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

Probing ECG-based mental state monitoring on short time segments

Raphaëlle N. Roy, Sylvie Charbonnier and Aurélie Campagne

Abstract— **Electrocardiography is used to provide features for mental state monitoring systems. There is a need for quick mental state assessment in some applications such as attentive user interfaces. We analyzed how heart rate and heart rate variability features are influenced by working memory load (WKL) and time-on-task (TOT) on very short time segments (5s) with both statistical significance and classification performance results. It is shown that classification of such mental states can be performed on very short time segments and that heart rate is more predictive of TOT level than heart rate variability. However, both features are efficient for WKL level classification. What's more, interesting interaction effects are uncovered: TOT influences WKL level classification either favorably when based on HR, or adversely when based on HRV. Implications for mental state monitoring are discussed.**

I. INTRODUCTION

In today's technological environment, mental state monitoring is a quickly expanding field. Its applications are numerous, ranging from gaming to education, including driving and security. What are called biocybernetic systems, or physiologically attentive user interfaces, are systems that take the user's covert aspects into account to adapt its functionality [1]. The assumption is that modulations in physiological features reflect modulations in the operator mental state (cognitive or emotional). Such covert aspects include workload and mental fatigue. First, mental workload (WKL) 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 [2], and depends on each individual's capabilities and effort [3]. Second, mental fatigue is a gradual and cumulative process associated with reduced alertness. It arises notably from growing time-on-task (TOT) [4].

Physiological measurements of an operator's cognitive state used in the mental state monitoring framework include direct measurements of mental activity such as electroencephalography (EEG), and indirect ones such as electrooculography, and electrocardiography (ECG) [4]. ECG has for main advantages of being noninvasive and convenient for daily living measurements. Although ECG has become widely used, the impact of TOT on ECG-based mental workload assessment has yet to be evaluated. Indeed, if such an influence exists, it may reduce the relevance of such ECG features in some conditions, and therefore call for features from other modalities such as EEG ones. Moreover, an assessment of ECG feature usability with very short time periods would benefit physiologically attentive user interfaces that could therefore prevent users from taking rapid and potentially erroneous decisions when detecting an abnormally high workload or mental fatigue. In this study, we assess the impact of working memory load and time-on-task on basic ECG features on short time periods and compare their relevance for mental state assessment using both statistical analyses and a classification performance evaluation. The results allow us to select the appropriate feature for each mental state for short time period assessment.

R. N. Roy is 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 Psychology and Neurocognition Laboratory (LPNC), Grenoble, France (e-mail: aurelie.campagne@upmf-grenoble.fr).

II. ECG FEATURES

ECG analyses are mainly based on the measurement of the RR interval. The R peak is the peak of the ECG wave with the highest amplitude, hence easier to detect than the other components. It is part of the well-known QRS complex and arises from ventricular depolarization [5, 6]. In order to compute the interval between two heart beats, it is therefore common practice to measure the time duration between two adjacent R peaks, as follows: $IBI_n = r_n - r_{n-1}$ with IBI as the interbeat interval and r_n the index of the n^{th} R peak. $\overline{IBI}_k$ is the mean IBI over the k^{th} analyzed time segment. The analysis of the interbeat interval allows for computing the mean heart rate (HR) over a period of time, expressed in beats per minute (bpm): $HR_k = \frac{60}{\overline{IBI}_k}$.

Heart rate variability (HRV) is another frequently used ECG feature. It reflects modulations in instantaneous heart rate over a given time period, usually several minutes. It is considered a good quantitative marker of the autonomic nervous system (ANS) activity [6]. HRV can be computed in the time domain, or in the frequency domain. In the time domain, a simple yet effective feature is the standard deviation of the IBI [6]:

$$HRV_k = \sqrt{\frac{1}{N-1}\sum_{n=0}^{N-1}(IBI_{k+n} - \overline{IBI}_k)^2}$$

with HRV_k as the HRV in the temporal domain for the k^{th} analyzed segment, and $\overline{IBI}_k$ as the mean IBI over the given segment. The HRV in the time domain has been shown to be most effective in WKL classification compared with other features [7].

