ceiling effect \( M_{25ct} = 8.74 \) (SD ± .92); \( M_{50ct} = 8.98 \) (SD ± .66) and \( M_{0ct} = 9.44 \) (SD ± .75). Furthermore, all groups were extremely successful in using the ERP-BCI for spelling and achieved an online average CRR of 99.26% (SD ± .30); \( M_{25ct} = 99.05 \) (SD ± .26); \( M_{50ct} = 99.42 \) (SD ± .13) and \( M_{0ct} = 99.30 \) (SD ± .147). The overall online CRR ranged between 93% and 100%. P300 amplitudes ranged between 2.68 and 11.00 μV (\( M = 5.90, SD = 1.90 \)) at Cz and between 2.37 and 9.68 μV at Pz (\( M = 5.29, SD = 1.77 \)). The average latencies were \( 305.44 \) ms at Cz (SD = 85.16, range: 200.22–445.31 ms) and \( 361.03 \) at Pz (SD = 90.58, range: 203.13–464.84 ms).

3.2. Effect of monetary reward on motivation and mood

Our first hypothesis was that motivation can be enhanced and mood can be positively influenced by monetary rewards. The money earned served as the independent variable the values in the QCM-BCI and the VAS motivation and mood as dependent variables. The MANOVA did not show any effects of monetary reward on incompetence fear, mastery confidence or challenge (\( F(6,58) = 1.73, p = .13 \)) nor did the Kruskal–Wallis H-test for the subscale of interest (\( \chi^2(2) = 2.04, p = .36 \)). Likewise, the Kruskal–Wallis H-test did not reveal significant differences between groups in the VAS motivation (\( \chi^2(2) = 4.48, p = .11 \)). Mood did not improve with monetary reward (Kruskal–Wallis H-test: SEL: \( \chi^2(2) = 4.10, p = .13 \); VAS mood: \( \chi^2(2) = .88, p = .64 \)).

3.3. Effect of monetary reward on P300 amplitude

The second hypothesis predicted that higher rewards would result in an increased P300 amplitude at locations Cz and Pz. No significant differences between groups were found (\( F(4,60) = 1.65, p = .17 \)). To ensure that no P300 amplitude differences between the groups existed prior to copy-spelling, a MANOVA with group as independent variable was calculated for the P300 amplitude during calibration runs. No group differences were found (\( F(4,60) = 1.26, p = .29 \)). To test whether the information about to which group participants were allocated affected the P300 amplitude between calibration and copy-spelling we used ANOVA for repeated measures and found significant changes in the P300 amplitudes for Pz (\( F(1,30) = 9.39, p < .01 \)). Post-hoc t-tests for dependent samples with calibration and experimental copy-spelling runs as independent variables for all three groups revealed a significant decrease of the P300 amplitude at Pz for the \( 0ct (t_{10} = -2.93, p < .05) \) and the \( 25ct \) group (\( t_{10} = -2.24, p < .05 \)) but not for the \( 50ct \) group (\( t_{10} = 67, p = .52, \) see Fig. 3).

To investigate whether there was a reward independent effect of motivation on P300 amplitude we pooled all participants and tested if there were correlations (Spearman) between motivation as measured with the VAS motivation and the four subcales of
the QCM-BCI and the P300 amplitude at Cz and Pz. We found a high positive correlation between self-reported motivation and P300 amplitude at Cz ($\rho = .50$, $p < .01$, Bonferroni corrected). Linear regression revealed that the VAS motivation significantly predicted the P300 amplitudes on Cz ($F_{1,31} = 8.50$, $p < .01$, $R^2 = .214$, see Fig. 4) and explained 21.40% of the variance. To further elucidate the positive correlation between motivation and P300 amplitude independent of monetary reward, we compared the first and fourth quartile of VAS motivation values (first: VAS $\leq 8.40$, N = 9; fourth: VAS $\geq 9.80$, N = 9). t-Test or Mann–Whitney U-test revealed significant differences between the groups at Cz ($t_{16} = -3.77$, $p < .01$, see Fig. 5a) and Pz ($Z = -2.34$, $p < .05$, see Fig. 5b) with the P300 amplitude in the lower quartile being significantly lower than the P300 amplitude in the upper quartile. To exclude that the rating of the VAS motivation was influenced by the CRR, we correlated the VAS motivation and the CRR. With $r = .22$ and $p = .23$ the correlation was low and non-significant. We did the same calculations for P300 latency and found no effects of motivation.

