Influence of Workload on Auditory Evoked Potentials in a Single-stimulus Paradigm

R. N. Roy [1,2], A. Breust[1], S. Bonnet[1], J. Porcherot[1], S. Charbonnier[2], C. Godin[1] and A. Campagne[3]

[1]*CEA LETI, Univ. Grenoble Alpes, Minatec Campus, Grenoble, France*
[2]*Gipsa-Lab, Univ. Grenoble Alpes and CNRS, Grenoble, France*
[3]*LPNC, Univ. Grenoble Alpes and CNRS, Grenoble, France*

Keywords: Workload, Mental State Monitoring, ERP, Audition.

Abstract: Mental workload can be assessed via neurophysiological markers. Temporal features such as event related potentials (ERPs) are one of those which are very often described in the literature. However, most of the studies that evaluate their sensitivity to workload use secondary tasks. Yet potentials elicited by ignored stimuli could provide mental state monitoring systems with less intrusive probing methods. For instance, auditory probing systems could be used in adaptive driving or e-learning applications. This study evaluates how workload influences auditory evoked potentials (AEPs) elicited by a single-stimulus paradigm when probes are to be ignored. Ten participants performed a Sternberg memory task on a touchpad with three levels of difficulty plus a view-only condition. In addition, they performed two ecological tasks of their choice, one deemed easy (e.g. reading novels), and the other difficult (e.g. programming). AEPs were elicited thanks to pure tones presented during the memory task retention period, and during the whole extent of the external tasks. Performance and AEPs were recorded and analyzed. Participants' accuracy decreased linearly with increasing workload, whereas the difference in amplitude between the P3 and its adjacent components, N2 and SW, increased. This reveals the relevance of this triphasic sequence for mental workload assessment.

1 INTRODUCTION

1.1 Workload and Mental State Monitoring

The aim of mental state monitoring (MSM) and neuro-ergonomics is to evaluate an operator's state in order to better supply her/him with help, information or safety measures. The applications are numerous, ranging from gaming to education, including driving and security. The new systems that allow this assessment are called passive Brain Computer Interfaces (pBCI; Zander et al., 2011), biocybernetic systems, or physiologically attentive user interfaces when they adapt their functionality to the user's covert state (Chen and Vertegaal, 2004). It is well known that mental state modulations are reflected by a variety of physiological signals, including neurophysiological signals. Hence, passive BCI systems use neural markers, such as spectral or temporal features to classify a given mental state. Most of those neural markers are derived from electro-encephalographical (EEG) signals (Blankertz

et al., 2010; van Erp et al., 2012).

One of the mental states that are of major interest to evaluate is workload. **Mental workload** has been extensively documented and can be defined either as the load in memory (i.e. number of items), the number of tasks to be performed in parallel, and more generally as a measure of the amount of cognitive and attentional resources engaged in a task. It is therefore considered to be close to task difficulty (Gevins and Smith, 2007), and to depend on each individual's capabilities and the effort put in performing the task (Cain, 2007). Therefore, a classical finding is that performance (e.g. response time and accuracy) decreases when workload increases (e.g. Sternberg, 1966; Natani et al., 1981; Gomarus et al., 2006).

1.2 EEG Markers of Workload

Several EEG features are known to react to an increase in workload. Spectral features such as power spectral density have been thoroughly described in the literature (e.g. Gomarus et al., 2006; Berka et al., 2007; Roy et al., 2013). For instance,

numerous indices have been created to better exploit the variation in the alpha and beta bands. One of such is the ratio of the theta activity at frontal sites with the alpha activity at parietal sites (Gevins and Smith, 2003; Holm et al., 2009).

Another commonly used feature is a temporal one: event related potentials (ERPs). Indeed, a stimulus, either from the task at hand or an external probe, can be used to elicit a neural response. Most studies that report variations of ERP components due to workload variations use visual stimuli. However, auditory or tactile ERPs should be considered for real-life applications. Indeed, systems that would make use of probes from other sensory modalities than vision would be less intrusive, less distracting and therefore less risky for the operator and its task. For instance, in a driving situation, ERPs elicited by tactile stimuli from the chair could be a good means to evaluate the drivers' workload (Sugimoto and Katayama, 2013). In the same manner, auditory probes can be good indicators of one's mental workload.