The HRV in the frequency domain can be computed from the beat index. The beat index, or RR interval tachogram, is the representation of the RR interval duration as a function of the number of progressive beats [6]. The spectrum of this beat index is then found using for instance Welch power spectrum density method. Typically, two main components of the spectrum can be defined: a low frequency (LF) component (0.04-0.15 Hz) and a high frequency (HF) one (0.15-0.40 Hz). The HF component, also known as vagal tone, is thought to reflect parasympathetic activity [5, 6]. Although the physiological correlates of the HF component are known, there is no consensus regarding the LF one. It may reflect sympathetic activity [6]. Therefore the ratio LF/HF is often used to observe fluctuations in the sympathetic-vagal balance [5]. Both HR and HRV are markers of arousal and engagement and have been used to characterize levels of alertness and WKl. Typical findings are reported in Table I.

The influence of the interaction between TOT and task demands on ECG features was recently studied by Fairclough et al. [9]. They showed that the higher portion of the LF component of HRV (0.08-0.13Hz) increases with TOT. They also showed that it decreases with task demands. But then, after a long period of time (one hour), the trend is reversed and it increases with task demands. Thus, the present study was designed to investigate joint effects of TOT and WKL on both features and classification performance and we expect TOT to impact WKL classification, as already shown for EEG features [17].

Table I Workload and Time-on-task effects on heart rate and heart rate variability as reported by the literature. *WKL: workload; TOT: Time-on-task*

	Increase in WKL	Increase in TOT
HR	Increase [8,9,10,11,13]	Decrease [4,8,14]
HRV	Decrease [7]	Increase [12]
HF	Decrease [9,10,14,15]	Increase [9,10,16]
LF	Increase [9,10] / Decrease [9]	Decrease[12] / Increase [9]

III. 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

Workload was manipulated using a modified Sternberg paradigm [18]. At each trial, the 19 participants (9 females; 24.9 +/- 3.7 years old in average) had to memorize a list of sequential digits visually presented on a computer screen. Then, a probe item flanked with question marks was 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 WKL level, were performed. Trials of low and high WKL were pseudo-randomly presented. In order to induce a mental fatigue between the two 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 two levels of mental fatigue depending on TOT (short/long).

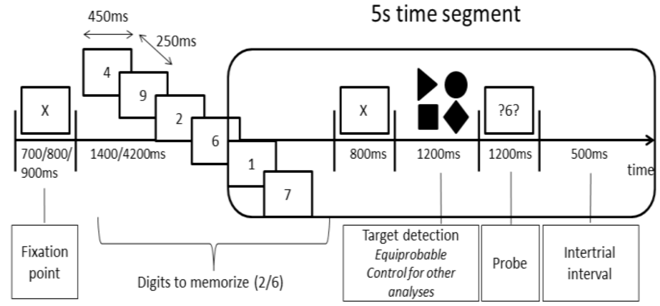

Figure 1. Trial structure. *The participants memorize a list of digits (2 or 6), and answer whether the probe was in the list. The circled segment is used for analyses. The next segment was added to perform other analyses that will be described in another paper.*

B. Data acquisition & preprocessing

WKl 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 [19]) before and at the end of the experiment, and between the 2 blocks. Participants' ECG signal was also recorded using Ag/AgCl electrodes positioned at the sternum and fifth intercostal space of the left ribcage. Data were processed and analyzed offline. The signal was sampled at 500 Hz and filtered between 1 and 40 Hz. We detected the R peaks using a threshold equal to half the amplitude of the segment's maximum. A refractory period of 200ms was also used, as it is physiologically impossible to have an adjacent QRS complex before this delay [20]. 5-s time segments were then selected (circled on Fig. 1). This epoching step was performed in order to analyze time segments of equal duration between conditions of WKL. Indeed, one should not compare segments of different duration when willing to perform HRV analyses [6]. Thus, a total of 160 segments were available per participant: 80 per WKL or TOT condition (high/low, short/long), 40 per WKL and TOT condition (high WKL & short TOT, low WKL and short TOT, etc.). Trial values were considered outliers and rejected when they exceeded 2 standard deviations from the mean.

C. Statistical analyses & classification

To assess the significance of each factor's effect, statistical analyses were performed in 2 different ways: First, we analyzed the HR and HRV features across all participants using repeated measures ANOVAs and Tukey post-hoc tests with WKL and TOT as factors. Behavioral and subjective data were analyzed in the same way. Secondly, participant-specific classifiers were built to detect TOT and WKL levels. The classifiers used were Fisher LDA, with either one (HR or HRV) or 2 features (HR and HRV: ALL). Using a 10-fold random cross-validation procedure, for each feature or feature combination we obtained results in classification performance for TOT and WKL levels independently, as well as for each feature or feature combination per TOT condition (STOT: short TOT ; LTOT long TOT). Classification performances were compared using single sample t-tests to test them against random (0.50). Then, we performed repeated measures ANOVAs with Tukey post-hoc tests to assess the significance of the influence of feature choice on TOT or WKL level classification, as well as the impact of TOT condition on WKL level classification using different features. Lastly, for illustration purposes, the averaged HRV in the frequency domain was computed per TOT level over the whole signal using an auto-regressive model (order 20).

