35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Mental fatigue and working memory load estimation: Interaction and implications for EEG-based passive BCI

Raphaëlle N. Roy, Stéphane Bonnet, Sylvie Charbonnier and Aurélie Campagne

Abstract— **Current mental state monitoring systems, a.k.a. passive brain-computer interfaces (pBCI), allow one to perform a real-time assessment of an operator's cognitive state. In EEG-based systems, typical measurements for workload level assessment are band power estimates in several frequency bands. Mental fatigue, arising from growing time-on-task (TOT), can significantly affect the distribution of these band power features. However, the impact of mental fatigue on workload (WKL) assessment has not yet been evaluated. With this paper we intend to help fill in this lack of knowledge by analyzing the influence of WKL and TOT on EEG band power features, as well as their interaction and its impact on classification performance. Twenty participants underwent an experiment that modulated both their WKL (low/high) and time spent on the task (short/long). Statistical analyses were performed on the EEG signals, behavioral and subjective data. They revealed opposite changes in alpha power distribution between WKL and TOT conditions, as well as a decrease in WKL level discriminability with increasing TOT in both number of statistical differences in band power and classification performance. Implications for pBCI systems and experimental protocol design are discussed.**

I. INTRODUCTION

Brain-computer interfaces (BCIs) are information transfer systems between one's brain and a machine. They were initially designed to provide surrogate communication and motor abilities to handicapped people [1]. BCIs are mainly applied on neural correlates such as sensorimotor rhythm modulations and evoked visual potentials. Lately, the BCI framework has been applied to monitor an operator's cognitive and emotional state, as well as to implicitly adapt systems in real-time using only brain waves as an input [2]. Those interfaces, now known as 'passive' BCIs (pBCIs) [3], are the new means to answer neuro-ergonomics issues and mental state monitoring (MSM) systems.

Most of current pBCI systems are electro-encephalography (EEG)-based and include measures of workload and/or mental fatigue [4, 5]. Mental workload (WKL) has been extensively documented in the MSM literature and can be defined either as the load in working memory (i.e. number of items), the number of tasks to be performed simultaneously and more generally as a measure of the amount of mental resources engaged in a task. Thus, it is considered a measure of task difficulty [6], and depends on each individual's capabilities and effort [7]. Regarding those factors' neural correlates, previous studies showed that the band power in the theta and delta frequency bands at frontal sites increases with workload, while the band power in the alpha band at parietal sites decreases [8, 9, 10]. Regarding mental fatigue, or reduced alertness, it arises from growing time-on-task (TOT) and induces an increase of band power in the low frequency (LF; <12 Hz) bands, coupled with a decrease in the high frequency (HF) bands [11].

In real life situations, BCI systems are mainly calibrated only once before being used for a long period of time (e.g. for driving); i.e. when the participant is awake and fully responsive. Therefore, they do not take into account the possible impact of mental fatigue and reduced alertness. To date, no study has objectively evaluated the interaction of those factors at the electrophysiological level for such EEG-based systems. Yet, this interaction might prevent systems from clearly distinguishing WKL levels and thus reduce classification accuracy. This interaction could also reveal a major involvement of mental fatigue in generating non-stationarity in the EEG signal. Indeed, although such non-stationarity is known to appear between training and test sessions, physiological causes have never been thoroughly investigated.

In the present study, we focus on this undocumented interaction between WKL and mental fatigue and its impact on classification performance. In order to do so, we designed an experiment that manipulates both working memory load and TOT. We hypothesized that an increase in mental fatigue with TOT would create band power feature interferences and drown out WKL level differences. It would therefore negatively impact WKL level classification performance.

II. METHODS

This research was promoted by Grenoble's hospital (France) and was approved by the French ethics committee (ID number: 2012-A00826-37).

