A P300 Auditory Brain-Computer Interface based on Mental Repetition

Alessandro Marassi[1], Riccardo Budai[2], Luca Chittaro[3]

[1]INAF – Osservatorio Astronomico di Trieste

[2]Department of Neurology, Azienda Ospedaliero-Universitaria Santa Maria della Misericordia, Udine

[3]Human-Computer Interaction Lab,Department of Mathematics and Computer Science, University of Udine, Italy

Corresponding author:

Alessandro Marassi

INAF – Osservatorio Astronomico di Trieste, via Tiepolo 11, 34143 Trieste, Italy

Tel: +39-0403199343

Fax: +39-040210036

E-mail: marassi@oats.inaf.it

Keywords: Brain-Computer Interface, EEG, P300, Auditory, Oddball mental task, mental workload.

Highlights

- We evaluate an auditory brain-computer interface (BCI) protocol that allows users to mentally choose among 6 options
- We contrast two ways of using the BCI: passively counting the presented target auditory stimuli or simply mentally repeating them when they occur
- The mental repetition task generated an ERP response similar to the traditional counting task
- The mental repetition task generated a mental workload smaller than the traditional counting task

Abstract

Objective: The current study evaluates an auditory brain-computer interface (BCI) protocol that allows users to mentally choose among 6 options.

Methods: The protocol is based on an oddball P300 paradigm. To reduce mental workload, we introduce a change in the typical oddball paradigm task: instead of passively counting the presented target auditory stimuli, we ask participants to simply mentally repeat them when they occur.

In the study, ten healthy participants carried out two calibration sessions respectively with traditional mental count and with the proposed mental repetition and then three free item selection sessions using mental repetition. A comparison has been conducted between off-line count and mental repetition classification accuracies achieved by each participant during the calibration sessions. The mental workload difference between the count and repeat calibration sessions of each participant was evaluated by computing alpha (at Po8) and theta (at Fz) spectral power density (SPD) curves.

Results: With the proposed protocol we got an average on-line item-selection information transfer rate (ITR) over 2 epochs of 2.35 bits/min, and an average on-line accuracy over 10 epochs of 81.7%. Nine out of ten participants showed a higher mental workload with the traditional mental count. The repeat activity was preferred by 8 out of 10 participants. The comparison conducted between off-line count and mental repetition classification accuracies shows a slightly worse average behavior for the repeat protocol in the first 8 averaged epochs.

Conclusions: Although off-line classification based on mental count data got better results, the proposed auditory BCI protocol with mental repetition achieved an on-line performance similar to the traditional counting oddball paradigm task but with a lower mental workload.

Significance: The results obtained with healthy subjects suggest that the proposed protocol can be a simpler alternative to the mental count, with comparable performance and lower mental load.

Introduction

A brain-computer interface (BCI) can be defined as "a communication system that does not depend on the brain's normal output pathways of peripheral nerves and muscles"[1].

Several electrophysiological characteristics can be extracted from the human EEG to control a BCI system. A frequent choice is to exploit P300, i.e. an Event-Related Potential (ERP) component that appears as a positive deflection over central and parietal scalp areas with a latency of 200 to 700 ms after the presentation of a rare or salient stimulus [2]. The stimulus can be visual, auditory, or somatosensory. The typical ERP-BCI setup uses an oddball-paradigm [3],[4] which consists of a sequence of standard irrelevant stimuli with interspersed rare deviant but relevant stimuli (the oddballs). When the attended stimuli appear, the user has to perform some activity such as increment a mental counter in order to elicit (enhance) a P300 response.

Several visual BCIs based upon the P300 oddball-paradigm have been proposed in the last decades, e.g.[5],[6], [7],[8],[9]. Unfortunately, the efficacy of such solutions can become unacceptable in patients with serious oculomotor impairments [10],[11],[12]. This motivates research on paradigms that are not based on sense of sight [13],[14],[15],[16],[17],[18].

While the traditional oddball approach in auditory BCIs is based on having the user passively count the number of occurrences of the target auditory stimulus, in this paper we aim at reducing mental workload by having the user simply mentally repeating the target stimulus when it occurs. We also aim at using a small number of electrodes to acquire the EEG signals.

The paper is organized as follows. First, we introduce related work on P300 paradigms and EEG-based workload measurement. Second, we illustrate in detail the proposed auditory BCI and the user study we carried to evaluate its performance, and to compare the traditional mental count approach with the proposed mental repetition. Third, we present and discuss the results of the study.

P300 paradigms for BCIs

The P300 speller, first described by Farwell and Donchin [5], uses visual P300 responses to choose letters from a 6x6 matrix presented on a computer screen. P300 speller with its variants as well as other visual BCIs have been described and tested both with healthy participants [7] and disabled subjects [8],[9],[6], [7] demonstrating that individuals with severe paralysis can use a visual P300-based BCI, provided that they are able to control their eyes to gaze and focus the target.

Unfortunately, patients with locked-in syndrome (LIS) are in a state of almost complete paralysis with residual voluntary gaze direction control or eye blink control which may hinder their ability to use a visual P300 speller in the later stages of the disease. Treder and colleagues [11] state that extreme cases of oculomotor impairment may pose problems even for gaze-independent BCIs, underlining that some patients suffer from involuntary gaze drift which could damage BCI performance. Although eye movements are not necessarily required to operate a visual P300 BCI, involuntary eye movements may make it difficult to orient visual attention to a specific location[10].The study by Kaufmann et al. [12] seems also to question the effectiveness of current gaze-independent approaches to BCI, while indicating a tactile approach as the most accurate.

Patients with complete locked-in syndrome (CLIS) lose volitional control over all their muscles and are not able to use a visual ERP paradigm, motivating research on alternative ones (auditory or tactile). Moreover, some researchers [19],[20],[21] are trying to apply auditory BCI protocols in cases of Minimally Conscious State (MCS) patients, a condition of severely altered consciousness where there is minimal evidence of any form of awareness. In such cases, a BCI might bear the potential to employ a ‘‘yes–no’’ spelling device to confirm an MCS diagnosis and offer the hope of functional interactive communication.

Since visual BCI paradigms are not eligible for LIS patients affected by (severe) oculomotor impairment [11] and CLIS patients, alternative paradigms based on auditory stimuli were investigated to allow both healthy and disabled users to make binary decisions [14],[15], [13] or to operate auditory P300 spellers [16], [18], [17] [22] [23] [24]. Käthner’s study further aimed also at quantifying the differences in subjective workload between the auditory and the visual P300 spelling application.

The auditory BCI paradigms we surveyed above confirm the possibility of using auditory stimuli in a BCI system; nevertheless, auditory-only BCIs suffer from issues in achievable accuracy, ITR and ease of use which are lower than in the visual BCI protocols for healthy subjects and patients as well. Some of the above mentioned studies include visual stimuli in their protocols, which are unfortunately unusable by visually impaired and CLIS patients. Other auditory-only studies are very limited in subjects' possibility of decision which is only binary. Approaches that rely on spatially distributed sound sources are limited by the human ability to discriminate sound directions and by the mental workload such an activity presents.

Mental workload

The ability to directly detect mental over- and under-load in humans is an essential feature in human interfaces evaluation and in complex monitoring and control processes [25]–[27]. Workload measurements can be categorized into three classes [28]:

Subjective: subjective measurement of levels of workload based on the rankings or scales to measure the amount of workload a person is feeling; it can consists of just one question (e.g., “Please rate the amount of mental effort invested in the task”), and the responses range from “very low mental effort” to “very high mental effort” as in the widely used mental effort scale by Paas [29].

Performance Based : performance measurement of workload relies on examining the capacity of an individual by means of a primary or secondary task. By measuring how well a person performs the task with increasing workload, an estimate of mental workload can be determined [30].

Physiological: physiological measurement relies on evidence that increased mental demands lead to increased physical response from the body. This type of measures relates the continuous physical reactions of the body to the amount of mental workload a person is experiencing. Unlike subjective measures, it does not require a direct response from the individual, [31].

Research focuses on five main physiological areas to measure mental workload: *cardiac activity, respiratory activity, eye activity, speech* and *brain activity* [30].

To measure *brain activity, EEG* is the most frequent choice [32]. Other physiological techniques that are used in neuroscience such as functional magnetic resonance imaging (*fMRI*), positron emission tomography (*PET*) can be employed, but they are able to capture exclusively metabolic changes in determined brain areas; unlike fMRI and PET, EEG can noninvasively measure brain activity in authentic, real-world settings with high temporal resolution, enabling it to measure changes in mental activity on the millisecond scale. Therefore, EEG measurements are continuously reflective of a participant's cognitive states [33].