Auditory evoked potentials (AEPs) are typically studied using oddball paradigms. The elicited components include a typical triphasic sequence (Smith et al., 1990): N2, P3 and Slow Wave (SW). When stimuli are ignored, the P3 component is anterior (P3a) and reflects an involuntary capture of attention, whereas when stimuli are attended to, it has a posterior distribution (P3b; Squires et al., 1975; Näätänen and Gaillard, 1983; Halgren et al., 1998; Strobel et al., 2008; Allison and Polich, 2008). It is well known that the P3 component has an amplitude reduction when workload increases, for both visual and auditory probes (Natani and Gomer, 1981; Kok, 2001; Schultheis and Jameson, 2004; Holm et al., 2009; Miller et al., 2011; Brouwer et al., 2012). Earlier components such as the N1, P2 or the N2 have also been reported to be sensitive to variations in workload (Ullsperger et al., 2001; Allison and Polich, 2008; Miller et al., 2011). As for later slow waves, to our knowledge no effect of workload has been reported on the negative component appearing just after the P3.

All those studies on the impact of workload on AEPs have been conducted using classical oddball paradigms in which participants had to detect (and/or count) a target infrequent item amongst distractors or novel sounds. However, for real-life applications of mental state monitoring systems, a less intrusive and distracting probing method should be used. That is to say that the use of a secondary task should be avoided in order to keep the operator focused on its primary task. Hence, Allison and

Polich (2008) have introduced the **single-stimulus paradigm** for assessing mental workload in an immersive environment in a less distracting way. In this paradigm, there are no non-target stimuli, they are replaced by silence, and only target stimuli are presented, at irregular intervals. As the authors point out, this is a stimulation method that is operationally easy to implement. In their study, participants had to either count or ignore these auditory stimuli while playing a video game. The authors indicated amplitude modulations for the P2, N2 and P3 components. However, the modulations were not the same depending on the contrast. For instance, it decreased when participants were playing in a difficult condition vs. a medium one. But, it increased from an easy one to a medium one.

1.3 Current Study

As we saw earlier, most of the studies that assessed the impact of workload on AEPs have used secondary tasks. However, this is less realistic to implement for real-life applications. It seems that the single-stimulus paradigm is the best way to elicit AEPs by interfering as little as possible in the operator's task. However, to our knowledge, only one study has paved the way to evaluating the usability of potentials evoked by this paradigm for mental workload assessment (Allison and Polich, 2008), and it gave no definite conclusion as to a robust amplitude modulation of ERP components. Therefore, these results need to be supported and extended. We intend to go further by examining how AEPs elicited by a single-stimulus paradigm are influenced by workload for a laboratory task, the Sternberg memory task, in which participants have to memorize a varying number of items, and also for work-related ecological tasks, i.e. reading and computer programming. In order to do so, we used pure tones of different frequencies, and we extracted the amplitude and latency of the triphasic AEP sequence N2, P3, SW, as well as the LPP elicited by these probes. A potentially interesting new feature was also evaluated, i.e. the difference in amplitude between these adjacent components.

2 MATERIALS AND METHODS

2.1 Experimental Protocol

Ten healthy right-handed participants (7 males; m = 28.89 years, s.d. = 7.02 years) volunteered for the experiment.

Workload was manipulated using a modified Sternberg task (Sternberg, 1966) and ecological external tasks. The first task was performed on a Windows Surface touchpad and was implemented in C++ using Qt/qwt and Visual Studio C++. As for the ecological tasks, they were performed by participants on their work computer or on their desk.