IV. RESULTS

A. *Behavioral, subjective & ECG results*

Participants reported feeling increasingly tired with TOT ($p<0.001$). As regards the WKL effect, participants were slower to respond and had a lower accuracy in the high WKL condition than in the low one ($p<0.001$). Across all participants, ANOVAs and Tukey post-hoc tests revealed that HRV significantly increased with TOT for the low load condition ($p<0.01$), whereas no significant TOT effect was observed for HR. As for WKL, HR significantly decreased with an increase in WKL, and there was also a tendency for a decrease in HRV, both in the long TOT condition ($p<0.01$ and $p=0.08$ respectively).

B. *Classification results*

The classification performance results are given in Table II.

TABLE II Classification performance for each factor and feature (mean proportion of correct classification and sd) and statistical results. *TOT: Time-on-task (S: short, L:long); WKL: workload; HR: heart rate; HRV: heart rate variability in the time domain; ALL: HR & HRV.*

Factors & Features		M	SD	T-test single sample against 0.5	ANOVA & Tukey post-hoc test
TOT	HR	0.64	0.13	* p<0.001	TOT > WKL (p=0.06) HR & ALL > HRV (*p<0.001), but only for TOT (*p<0.001)
	HRV	0.56	0.12	Tendency p=0.06	
	ALL	0.65	0.13	*p<0.001	
WKL	HR	0.56	0.05	*p<0.001	
	HRV	0.54	0.06	*p<0.05	
	ALL	0.57	0.05	*p<0.001	
WKL HR	STOT	0.54	0.09	Tendency p=0.09	Opposite pattern: HR STOT<LTOT, HRV STOT>LTOT but n.s.
	LTOT	0.58	0.07	*p<0.001	
WKL HRV	STOT	0.57	0.10	*p<0.01	
	LTOT	0.54	0.07	*p<0.05	
WKL ALL	STOT	0.57	0.08	*p<0.001	
	LTOT	0.57	0.08	*p<0.01	

1) *TOT classification results*

We can see that classification performances were almost all significantly different from random. TOT classification performance reached 65 % of correct classification when both HR and HRV were used as input features, with HR as the most discriminative one. Fig. 2, a Poincaré plot that displays IBI fluctuations per TOT condition, illustrates this phenomenon for one participant. For this participant, the shift with TOT of the feature distribution is quite clear. Although classification of TOT is less accurate with HRV (56 %), TOT effect is clearly visible in the frequency domain, as illustrated by Fig. 3. Indeed, there was a clear increase in the HF component with increasing TOT. Thus, the ratio LF/HF decreased with TOT.

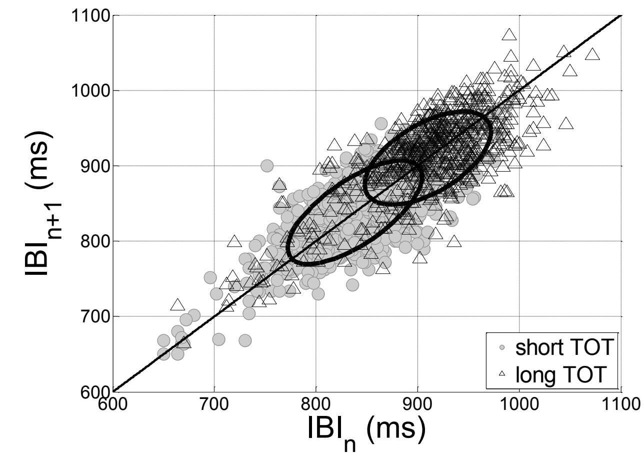

Figure 2 Poincaré plot that displays IBI fluctuations per TOT condition (1 participant). Conditionwise covariance matrices indicate a fitted Gaussian distribution by equidensity contours at 1.5 s.d.

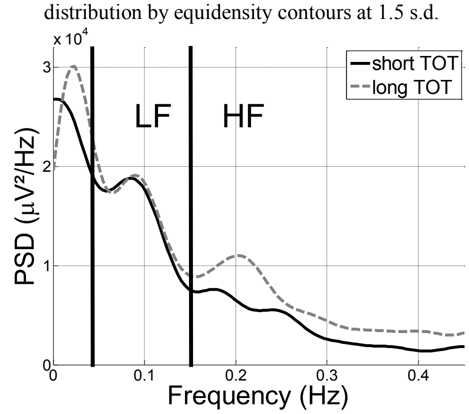

Figure 3 Average PSD of HRV per TOT condition across all participants.

