# A COGNITIVE BRAIN-COMPUTER INTERFACE PROTOTYPE FOR THE CONTINUOUS MONITORING OF VISUAL WORKING MEMORY LOAD.

*Aldo Mora Sánchez*[1,2], *Antoine Gaume*[1,2], *Gérard Dreyfus*[2], *François-Benoît Vialatte*[1,2]

1 - Brain plasticity Unit. CNRS, UMR8249, Paris, 75005, France
2 - ESPCI-ParisTech, PSL Research University, Paris, 75005, France

## ABSTRACT

We investigate the neural correlates of visual working memory using electroencephalography (EEG). Our objective is to develop a cognitive Brain-Computer Interface (BCI) able to monitor visual working memory load in real-time. A system with these properties would eventually have different applications, such as training, rehabilitation, or safety while operating dangerous machinery. The BCI performances were evaluated using cross-validation. With an appropriately chosen classification threshold, it is possible to detect high working memory load with a sensitivity of 68% and a specificity of 72%. However, it is well known that some subjects are BCI *illiterate*, meaning that up to 30% of the users have too high signal variability to use EEG-based BCI systems. If we analyse each subject individually, it is possible to detect high working memory load with a sensitivity of 78% and a specificity of 81% (accuracy = 81%) for a typical good subject. Changes due to visual working memory load were observed in frontal, parietal, and occipital regions.

***Index Terms*—** Brain-computer interfaces, cognitive information processing, pattern recognition, classification.

## 1. INTRODUCTION

Humans can interact directly with machines using their brain activity. This is, in general terms, a Brain-Computer Interface (BCI) [1]. Such devices take brain activity as an input, and transform it into an output via a translation algorithm, bypassing the motor system. The outcome is often a command, for instance, choosing a letter or a certain movement. Electroencephalography (EEG) is the most frequently used technique due to its non-invasive nature and low acquisition and operational costs. Furthermore, EEG is also a signal of interest because electrical brain signals provide access to neural dynamics with a very high temporal resolution.

In healthy subjects, regular speech is an efficient channel of communication. Besides, the motor system provides a fine-tuned means of control with several degrees of freedom. Both the motor system and speech demand relatively low cognitive effort as compared to BCI. Therefore, BCIs were initially thought for patients whose conditions prevent them from adequately using those systems. For instance, locked-in patients [2]. However, BCI systems can go beyond communication, for instance, they can be useful for cognitive monitoring as suggested by [3].

Zander and Kothe [4] propose a re-thinking of BCI usage. They suggest a new classification according to the BCI functional mechanisms, and outline usability according to the target (healthy or disabled users). BCI are classified into active, reactive and passive. This paper deals with passive interfaces.

### 1.1. Passive BCI

The current state of BCI performance can hardly compete with the above mentioned mechanisms available to healthy users. Therefore, it is proposed to combine BCI technology with cognitive monitoring in a new approach: passive BCI. Cognitive monitoring refers to the analysis of brain signals in order to infer information about the cognitive state of the user. The idea of a passive BCI is to feed this information to a system to improve its performance, in a way which is non-voluntarily driven by the user. Contextual information about the cognitive state could be critical in safety-related tasks such as driving or in industrial environments. Cognitive load or attention, for instance, have a high impact on the performance of these activities. By including this information, certain commands can be triggered depending on the cognitive state, allowing the system to adapt to the users, without them intentionally triggering the actions.

### 1.2. Working Memory Load

Working memory is regarded as a system that keeps information (with storage and time limitations) while it is being manipulated. It works as an interface between perception, long-term memory and action [5].

978-1-4673-7454-5/15/$31.00 

There are precedents for the assessing of working memory load using EEG spectral features, particularly, in prefrontal and parietal regions of the brain. Furthermore, after training working memory it is possible to observe and measure certain changes [6].

