Neuronal Mental Workload Registration during Execution of Cognitive Tasks

Thea Radüntz

Federal Institute for Occupational Safety and Health,
Unit. 3.4 'Mental Health and Cognitive Capacity'
Nöldnerstr. 40/42, D-10000 Berlin, Germany
raduentz.thea@baua.bund.de

Abstract. Neuronal workload measurement is a key-technology for optimizing work conditions in human-machine systems. Specific aims are the identification of neurophysiological parameters indicative for workload and their validation by systematic variation of external load conditions.

The battery consists of tasks with diverse complexity and difficulty. The sample consists of 34 people and shows high variability in respect to the cognitive capacity and hence to the experienced mental workload. The electroencephalogram (EEG) as well as further workload relevant bio signal data and the NASA-TLX as a subjective questionnaire method are registered.

Results from the NASA-TLX questionnaire reveal the predominant role of the mental dimension at the implemented task battery. Furthermore, the NASA-TLX indicates the existence of diverse levels of difficulty with several tasks per level. Analysis of EEG spectra demonstrates an increase of frontal theta band power and a decrease of alpha band power with increasing task difficulty level.

Keywords: mental workload, electroencephalogram (EEG), signal processing, pattern recognition.

1 Introduction

The development of advanced information and communication technology as well as of highly interactive work environments and work assistance systems is unstoppable. Although the main goal of this development is to simplify the work, employees complain about high mental workload and stress. Problems arise from information overload, frequent work interruptions or from a multitude of irrelevant information [9], [10], [13].

On the other hand work associated with automation and supervisory control can be linked to repetitive and monotonous tasks that may be accompanied by complacency, fatigue, reduced vigilance [14], [15], [3], [2], [12] and increased error rates, hence a safety risk for further persons [17].

An objective method for mental workload registration is absolutely essential particularly with regard to the long-term negative consequences of inappropriate

D. Harris (Ed.): EPCE 2014, LNAI 8532, pp. 59–69, 2014.

workload on the individual's health as a serious problem of modern society. Appropriate work efficiency is only possible in an optimal workload range that can most efficiently be measured where information processing takes place, i.e. in the brain. Neuronal workload measurement is hence a key-technology for optimizing work conditions in human-machine systems.

Hence, the overall goal is continuous, online registration and monitoring of mental workload on neuronal basis. The theoretical background is given by the variability of the EEG frequency bands according to attention, fatigue and mental workload. Specific aims are the identification of neurophysiological parameters indicative for workload and their validation by systematic variation of external load conditions.

2 Methods

The tests took place in the shielded lab of the Federal Institute for Occupational Safety and Health in Berlin.

The electroencephalogram (EEG), as a direct signal of bioelectrical brain activity, as well as further workload relevant bio signal data (e.g. heart rate, blood pressure) and the NASA-TLX as a subjective questionnaire method are registered. Hence, subjective and objective methods can be combined and contribute to the validation of the mental workload registration.

2.1 Procedure

The experiment was fully carried out with each subject in a single day. It consisted of two parts: a training phase and the main experiment. During the training phase subjects were familiarized with the cognitive tasks. The cognitive tasks were the same as these of the main experiment but shorter in time. They were repeated until the subject reached an accuracy index of at least 80%. The training phase should create similar individual starting conditions in respect to the performance, so that we can investigate the workload's effect independent from learning effects.

The main experiment started after a short break subsequent to the training phase. The tasks were presented in the same counterbalanced order as presented during the training phase. They took place in the shielded lab of the Federal Institute for Occupational Safety and Health and were controlled remotely through a remote desktop connection, an intercommunication system and a video monitoring system.

2.2 Subjects

The sample consists of 57 people in paid work and shows high variability in respect to the cognitive capacity and hence to the experienced mental workload. At the time of writing only the data of the first 34 people has been analyzed. Table 1 describes the sample set used.

