Monitoring of mental workload levels

Cinaz, Burcu ; La Marca, Roberto ; Arnrich, Bert ; Tröster, Gerhard

MONITORING OF MENTAL WORKLOAD LEVELS

Burcu Cinaz, Roberto La Marca, Bert Arnrich and Gerhard Tröster
ETH Zurich, Electronics Laboratory
Gloriastrasse 35, 8092 Zurich, Switzerland

ABSTRACT

Mobile healthcare applications offer new opportunities to prevent long-term health damage due to increased mental workload by continuously monitoring physiological signs related to prolonged high workload and providing just-in-time feedback. In order to achieve a day-by-day quantification of mental load, first different load levels which occur during a workday have to be discriminated. This work goes one step towards this goal: we present our experiment design and preliminary results in discriminating different levels of mental workload based on heart rate features obtained from a mobile ECG system. Based on the subjective ratings of the participants under study, we show that all participants perceived the induced load levels as intended from the experiment design. The heart rate variability (HRV) features under investigation could be classified into two distinct groups. Features in the first group, representing markers associated with parasympathetic nervous activity, show a decrease in their values with increased workload. Features in the second group, representing markers associated with sympathetic nervous activity, show an increase of their values with increased workload. These results provide evidence that a mobile ECG system is suited to discriminate different levels of mental workload. This would enable the development of mobile applications to monitor mental workload levels and to prevent long-term damage by giving early warning signs in case of prolonged high workload.

1. INTRODUCTION

Recently, the European Foundation for the Improvement of Living and Working Conditions called the attention on the increasing level of mental disorders due to work-related stress. The workplace has changed due to globalization, use of new information and communication technology, resulting in an increased mental workload. Work-related stress was found to be the second most common work-related health problem across the EU15 [European Foundation, 2007]. Work-related stress occurs when there is a mismatch between job load and the capabilities, resources or needs of the worker. If the worker is not able to recover, long-term damage may result in the development of mental disorders [Van Daalen et al., 2009]. Mobile healthcare offers new opportunities to prevent long-term damage by continuously monitoring mental workload and providing just-in-time feedback increasing workers awareness for improving self-management of mental workload.

Mobile monitoring of work-related stress or mental workload is still in an exploratory stage. One example is the exploratory research project "Mobile Heart Health", which aims to detect early signs of stress triggered by physiological or contextual changes and provide just-in-time mobile coaching [Morris & Guilak, 2009]. Most of the existing studies often try to discriminate a state of mental load from a resting condition. In [Setz et al., 2010], [Arnrich et al., 2010] and [La Marca et al., 2010] two stress factors relevant at the workplace were under investigation: high cognitive load under time pressure and social-evaluative threat. In all three studies mild cognitive load was discriminated from a constant high stress level but different stress intensities were not investigated. [Soga et al., 2007] used a mental arithmetic task to induce mental workload and investigated the recovery patterns of physiological responses as indicators of stress. [Kim et al., 2008] studied heart rate variability (HRV) features of subjects under chronic stress. Subjects were divided into a high and a low stress group based on their self-reporting stress scores. [Henelius et al., 2009] investigated the ability of short-term HRV metrics to discriminate between low and high level of mental workload.

In an office workplace scenario however, a worker is confronted with different levels of mental load during an office day. In order to achieve a day-by-day quantification of the mental load, first the different load levels have to be discriminated, and in a second step, the overall load can be estimated by accumulating these levels accordingly. This work goes one step towards this goal: we present our experimental results in discriminating different levels of mental workload. For an "everyday life application", a minimal sensor setup is desired for comfort reasons. This work therefore focuses on a single sensor modality: a mobile ECG system to measure heart rate (HR). The analysis of the HRV was chosen, because it represents a sensitive stress and mental load measure. Increased stress leads to an activation of the sympathetic nervous system and withdrawal of the parasympathetic nervous system [Veltman & Gaillard, 1998]. In this work, we investigate HRV features in the time as well as in the frequency domain.