## 2. MATERIALS AND METHODS

### 2.1. The task

According to Cowan's model [7], working memory acts as an activation buffer for items previously stored in the long term memory. This accounts for the remarkable consistency of the number of items that can be stored at a given time, regardless the size of the items themselves. The number of items is then limited, however, the human brain is extraordinarily good at finding associations between items in working memory, in order to reduce the number of occupied items (mnemonics, for instance). This was a key point while designing the task, that can be described as follows:

- Subjects sit in front of a computer screen and are presented with a collection of figures that will be used during the experiment. They are asked to give a short name for them. There are different sets of figures, and each set corresponds to a semantic field: animals, transport, etc.
- The *target*, a specific sequence of figures, containing either a small number (2) or a large number (determined by the calibration, typically 5) of elements is presented to the subject who is asked to keep it in memory. An example of target corresponding to low visual working memory load would be, for instance, a train followed by a bicycle.
- A random sequence of figures generated from the total set slides from right to left. The subjects simply have to press a button whenever they find the target. This is considered one trial. Trials last on average 25 seconds.
- Subjective feedback is collected on every successful trial: reported cognitive state (bored, stressed or challenged) and reported number of figures actually remembered.

The distance between the subjects and the screen was 60 centimeters, the screen model was ProLite E2208HDD. Lighting conditions were normal room contitions. The size of the figures was 100x100 pixels.

The fact that subjects were asked to verbalize the name of the figures induces in them a homogeneous storage-retrieval technique: simply to repeat the names of the sequence of stored items and to compare it with the observed sliding items. Subjects were asked, after the experiment, which technique they used. They reported to have used precisely this method most of the time. This has several advantages. Firstly, all the items are kept as separate items. Secondly, visual comparisons guarantee that we are measuring visual working memory. Furthermore, this comparison is made in a sufficiently continuous manner. Finally, the fact that a given group of objects belongs to an evident semantic field, reduces the likelihood of item merging. In fact, the set of figures is changed several times during the experiment to prevent subjects from developing compression strategies. Or, at least, to reduce them.

As it is unlikely that patterns similar to the target, or distractors, appear by chance, parameters are tuned so that exactly fifty percent of the time a distractor appears. The (low) probability of a distractor appearing by chance was also taken into account. Distractors are important to prevent subjects from storing only a smaller amount of items (for instance, first and last, or first few). Distractors appear half of the time so that subjects do not learn that it is more likely to find, for instance, first the distractor and then the pattern, or vice-versa.

The subjective data was not included in the analysis, however, it was used to verify that the task was indeed inducing a change in the perceived cognitive state.

Finally, the small window size prevents eye movements, that are known to produce eye-related artifacts. The experiment was written in Matlab® 2015a using Psychophysics Toolbox extensions [6].

### 2.2. Data acquisition and pre-processing

Brain activity was recorded using a 16 channel EEG device (Brain Products V-Amp) at a sampling rate of 2000 Hz. The electrode set-up is shown in figure 1. 12 healthy subjects of age between 21 and 31 were recorded, 6 males and 6 females, with normal or corrected to normal vision, and absence of any brain disorder or drug consumption. The study followed the principles outlined in the Declaration of Helsinki. All participants were given explanations about the nature of the experiment and signed an informed consent form before the experiment started.

Data was filtered in the range of 1 to 90 Hz with a 3rd order butterworth filter. A notch filter at 50 Hz was used to remove the line noise as well.

In order to obtain a clean marker that can be further used even in noisy, real-life conditions, special attention was paid

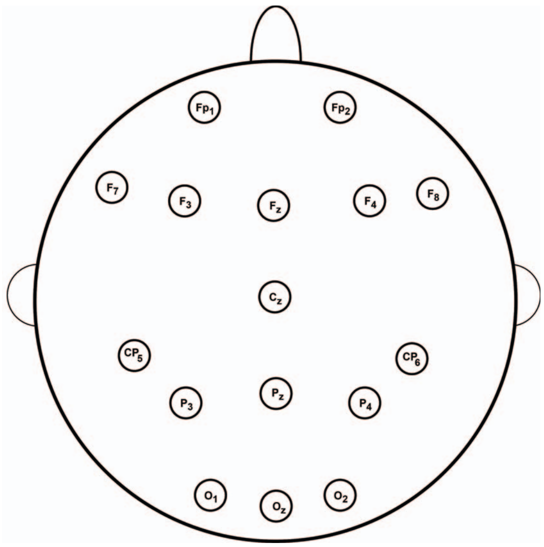


**Fig. 1**. Eelctrodes setup.

to the pre-processing and cleaning procedure. All the trials were visually inspected and those that were heavily artifacted were rejected from the study. Approximately 15 % of the trials were rejected for this reason. Recordings corresponding to the part of the task where distractors appeared were removed too, in order to avoid arousal effects. Eye blinks were removed using Independent Component Analysis (ICA) trial-wise, in order to fulfill the stationarity requirement. The results of this cleaning procedure were visually inspected for each trial. ICA decomposition was performed using the runICA script included in EEGLAB [7]. Finally, trials were cut into 10 seconds epochs.