Spectral features derived from the EEG are used for measuring the workload [34] [34].

Alpha is, with the exception of irregular activity in the delta range and below, the dominant frequency in the human scalp EEG of adults. Alpha and theta respond in different and opposite ways to mental workload: theta synchronizes with increasing task demands, whereas alpha desynchronizes. If EEG power in a resting condition is compared with a test condition, alpha power decreases (desynchronizes) and theta power increases (synchronizes), as shown by a variety of studies [35]–[40].

Study Goal

The central goal of our research is twofold. First, we wanted to explore the feasibility of an auditory P300-based BCI protocol, which could be of interest in cases where the user is unable to take advantage of visual BCIs. In particular, we created a proof-of-concept auditory BCI protocol that requires only 8 EEG electrodes, and supports 6 possible choices, represented by auditory stimuli consisting of the spoken digits from 1 to 6. This focus on 6 digits (6-target stimuli detection) aims at enabling a range of selections which in principle could drive a “hex-o-spell”-like [41] interface.

Unlike other protocols which use more demanding mental tasks, such as spatial distributed sound direction or other target property mental discrimination, to elicit stronger ERP responses [22], [42], [43], our approach tries to reduce mental workload by introducing a change in the typical oddball paradigm counting task commonly employed in the literature [3], [44]–[46]: instead of passively counting how often the target auditory stimulus is presented, we ask users simply to mentally repeat it when it occurs.

The second goal was to evaluate the effects of the proposed auditory P300 task by comparing user’s mental workload and performance with mental repetition and with the traditional mental counting (see details in Data Analysis section).

In particular, we assess performance with the two mental activities in terms of real-time accuracy and Information Transfer Rate (ITR)(see ITR computation section).

Methods

Participants

Ten healthy subjects (6 female, 4 male, M=35.1 years, SD=7.95 years, range 26-54) participated in the study. Participants enrolled in the study on a voluntary basis, receiving no compensation, and were informed in detail about the nature of the study to obtain their informed consent.

Data Acquisition

The EEG was recorded with Ag/AgCl electrodes over 8 channels (Fz, Cz, Pz, Oz, P3, P4, PO7, PO8) based on the modified 10–20 system of the American Electroencephalographic Society [47]. Each channel was referenced to the right and grounded to the left mastoid. The signals from the electrodes were amplified using a Micromed 32-channel amplifier SAM 32FO FC1, sampled at 128 Hz. Impedances were kept below 5 kΩ.Stimulus presentation, data acquisition, processing, storage, and on-line display of the participants' EEG were handled with BCI2000 on a HP G62 Notebook PC.

Design and Procedure

In our auditory protocol, the participant was presented with audio stimuli (digital recordings of single spoken digits) delivered in random order with rare intersparsed attended stimuli to generate target-related oddball events.

Farwell and Donchin [5] examined the effects of ISI on signal to noise ratio and users' performance and found that a longer ISI provided a higher rate of communication for three of the four participants in their study. Considering Farwell and Donchin's results and the fact that auditory stimuli presentations require a larger amount of time than visual stimuli, in this study we increased stimulus presentation time and ISI respectively to 375 ms and 1250 ms. The duration of each audio stimulus was 300 ms on average, allocated in the 375 ms stimulus interval.

The experimental procedure was organized in two parts: a calibration session, also to acquire user data to be later used to train the classifier, and an experimental item-selection task. The calibration session lasted about 10 minutes and was organized into 6 subsequent temporal intervals (called *subsessions* in the following). At the beginning of each subsession, the participant was verbally given by the system a different target number chosen from the set {1,2,3,4,5,6}. During each subsession, all 6 numbers {1,2,3,4,5,6} were auditorily presented for 10 times (called *epochs* in the following), for a total amount of 60 number vocalizations per subsession. In each epoch, the 6 numbers were presented in random order. The participant's task was to listen to the presented numbers and mentally repeat the target number whenever it occurred. Therefore, the whole calibration session consisted of 360 vocalizations, of which 60 attended and 300 not attended.

An analogue calibration session with mental count was also performed to check whether participants found it easier than repetition, so that each participant performed one calibration session in the "repeat" mode and one calibration session in the "count" mode, and the order was counterbalanced. Then, the experimental task was carried out only in the "repeat mode" to prove the feasibility of a BCI based upon the new protocol.

The task was organized into three independent item-selection sessions of respectively 2, 5 and 10 epochs. During each epoch, the six spoken digits, organized in a six-stimulus oddball paradigm, were presented to the participant who had to attend to a predetermined sequence of designated targets, by simply mentally repeating them whenever they occurred, while ignoring all the other digits. The predetermined sequence of designated targets was organized as six quadruples of digits (1111, 2222, 3333, 4444, 5555, 6666) and was the same for the three item-selection sessions.

The study of different item-selection sessions consisting of different numbers of epochs was exploratory in nature to determine the smallest number of averaged P300 trials as the best compromise between efficiency and accuracy [5], [15].

Participants sat on a comfortable chair, wearing a pair of headphones connected to the PC audio output. Volume control was adjusted at the level that the participant judged comfortable, while (s)he heard the same auditory stimuli later used in the experiment. Then, we briefly explained the procedure (one system training and three item-selection sessions) and what the participant was expected to do during the sessions. We recommended participants not to move during the sessions in order to minimize artifacts, not to yawn or blink, and possibly avoiding eye motion or face/head muscle contraction. To confirm that participants had fully understood the procedure, we conducted preliminary tests during which they were allowed to see their own EEG signals on a display to let them visually appreciate effects of artifacts.

The two calibration sessions in count mode and repeat mode were then performed. The repeat mode calibration data (recorded in BCI2000 *.dat format) [48], [49] were then processed using the BCI2000 tool P300Classifier [48] to determine a set of optimal features to be later used in the analysis and in the on-line classifier during the experimental item-selection task. On the basis of the features and weights obtained by the analysis of the “repeat mode” calibration data, three item-selection sessions (respectively on 2, 5 and 10 epochs) were performed, in which each participant was asked to "write" (choose) six preset 4-symbol sequences which were to be on-line classified by the BCI system (specifically: 1111, 2222, 3333, 4444, 5555, 6666). After each symbol selection the system gave auditorily feedback to the user on the performed classification, so that s(he) could immediately realize if s(he) had made the correct choice.

Data Analysis

Mental workload evaluation

In order to evaluate the mental workload differences between the two count and repeat activities experimented by each participant in the two counting and repeat calibration sessions, we have computed alpha (band 8-13 Hz at Po8) and theta (band 4-7 Hz at Fz) spectral power density (SPD) curves in the two cases over the entire sessions. The area included between the two curves has then been determined as their integrals' difference (count-repeat).

Classification

Stepwise linear discriminant analysis (SWLDA), an extension of Fisher’s linear discriminant analysis, implemented in BCI2000 P300Classifier, was used to classify the calibration and item-selection data.

ITR computation

Accuracy was defined as the percentage of correctly classified numbers. We compared the results obtained over homogeneous data (sequences) in the three item-selection sessions.

Different methods of ITR calculation [1], [50] are used in the literature: as Käthner and colleagues [51] as well as Höhne [24] acknowledge, this makes it difficult and tedious to compare ITR among different studies. As reported by Kronegg [52] and Schlögl [50], the

most popular definition of ITR is the one by Wolpaw [1], which is reasonably simple and has often been used. A more generic definition is the one by Nykopp [50]. The different approaches and their implications are discussed by Kronegg [52], who shows that, depending on the number of classes and on the SNR, Wolpaw's bit-rate can be higher, lower or equal to Nykopp's bit-rate. Two BCIs with bit-rates computed using different definitions should therefore not be directly compared. Since the great majority of the studies use the formula suggested by Wolpaw [1] from Pierce [53], it was used here to compute the number of bits transmitted per trial (bit rate):

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

where N is the number of possible targets and P is the probability of the target. The data from the two different calibration sessions (count and repeat) were analyzed and compared to highlight possible performance differences.

EEG analysis

EEG analysis was carried out with MatLab [54] and EEGLab toolbox [55].

As already described, each calibration session consists of 6 subsessions (one for each considered symbol) during which the 6 numbers are auditorily presented to the participant in a random manner over 10 epochs, for a total number of 60 presentations for each symbol to be recognized. Therefore, the whole calibration session consists of 360 presentations.