In the modified Sternberg task, the 20 French consonants were used as visual stimuli (vowels were excluded to avoid chunking strategies). In addition, for both the Sternberg task and the external tasks, the auditory probes were six pure tones ranging from 750 to 2000 Hz with a 250 Hz step with a 100 ms duration, including a 10 ms rise and a 10 ms fall. They were presented binaurally using a Logitech PC headset.
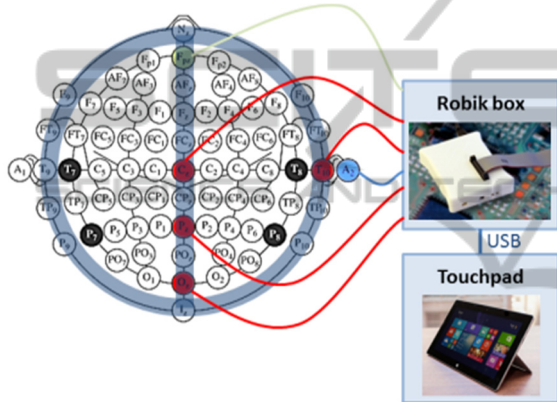

Figure 1: EEG electrode and acquisition system setup.

During the Sternberg task, participants had to memorize a list of sequential consonants visually presented on the touchpad screen. Then, a keyboard was presented (Figure 2). The participants had to retrieve the consonants and the order in which they were presented, and answer as accurately as possible by touching the screen. Three levels of workload were considered, i.e. 3, 5 and 7 consonants to memorize (easy, medium and high workload respectively), as well as a 'view-only' condition, or idle state, in which they only focused their attention on the fixation point and had no item to memorize or retrieve. All trials were pseudo-randomly presented. The six auditory probes were presented during the retention period, with an inter-tone interval of 2005 + i * 1000 ms, i ranging from 0 to 5. Participants performed a training session of 3 easy and 3 difficult trials for 4.5 minutes. Then, they performed 3 trials of the 4 conditions: easy, medium, hard and view-only, for 9 minutes.

Then, the participants had to choose two tasks amongst several ecological tasks: an easy and a difficult one (as defined by us). For the easy task, 4 participants watched YouTube movies without the sound or read funny stories on 9gag.com, 4 read novels or newspapers and 2 surfed on the internet. For the difficult task, 5 participants who are computer scientists or students in computer science wrote code on their computer, 2 read scientific publications in English (not their mother tongue), and the last 3 played difficult Sudoku grids online. Each external task lasted for 15 minutes. In total, the whole experiment lasted for 44 minutes.

2.2 Measures and Analyses

The accuracy of participants' answers to the implemented task was recorded, as well as their EEG activity for all tasks using the Robik box acquisition system (Filipe et al., 2011). EEG activity
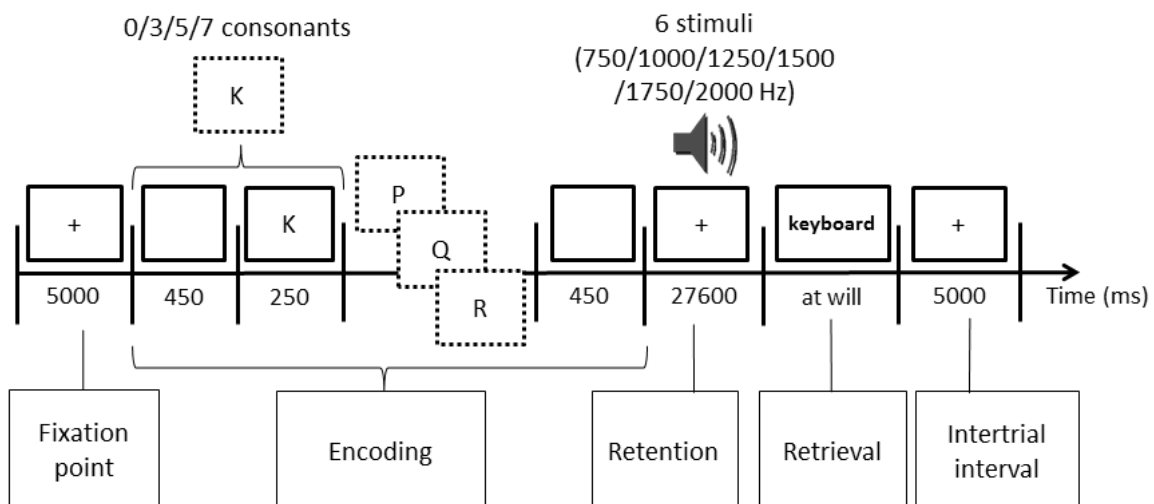

Figure 2: Trial structure for the Sternberg memory task.

was recorded from four passive Ag/AgCl electrodes: Oz, Pz, Cz and T10 positioned according to the 10-20 system. These electrodes were fastened on a double strap headband and kept moist with physiological serum. The reference was set at Fpz and the ground electrode at the right earlobe (Figure 1). The data were sampled at 390 Hz, band-pass filtered between 1 and 20 Hz, and re-referenced to a common average reference.