Table 1. Sample set

Age	Male	Female	Total
30 - 39	4	5	9
40 - 49	7	8	15
50 - 59	1	5	6
60 - 69	3	1	4
Total	15	19	34

2.3 Tasks

The simulation of miscellaneous cognitive task requirements is realized through the implementation of a task battery in the E-Prime application suite. The battery consists of tasks with diverse complexity and difficulty inducing different levels of mental workload. The implemented tasks are listed in Table 2.

Table 2. Task battery

Task duration [m]	0nb 5	2nb 5	Sternberg 10	Seriell sternberg 10	Stroop 5
Task Duration [m]	Switch PAR 5	Switch NUM 5	Switch XXX 10	AOSPAN 20	

In this paper we concentrate on the analysis and evaluation of three tasks: 0-back as the easiest one, 2-back as a working memory task with moderate workload, and aospan as a demanding dual task (see Figures 1, 2, 3). The latter was the self adapted and translated version of the AOSPAN task developed by [18].

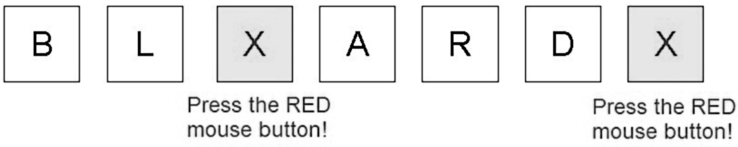


Fig. 1. 0-back task: Press the mouse button if the presented letter is 'X'

2.4 Subjective Ratings

Subjective workload was captured with a computerized version of the NASA-TLX [6]. After each task during the training phase, subjects were asked to rate the workload sources in 15 pairwise comparisons of NASA-TLX's six workload dimensions: mental demand, physical demand, temporal demand, performance,

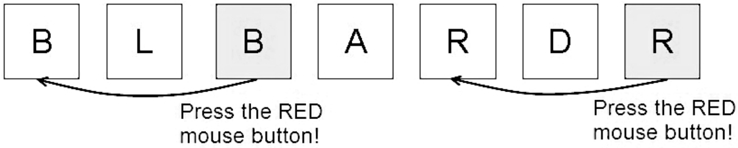


Fig. 2. 2-back working memory task: Press the mouse button if the presented letter is the same as the next to last letter seen.

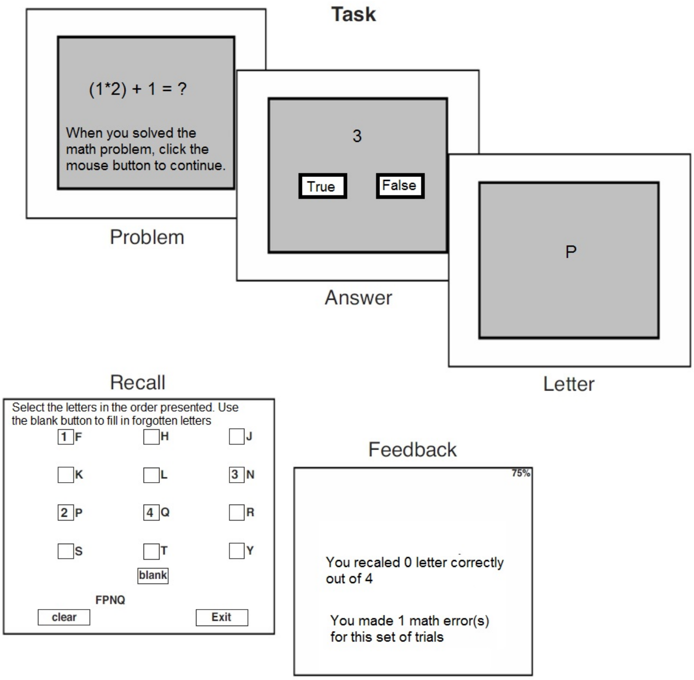


Fig. 3. AOSPAN dual task: memorize letters in the order presented while simultaneously solving math problems. Trials consist of 3 sets of each set-size, with the set-sizes ranging from 3-7.

effort, frustration. This required the subject to choose which dimension is more relevant to workload in the specific task. Hence, we gained an individual weighting of these subscales based on their perceived importance.