2. METHODS

2.1 Experiments

Seven healthy subjects participated in this study (age between 25 and 34 years). Due to the effects of oral contraceptives and menstrual cycle phase on HRV, we decided to restrict the sample to male subjects. The experiment was designed to investigate different levels of mental workload. Three sessions with low, medium and high workload were chosen. Each session consisted of a baseline (10 minutes), workload (20 minutes) and recovery (15 minutes) period. Subjects performed each session on separate days in the afternoon, while the different sessions were randomly assigned for each subject in order to avoid sequence effects and, therefore, to counterbalance learning effects. To induce different levels of mental workload, we used the N-Back Test [Jaeggi et al., 2008]. Three variants of this task were used to induce low, medium and high workload which are likely to be present during an office work day:

   1.   Position 1 Back (Low Workload; very easy task with visual stimuli): A square appears every 4.5 seconds in one of eight different positions on a regular grid on the screen. The subject has to respond by using the keyboard if the position of the currently shown square is the same as the one that was presented just before. This kind of workload is comparable to monotonous monitoring tasks where the subject has to sustain his attention at the same level.

   2.   Arithmetic 1 Back (Medium workload; easy task with combined visual and auditory stimuli): An integer number between 0 and 9 appears every 4.5 seconds on the screen. For each number a math operator (add, subtract, multiply or divide) is presented via an audio message. The subject has to apply the math operation on the currently shown number and the one that was presented just before. The result of the calculation has then to be entered on the keyboard. This task reflects medium cognitive load since the subject has to memorize one number and to perform a math task in the given time.

   3.   Dual Arithmetic 2 Back (High Workload; demanding task with combined visual and auditory stimuli): In this mode, the two former position and arithmetic tasks are combined. An integer number between 0 and 9 appears every 4.5 seconds in one of eight different positions on a regular grid. For each number a math operator (add, subtract, multiply or divide) is presented via an audio message. The subject has to respond if the position of the currently shown number is the same as the one that was presented 2 positions back. In addition, the subject has to apply the math operation on the currently shown number and the one that appeared 2 positions back. The result of the calculation has then to be entered on the keyboard. This task represents a high cognitive load since the subject has to memorize the position and the value of two numbers and has to perform a math task in the given time.

   Directly after the workload period, the subject was asked to assess his perceived workload. For this subjective rating we employed the NASA Task Load Index (TLX) [Hart & Stavenland, 1988]. First, the subject has to rate 6 items on a scale from 1 to 20 that best indicate his experience in the task. The rating consists of the following items: mental demand, physical demand, temporal demand, own performance, effort and frustration. Next the subject is systematically asked which of the items represents the more important contributor to the workload. Based on these comparisons, the total workload is computed as a weighted average of the ratings. In addition, the individual performance for each workload task is recorded. The physiological responses were measured with the Zephyr BioHarness chest belt [Zephyr, 2010]. The monitoring belt consists of three smart fabric sensors to acquire cardiac activity, breathing rate and skin

temperature. The ECG data was sampled with 250Hz. In this work, we focus on the analysis of heart rate variability features in the time and frequency domain.

2.2 Data Analysis

We investigated the subjective ratings of the total workload obtained with the NASA Task Load Index by comparing the individual ratings for each workload period. Based on this data we performed an ANOVA test to investigate whether the perceived ratings differed significantly between the workload periods.

For the analysis of the heart rate data, we first removed RR intervals which differ more than 20% from their predecessors in order to remove artifacts. Next, we calculated a set of time and frequency HRV features following the guidelines of the European Task Force [Malik et al., 1996]: mean heartbeat intervals (Mean RR), standard deviation of RR intervals (SDNN), root mean square of successive differences (RMSSD), and the percentage of intervals that vary more than 50ms from the previous interval (pNN50). In addition, the HRV index (bin width 1/128 sec.), and the triangular interpolation of the R peak interval histogram (TINN) were extracted as geometric parameters. All features were calculated on the overall workload periods (20min) for each session. The analysis of HRV features in the frequency domain was done using the Lomb periodogram since it does not require resampling of unevenly sampled signals such as RR data [Clifford, 2002]. We used two frequency bands defined as follows: low frequency (LF):0.04-0.15 Hz and high frequency (HF): 0.15-0.4 Hz [Malik et al., 1996]. Next we calculated the normalized values of LF, HF and LF/HF which represents the relative value of each power component in proportion to the total power minus the very low frequency component. We obtained all HRV features for each phase of the experiment (baseline, workload and recovery). We compared these features obtained for the three workload periods by using the ANOVA test. As significance level, $p<0.05$ was considered.