EEG analyses were only performed on the signal acquired from the Cz electrode, as it proved relevant for workload estimation from both early and late components (Allison and Polich, 2008). ERPs elicited by the auditory probes during the different tasks were extracted by subtracting a 200 ms pre-stimulation baseline to the 600 ms post-stimulation signal. Trials with a maximum over 100 µV were rejected. Absolute peak amplitude and latency were extracted for the 4 following components: the N2, negative deflection between 150 and 250 ms, the P3, positive deflection between 220 and 350 ms, the SW, negative deflection between 350 and 450 ms, and the LPP, positive deflection between 400 and 600 ms. The differences in amplitude between P3 and N2 (P3N2), P3 and SW (P3SW), and LPP and SW (LPPSW) were also extracted.
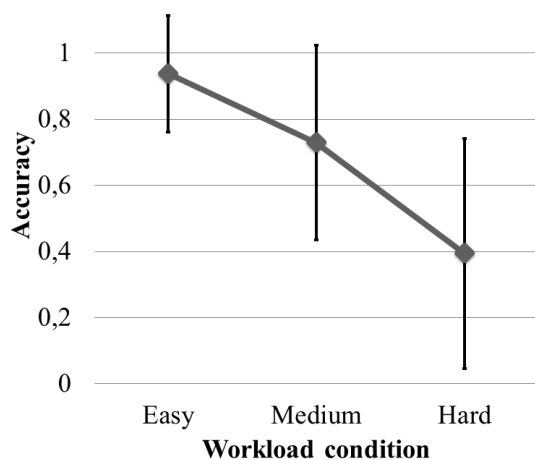

Figure 3: Accuracy of participants' answer in the Sternberg task depending on workload condition – Average and standard deviation across participants.

To statistically assess workload's impact on both accuracy and AEPs (amplitude and latency), we performed repeated measure ANOVAs with Tukey post-hoc tests. The analyses of accuracy were performed using Statistica, and all the EEG analyses were performed using Matlab. Performance in the Sternberg task was only evaluated for the conditions in which an answer was expected, i.e. the easy,

medium and hard conditions. As for EEG data analysis, in the Sternberg task, we had 4 levels of workload: easy, medium, hard and view-only. Only trials for which a correct answer was given were kept, in order to effectively evaluate the impact of an increasing number of items in memory. For the external tasks, we had 2 workload levels: easy and hard.

3 RESULTS

3.1 Task Performance

Participants were less accurate when workload increased ($p<0.001$), and this effect was linear (linear polynomial p<.01; quadratic polynomial n.s.; Figure 3). We can observe a large variability between participants, mostly for the medium and hard conditions.

3.2 Auditory Evoked Potentials

Grand averages across participants of the ERPs elicited by the pure tones in the Sternberg and the external tasks depending on workload condition are respectively given in Figures 4 and 5. It should be noted that there were two peaks to the P3 component for the different conditions, but for the view-only one.

Although the averaged signal shows modulations of components' amplitude depending on workload condition for both the implemented task and the external tasks (e.g. increased N2 and SW amplitudes, decreased P3 amplitude), these differences, as well as latency differences were not significant. This is most certainly due to too much variance between participants resulting in a levelling at the group level. This variability is illustrated by Figure 6 which displays the number of participants that show a significant difference in voltage across trials depending on time and condition comparison (e.g. E vs. M: Easy vs. medium difficulty levels). Indeed, on our total number of ten subjects, only a maximum of three participants have a significant difference congruent in time. Nevertheless, we can see that the time periods of the N2, P3, SW and LPP components are somewhat relevant at the participant level.