After each task during the main experiment, subjects were asked again to rate the task within a 100-points range with 5-point steps. They indicated their rating by clicking on a 5-point step box with an optical mouse.

2.5 Physiological Measures

The electroencephalogram (EEG) as well as the blood pressure (BP), the heart rate (HR) and the inter-beat interval (IBI) were digitally recorded during the main task only. Digital signal processing and calculation of mean values were done with MATLAB.

EEG. The EEG was captured by 25 electrodes placed at positions according to the 10-20-system and recorded with reference to Cz and at a sample rate of 500 Hz. For signal recording we used an amplifier from BrainProducts GmbH and their BrainRecorder software.

For workload calculations we implemented a modular MATLAB-Toolbox. Figure 4 describes the signal processing pipeline.

The pre-processing module reads the recorded EEG signal. The signal is multiplied with a Hamming window function and filtered with a bandpass filter (order 100) between 0.5 and 40 Hz. Subsequently, independent component analysis (ICA) is applied to the signal and the calculated independent components are visually inspected and classified as either an artifact or signal component. The signal components are projected back onto the scalp channels and the now artifact-corrected EEG signal is passed over to the next module. There it is cut into segments of 10 seconds length, overlapping by 5 seconds. The segments are then transformed to frequency domain using Fast Fourier Transformation (FFT) and workload relevant frequency bands are computed (θ: 4-8 Hz, α: 8-12 Hz).

The combination of the θ- and α-band provides the basis for the indexing, training and classification of mental workload according to Lei's Logistic Function Model [11]:

$$W = \frac{1}{1 + e^{-(b0+b1*\theta+b2*\alpha)}} \ . \tag{1}$$

The model allows the identification of individual workload.

Cardiovascular Parameters. Blood pressure was recorded continuously by the FMS Finometer Pro device. A finger cuff was placed around subject's finger and systolic and diastolic blood pressure as well as the heart rate and the inter-beat interval were detected automatically. The recorded data was processed in the time domain.

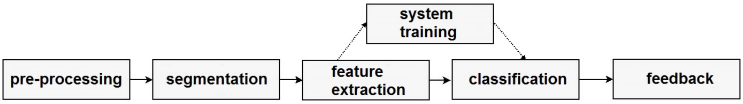


Fig. 4. Signal processing pipeline

2.6 Performance

Mean reaction time computation was meaningful only for 0-back and 2-back, where subjects were asked to respond quickly after the presented stimuli. During AOSPAN's recall slide there was no time pressure and the reaction time depended also on the set size presented and on mouse movements. In addition, subjects were allowed to correct themselves several times. The time alloted for solving the math tasks was computed individually during the training phase by calculating the mean time of the responses. In the main experiment, when this individually calculated time was exceeded, the next slide was shown and a math error registered for the current operation.

For all three tasks the individual percentage of false responses was calculated. For AOSPAN, false responses include the number of sets in which the letters are not recalled in correct serial order, math errors and the above mentioned math speed errors.

2.7 Analysis

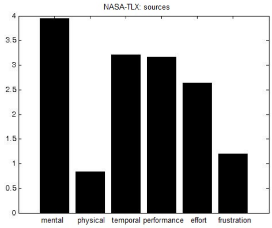
Physiological measures were obtained and calculated for the following three tasks: 0-back, 2-back and AOSPAN. Six ANOVAs were carried out utilizing repeated measures design, one within-subject factor (θ, α, systolic BP, IBI, percentage of errors, NASA-TLX) and 3 levels (the tasks), where the differences between the tasks were examined and tested with a post-hoc test (Bonferroni). A paired-samples t-test was computed for the 0-back and 2-back reaction times.

3 Initial Results

First results computed over 34 subjects and three tasks will be presented in the following section. They comprise the obtained subjective ratings and task performance as well as the workload relevant θ- and α-band from the EEG and the systolic BP and IBI.