3. RESULTS

As shown in **Figure 1** (left), all subjects perceived the induced load levels by the three variants of the N-Back as intended from the experiment design (ANOVA, $p<0.01$). In **Figure 1** (right) it is shown that also the individual performance reflects the three different workload levels.

The mean values including standard errors of all HRV features are listed in Table 1. It can be observed that the HRV features can be classified into two distinct groups. Features in the first group show consistently a decrease in their values with increased workload. A statistically significant decrease can be observed for the features RMSSD, pNN50 and HF ($p<0.05$) while Mean RR, SDNN, HRV Index and TINN show a consistent but non-significant decrease. In contrast, features in the second group show an increase of their values with increased workload. A statistically significant increase can be observed for the features LF and LF/HF ratio ($p<0.05$).
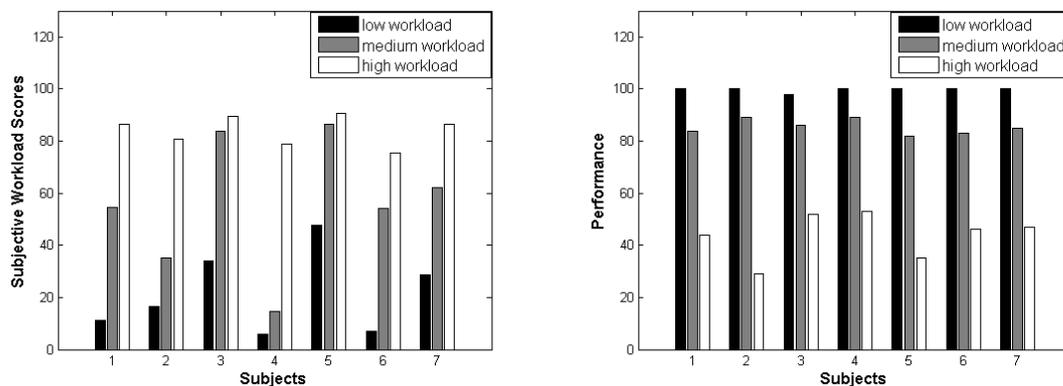

Figure 1 Perceived workload obtained from the NASA task load index and performance scores for each task.

Table 1. Comparison of Mean HRV Features ± standard error during low, medium and high workload periods.

HRV Features	Low Workload	Medium Workload	High Workload	F; p > F
Mean RR [ms]	875.3 ± 32.2	803.2 ± 36.5	769.1 ± 43.0	2.09; 0.15
SDNN [ms]	72.2 ± 8.4	58.7 ± 7.8	51.5 ± 6.4	1.89; 0.18
RMSSD [ms]*	51.6 ± 5.2	38.7 ± 4.4	31.2 ± 4.6	4.65; 0.02
pNN50 [%]*	30.7 ± 4.8	19.3 ± 3.6	12.4 ± 3.2	5.48; 0.01
HRV Index	19.5 ± 2.4	14.9 ± 1.8	13.0 ± 1.5	2.86; 0.08
TINN [ms]	462.8 ± 45.7	385.7 ± 53.1	385.1 ± 53.7	0.77; 0.48
LF [n.u.]*	64.3 ± 2.9	70.1 ± 2.7	77.8 ± 4.6	3.66; 0.04
HF [n.u.]*	35.6 ± 2.9	29.8 ± 2.7	22.2 ± 4.6	3.66; 0.04
LF/HF*	1.9 ± 0.2	2.5 ± 0.3	4.6 ± 1.0	4.59; 0.02

Mean ± Standard Error

* p < 0.05

4. CONCLUSION

We have presented an experiment design to induce three different levels of mental workload and to discriminate the workload levels based on heart rate features obtained from a mobile ECG system. According to the subjective ratings and the performance of the participants, we could show that all participants perceived the induced load levels as intended from the experiment design. In accordance, the performance decreased with increasing workload. The investigated HRV features could be classified into two distinct groups with respect to their response: with increasing workload, features in the first group showed a decrease in their values, while features in the second group showed an increase of their values. The features RMSSD, pNN50 and HF showed a statistically significant decrease while LF and LF/HF ratio showed a statistically significant increase with increased workload. The remaining features showed a consistent but non-significant increase or decrease, what might be explained by the limited number of subjects. Therefore, an increase in workload seems to be associated with a decrease in parasympathetic nervous activity and probably a concomitant increase in sympathetic activity. In conclusion, our experimental results show that a mobile heart rate sensor is suited to discriminate different levels of mental workload induced by cognitive tasks. In future work we are going to employ the mobile heart rate sensor in monitoring mental load during real office tasks.