Sixty single trial classification results, related to the 60 attended stimuli (10 for each symbol), were extracted from each calibration session by running the classifier with the previously calculated participant's features and were used to build the corresponding confusion matrices, showing the classification results compared with the actual classes. Each column of the matrix represents the instances in a predicted class, while each row represents the instances in an actual class.

Confusion matrix elements n_{ij} indicate how many class i samples have been classified as belonging to class j. The total number of correct classifications is given by the sum of the elements of the main diagonal n_{ii}, while n_{ij} elements, with i different from j, represent the wrong classified samples. As a result, 6x6 confusion matrices were produced, as in Fig. 1, where the i-th row is related to the i-th symbol, while the j-th column refers to the j-th calibration subsession results including 10 j-th symbol single trial presentations. It follows that the sum of the elements of each column is 10.

		Calibration subsessions					
		1	**2**	**3**	**4**	**5**	**6**
Symbols	**1**	7	0	0	1	0	0
	2	1	7	1	0	0	0
	3	0	0	8	0	0	0
	4	1	1	1	8	0	0
	5	0	1	0	1	10	1
	6	1	1	0	0	0	9

Fig. 1 Single trials classification confusion matrix

Considering the sum of the elements of the main diagonal as a characteristic parameter of each confusion matrix computed for every participant and for each performed mental activity:

$$\sum_{i=1}^{6} n_{ii}$$

it is possible to compare each participant's single trial performance in the 2 cases 'repeat' and 'count'.

A comparison has been conducted between off-line count and mental repetition classification accuracies achieved by each subject during the calibration sessions as a function of the number of averaged epochs to highlight possible performance differences between the two protocols.

Statistical Analysis

We performed a Friedman test to determine if there was a significant difference in the average accuracies achieved in the 10, 5 and 2 epochs item-selection sessions.

Chi-square tests were used to investigate multiclass errors in single trial analysis. Wilcoxon signed rank test was applied on target-disaggregated single trial hits in the Count and Repeat tasks. R^2 values were computed to quantify the differences between attended and unattended symbol responses. Wilcoxon signed rank test was applied on participants' SPD curve mean values in the Count and Repeat tasks. Cross-correlation coefficients were estimated to compare attended symbol responses in the count and repetition tasks. Statistical analysis was performed with R [56] and GNU PSPP [57].

Results

On-line classification accuracy and bit rate

Table 1 reports on-line accuracies and information transfer rates (expressed in bits/min) achieved with the proposed on-line protocol by each single participant on respectively 2, 5 and 10 epochs. As expected, the quality of the detection increases with the number of trials.

The obtained mean on-line accuracy (81.7% achieved in 10 epochs item-selection sessions, 72.9% in 5 epochs and 57.1% in 2 epochs) is comparable with the accuracy of previously proposed protocols (see Table 10).

As shown in Table 1, the on-line information transfer rate (ITR) obtained with the proposed protocol ranged from an average of 1.06 (10 epochs) to 2.35 bits/minute (2 epochs). One participant (S04) achieved an on-line ITR of 6.09 over 2 epochs.

Participant	On-line classification accuracy (%) and Bit Rate (bits/min)					
	2 epochs %	bits/min	5 epochs %	bits/min	10 epochs %	bits/min
S01	59.0	1.96	71.0	1.27	100.0	1.59
S02*	54.0	1.62	71.0	1.27	83.0	0.96
S03	25.0	0.1	67.0	1.11	88.0	1.08
S04	91.7	6.09	91.7	2.44	100.0	1.59
S05*	62.5	2.34	54.2	0.65	54.2	0.32
S06*	29.2	0.22	58.3	0.78	54.2	0.32
S07*	37.5	0.55	41.7	0.31	41.7	0.16
S08	45.8	1.02	91.7	2.44	100.0	1.59
S09	87.5	5.39	100.0	3.18	100.0	1.59
S10	79.2	4.19	83.0	1.91	95.8	1.38
Mean	**57.1**	**2.35**	**72.9**	**1.53**	**81.7**	**1.06**
St.Dev.	**23.5**		**18.7**		**22.8**	

Table 1 On-line accuracies and information transfer rates achieved by each participant with the proposed protocol. Asterisks indicate male participants.

Features and waveforms analysis

Every electrode, both in the counting and in the repeat tasks, with the only exception of Fz, shows a clear difference between attended/unattended symbol plots (see Fig. 2). Pz, Oz, P3 and PO7 show the largest signal amplitudes in the attended epochs.

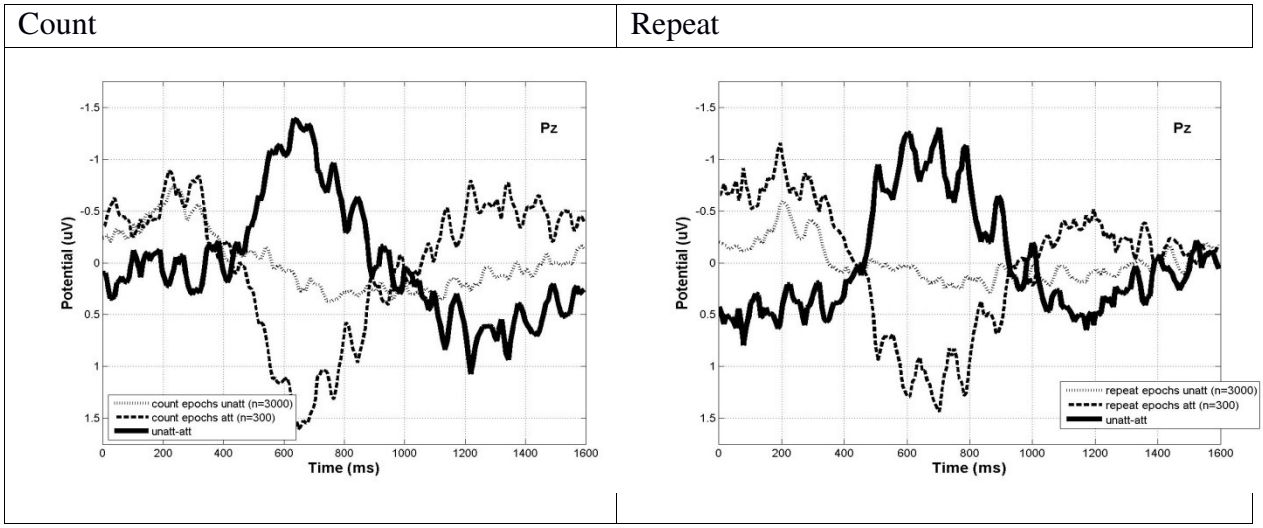

Fig. 2 Attended/Unattended Pz grand average plots comparison in the 'count' and 'repeat' all subjects conditions

Such differences can be quantified considering R^2 values between attended and unattended symbol responses in the 'count' and 'repeat' all subjects conditions (see Fig. 3).