3.4. Effect of monetary reward on performance and communication speed

We used CRR as a measure of performance. Twenty-nine participants achieved a CRR of 100% and average CRR was 99% with 15 sequences (=30 target flashes per letter selection). Four participants (one participant of the control group, two of the 25ct group and one of the 50ct group) achieved a CRR of 97%. Therefore, no differences in CRR as a function of monetary reward were found between the three groups when comparing online performance with 15 sequences. Participants achieved an average ITR of 6.33 bits/min. The entire sample of N = 33 achieved offline accuracies of 70% with 2.0 sequences (SD $\pm .79$; 40.98 bits/min); of 90% with 3.3 sequences (SD $\pm 1.59$; 24.63 bits/min) and of 100% with 6.5 sequences (SD $\pm 3.84$; 14.44 bits/min). Repeated measures ANOVA revealed a significant difference in ITR rates ($F_{2,31} = 85.97$, $p < .001$). Post-hoc pair wise comparisons were in the expected direction (70% < 90% < 100%, all $p < .01$).

3.5. Effect of motivation on performance and communication speed independent of reward

When comparing the two quartiles of the participants being motivated the least and the most we found a trend ($Z = -1.82$, $p = .08$) for the upper motivation quartile to achieve 100% correctness with an average of 4 sequences (SD $\pm 1.94$) and of 7.67 sequences in the lowest quartile (SD $\pm 4.62$) corresponding to 22.66

---

**Fig. 4.** Correlation between P300 amplitude at Cz and values of the VAS motivation. The amplitude at Cz is plotted against the values of the VAS motivation, black line: linear regression.

**Fig. 5.** (a and b) Event-related potential P300 at Cz (a) and Pz (b) for the least and most motivated subjects (N = 9 per group).

**Fig. 6.** Accuracy reached by participants after reclassification on the basis of a single trial.
and 13.38 bits/min. We compared offline the participants being least and most motivated according to the sequences needed to spell for 70% or 90% correct. No significant differences between the groups were found. We also calculated the single trial accuracy by reclassifying the data using only a single sequence per trial (=2 target flashes per letter selection). This resulted in normally distributed data (see Fig. 6) and therefore successfully eliminated the ceiling effect. We then found a significant correlation between the single trial accuracy and the motivation subscale of challenge (Spearman's $\rho = .40, p < .05$).

However, when comparing single trial accuracies of the least motivated ($M = 51.47, SD = 12.87$) and the most motivated participants ($M = 57.83, SD = 13.71$) no significant differences between groups were found ($t_{(16)} = -1.02, p = .33$).

### 4. Discussion

The aim of this study was to investigate the effect of motivation on ERP-BCI performance. We firstly hypothesized that motivation to spell with the ERP-BCI could be manipulated by monetary reward, secondly that the P300 amplitude would vary in relation to motivation, and finally that performance and communication speed could be increased by higher motivation. We could not confirm our first hypothesis as self-reported motivation was not affected by monetary reward. Motivation was high in all groups independent of reward and performance was at ceiling with 99% average accuracy also independent of reward and self-reported motivation. It could be argued that the high performance caused the high motivation values since the visual analogue scale was handed out to participants after the copy-spelling runs. But the QCM-BCI was used before the participants received any feedback about their performance. In the subscales mastery confidence and interest all participants showed very high values even though all of them were BCI novices. Nevertheless, the provided feedback of results could have contributed to keeping up the high motivation across all groups as all participants were continuously positively reinforced by correct letter selections.

After the session, participants reported verbally that they were highly interested in this BCI study (qualitative statement) which supports the high values for interest in the QCM-BCI subscale even though the QCM-BCI values were independent from monetary reward. Thus, interest could have been an important motivator for participation. Seventeen participants stated that they would have participated in the study even without being rewarded. Schiefele (2001) emphasized interest as being a prerequisite for motivation. He postulated two components of interest: the intrinsic feeling related and the value related. The feeling related component refers to emotional states, like personal involvement or excitement experienced by the task. The value related component refers to how important the task is considered by an individual for satisfaction of his or her personal values, like enhancement of self esteem. In our study both of these components could have had impact, because the BCI task was feeling related as it was new and exciting. All the participants reported to have made an effort to do their best to spell the given sentence correctly, which could be interpreted as value related. According to Ainley (2006) being interested in learning tasks leads to such high a motivation that no other rewards are needed. Therefore, the overall high motivation and performance and the ineffectiveness of monetary reward in this study may be ascribed to high interest. This assumption is further confirmed by Johnson’s triarchic model of the P300 amplitude (1986) according to which interest influences the component stimulus meaning such that stimuli within a task of interest are valuable and capture attention which in turn could have contributed to a higher P300 amplitude.

Finally, it has to be noted that participants were students of psychology and stated that they found the BCI experiment absolutely exciting compared to other studies in which they often have to fill out questionnaires only. Thus, the group of participants may not have been representative for real world BCI users.