A. *Experimental design*

WKL was manipulated using a modified Sternberg paradigm [12]. At each trial, the 20 healthy participants (9 females; M = 25, $S.D.$ = 3.5 years) had to memorize a list of sequential digits visually presented on a computer screen. Then, a probe item flanked with question marks was

R. N. Roy and S. Bonnet are with the CEA/LETI/DTBS, MINATEC Campus, Grenoble, France (e-mail: raphaelle.roy@cea.fr, stephane.bonnet@cea.fr)
S. Charbonnier is with the Gipsa-Lab, Grenoble, France (e-mail: sylvie.charbonnier@gipsa-lab.grenoble-inp.fr)
A. Campagne is with the Laboratoire de Psychologie et Neurocognition (LPNC), Grenoble, France (e-mail: aurelie.campagne@upmf-grenoble.fr).

presented (Fig. 1). The participants had to answer as quickly as possible whether the probe was present or not in the memorized list using a response box. Two levels of WKL were considered, i.e. 2 and 6 digits to memorize (low and high WKL respectively). Two 10-minute blocks, each including 40 trials of each workload level, were performed. Trials of low and high WKL were pseudo-randomly presented. In order to induce a mental fatigue between the 2 blocks, participants also carried out an intermediate similar task during 50 minutes. Given that the task was repetitive and stimulus poor, this delay allowed us to presuppose 2 levels of mental fatigue depending on TOT (short/long).

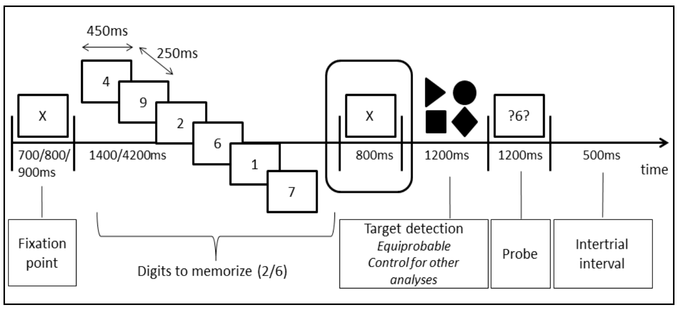

Figure 1. Trial structure. *Participants memorize a list of 2 or 6 digits, and answer whether the probe item was in the list. The circled segment is used for analyses. The next segment was added to perform other analyses that will be described in another paper.*

B. Data acquisition and pre-processing

Workload and mental fatigue manipulation was confirmed thanks to behavioral and subjective measures. Participants' reaction times (RTs) and accuracy were measured, as well as their answers to a mental fatigue questionnaire (Karolinska questionnaire) before and at the end of the experiment, and between the 2 blocks. In addition, we recorded participants' EEG activity using a BrainAmp^{TM} system (Brain Products, Inc.) and an Acticap® equipped with 32 Ag-AgCl active electrodes that were positioned according to the extended 10-20 system. The reference and ground electrodes used for acquisition were those of Acticap, i.e. FCz for the reference electrode and AFz for the ground electrode. The data were sampled at 500 Hz. The electro-oculographic (EOG) activity was also recorded using 2 electrodes positioned at the eyes outer canthi, and 2 respectively above and below the left eye. The EEG signal was band-pass filtered between 1 and 40 Hz, re-referenced to a common average reference and corrected for ocular artifacts using the signal recorded from the EOG electrodes and the SOBI algorithm [13]. Time segments of 800 ms of signal were then selected (circled on Fig. 1). This epoching step was performed to only analyze time segments in which participants were loaded and had not yet performed the recognition task. This way, we avoided analyzing neural correlates associated with memory-encoding and memory-scanning processes.