A total of 235 epochs were analysed, 55 percent corresponding to low working memory load, and 45 percent to high working memory load.

### 2.3. Feature extraction, calibration and classification

For each epoch, and for each channel, a collection of features were extracted in order to feed the classifier that will distinguish between epochs corresponding to low working memory load and high working memory load. Classical spectral features were extracted using Matlab p-welch function, with a hamming window of 0.5 seconds. Spectral features included absolute and relative power in the delta (1 - 4 Hz), theta (4 - 8 Hz), alpha (8 - 12 Hz), lower beta (12 - 20 Hz), upper beta (20 - 30 Hz), lower gamma (30 - 45 Hz) and upper gamma (55 - 90 Hz) ranges. Relative power is the fraction of the total power corresponding to a particular band. This has the advantage of reducing inter-subject variability. However, we should be careful while drawing conclusions about changes in relative power associated to a cognitive state, given that the relative power in a particular band can be affected by changes in other bands.

For each subject, half of the data (calibration data) was attached to the existing users dataset, in order to predict the remaining half. Features, and their corresponding cross terms, were ranked by Gram-Schmidt Orthogonalization (GSO) procedure according to their classification power. The best features were then fed to a Linear Discriminant Analysis classifier. Cross-validation was performed by repeating this process for all subjects. The ideal number of features was set as the one that minimized the classification error.

A diagram showing the overall algorithm performed on every subject is shown in figure 2.

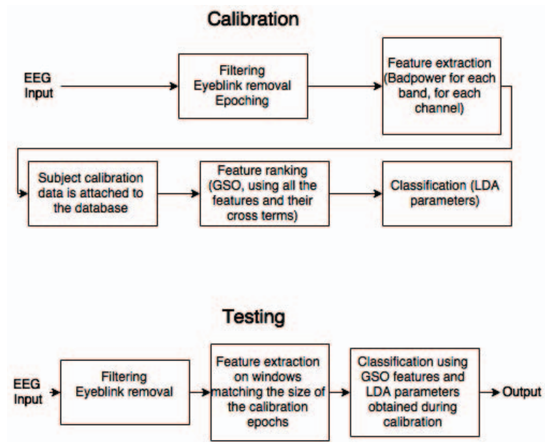


**Fig. 2**. Overall algorithm performed on each subject.

### 2.4. Visualization: ROC curves

When a new instance is used as input in a classifiers, the output is often a *score*, the posterior probability. A high score simply means that the instance is likely to belong to the positive class. A low score, in turn, means that the instance is likely to belong to the negative class. Varying the threshold required to belong to a class, allows us to change the sensitivity-specificity of the classification process. A receiver operating characteristic (ROC) curve is a useful tool to analyse this behaviour visually. A ROC curve plots false positive rate vs. true positive rate for different threshold values. As a reference, a random classifier has a ROC curve of unitary slope starting at the origin, spanning a surface of 0.5. Better classifiers should be then plotted above the diagonal, having an area under the curve greater than 0.5.

## 3. RESULTS

The classifier performed better when six features were used, as can be observed in figure 3. It is important to remember that those features were actually cross terms, i.e., the product of two features. All the following estimations were obtained by setting the number of features to 6. Figure 3 shows that the mean error rate obtained using cross-validation was 32%. However, due to the problem of BCI literacy, it is more interesting to analyse this error on a subject basis rather than globally. Nine subjects out of the twelve, had both an error rate lower than that of a random classifier, and a ROC curve with area bigger than 0.5. Those two values are evidently correlated, as we can see in figure 4. Subjects plotted in the top left square were considered as *good subjects*. Figure 5 shows the ROC curve of the classifier, together with the ROC curve of a typical good subject. For the whole set of subjects, we can expect to detect high working memory load 68% of the time with 72% of accuracy. In particular, for a single good subject, we can observe that we can detect high working memory load 77% of the time with 81% of specificity.