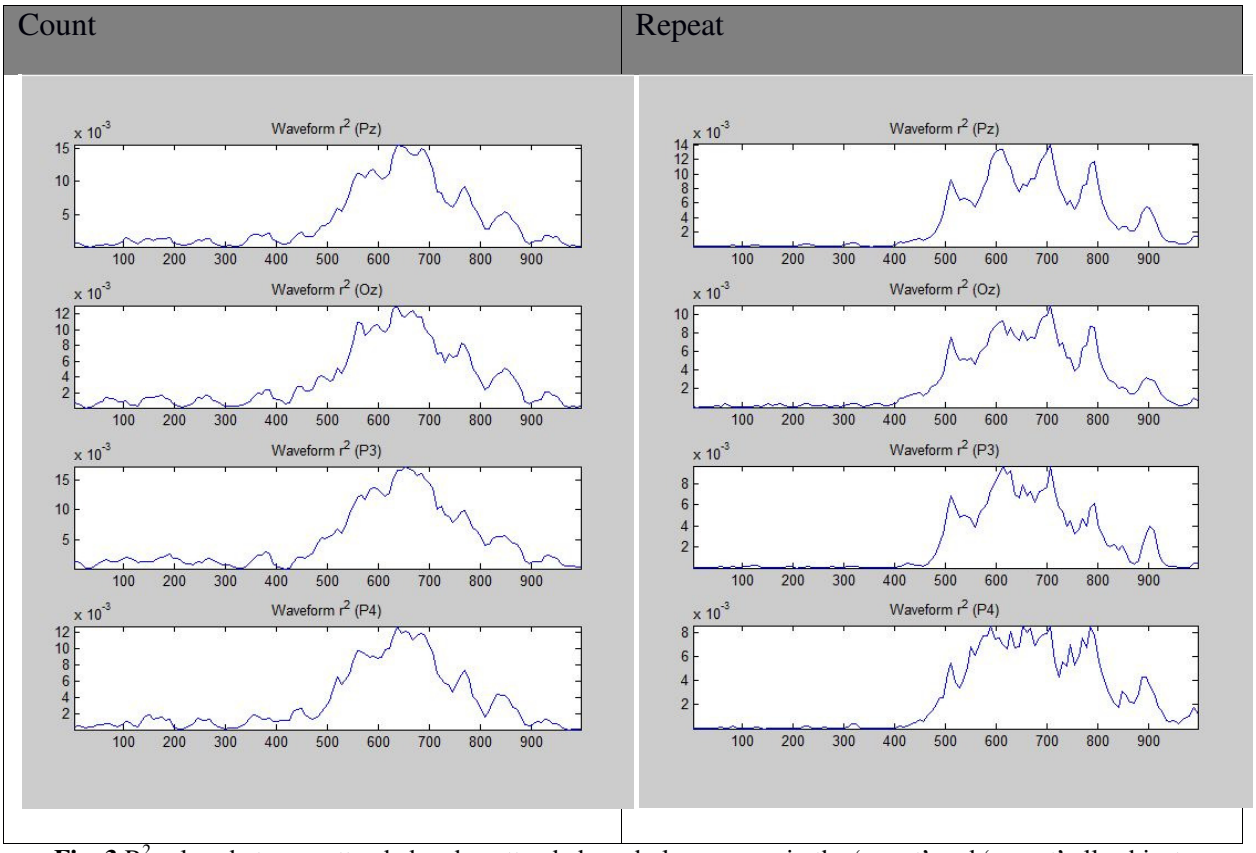

Fig. 3 R^2 values between attended and unattended symbol responses in the 'count' and 'repeat' all subjects conditions for channels Pz, Oz, P3, P4

Latency and amplitude differences may be noticed between the same electrodes attended symbols plots, following the count and repeat protocols (see Table 2). The count protocol shows a larger N2 latency on all electrodes and a smaller N2 negative amplitude on Cz, Pz, Oz and P3. Count protocol P300 latency is smaller on Pz, Oz and P3, while its P300 amplitude is larger on Cz, Pz, P3, P4, PO7, PO8.

Electrode	**N2 Latency (ms)**		**N2 Amplitude (µV)**		**P300 Latency (ms)**		**P300 Amplitude (µV)**	
	repeat	*count*	*repeat*	*count*	*repeat*	*count*	*repeat*	*count*
Fz	-	-	-	-	-	-	-	-
Cz	193.56	228.21	-1.0158	-0.6170	603.30	678.62	0.4674	0.8663
Pz	195.07	222.19	-1.1592	-0.8850	702.72	651.50	1.4396	1.6079
Oz	195.07	220.68	-0.7167	-0.6855	702.72	674.10	1.6203	1.3087
P3	195.07	320.10	-1.0470	-0.9660	702.72	653.01	1.3212	1.6577
P4	193.56	219.17	-1.0657	-0.8787	704.23	656.02	1.1031	1.3586
PO7	196.58	320.10	-0.6357	-0.7292	622.88	672.59	1.4957	1.7699
PO8	232.73	297.51	-0.5796	-0.8039	646.99	656.02	1.1654	1.3337

Table 2 All subjects grand average N2 and P300 Amplitudes and Latencies

Attended symbols plots, relative to the same electrodes, in the count and repeat trials may be directly compared in order to highlight waveform differences and similarities (see Fig. 4). Data and figures about amplitude and latency show that the two evaluated methods (repeat and count) produce almost superimposable results, but count P300 amplitudes are larger on 7 electrodes out of 8.

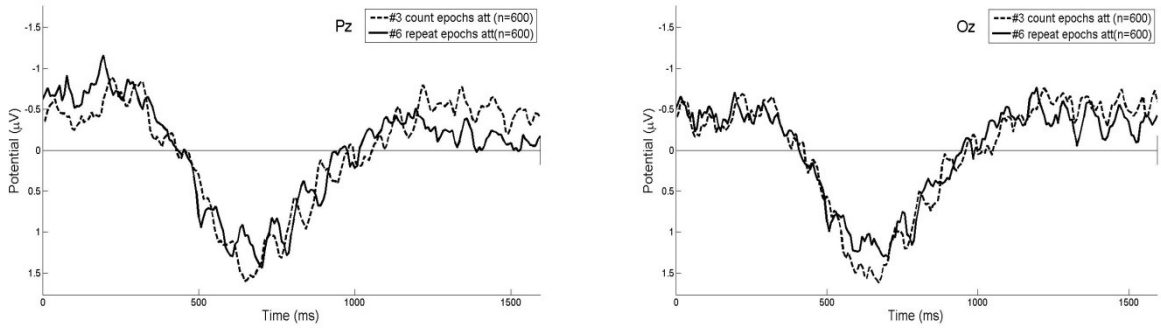

Fig. 4 Count/repeat Pz, Oz all subjects grand average attended symbols plots comparison

We computed the cross-correlation coefficient to compare the two curves of each plot (see Table 3). *Cz repeat* vs. *Cz count* obtained a poor resemblance (0.2511) while *Oz repeat* vs. *Oz count* was characterized by high similarity (0.9562).

Electrode	Correlation coefficient
Fz	0.0000
Cz	0.2511
Pz	0.8974
Oz	0.9562
P3	0.8894
P4	0.8892
PO7	0.9221
PO8	0.9295

Table 3 Cross-correlation coefficients computed on the signals recorded on all electrodes in the two 'count' and 'repeat' conditions

Single trials analysis

Considering the sum of the elements of the main diagonal of each confusion matrix computed for every participant and for each performed mental activity:

$$\sum_{i=1}^{6} n_{ii}$$

it is possible to compare each participant's performance in the 2 conditions. Results are reported in Table 4, together with each participant's expressed preference.

Participant	Repeat	Count	Preference (R|C)
S01	37	47	R
S02	25	22	R
S03	35	36	R
S04	51	49	R
S05	25	40	R
S06	25	19	R
S07	31	40	C
S08	39	47	C
S09	50	46	R
S10	43	48	R
Mean	**36.1**	**39.4**	**8R, 2C**

Table 4 Participants' single trial hits and Repeat/Count preference

The majority of participants (8 out of 10) reported to prefer the repeat mental activity as it was less demanding for them, while the remaining two participants clearly favored mental counting, commenting that it helped them to better concentrate. To further test and compare their mental counting behavior, participants S07 and S08, the only ones preferring mental counting, have also performed another item-selection test on 2 epochs with mental counting, both achieving an accuracy of 75% instead of respectively 37.5% and 45.8% obtained in their preceding repeat mental activity tests.

The above reported difference in participants' preferences was analyzed with a Chi-square test which revealed that it is close to significance (χ^2=3.60, p=0.058). Performance in single trials Repeat and Count classification results was analyzed with Wilcoxon test, which revealed no statistically significant differences (Z=-1.27, p=0.203).
Single trials classification results were further analyzed to detect any sound bias and to check that there were no significant differences in the effects of the six different sounds.

Target/selected	1	2	3	4	5	6	hits	errors
1	68	7	6	6	4	9	68	32
2	10	70	11	8	10	5	70	44
3	6	7	66	9	10	4	66	36
4	5	7	4	65	10	7	65	33
5	6	5	5	5	62	12	62	33
6	5	4	8	7	4	63	63	28

Table 5 Count. Participants' cumulative hits and errors

Target/selected	1	2	3	4	5	6	hits	errors
1	63	11	5	11	9	11	63	47
2	9	66	8	5	5	10	66	37
3	5	4	61	8	8	4	61	29
4	8	9	7	60	12	12	60	48
5	8	7	9	9	56	8	56	41
6	7	3	10	7	10	55	55	37

Table 6 Repeat. Participants' cumulative hits and errors

A Chi-square goodness of fit test performed on participants' cumulative errors in the Count and Repeat protocols (Table 5 and Table 6) confirmed the null hypothesis of equal expected errors frequencies in the two cases (χ^2= 4.233, df = 5, p-value = 0.516 and χ^2 = 6.3473, df = 5, p-value = 0.274).