C. Analyses

To assess the significance of each factor's effect, statistical analyses were performed in 2 different ways: First, the averaged power of the EEG signal in the 5 frequency bands: delta [1-4 Hz], theta [4-8 Hz], alpha [8-12 Hz], beta [12-30 Hz], and gamma [30-40 Hz] was estimated using Welch's power spectral density estimation. The changes in band power on midline electrodes were then analyzed across participants using repeated mesures ANOVAs and Newman-Keuls posthoc tests. Behavioral and subjective data were analyzed in the same way. Secondly, participant-specific classifiers were built to detect TOT and WKL levels. The processing chain used to perform classification is as follows: first, the EEG signal was divided into the 5 frequency bands. For each band, 15 electrodes were then selected using a method based on Riemannian geometry [14]. Next, a spatial filtering step was executed using 6 common spatial pattern (CSP) filters. Lastly, a binary classification was performed across experimental conditions using Fisher's linear discriminant analysis (FLDA) with 30 features (6 spatial filters x 5 frequency bands). Using a 10-fold random cross-validation procedure, we obtained results in classification performance for TOT and WKL levels independently, as well as for WKL level per mental fatigue state (WKL|sTOT and WKL|lTOT respectively for the short and long TOT conditions). In order to put ourselves in a more realistic classification situation, we also added a condition in which we trained our workload classifier on the short TOT data, and tested it on the long TOT data (WKL|diff). This condition is supposed to mirror a system that would be calibrated before use, and then be used during a long period of time. Classification performances were compared using t-tests.

III. RESULTS

A. Behavioral & subjective results

Participants reported feeling increasingly tired with TOT ($p<0.001$, Fig. 2). As regards the WKL effect, participants were slower to respond and had a lower accuracy in high WKL conditions than in low ones ($p<0.001$).

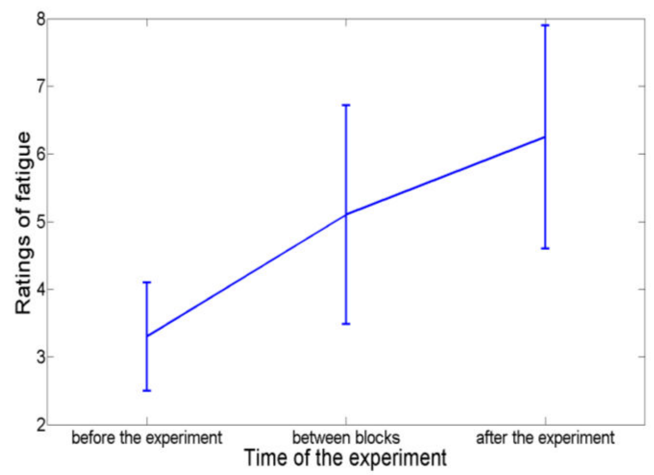

Figure 2. Subjective ratings of mental fatigue at different times of the experiment ($p<0.001$) *Scale ranging from 1 'highly alert state' to 10 'very drowsy with great effort to stay awake, fighting against sleep'.*

B. EEG results

Across all participants, the band power in the alpha and beta frequency bands significantly decreased at all midline electrodes with increasing load in working memory (p<0.05). Fig. 3 is a topographic map of signed r² [15] that illustrates WKL level discriminability for the alpha band (grand average). Here, it can be seen that the relevant electrode locations for WKL discriminability are mostly centro-parietal ones, as those electrode locations are highly correlated with the low WKL condition.

Moreover, the power in the alpha frequency band – especially the low alpha- increased with growing TOT for all midline electrodes (p<0.05), and the power in the delta, theta and beta bands also significantly increased for electrodes Cz, CPz and Pz (p<0.05). Those modulations are illustrated by Fig. 4 in which we can see a clear distinction between conditions around 8 to 12 Hz (alpha band) and by Fig. 5 that displays the alpha power scalp distribution per WKL and then per TOT condition. In this last figure we can see that opposite phenomena take place: alpha power decreases with WKL (mostly at centro-parietal sites) whereas it increases with TOT (mostly at fronto-central sites).