2) *WKL classification results*

Regarding WKL classification performance, it reached 57 % of correct classification when both features were used, although no significant difference was found between feature choice conditions. Fig. 4 illustrates the feature distribution for 1 participant. In addition, we observed opposite patterns of performance depending on feature and TOT condition: performance was higher for HR-based WKL estimation in the LTOT condition, whereas they were higher for HRV-based estimation in the STOT condition.

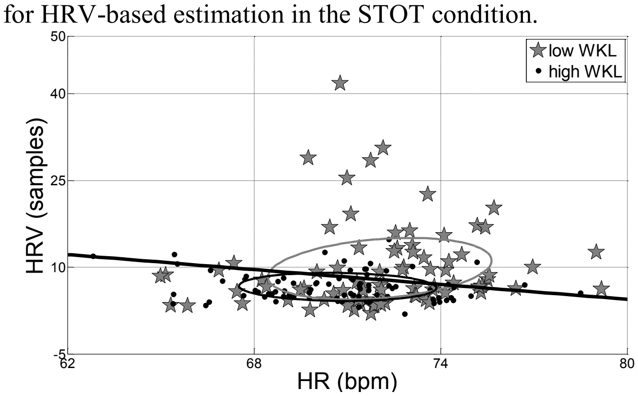

Figure 4 Feature distribution per WKL condition (1 subject). Conditionwise covariance matrices indicate a fitted Gaussian distribution by equidensity contours at 1.5 s.d. The line is the FLDA separating hyperplane. *HR: mean heart rate; HRV: heart rate variability (time domain).*

V. DISCUSSION

As expected, statistical analyses revealed a negative influence of WKL and TOT condition on both behavioral performance and subjective feeling. Regarding TOT, in keeping with the literature, the HF component of the HRV increased with it [9, 10, 16]. Moreover, although classification was performed on very short time segments (5s), good performance was achieved using HR and HRV for TOT assessment (65 %). Compared to HRV, the HR feature gave best results for the classification of this factor. What is more, the TOT effect was found to be conditional on the WKL level. Indeed, contrary to what could be expected from [9], we did not find the trend reversal in the WKL modulation of HRV after a long TOT. But it appeared that the classical increase of HRV with TOT was only present in the low load condition. This phenomenon was never reported in the literature, and may be due to an effort of the participants to stay alert for the more demanding trials, hence increasing their amount of recruited cognitive resources and diminishing the impact of TOT on HRV.

As for WKL, HRV was as useful as HR for its estimation, and their combined use allowed us to reach a rate of 57 % of good WKL classification. Yet, the WKL effect was also found to be conditional on the TOT level. Indeed, HR and HRV both decreased with an increase in WKL in the LTOT condition, a phenomenon reflected by a better HR-based WKL estimation in the LTOT than in the STOT condition (58 % vs. 54 % respectively). On the other hand, HRV-based estimation was better in the STOT then in the LTOT condition (57 % vs. 54 % respectively). It should be noted that the decrease of HR with WKL in the LTOT condition may be due to the monotony of the digit presentation in the high load condition, as opposed to the rapidly engaging low load trials.

This study has allowed us to investigate interaction effects that occur between TOT and WKL and that affect ECG features, hence potentially distorting the classifiers' output. In the future, we plan on focusing on these effects and on analyzing individual differences more thoroughly. Indeed, when asked at the end, several participants reported feeling more tired at the beginning of the experiment than at the end, even though they had rated their mental fatigue otherwise. If they indeed presented this pattern opposite to what we expected, their features, when averaged with the others', may have cancelled out major effects. Moreover, even though some discrepancies with the literature have been found, this study has assessed the usability of HR and HRV features for mental state monitoring on very short time segments. Yet, they can be subject to interactions and are nonspecific to a given mental state. This calls for in-depth research of more robust features and for modality fusion. By now, some real-time assessment systems that include several recording modalities have been designed [21].

ACKNOWLEDGMENT

The authors thank Stéphane Bonnet, PhD, from the CEA-LETI for his strong implication in this study.

REFERENCES

[1] D. Chen and R. Vertegaal, "Using mental load for managing interruptions in physiologically attentive user interfaces", in *Human factors in computing systems*, pp. 1513-1516, 2004.