A Chi-square goodness of fit test on participants' cumulative hits in the Count and Repeat protocols (Table 5 and Table 6) confirmed the null hypothesis of equal expected hits frequencies in the two cases (χ^2= 0.6904, df = 5, p-value = 0.983 and χ^2= 1.4432, df = 5, p-value = 0.919).It follows that there is no significant bias in hits or errors towards any of the six sounds in the two different protocols.

A Wilcoxon signed rank test applied on all participants' target-disaggregated cumulative hits in the Count and Repeat protocols (Table 5 and Table 6), highlighted a significant difference (V = 21, p-value = 0.034) between the two.

A Wilcoxon signed rank test applied on each participants' target-disaggregated hits in the Count and Repeat protocols (see Table 7), produced the results reported in Table 8, revealing a significant difference only for S06(*).

Target	S01		S02		S03		S04		S05		S06		S07		S08		S09		S10	
	C	R	C	R	C	R	C	R	C	R	C	R	C	R	C	R	C	R	C	R
1	9	8	6	4	7	8	7	7	6	7	4	5	4	5	10	5	7	7	8	7
2	10	5	2	4	6	4	7	10	6	4	5	6	9	5	7	9	8	10	10	9
3	6	7	3	4	5	7	8	10	9	5	2	4	8	6	8	5	9	8	8	5
4	7	6	4	5	6	4	8	8	7	2	3	4	8	6	8	8	5	8	9	9
5	6	5	5	4	7	8	10	8	6	3	3	3	5	5	7	5	8	8	5	7
6	9	6	2	4	5	4	9	8	6	4	2	3	6	4	7	7	9	9	8	6
	47	**37**	**22**	**25**	**36**	**35**	**49**	**51**	**40**	**25**	**19**	**25**	**40**	**31**	**47**	**39**	**46**	**50**	**48**	**43**

Table 7 Participants' target-disaggregated hits in the Count and Repeat protocols

Participant	Wilcoxon V	Wilcoxon p-value
S01	18.5	0.106
S02	7	0.520
S03	12	0.830
S04	3.5	0.713
S05	20	0.058
S06*	0	0.048
S07	14	0.099
S08	8.5	0.269
S09	1	0.423
S10	11.5	0.341

Table 8 Wilcoxon signed rank test results on each participants' target-disaggregated hits in the Count and Repeat protocols

Off-line count and mental repetition classification accuracies comparison

Considering the individual results, 6 subjects out of 10 (S01, S03, S05, S07, S08, S10) showed a better 'count' performance in the first averaged epochs, while 4 of them (S02, S04, S06, S09) achieved better results with 'repeat' protocol.

As an example Fig. 5 shows off-line count and mental repetition classification accuracies achieved by subjects S07 and S02 during the calibration sessions as a function of the number of averaged trials.

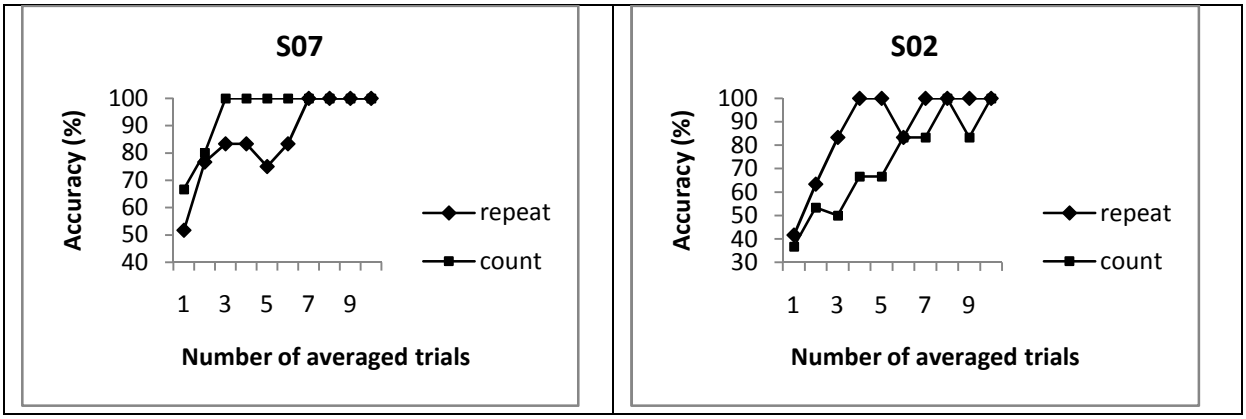

Fig. 5 S07 and S02 participant's classification accuracies as a function of the number of averaged trials

Fig. 6 shows the average off-line count and mental repetition classification accuracies achieved by all subjects during the calibration sessions as a function of the number of averaged trials.

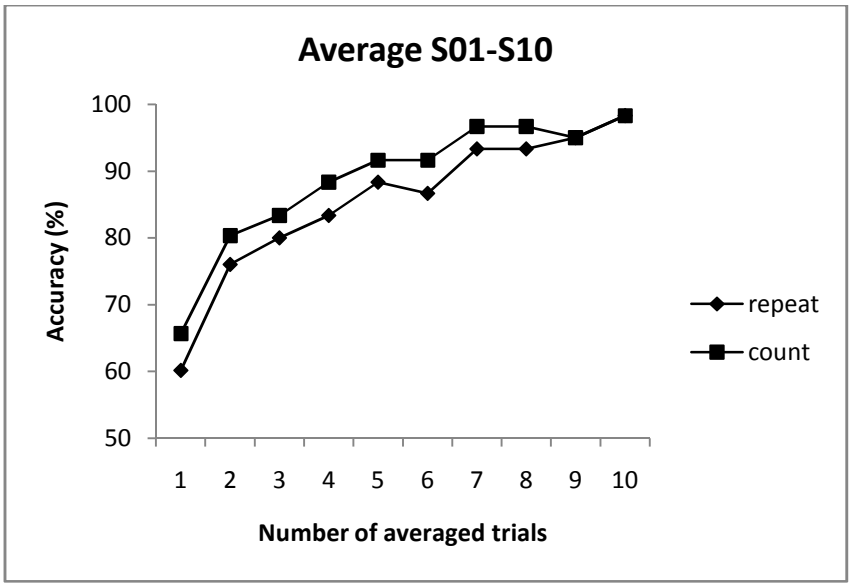

Fig. 6 All participants' grand average classification accuracies as a function of the number of averaged trials

It outlines a worse average behavior for the repeat protocol from 1 to 8 averaged epochs (ranging from -8.38 to -3.45%). Results on 9 and 10 averaged epochs appear equivalent.

Mental workload

Alpha (at Po8) and theta (at Fz) spectral power density (SPD) curves (see fig. 7 and fig. 8) have been computed in the two cases over the entire count and repeat calibration sessions. The area included between the two curves has then been determined as their integrals' difference (count-repeat).

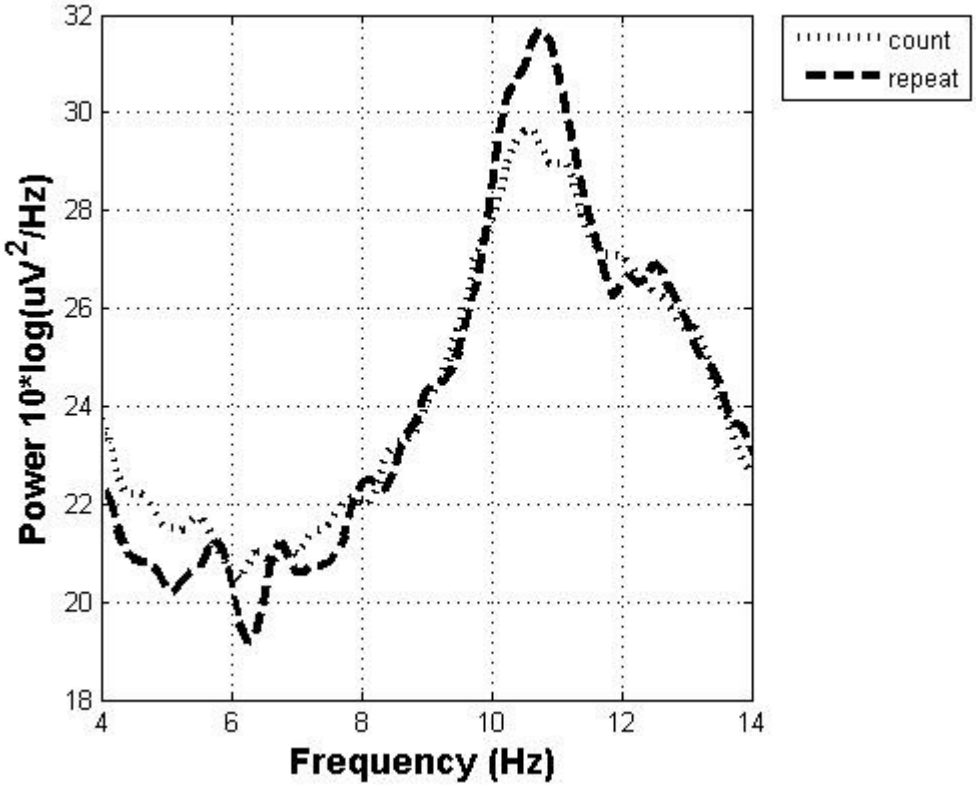

Fig. 7 Po8 Spectral Power Density (SPD) curves including Alpha band (8-13 Hz)