Interestingly, it so appeared that the differences in amplitude between adjacent components were more robust to this inter-participant variance. Indeed, for the Sternberg task, the P3N2 significantly increased with workload ($F(3,27) =$

8.39 , p<.001), and this effect was linear between the easy, medium and hard conditions (linear polynomial p<.01; quadratic polynomial n.s.). The same workload effect was observed on the P3SW (F(3,27) = 4.52, p<.05). It was also linear between the easy, medium and hard conditions (linear polynomial p<.05; quadratic polynomial n.s.).
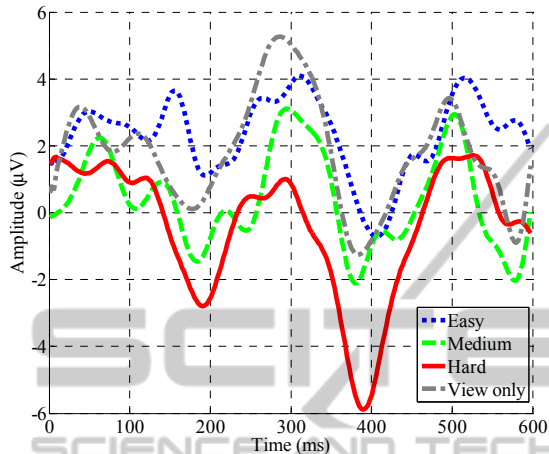

Figure 4: Impact of workload condition on the AEPs at the Cz electrode during the Sternberg task. Data smoothed using a 5-sample moving average.

Figure 7 displays the difference waveforms between the hard and easy, and the medium and easy conditions. It illustrates well the increase in difference for these adjacent components with increasing workload. Figure 8 details those amplitude differences for the Sternberg task through box plots. For the LPPSW, no significant difference was observed.
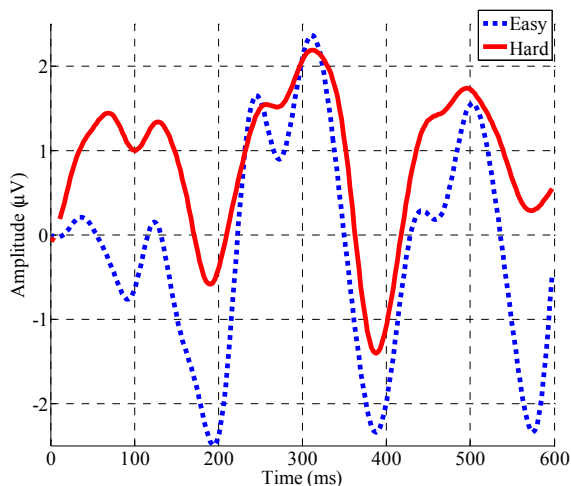

Figure 5: Impact of workload condition on the AEPs at the Cz electrode during the ecological tasks. Data smoothed using a 5-sample moving average.

As regards the external tasks, although the signal at the Cz electrode seemed greatly modulated by workload, the impact of this factor was in fact smaller on peak amplitudes, and was not significant.
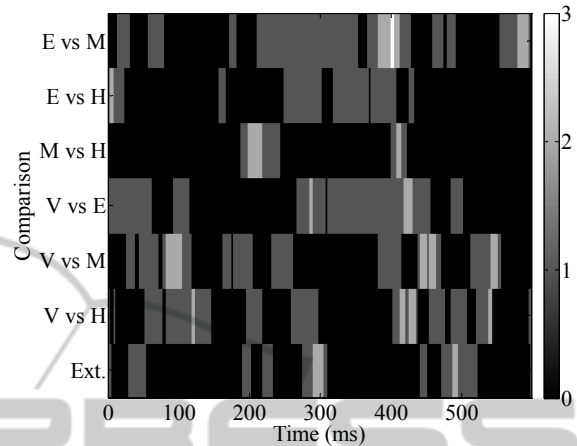