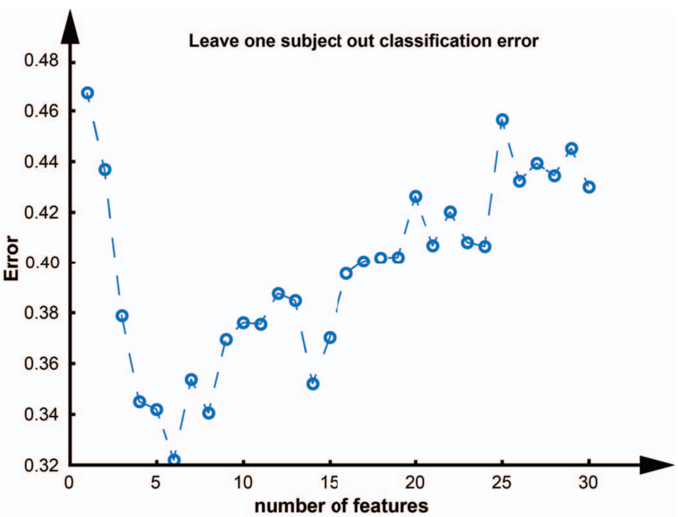


**Fig. 3**. Error rate as a function of the number of spectral features

Feature selection was performed on a subject basis, because GSO algorithms are fast and it can be implemented online. However, for a descriptive characterization of the underlying process, we performed *bagging* with the selected features of each individual, in order to choose the most representative ones. For each subject a set of features obtained by GSO was formed. The size of the set being the number of features that minimized the classification error for that individual. All these sets were merged into one, and features were ranked according to the number of times they appeared. As a result, most of the chosen features were present in most of the subjects. The final features were the following:

- relative theta power, electrode Fz
- relative upper gamma power, electrode Pz

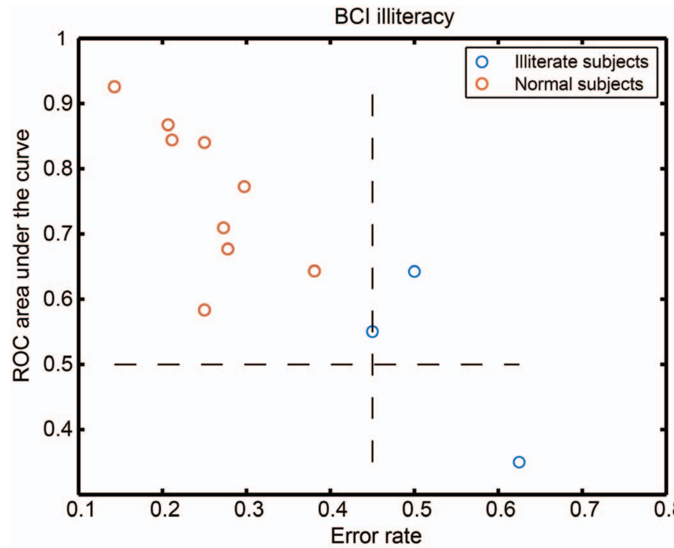


**Fig. 4**. Usability of the BCI prototype. Three subjects over twelve (25%) were considered illiterate.

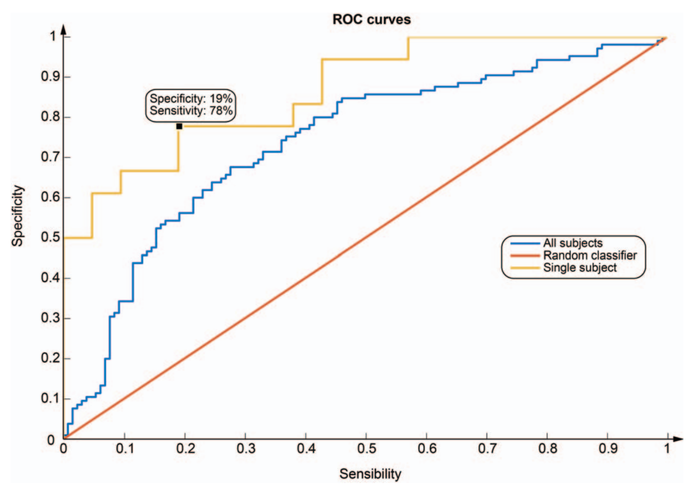


**Fig. 5**. ROC curves for the whole set of subjects, and a typical good subject

- upper beta power, electrode Fz
- relative lower beta power, electrode F4
- relative delta power, electrode Fp1
- relative alpha power, electrode O1
- lower gamma power, electrode Fz
- relative lower gamma power, electrode Fz
- relative upper beta power, electrode Pz

There are 9 features instead of 12 because 3 of the single features were in more than one cross term. It is important to note that although both, linear features and cross terms were included in the GSO procedure, all the selected features for all the subjects were cross terms. Single features then did not convey as much information as cross-terms. Figure 6 shows that for most of the selected features, classification power was rather poor when they were used alone.