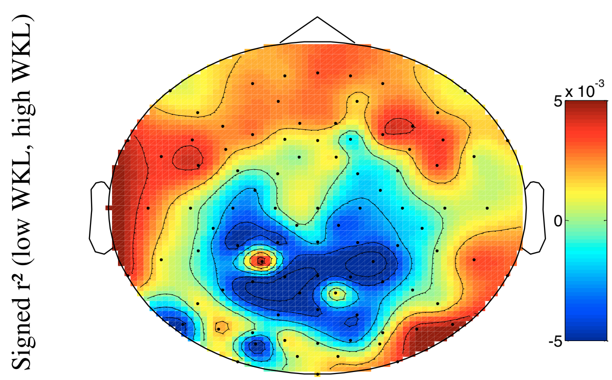

Figure 3. Scalp topography of the signed coefficients of determination (SCDs, or signed r² values) for the alpha band power and workload levels. *Indicates discriminability power of alpha band power between conditions of WKL (Grand average).*

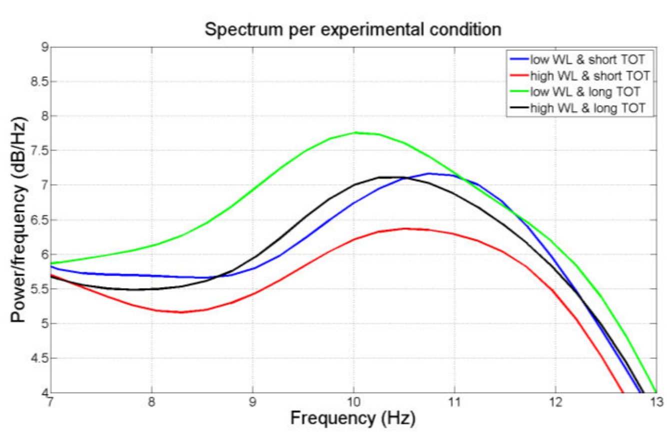

Figure 4. Power density spectrum on electrode Pz *(Grand average; wkl: workload; TOT: time-on-task)*

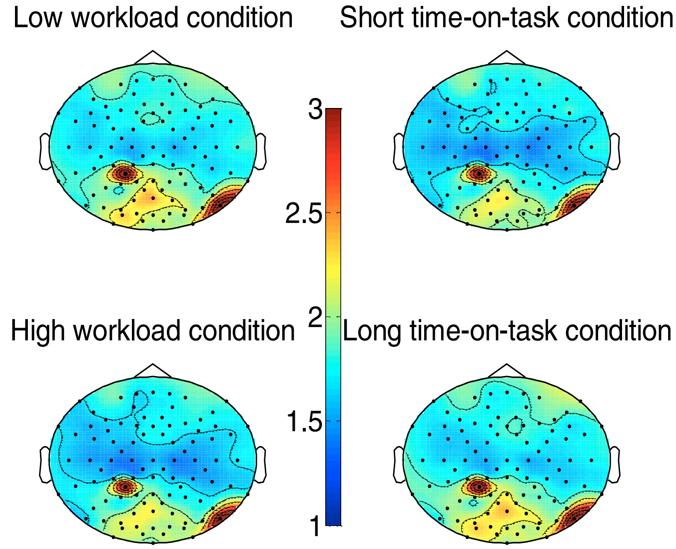

Figure 5. Averaged alpha power scalp distributions *(µV²)* depending on WKL (Left: low/high WKL) or TOT condition (Right: short/long TOT)

Regarding the interaction between WKL and TOT (Fig. 6), there were more significant differences between WKL conditions per frequency band and electrode in the short TOT condition than in the long TOT condition. For instance, we can see that differences in theta power at Fz disappear when TOT increases, as well as differences in high alpha power for anterior electrode sites, and in beta power at the CPz, Pz and Oz electrodes. The disappearance of differences in high alpha power is reflected in Fig. 4 by a peak shift between TOT conditions.