[2] 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, pp. 15-31, 2007.

[3] 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

[4] S. K. L. Lal and A. Craig, "A critical review of the psychophysiology of driver fatigue", *Biol. Psych.*, vol. 55, nº 3, pp. 173−194, 2001.

[5] G. G. Berntson, J. T. Bigger Jr, D. L. Eckberg, P. Grossman, P. G. Kaufmann, M. Malik, H. N. Nagaraja, S. W. Porges, J. P. Saul, P. H. Stone, and M. W. van der Molen, "Heart rate variability: origins, methods, and interpretive caveats", *Psychophysiology*, vol. 34, nº 6, pp. 623-648, 1997.

[6] M. Malik, J. T. Bigger, A. J. Camm, R. E. Kleiger, A. Malliani, A. J. Moss, and P. J. Schwartz. "Heart rate variability: Standards of measurement, physiological interpretation, and clinical use". *European Heart J.*, vol. 17, n°3, pp. 354-381, 1996.

[7] A. N. Kalinichenko and O. D. Yur'eva, "Assessment of human psychophysiological states based on methods for heart rate variability analysis", *Patt. Recog. and Image Anal.*., vol. 22, pp. 570-575, 2012.

[8] G. Borghini, G. Vecchiato, J. Toppi, L. Astolfi, A. Maglione, R. Isabella, C. Caltagirone, W. Kong, D. Wei, Z. Zhou, L. Polidori, S. Vitiello, and F. Babiloni, "Assessment of mental fatigue during car driving by using high resolution EEG activity and neurophysiologic indices", in *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, San Diego, 2012, pp. 6442-6445.

[9] S. H. Fairclough, L. Venables, and A. Tattersall, "The influence of task demand and learning on the psychophysiological response", *Int. J. of Psychophysiology*, vol. 56, nº 2, pp. 171-184, 2005.

[10] A. R. Subhani, X. Likun, and A. Saeed Malik, "Association of Autonomic Nervous System and EEG Scalp Potential During Playing 2D Grand Turismo 5", in *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, San Diego, 2012, pp. 3420-3423.

[11] L. Salahuddin, J. Cho, M. G. Jeong, and D. Kim, "Ultra short term analysis of heart rate variability for monitoring mental stress in mobile settings ", in *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, Lyon, France, 2007, pp. 4656-4659.

[12] N. Pattyn, X. Neyt, D. Henderickx, and E. Soetens, "Psychophysiological investigation of vigilance decrement: Boredom or cognitive fatigue?", Physiology & Behavior, vol. 93, nº 1‑2, pp. 369-378, 2008.

[13] G. Larue, A. Rakotonirainy and A. N. Pettitt, "Driving performance on monotonous roads", in *Proc. of 20th Canadian Multidisciplinary Road Safety Conf.*, Niagara Falls, Ontario, June, 2010.

[14] S. H. Fairclough and L. Venables, "Prediction of subjective states from psychophysiology: A multivariate approach", *Biol. Psych.*, vol. 71, nº 1, 2006.

[15] L. R. Fournier, G. F. Wilson, and C. R. Swain, "Electrophysiological, behavioral, and subjective indexes of workload when performing multiple tasks: manipulations of task difficulty and training", *Int. J. of Psychophysiology*, vol. 31, nº 2, pp. 129−145, 1999.

[16] C. Zhao, M. Zhao, J. Liu, and C. Zheng, "Electroencephalogram and electrocardiograph assessment of mental fatigue in a driving simulator", *Accident Analysis & Prevention*, pp. 83–90, 2012.

[17] R. N. Roy, S. Bonnet, S. Charbonnier and A. Campagne, "Mental fatigue and working memory load estimation: Interaction and implications for EEG-based passive BCI", to appear in *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, 2013.

[18] S. Sternberg, "High-speed scanning in human memory", *Science*, vol. 153, nº 3736, pp. 652−654, 1966.

[19] T. Akerstedt and M. Gillberg, "Subjective and objective sleepiness in the active individual", Int. J. Neurosci., vol. 52, n°12, pp. 29 37, 1990.

[20] L. Sörnmo and P. Laguna, *Bioelectrical signal process. in cardiac and neurological applicat.*. Amsterdam; Elsevier Ac. Press, 2005.

[21] J. Jarvis, F. Putze, D. Heger, and T. Schultz, "Multimodal person independent recognition of workload related biosignal patterns", in *Proc. of the 13th Int. Conf. on Mult..Interfaces*, 2011, pp. 205-208.