Figure 6: Number of participants that show a significant difference in voltage across trials depending on condition comparison and time. E: easy; M: medium, H: hard, V: view-only, Ext: external tasks.
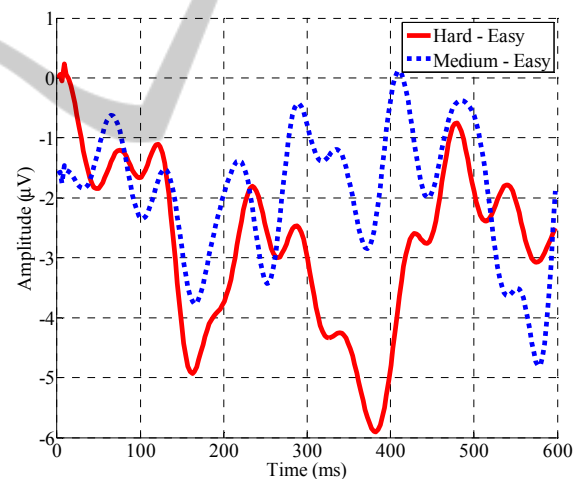

Figure 7: Difference waveforms at the Cz electrode for the Sternberg task. Data smoothed using a 5-sample moving average.

4 DISCUSSION

The aim of this study was to assess the impact of workload on auditory evoked potentials elicited by a single-stimulus paradigm with probes ignored by the participants. Workload was efficiently manipulated using a Sternberg paradigm with decreasing performance when the task increased in difficulty.
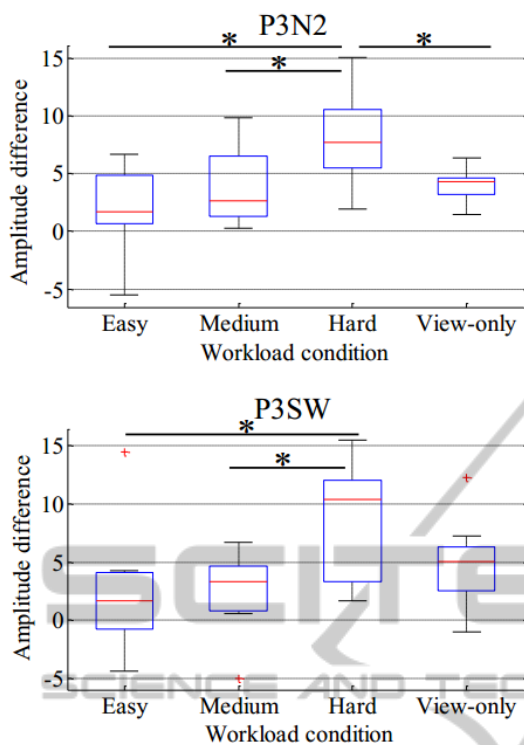
Figure 8: Box plots of the amplitude differences between the P3 and N2 components, and the P3 and SW components, at the Cz electrode for the Sternberg task. Average across participants.*p<0.05.

The main findings of this study are a significant modulation of the N2, P3 and SW triphasic AEP sequence for the Sternberg task. Indeed, the difference in amplitude between the P3 and each of the two adjacent components N2 and SW increased with workload. This phenomenon has never been described in the literature and could provide an interesting feature for mental workload estimation as it seems more robust to inter-participant variability than components' amplitude.

Indeed, no significant effect of workload was found for these amplitudes, nor for the latencies of the components at the group level. Although not significant, our components' amplitude results are in line with the work of Allison and Polich (2008). They introduced the use of the single-stimulus paradigm and workload impact on ignored probes, and reported larger components for the view-only condition. They also indicated that an increase in components' amplitude could be observed when workload increases, as we found for the N2 and SW components. This also explains the increase in amplitude difference for the P3N2 and the P3SW. Also, similarly, they did not report significant modulations of AEPs' amplitude between the easy

and medium, and easy and hard conditions with their experimental paradigm. Lastly, the absence of any significant modulation of our triphasic sequence for the ecological tasks may be due to a bad choice of ecological tasks, or to an insufficient engagement from the participants. These issues are critical for real-life experimentations.