3.1 Subjective Ratings

Results of the subjective ratings are presented in Figure 5. Figure 5(a) indicates the predominant role of mental demands at the implemented task battery. Hence, the induced workload originates from information processing and should be reflected in the EEG. Figure 5(b) shows the average workload index for the selected tasks 0-back, 2-back and AOSPAN as representatives of a low, moderate and high workload tasks. Workload means changed significantly during the experiment (Greenhouse-Geisser F(1.9;65.5) = 57.1, p<0.001). Post-hoc analysis showed that the mean workload was significantly lower during the 0-back task than in the other two tasks.

(a) 

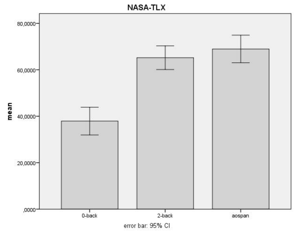
(b) 

Fig. 5. (a) NASA-TLX: source rating over all tasks and 34 subjects. (b) NASA-TLX: workload index computed for 0-back, 2-back and AOSPAN over 34 subjects.

3.2 Physiological Measures

EEG. Analysis of the EEG spectra at the Fz and Pz electrode demonstrates an increase of the θ-band power and a decrease of the α-band power with increasing task difficulty level. This fact assures successful system training and an individual parametrization in the context of Equation 1.

Results obtained from the Fz electrode are presented in Figure 6. They are consistent with previous observations of several other authors (e.g. [16], [5], [4]). θ- and α-band means changed during the experiment. The θ-band changed significantly whereas the α-band revealed no significant changes (Greenhouse-Geisser F(1.6;46.2) = 8, p<0.01; Greenhouse-Geisser F(1.9;63) = 3.6, p=0.04). Post-hoc analysis of the θ-band showed that the means were significantly larger during the AOSPAN-task than in the 0-back task.

Cardiovascular Parameters. Both systolic BP and IBI differed between the three tasks significantly (Greenhouse-Geisser F(1.3;45) = 15, p<0.001; Greenhouse-Geisser F(1.4;46.5) = 6.5, p<0.01). IBI in the 0-back task were, according to post-hoc analysis, lower than in the 2-back and AOSPAN task. Systolic BP means were significantly larger during the AOSPAN-task than in 0-back and 2-back tasks. Results of systolic BP and IBI are presented in Figure 7(a) and Figure 7(b).

3.3 Performance

Mean reaction times (RT) and error rates are presented in Figure 8. Figure 8(a) demonstrates differences in RT between the 0-back and the 2-back task (t-test

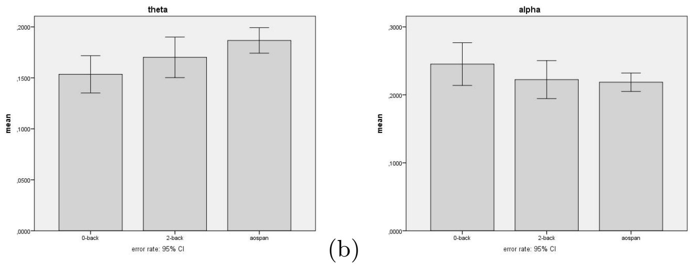


Fig. 6. EEG - Fz: θ-band (a) and α-band (b) computed for 0-back, 2-back and AOSPAN over 34 subjects

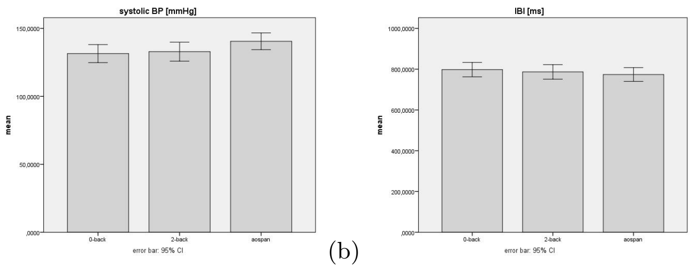


Fig. 7. Systolic BP (a) and IBI (b) computed for 0-back, 2-back and AOSPAN over 34 subjects