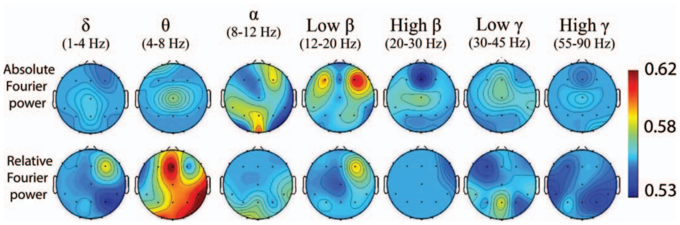


**Fig. 6**. Accuracy of classification for each single feature

## 4. DISCUSSION

By looking at the ROC curves, we can observe that the classifier consistently performs much better than a random classifier across the whole range of possible thresholds, and that this improves even more when we consider the classification on a subject basis. Furthermore, the threshold value can be adjusted depending on the expected applications. For instance, in a learning environment, we can afford skipping some true positives (high memory load), as long as we can be certain about the detected positives.

In [8], authors conclude that the last 30 years of BCI research have shown that between 20% and 30% of the users cannot use BCI systems with enough accuracy to achieve control. However, that paper addresses BCI for communication, not for monitoring as it is the case of this work. Hence, their claims, although useful for establishing a comparison, should not be blindly generalized. In our case, even though we kept illiterate subjects in our database, for a typical non-illiterate subject, we can observe that, if we are willing to accept only a 19% false positive rate, we can still accurately detect high memory load approximately 78% of the time.

The main features found in this study are consistent with literature reports of working memory. In [9], it is discussed that changes due to working memory load are often observed in the theta power of midline electrodes, and alpha power in occipital electrodes.

## 5. REFERENCES

[1] Jonathan R Wolpaw, Niels Birbaumer, William J Heetderks, Dennis J McFarland, P Hunter Peckham, Gerwin Schalk, Emanuel Donchin, Louis A Quatrano, Charles J Robinson, Theresa M Vaughan, et al., "Brain-computer interface technology: a review of the first international meeting," *IEEE transactions on rehabilitation engineering*, vol. 8, no. 2, pp. 164–173, 2000.

[2] Thilo Hinterberger, Andrea Kübler, Jochen Kaiser, Nicola Neumann, and Niels Birbaumer, "A brain–computer interface (bci) for the locked-in: comparison of different eeg classifications for the thought translation device," *Clinical Neurophysiology*, vol. 114, no. 3, pp. 416–425, 2003.

[3] Klaus-Robert Müller, Michael Tangermann, Guido Dornhege, Matthias Krauledat, Gabriel Curio, and Benjamin Blankertz, "Machine learning for real-time single-trial eeg-analysis: from brain–computer interfacing to mental state monitoring," *Journal of neuroscience methods*, vol. 167, no. 1, pp. 82–90, 2008.

[4] Thorsten O Zander and Christian Kothe, "Towards passive brain–computer interfaces: applying brain–computer interface technology to human–machine systems in general," *Journal of Neural Engineering*, vol. 8, no. 2, pp. 025005, 2011.

[5] Alan Baddeley, "Working memory," *Science*, vol. 255, no. 5044, pp. 556–559, 1992.

[6] Pernille J Olesen, Helena Westerberg, and Torkel Klingberg, "Increased prefrontal and parietal activity after training of working memory," *Nature neuroscience*, vol. 7, no. 1, pp. 75–79, 2004.

[7] N. Cowan, *Attention and memory: an integrated framework*, Oxford University Press, Oxford, Oxfordshire, 1995.

[8] Benjamin Blankertz, Claudia Sanelli, Sebastian Halder, Eva-Maria Hammer, Andrea Kübler, Klaus-Robert Müller, Gabriel Curio, and Thorsten Dickhaus, "Predicting bci performance to study bci illiteracy," 2009.

[9] Lars Michels, Kerstin Bucher, Rafael Lüchinger, Peter Klaver, Ernst Martin, Daniel Jeanmonod, and Daniel Brandeis, "Simultaneous eeg-fmri during a working memory task: modulations in low and high frequency bands," *PLoS One*, vol. 5, no. 4, pp. e10298, 2010.