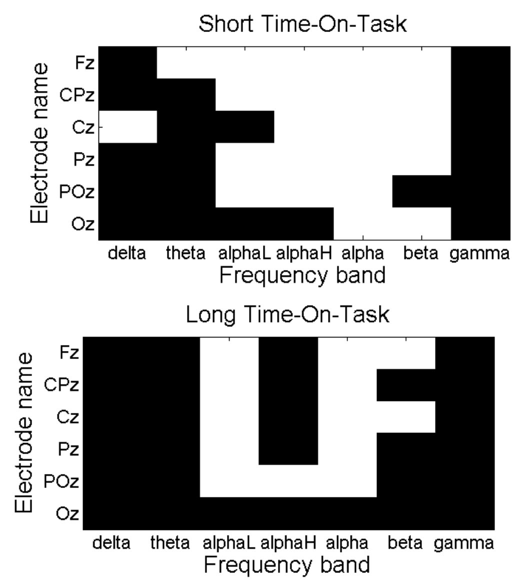

Figure 6. Significant differences across participants (in white) between WKL conditions per midline electrode, frequency band and TOT condition *(p<0.05)*

C. Classification results

It can be seen on Table I that our processing chain achieved 98.04 % of correct classification of the TOT condition, on average for the 20 subjects. As for the WKL condition, the performance was 65.51 % independently of the TOT condition. This classification performance decreased to 59.76 % and 58.75 % of correct classification when we trained and tested our WKL classifier respectively with the short TOT data (WKL|sTOT) and the long TOT data (WKL|lTOT). Moreover, when we trained our workload classifier on the short TOT data and tested it on the long TOT data, we dropped to 50.10 % of correct classification (WKL|diff), which is not different from random ($p = 0.11$). An interesting result is that classification performance significantly dropped from 59.76 to 50.10 % between the condition in which the classifier is trained and tested on data recorded in the same condition (WKL|sTOT), and the condition in which it is trained on WKL|sTOT and tested on WKL|lTOT (WKL|diff ; $p<0.001$).

TABLE I Classification performances for each factor and condition (mean and sd). *Averaged results of a 10-fold random cross-validation procedure repeated 10 times*

Condition	TOT	WKL	WKL|sTOT	WKL|lTOT	WKL|diff
% of correct classifi-cation	**98.04 (2.01)**	**65.51 (5.93)**	**59.76 (7.18)**	**58.75 (6.65)**	**50.10 (0.29)**

IV. CONCLUSION & OUTLOOK

High classification performance was obtained for mental fatigue state (98.04 %), with a standard amplification of band power for LF bands with increasing TOT. However, we did not observe a significant decrease in band power for HF bands. Besides, given the length of the analyzed time period (800 ms), we achieved workload classification performances comparable to those reported in the literature for 2.5 s windows [16]. As reported in the literature, we do observe a decrease in alpha power with an increase in working memory load, also mostly at the centro-parietal sites. This decrease in alpha power is also reported by [10] who speak about it in terms of increasing alpha band desynchronization with ascending cognitive load. In our study we also reported an expected rise in theta power at frontal sites with growing workload, but only in the short TOT condition. This result contributes to illustrate the significant impact of mental fatigue on the EEG power feature data, impact that resulted in a significant degradation of classification performances as expected. What's more, even when the classifier was trained and tested on data recorded in the same condition, the long TOT one, we could observe a slight decrease in classification performance compared with the short TOT one. This can be explained by an increase in power in the LF bands, increase that drowns out the subtle differences that would allow a good classification of workload level.

These results bring to light a real flaw in using band power features such as alpha band desynchronization for cognitive state assessment as those are subject to major arousal related fluctuations. This calls for the development of adaptive systems and for in-depth research of more robust features of cognitive state. Lately, several methods to alleviate non-stationarities have been proposed, such as work on common spatial patterns [17]. However, there is still upstream research to be performed in order to evaluate how cognitive states interact at both the source and the feature level. This research would also benefit context-aware computer systems that make use of covert aspects of the ongoing user state [18].

REFERENCES

[1] J. R. Wolpaw and E. W. Wolpaw, "Brain-Computer Interfaces: Something new under the sun", in *Brain-Computer Interfaces: Principles and Practice*. New York: Oxford University Press, 2012.