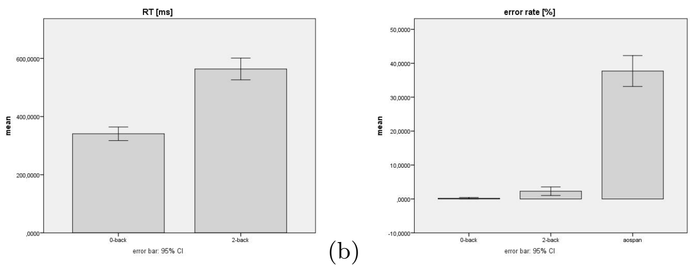


Fig. 8. (a) Reaction time computed for 0-back and 2-back over 34 subjects. (b) Percentage of false responses computed for 0-back, 2-back and AOSPAN over 34 subjects.

$p<0.01$). Figure 8(b) shows the average error rate for the selected tasks 0-back, 2-back and AOSPAN. Error rate means changed significantly during the experiment (Greenhouse-Geisser $F(1.1;39.2) = 271.8$, $p<0.001$). Post-hoc analysis revealed significant changes of the mean percentage of false responses between all three tasks.

4 Discussion

The central issue addressed by this paper is the registration of workload by means of neurophysiological parameters. Therefore a task battery of miscellaneous cognitive task requirements was implemented, including various complexity and difficulty, hence inducing different levels of mental workload. In this paper we concentrated on the 0-back, 2-back and AOSPAN tasks as representatives for an easy, moderate and difficult task. Subjective ratings derived from the NASA-TLX questionnaire demonstrate significant workload differences only between the easy 0-back task and the other two.

The RT between 0-back and 2-back task increased significantly and also the error rate between the three tasks grew significantly. In contrast to the subjective ratings that could distinguish only between two levels, the performance score indicates a difference between all task difficulty levels. Furthermore, the computed error rate for the AOSPAN task is remarkably high.

The IBI, similar to the subjective ratings, shows significant differences only between the 0-back task and the other two, while the systolic BP means were significantly higher between the AOSPAN and the other two tasks and coincide with AOSPAN's extremely high error rate.

The EEG as a direct signal of brain activity and the frequently observed variability of the θ- and α-band according to attention, fatigue and mental workload, constitute the theoretical background for the implementation of an objective method for neuronal mental state monitoring. Initial results of our EEG signal processing demonstrate that the θ-band increases with advanced task difficulty. It can significantly distinguish between an easy task like 0-back and a difficult task like AOSPAN. The α-band shows the tendency to decrease with task difficulty increase but did not obtain significance. These observations are consistent with several other studies. There the θ-band is a reliable parameter which is enhanced with increasing difficulty, whereas the α-band seems to be less reliable with respect to the decrease which is normally expected. In some studies this behavior is linked to different forms of attention (internal vs. external) and other task requirements [8], [7]. However, the expected tendencies for an increase of the θ- and a decrease of the α-band are given in our study.

In addition and with respect to the gained error rates, we have to ask to which extent subjects remained sufficiently motivated during the AOSPAN task. A second question would be whether the subjects continued to invest enough effort in problem-solving, even if they realized that the demands on them exceeded their own processing capacity. That could also be a reason why subjects' ratings between 2-back and AOSPAN task did not reach significance. The systolic BP shows a significant increase during the AOSPAN task, but this could be linked to emotional reactions like frustration. It is commonly known that BP is influenced by this [1]. Furthermore, we quickly checked the NASA-TLX frustration subscale and noted that AOSPAN received the largest mean value there. The analysis of the battery's further tasks could help solving these questions. The same applies to the processing of the further subjects and the detailed analysis of NASA-TLX subscales.

Based on such consolidated findings of neuronal brain states an optimal task sharing between human and machine with efficient cognitive processing for the operator could be defined. The benefits for older employees are maintenance of autonomy and working ability due to the moderate mental workload when working with the human-machine system. An additional important benefit is the prevention of negative impacts of sustained over- or underload on the mental health and cognitive capacity of the working population.