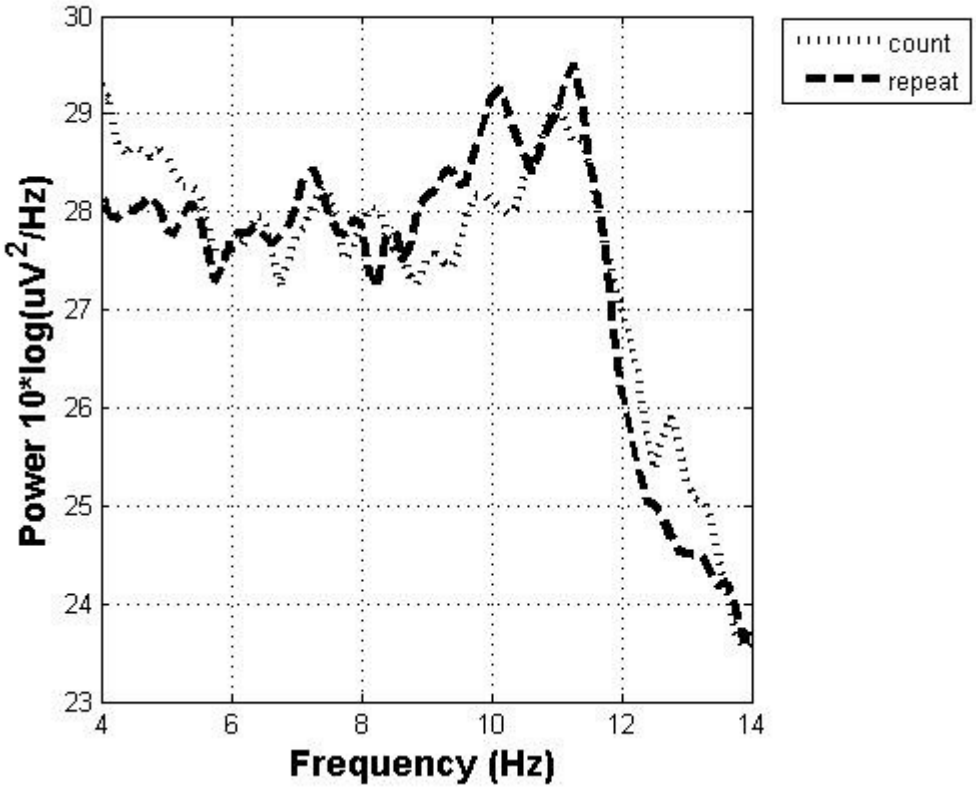

Fig. 8 Fz Spectral Power Density (SPD) curves including Theta band (4-7 Hz)

Table 9 reports the integrals' difference between the two curves (count-repeat) computed for each participant in the alpha (8-13 Hz) and theta (4-7 Hz) bands.

Participant	Alpha (8-13 Hz) count-repeat	Theta (4-7 Hz) count-repeat
S01	-1.9148	0.1601
S02	-2.0779	1.3351
S03	-2.7738	0.6658
S04	-2.3377	0.3683
S05	-2.5053	2.0577
S06	-3.0367	0.2691
S07	-0.4489	0.4606
S08	+2.1499	-2.8202
S09	-4.8038	1.2712
S10	-15.4641	19.0807

Table 9 *Difference between the two curves (count-repeat) computed for each participant in the alpha (8-13 Hz) and theta (4-7 Hz)*

Nine out of ten participants show a count alpha level lower than the corresponding repeat alpha level; the same nine participants show a count theta level higher than the corresponding repeat theta level. On the contrary, participant S08 shows a higher count alpha level and a lower alpha level; incidentally S08, together with S07, reported to prefer mental counting. It can also be noticed that S07 shows the smallest negative count-repeat alpha level among all participants.

Wilcoxon signed rank test applied on SPD curve mean values in the Count and Repeat protocols, highlighted: (i) a significant difference (V = 21, p-value = 0.01) in the 8-12 Hz band (ii) a not significant difference (V = 21, p-value = 0.059) 4-7 Hz band.

Discussion

Classification accuracy and bit rate

The on-line information transfer rate (ITR) obtained with the proposed protocol ranging from an average of 1.06 (10 epochs) to 2.35 bits/minute (2 epochs) with one participant (S04) who achieved an on-line ITR of 6.09 over 2 epochs (see Table 1), positions itself on the average of the previously proposed protocols (see Table 10).

Average on-line accuracy (81.7% achieved in 10-epochs item-selection sessions, 72.9% in 5-epochs and 57.1% in 2-epochs) is to be compared with the accuracy of the previously

proposed protocols (see Table 10, indicating also -when available- the ITR computation method used).

Task	ITR (bits/min)	ITR Comp.Method	Accuracy(%)	On/Off-line
3-Stimuli (2-25 sequences) [13]	2.46-0.39	Wolpaw	78.54-93.33	Off
Auditory speller 1 [16]	1.54	Wolpaw	65	Off
Auditory speller 2 [17]	1.9	Wolpaw	59.38	On
4-Choice [15]	0.43-1.80	Wolpaw	65.4	Off
Audiostream [14]	1.4	N/A (possibly Wolpaw)	72.6	Off
Audiostream (with ICA) [14]	3.03	N/A (possibly Wolpaw)	82.4	Off
Spatial multiclass BCI [22]	17.39	Wolpaw	68.7	Off
Spatial multiclass BCI [58]	2.84	N/A (possibly Nykopp)	77.4	On
PASS2D [24]	3.4	Wolpaw	89.37	On
Spatial multiclass BCI [51]	2.76	Wolpaw	65	On

Table 10 Average bit rates and accuracies achieved in previous studies

In order to verify if the use of quadruples of the same digit (1111, 2222, 3333, 4444, 5555, 6666) in the item-selection task might improve accuracy in each quadruple due to the repetition of the same number, we employed Friedman test to determine if there was a significant difference among the accuracies achieved on the first, second, third and fourth digits. The null hypothesis was that the distribution of the ranks of each type of score (i.e. first, second, third and fourth digits) was the same. Differences among the four accuracies were not statistically significant (2 epochs: $\chi2=1.1512$, df = 3, p = 0.765; 5 epochs: $\chi2 = 0.6094$, df = 3, p = 0.894; 10 epochs: : $\chi2 = 2.1346$, df = 3, p = 0.545).

The comparison conducted between off-line count and mental repetition classification accuracies achieved during the calibration sessions shows a worse average behavior for the repeat protocol from 1 to 8 averaged epochs (ranging from -8.38 to -3.45%). Results on 9 and 10 averaged epochs appear equivalent. This result is in accordance with previous studies (Xu, Zhang, Ouyang, & Hong, 2013) which showed better a classification performance with higher difficulty tasks.

If we consider the individual results obtained in this study, 6 subjects out of 10 (S01, S03, S05, S07, S08, S10) show a better 'count' performance in the first averaged epochs, while 4 of them (S02, S04, S06, S09) seem to behave better with 'repeat'. In particular S05, S07 and

S08 show a pronounced worse behaviour in the ‘repeat’ task; incidentally S07, together with S08, reported to prefer mental counting.

Mental workload

Käthner and colleagues [51] study aimed also at quantifying the differences in subjective workload between the auditory and the visual P300 spelling application. A significantly higher workload was reported for the auditory speller compared to the visual paradigm.

Unlike other protocols which use more demanding mental tasks, such as spatial distributed sound direction or other target property mental discrimination, to elicit stronger ERP responses [22], [42], [43], our approach tries to reduce mental workload by introducing a change in the typical oddball paradigm task commonly employed in the literature [3], [44]–[46]: instead of passively counting how often the target auditory stimulus is presented, the participant is requested to simply mentally repeat it whenever it appears.