ACKNOWLEDGEMENT

REFERENCES

Arnrich *et al.*, 2010] Arnrich, Bert and Setz, Cornelia and La Marca, Roberto and Tröster, Gerhard and Ehlert, Ulrike. 2010. What does your chair know about your stress level? *IEEE Transactions on Information Technology in Biomedicine*, 14(2), 207–214.

Clifford, 2002] Clifford, Gari D. 2002. *Signal Processing Methods for Heart Rate Variability Analysis*. Ph.D. thesis, St Cross College.

European Foundation, 2007] European Foundation for the Improvement of Living and Working Conditions. 2007. *Work-related stress*. http://www.eurofound.europa.eu/.

Hart & Stavenland, 1988] Hart, S. G. and Stavenland, L. E. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Chap. 7, pages 139–183 of:* Hancock, P. A., & Meshkati, N. (eds), *Human Mental Workload*. Elsevier.

Henelius *et al.*, 2009] Henelius, A. and Hirvonen, K. and Holm, A. and Korpela, J. and Muller, K. 2009. Mental workload classification using heart rate metrics. *Conf Proc IEEE Eng Med Biol Soc*, 1, 1836–1839.

Jaeggi *et al.*, 2008] Jaeggi, Susanne M. and Buschkuehl, Martin and Jonides, John and Perrig, Walter J. 2008. Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences*, 105(19), 6829–6833.

Kim *et al.*, 2008] Kim, Desok and Seo, Yunhwan and Salahuddin, Lizawati. 2008. Decreased Long Term Variations of Heart Rate Variability in Subjects with Higher Self Reporting Stress Scores. *In: Pervasive Healthcare*.

La Marca *et al.*, 2010] La Marca, R. and Waldvogel, P. and Thörn, H. and Tripod, M. and Pruessner, J. and Ehlert, U. 2010. Cold face test response indicates biopsychological adaptability to acute stress. *Revised manuscript under revision*.

Malik *et al.*, 1996] Malik, Marek and Bigger, J. Thomas and Camm, A. John and Kleiger, Robert E. and Malliani, Alberto and Moss, Arthur J. and Schwartz, Peter J. 1996. Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. *Eur Heart J*, 17(3), 354–381.

Morris & Guilak, 2009] Morris, Margaret and Guilak, Farzin. 2009. Mobile Heart Health: Project Highlight. *IEEE Pervasive Computing*, 8(2), 57–61.

Setz *et al.*, 2010] Setz, Cornelia and Arnrich, Bert and Schumm, Johannes and La Marca, Roberto and Tröster, Gerhard and Ehlert, Ulrike. 2010. Discriminating Stress from Cognitive Load Using a Wearable EDA Device. *IEEE Transactions on Information Technology in Biomedicine*, 14(2), 410–417.

Soga *et al.*, 2007] Soga, C. and Miyake, S. and Wada, C. 2007 (Sept.). Recovery patterns in the physiological responses of the autonomic nervous system induced by mental workload. *Pages 1366–1371 of: SICE, 2007 Annual Conference*.

Van Daalen *et al.*, 2009] van Daalen, G. and Willemsen, T.M. and Sanders, K. and van Veldhoven, M.J.P.M. 2009. Emotional exhaustion and mental health problems among employees doing people work: the impact of job demands, job resources and family-to-work conflict. *Int Arch Occup Environ Health*, 82, 291–303.

Veltman & Gaillard, 1998] Veltman, J. A. and Gaillard, A. W. 1998. Physiological workload reactions to increasing levels of task difficulty. *Ergonomics*, 41(May), 656–669.

Zephyr, 2010] Zephyr. 2010. *Zephyr*. http://www.zephyr-technology.com/.