Acknowledgments. I would like to thank my student assistant Jon Scouten for daily operational support and proofreading. I would also like to express my sincere appreciation to Dr Gabriele Freude and Dr Uwe Rose for their valuable and constructive suggestions.

References

1. Brave, S., Nass, C.: Emotion in human-computer interaction. In: The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications, pp. 81–96 (2002)
2. Debitz, U., Gruber, H., Richter, G.: Psychische Gesundheit am Arbeitsplatz. Teil 2: Erkennen, Beurteilen und Verhüten von Fehlbeanspruchungen, 3rd edn. InfoMedia Verlag (2003)
3. Hacker, W., Richter, P.: Psychische Fehlbeanspruchung. Psychische Ermüdung, Monotonie, Sättigung und Stress (Spezielle Arbeits- und Ingenieurpsychologie in Einzeldarstellungen), 2nd edn. Springer, Berlin (1984)
4. Gevins, A., Smith, M.E.: Neurophysiological measures of working memory and individual differences in cognitive ability and cognitive style. Cerebral Cortex 10(9), 829–839 (2000)
5. Hagemann, K.: The alpha band as an electrophysiological indicator for internalized attention and high mental workload in real traffic driving. Ph.D. thesis, University of Düsseldorf, Germany (2008)
6. Hart, S.G., Staveland, L.E.: Development of the NASA TLX: results of empirical and theoretical research. In: Hancock, P.A., Meshkati, N. (eds.) Human Mental Workload, pp. 139–183. North Holland, Amsterdam (1988)
7. Kelly, S.P., Lalor, E.C., Reilly, R.B., Foxe, J.J.: Increases in Alpha Oscillatory Power Reflect an Active Retinotopic Mechanism for Distracter Suppression During Sustained Visuospatial Attention. Journal of Neurophysiology 95(6), 3844–3851 (2006), doi:10.1152/jn.01234.2005
8. Klimesch, W.: EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. Brain Research Reviews 29(2-3), 169–195 (1999)
9. Kompier, M.A.J., Kristensen, T.S.: Organisational work stress interventions in a theoretical, methodological and practical context. In: Dunham, J. (ed.) Stress in the Workplace: Past, Present and Future, pp. 164–190. Whurr Publishers, London (2001)
10. Landsbergis, P.A., Cahill, J., Schnall, P.: The changing organisation of work and the safety and health of working people: a commentary. Journal of Occupational Environmental Medicine 45(1), 61–72 (2003)
11. Lei, S.: Driver mental states monitoring based on brain signals. Ph.D. thesis, TU Berlin, Germany (2011)

12. May, J.F., Baldwin, C.L.: Driver fatigue: The importance of identifying causal factors of fatigue when considering detection and countermeasure technologies. Transportation Research, Part F 12(2009), 218–224 (2008)
13. NIOSH - NORA Organization of work team members. The changing organization of work and the safety and health of working people. NIOSH-Publications Dissemination, Cincinnati (April 2002)
14. Parasuraman, R., Molloy, R., Singh, I.L.: Performance consequences of automation induced complacency. International Journal of Aviation Psychology 3(1), 1–23 (1993)
15. Parasuraman, R., Mouloua, M., Molloy, R.: Monitoring automation failures in human machine systems. In: Mouloua, M., Parasuraman, R. (eds.) Human Performance in Automated Systems: Current Research Trends, pp. 45–49. Earlbaum, Hilsdale (1994)
16. Posner, M.E., Peterson, S.E.: The attentional system of the human brain. Annual Review of Neuroscience 13, 25–42 (1990)
17. Sträter, O.: Warum passieren menschliche Fehler und was kann man dagegen tun?, Forum Prevention, AUVA - Allgemeine Unfallversicherungsanstalt, Wien (2001)
18. Unsworth, N., Heitz, R.P., Schrock, J.C., Engle, R.W.: An automated version of the operation span task. Behavior Research Methods 37, 498–505 (2005)