The new proposed ‘repeat’ protocol has achieved good real-time results and its performance is comparable with previously proposed protocols (see Tables 1 and 10).

Off-line all subjects’ ‘repeat protocol’ shows a worse average behavior than ‘count’ protocol from 1 to 8 averaged trials (ranging from -8.38 to -3.45%). On the contrary, results on 9 and 10 averaged trials appear to be equivalent. Eight participants out of ten said that mental counting was heavier than the mental repetition activity. This subjective result is consistent with the difference of alpha (at Po8) and theta (at Fz) spectral power density (SPD) curves computed in the ‘count’ and ‘repeat’ cases over the entire sessions (see Table 9).

Nine out of ten participants show a count alpha level lower than the corresponding repeat alpha level; the same nine participants show a count theta level higher than the corresponding repeat theta level. On the contrary, participant S08 shows a higher count alpha level and a lower theta level; incidentally S08, together with S07, reported to prefer mental counting. It can also be noticed that S07 shows the smallest negative count-repeat alpha level among all participants (see Table 9).

Inter stimulus interval (ISI)

As already remarked, considering Farwell and Donchin's results [5] on ISI duration and rate of communication and the fact that auditory stimuli presentations require a larger amount of time than visual stimuli, in this study we increased stimulus presentation time and ISI respectively to 375 ms and 1250 ms.

For a comparison, in previously cited auditory protocols used stimulus duration ranges from 40 to 600 ms and ISI from 125 to 1400 ms.

Electrodes configuration

The previously cited auditory protocols have used a number of electrodes ranging from 16 to 67, Halder [13], [16], [10].

We used the eight channels (Fz, Cz, Pz, Oz, P3, P4, PO7, PO8) of the Krusienski configuration [46]. Each channel was referenced to the right mastoid and grounded to the left. This solution was able to give effective results while using a reduced number of electrodes.

Conclusion

The proposed auditory protocol has achieved good on-line results with healthy subjects, showing a performance comparable, and sometimes higher, than previously proposed protocols. Moreover, it uses a small number of electrodes and requires only a 10-minute calibration session. The obtained results suggest that the proposed auditory BCI protocol might be promising for application with users who are unable to take advantage of visual BCIs.

The main benefit of the new method is the introduction of a change in the typical oddball paradigm task to make it simpler for users: instead of incrementally counting the presented target stimuli, they simply mentally repeat the stimuli as they are when they occur.

The repeat activity appeared to be mentally less demanding and preferred by the majority of the participants. Nine out of ten participants showed a higher workload in the count sessions.

However, the individual differences in performance and preference reported in the paper suggest that future tests, especially if they involve patients, should take into account user's preferences and characteristics in order to allow him/her to achieve the best results with the least demanding task. The specific patient's communication requirements could also be taken into account: the protocol should allow binary communication as well as multiple choices, scaling ITR as a function of each participant's capability. Attempts to decrease the number of electrodes will be also made to minimize setting times and patient's stress due to preparation [59].

The auditory protocol presented in the current study could be easily adapted to be employed as a ''yes–no'' spelling device. In this way, in addition of using it as a communication channel, it could be interesting to test it as a possible tool to distinguish among different clinical states of consciousness for patients with altered states of consciousness.

Acknowledgements

The research project was co-financed by the Friuli Venezia Giulia region with a grant named "Advanced emergency medical services for the disabled based on innovative ICT technologies".

We thank Jürgen Mellinger of the Institute of Medical Psychology and Behavioral Neurobiology of the University of Tübingen for his valuable advice and support in the use of BCI2000 modules, files and information retrieval.

Moreover, we thank Barbara Porcella of the Comunità Piergiorgio NPO (Non Profit Organization), Reference Centre for Disability Products and Assistive Devices and Technologies for the Disabled in Udine, and Alberto Catto for the useful feedback they provided about the tested techniques.

References

[1] J. R. Wolpaw *et al.*, 'Brain-Computer Interface Technology: A Review of the First International Meeting', *IEEE Trans Rehabil Eng 2000*, pp. 164–173, 2000.

[2] J. Polich, 'Updating P300: An integrative theory of P3a and P3b', *Clin. Neurophysiol.*, vol. 118, no. 10, pp. 2128–2148, Oct. 2007.
[3] M. Fabiani, G. Gratton, D. Karis, and E. Donchin, 'Definition, identification, and reliability of the P300 component of the event-related brain potential', *Adv. Psychophysiol.*, vol. 2, pp. 1–78, 1987.
[4] S. Sutton, M. Braren, J. Zubin, and E. R. John, 'Evoked-Potential Correlates of Stimulus Uncertainty', *Science*, vol. 150, no. 3700, pp. 1187–1188, Nov. 1965.
[5] L. A. Farwell and E. Donchin, 'Talking off the top of your head: A mental prosthesis utilizing event-related brain potentials', *Electroencephalogr Clin Neurophysiol.*, pp. 510–523, 1988.
[6] E. Donchin, K. M. Spencer, and R. Wijesinghe, 'Assessing the Speed of a P300-Based Brain-Computer Interface', *IEEE Trans Rehabil Eng*, vol. 8, no. 2, pp. 174–9, 2000.
[7] H. Serby, E. Yom-Tov, and G. Inbar, 'An Improved P300-Based Brain-Computer Interface', *IEEE Trans Neural Syst Rehabil Eng*, vol. 13, no. 1, pp. 89–98, 2005.
[8] F. Piccione *et al.*, 'P300-based brain computer interface: Reliability and performance in healthy and paralysed participants', *Clin. Neurophysiol.*, vol. 117, pp. 531–537, 2006.
[9] F. Nijboer *et al.*, 'A P300-based brain-computer interface for people with amyotrophic lateral sclerosis', *Clin Neurophysiol*, vol. 119, no. 8, pp. 1909–1916, 2008.
[10] E. W. Sellers and E. Donchin, 'A P300-based brain–computer interface: Initial tests by ALS patients', *Clin. Neurophysiol.*, pp. 538–548, 2006.
[11] M. S. Treder, N. M. Schmidt, and B. Blankertz, 'Gaze-independent brain-computer interfaces based on covert attention and feature attention', *J. Neural Eng.*, vol. 8, pp. 1–12, 2011.
[12] T. Kaufmann, E. M. Holz, and A. Kuebler, 'Comparison of tactile, auditory, and visual modality for brain-computer interface use: a case study with a patient in the locked-in state', *Front. Neurosci.*, vol. 7, 2013.
[13] S. Halder *et al.*, 'An auditory oddball brain–computer interface for binary choices', *Clin. Neurophysiol.*, pp. 516–523, 2010.
[14] N. J. Hill, T. N. Lal, K. Bierig, N. Birbaumer, and B. Scholkopf, 'An auditory paradigm for brain-computer interfaces', *Adv. Neural Inf. Process. Syst.*, pp. 569–576, 2005.
[15] E. W. Sellers and E. Donchin, 'A P300-based brain–computer interface: Initial tests by ALS patients', *Clin. Neurophysiol.*, vol. 117, no. 3, pp. 538–548, Mar. 2006.
[16] A. Furdea *et al.*, 'An auditory oddball (P300) spelling system for brain-computer interfaces', *Psychophysiology*, pp. 1–9, 2009.
[17] D. S. Klobassa, T. M. Vaughan, P. Brunner, N. E. Schwartz, and J. R. Wolpaw, 'Toward a high-throughput auditory P300-based brain-computer interface', *Clin. Neurophysiol.*, pp. 1252–1261, 2009.
[18] A. Kuebler, A. Furdea, S. Halder, E. M. Hammer, F. Nijboer, and B. Kotchoubey, 'A Brain–Computer Interface Controlled Auditory Event-Related Potential (P300) Spelling System for Locked-In Patients', *Ann N Acad Sci*, vol. 1157, pp. 90–100, 2009.
[19] M. Risetti, R. Formisano, J. Toppi, L. Astolfi, and D. Mattia, 'Approaching rehabilitation of severe Acquired Brain Injuries by means of a neurophysiological screening', *Int J Bioelectromagn*, vol. 13, pp. 146–147, 2011.
[20] D. H. Coyle, A. Carroll, J. Stow, A. McCann, A. Ally, and J. McElligott, 'Enabling Control in the Minimally Conscious State in a Single Session with a Three Channel BCI', presented at the The 1st international DECODER Workshop, Paris, 2012.
[21] D. Lulé, Q. Noirhomme, S. Kleih, C. Chatelle, S. Halder, and A. Demertzi, 'Probing command following in patients with disorders of consciousness using a brain–computer interface', *Clin Neurophysiol*, vol. 124, no. 1, pp. 101–106, 2013.