Another reason why the manipulation of motivation by using a monetary reward did not influence self-reported motivation might have been the between-subjects study design. Previous research suggesting monetary reward as positive reinforcer to increase motivation (Hömberg et al., 1981; Begleiter et al., 1983; Yeung and Sanfey, 2004; Goldstein et al., 2006) used a within-subjects design. The within-subjects design might have caused the different monetary rewards to be more salient to the participants compared to the between-participants design used in the current study. Another reason might have been the different way of rewarding the participants in comparison to preceding studies (Begleiter et al., 1983; Yeung and Sanfey, 2004; Sato et al., 2005). In these studies participants were punished for reacting wrongly or not reacting at all which was not implemented in our study. The most straightforward explanation, however, is the high interest of all subjects independent of reward which lead to a low variance of self-reported motivation.

Our second hypothesis which suggested that reward and motivation would affect the P300 amplitude could be confirmed threefold. Firstly, after the calibration runs, participants were informed to which group they were assigned. In the subsequent copy-spelling runs P300 amplitude was significantly lower in the control group and in the group with 25 cent reward. In the group with the highest reward no such difference was found. Secondly, for the copy-spelling runs, we found a high positive correlation between P300 amplitude at Cz and self-reported motivation (VAS). The linear regression showed that the values in the visual analogue scale of motivation significantly predicted the P300 amplitudes on Cz and that 21.40% of the variance could be explained by the VAS motivation. Furthermore the additional single trial analysis revealed a correlation between single trial accuracy and the QCM-BCI subscale of challenge, a relation that was covered by the ceiling effect during online performance. Finally, after a redistribution of participants into quartiles according to the VAS motivation and a subsequent comparison between the lowest and the highest quartile, the least motivated subjects presented with a significantly lower P300 amplitude as compared to the most motivated subjects. These findings clearly illustrate the influence of motivation on the P300 amplitude as presumed by Kübler and colleagues (2001a). However, a lower P300 amplitude did not affect the classification results. The overall high performance which was found despite a motivation dependent reduction in the P300 amplitude confirmed that the SWLDA (Donchin et al., 2000; Sellers and Donchin, 2006; Krusienski et al., 2008) is a robust classification method for ERP-BCI. Concerning the P300 amplitude measurement it might be argued that the ISI interval was too short in this study and therefore might have caused overlapping ERPs that impaired the ERP measurements. There are two reasons why we chose this short ISIs and why we believe that the measurement was not impaired by overlapping ERPs: Firstly, it was proven that BCI that are based on relatively short ISIs yield high an accuracy despite possibly overlapping ERP effects. Sellers and his colleagues (2006) found higher classification accuracies for a visual speller when using short ISIs (175 ms) in comparison to longer ISIs (350 ms). Nijboer and coworkers (2008a,b) found an average online performance of 79% in ALS patients while using a 175 ms ISI, likewise Serby and her colleagues (2005) found an average accuracy of 79.5% with an ISI of 125 ms. We also avoided the possibility of constantly overlapping ERPs by ensuring that the target stimulus was highlighted no more than twice consecutively (probability 2.04%). In all other cases our ISI for the target stimulus was 312.5 ms or more which...
follows the suggestion of Martens and colleagues (2009) who advise to minimize the number of trials in which the target to target intervals are less than 200 ms to avoid overlapping ERPs. Secondly, short ISIs increase the information transfer rate, and high ITRs are desirable having in mind real world applications of BCI.

Finally, our third hypothesis stated that communication accuracy and speed would increase with motivation. As online performance with 15 sequences was at ceiling level we reanalyzed the data offline. We then found a trend that the highest motivated subjects would have needed fewer sequences (4) to achieve 100% CRR as compared to the lowest motivated participants who would have needed 7 sequences. But even with 4–7 sequences, CRR values were not normally distributed indicating that the ceiling effect was still present. Data were normally distributed when using 1 sequence only, but the trend could not be confirmed. For this reason, we could not conclusively confirm or reject our third hypothesis.

5. Conclusion

With the current study we were able to demonstrate that motivation affected the P300 amplitude within an ERP-BCI. Linear regression revealed that motivation accounted for 21.40% of P300 amplitude variance at Cz. When calculating single trial accuracy, we found a correlation between accuracy and the motivation subscale of challenge. Therefore, we conclude that if performance is not at ceiling due to higher task difficulty, as for example in BCI based on non-visual stimulation (Cincotti et al., 2007; Nijboer et al., 2008b; Furdea et al., 2009), or due to altered brain activation, as found in individuals with brain injury or neurological disease, motivation may become even more influential. To further elucidate this issue, it will be necessary to investigate the influence of motivation on BCI performance with the target patient groups as for example individuals with amyotrophic lateral sclerosis or patients in the locked-in state after stroke. Future studies also have to address the effect of motivation when regulation of a brain response, such as in motor imagery based BCIs, is required by participants with disease. We recommend assessing motivation quantitatively in BCI studies as it may explain some of the variance in BCI performance.

Acknowledgements

This work is supported by the European ICT Programme Project FP7-224631. 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