[2] L. George and A. Lécuyer, "An overview of research on "passive" brain-computer interfaces for implicit human-computer interaction", in *Int. Conf. on Appl. Bionics and Biomechanics*, Venice, Italy, 2010.

[3] T. O. Zander and C. Kothe, "Towards passive brain–computer interfaces: applying brain–computer interface technology to human–machine systems in general", *J. of Neural Eng.*, vol. 8, 2011.

[4] B. Blankertz, M. Tangermann, C. Vidaurre, S. Fazli, C. Sannelli, S. Haufe, C. Maeder, L. Ramsey, I. Sturm, G. Curio, and K.-R. Müller, "The Berlin Brain–Computer Interface: Non-Medical Uses of BCI Technology", *Frontiers in Neuroscience*, vol. 4, 2010.

[5] J. van Erp, F. Lotte and M. Tangermann, "Brain-Computer-Interfaces: Beyond Medical Applications", *Computer*, vol. 45, pp.26-34, 2012.

[6] A. Gevins and M.E. Smith "Electroencephalography (EEG) in Neuroergnomics", in *Neuroergonomics: The brain at work,* R. Parasuraman, R and M. Rizzo Eds. New York, NY: Oxford University Press, 2007, pp. 15-31.

[7] B. Cain. (2007) "A review of the mental workload literature", Defense Research and Development, Toronto, Canada, [Online]. Available: http://ftp.rta.nato.int/public//PubFullText/RTO/TR/RTO-TR-HFM-121-PART-II///TR-HFM-121-Part-II-04.pdf

[8] F. Schober, R. Schellenberg, and W. Dimpfel, "Reflection of mental exercise in the dynamic quantitative topographical EEG", *Neuropsychobiology*, vol. 31, pp. 98-112, 1995.

[9] A. Holm, K. Lukander, J. Korpela, M. Sallinen, and K. M. I. Müller, "Estimating Brain Load from the EEG", *TheScientificWorldJOURNAL*, vol. 9, pp. 639-651, 2009.

[10] A. Stipacek, R. H. Grabner, C. Neuper, A. Fink, and A. C. Neubauer, " Sensitivity of human EEG alpha band desynchronization to different working memory components and increasing levels of memory load", *Neuroscience Lett.*, vol. 353, n° 3, p. 193 - 196, 2003.

[11] W. Klimesch, "EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis", *Brain Research Reviews*, vol. 29, pp. 169-195, 1999.

[12] S. Sternberg, "High-speed scanning in human memory", *Science*, vol. 153, pp. 652-654, 1966.

[13] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso and E. Moulines, "A blind source separation technique using second-order statistics", *IEEE Trans. Signal Process.*, vol. 45, n°2, 1997, pp. 434-444.

[14] A. Barachant and S. Bonnet, "Channel selection procedure using Riemannian distance for BCI applications", in *Proc. IEEE Conf. on Neural Eng.*, Cancun, 2011, pp. 348-351.

[15] B. Blankertz, S. Lemm, M. Treder, S. Haufe, and K.-R. Müller, "Single-trial analysis and classification of ERP components — A tutorial", *NeuroImage*, vol. 56, n° 2, pp. 814-825, 2011.

[16] A.-M. Brouwer, M. A. Hogervorst, J. B. F. van Erp, T. Heffelaar, P. H. Zimmerman, and R. Oostenveld, " Estimating workload using EEG spectral power and ERPs in the n-back task ", *J. of Neural Eng.,* vol. 9, n° 4, 2012.

[17] B. Blankertz, M. Kawanabe, R. Tomioka, F. U. Hohlefeld, V. Nikulin and K.-R. Müller, "Invariant Common Spatial Patterns: Alleviating Nonstationarities in Brain-Computer Interfacing", in *Advances in Neural Inform. Process. Syst.*, vol. 20, pp. 113-120, Vancouver, Canada, 2008.

[18] T. O. Zander and S. Jatzev, "Context-aware brain–computer interfaces: exploring the information space of user, technical system and environment", *J. of Neural Eng.*, vol. 9, n°. 1, 2012.