[22] M. Schreuder, B. Blankertz, and M. Tangermann, 'A New Auditory Multi-Class Brain-Computer InterfaceParadigm: Spatial Hearing as an Informative Cue', *PLoS ONE*, pp. 1–14, 2010.
[23] M. Schreuder, T. Rost, and M. Tangermann, 'Listen, You are Writing! Speeding up Online Spelling with a Dynamic Auditory BCI', *Front. Neurosci.*, vol. 5, Oct. 2011.
[24] J. Höhne, M. Schreuder, B. Blankertz, and M. Tangermann, 'A novel 9-class auditory erp paradigm driving a predictive text entry system', *Front Neurosci*, vol. 5, no. 99, 2011.
[25] B. Xie and G. Salvendy, 'Review and reappraisal of modeling and predicting mental workload in single- and multi-task environments', *Work Stress*, vol. 14, no. 1, pp. 74–99, 2000.
[26] J. Kohlmorgen *et al.*, 'Improving human performance in a real operating environment through real-time mental workload detection', in *Toward Brain-Computer Interfacing*, MIT Press, 2007.
[27] D. Heger, F. Putze, and T. Schultz, 'Online Workload Recognition from EEG data during Cognitive Tests and Human-Computer Interaction', presented at the KI 2010: Advances in Artificial Intelligence 33rd Annual German Conference on AI, Karlsruhe, 2010, vol. 6359.
[28] R. Shriram, M. Sundhararajan, and N. Daimiwal, 'EEG Based Cognitive Workload Assessment for Maximum Efficiency.', *IOSR J. Electron. Commun. Eng. IOSR-JECE*, pp. 34–38, 2012.
[29] F. Paas, 'Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach', *J. Educ. Psychol.*, vol. 84, pp. 429–434, 1992.
[30] S. Miller, 'National Advanced Driving Simulator'. N01-006, 2001.
[31] D. Embrey, C. Blackett, P. Marsden, and J. Peachey, 'Development of a Human Cognitive Workload Assessment Tool'. Human Reliability Associates, 2006.
[32] R. M. Hope, Z. Wang, Z. Wang, Q. Ji, and W. D. Gray, 'Workload Classification Across Subjects Using EEG', presented at the Human Factors And Ergonomics Society 55th Annual Meeting, 2011.
[33] C. Berka, D. J. Levendowski, R. E. Olmstead, M. V. Popovic, M. Cvetinovic, and M. M. Petrovic, 'Real-time analysis of EEG indices of alertness, cognition, and memory with a wireless EEG headset', *Int. J. Hum.-Comput. Interact.*, vol. 17, pp. 151–170, 2004.
[34] A. M. Brouwer, M. A. Hogervorst, J. van Erp, T. Heffelaar, P. H. Zimmerman, and R. Oostenveld, 'Estimating workload using EEG spectral power and ERPs in the n-back task', *JNeural Eng*, vol. 9, pp. 1–14, 2012.
[35] W. Klimesch, 'EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis', *Brain Res. Rev.*, vol. 29, pp. 169–195, 1999.
[36] A. Gevins and M. E. Smith, 'Neurophysiological measures of cognitive workload during humancomputer interactions', *Theor. Issues Ergon. Sci.*, vol. 4, pp. 113–131, 2003.
[37] W. Klimesch, B. Schack, and P. Sauseng, 'The functional significance of theta and upper alpha oscillations for working memory: A review.', *Exp. Psychol.*, vol. 52, pp. 99–108, 2005.
[38] P. Antonenko, F. Paas, R. Grabner, and T. Van Gog, 'Using Electroencephalography to Measure Cognitive Load', *Educ Psychol Rev*, vol. 22, pp. 425–438, 2010.
[39] F. Chen, 'Neuro-ergonomic Research for Online Assessment of Cognitive Workload'. National ICT Australia (NICTA), 2011.
[40] C. Walter, G. Cierniak, P. Gerjets, W. Rosenstiel, and M. Bogdan, 'Classifying mental states with machine learning algorithms using alpha activity decline', in *ESANN 2011 proceedings*, Bruges, 2011.

[41] B. Blankertz *et al.*, 'The Berlin brain-computer interface presents the novel mental typewriter Hex-o-Spell', presented at the Proceedings of the 3rd International Brain-Computer Interface Workshop and Training Course 2006, Graz, 2006, pp. 108–109.
[42] J. Guo, S. Gao, and B. Hong, 'An auditory brain–computer interface using active mental response', *IEEE Trans Neural Syst Rehabil Eng*, vol. 18, no. 3, pp. 230–235, 2010.
[43] H. Xu, D. Zhang, M. Ouyang, and B. Hong, 'Employing an active mental task to enhance the performance of auditory attention-based brain–computer interfaces', *Clin Neurophysiol*, vol. 124, no. 1, pp. 83–90, 2013.
[44] W. T. Picton, 'The P300 wave of the human event-related potential', *J Clin Neurophysiol*, vol. 9, no. 4, pp. 456– 479, 1992.
[45] S. Huettel and G. McCarthy, 'What is odd about the odd-ball task? prefrontal cortex is activated by dynamic changes in response strategy', *Neuropsychologia*, vol. 42, pp. 379–386, 2004.
[46] D. J. Krusienski, E. W. Sellers, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw, 'Toward Enhanced P300 Speller Performance', *J Neurosci Methods*, pp. 15–21, 2008.
[47] 'American Electroencephalographic Society guidelines for standard electrode position nomenclature', *J. Clin. Neurophysiol. Off. Publ. Am. Electroencephalogr. Soc.*, vol. 8, no. 2, pp. 200–202, Apr. 1991.
[48] G. Schalk and J. Mellinger, *A Practical Guide to Brain-Computer Interfacing with BCI2000*. London: Springer-Verlag, 2010.
[49] G. Schalk, D. McFarland, and J. Mellinger, 'BCI2000', *BCI2000 | Schalk Lab [Internet]*, 2011. [Online]. Available: http://www.schalklab.org/research/bci2000. [Accessed: 21-Jun-2014].
[50] A. Schlögl, J. Kronegg, J. Huggins, and S. G. Mason, 'Evaluation criteria for BCI research', in *Towards Brain-Computer Interfacing*, 1 vols, Cambridge (MA): MIT Press, 2007, pp. 327–342.
[51] I. Käthner, C. A. Ruf, E. Pasqualotto, C. Braun, N. Birbaumer, and S. Halder, 'A portable auditory P300 brain–computer interface with directional cues', *Clin. Neurophysiol.*, vol. 124, pp. 327–338, 2013.
[52] J. Kronegg, S. Voloshynovskiy, and T. Pun, 'Analysis of bit-rate definitions for Brain-Computer Interfaces', presented at the 2005 Int. Conf. on Human-computer Interaction (HCI'05), Las Vegas, Nevada, USA, 2005.
[53] J. Pierce, *An introduction to information theory: symbols, signals & noise*. New York: Dover Publications, 1980.
[54] C. Moler, J. Little, and S. Bangert, 'MATLAB and Simulink for Technical Computing - MathWorks Italia', 2013. [Online]. Available: http://www.mathworks.it/. [Accessed: 21-Jun-2014].
[55] A. Delorme and S. Makeig, 'EEGLAB - Open Source Matlab Toolbox for Electrophysiological Research', *EEGLAB*, 2013. [Online]. Available: http://sccn.ucsd.edu/eeglab/. [Accessed: 21-Jun-2014].
[56] R. Gentleman and R. Ihaka, 'The R Project for Statistical Computing', 1997. [Online]. Available: http://www.r-project.org/. [Accessed: 21-Jun-2014].
[57] B. Pfaff, J. Darrington, and J. Stover, 'PSPP - GNU Project - Free Software Foundation', *PSPP*, 2001. [Online]. Available: http://www.gnu.org/software/pspp/. [Accessed: 21-Jun-2014].
[58] M. Schreuder, T. Rost, and M. Tangermann, 'Listen, You are Writing! Speeding up Online Spelling with a Dynamic Auditory BCI.', *Front Neurosci*, vol. 5, Oct. 2011.
[59] C. Guger, G. Krausz, B. Allison, and G. Edlinger, 'Comparison of dry and gel based electrodes for P300 brain–computer interfaces', *Front. Neurosci.*, pp. 1–7, 2012.