However, this study is just a preliminary study and the results need to be further examined and confirmed using more participants, as well as more trials per condition and additional ecological tasks. Also, a problem of variable lag between trigger and sound release due to the use of the touchpad has appeared. It could have brought more variance and therefore reduced the effects. Next time, we will record the audio signal along with the EEG signal to realign our ERPs. That being said, these significant differences in amplitude might appear when performing classification at the subject level. Therefore, the next step should be to try and estimate each participant's mental workload using these amplitudes, as well as the P3N2 and P3SW differences and compare their efficiency. This should be done for laboratory-type tasks as well as for ecological tasks, as we have done in this study.

Our study brings new light on the use of AEPs, as well as the single-stimulus paradigm for mental state monitoring. Robust features such as differences in amplitude could be used for workload assessment in a non-intrusive way by probing operators with pure tones irrelevant to their task at hand.

5   CONCLUSIONS

This study fits into the mental state monitoring growing research environment. Mental workload assessment is a new challenge that can be tackled by evaluating the relevance of several neuro-physiological features, such as auditory evoked potentials. Our results show that the amplitude of these potentials elicited by pure tones in a single-stimulus fashion are modulated by workload for laboratory-type tasks as well as ecological tasks, although not significantly at the group level. However, the differences in amplitude between the adjacent components of the triphasic AEP sequence N2, P3 and SW were significantly modulated for our laboratory task. These promising results should be taken to the next step by comparing their relevance with other features using classification algorithms. The use of more electrodes as well as other recording modalities should also be considered to improve mental workload assessment. Finally, with

the aim to get closer to real-life implementation, a thorough ecological task battery setup should be designed to better ascertain workload modulation in work settings.

REFERENCES

Allison, B. Z., and Polich, J. (2008). Workload assessment of computer gaming using a single stimulus event-related potential paradigm. *Biological Psychology*, 77(3), pp. 277-283.

Berka, C., Levendowski, D. J., Lumicao, M. N., Yau, A., Davis, G., Zivkovic, V. T., Olmstead, R. E., et al. (2007). EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks. *Aviation, Space, and Environmental Medicine*, 78, 231-244.

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

Brouwer, A.-M., Hogervorst, M. A. , van Erp, J. B. F., Heffelaar, T., Zimmerman, P. H. and Oostenveld, R. (2012). Estimating workload using EEG spectral power and ERPs in the n-back task. *Journal of Neural. Enineering*, 9(4).

Cain, B. (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.

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

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

Filipe, S., Charvet, G., Foerster, M., Porcherot, J., Bêche, J.-F., Bonnet, S., Audebert, P., Régis, G., Zongo, B., Robinet, S., Condemine, C., Mestais, C., Guillemaud, R. (2011). A wireless multichannel EEG recording platform. In Conf. Proc. of the IEEE Eng. in Med. and Biol. Soc., Boston, Massachusetts USA, pp. 6319-6322.

Gevins, A., and Smith, M. E. (2003). Neurophysiological measures of cognitive workload during human-computer interaction. *Theoretical Issues in Ergonomics Science*, 1, 113-131.

Gevins, A., and Smith, M. E. (2007). Electroencephalography (EEG) in Neuroergnomics. In *Neuroergonomics: The brain at work*, R. Parasuraman, R and M. Rizzo Eds. New York, NY: Oxford.

Gomarus, H. K., Althaus, M., Wijers, A. A., and Minderaa, R. B. (2006). The effects of memory load and stimulus relevance on the EEG during a visual

selective memory search task: An ERP and ERD/ERS study. *Clinical Neurophysiology*, 117, 871-884.

Halgren, E., Marinkovic, K., and Chauvel, P. (1998). Generators of the late cognitive potentials in auditory and visual oddball tasks. *Electroencephalography and clinical Neurophysiology*. 106, pp. 156-164.

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

Kok, A. (2001). On the utility of P3 amplitude as a measure of processing capacity. *Psychophysiology*, 38, 557–577.

Miller, M. W., Rietschel, J. C., McDonald, C. G., and Hatfield, B. D. (2011). A novel approach to the physiological measurement of mental workload. International *Journal of Psychophysiology*, 80, 75-78.

Näätänen, R. Gaillard, A. W. K., and Varey, C. A. (1981). Attention effects on auditory EPs as a function of inter-stimulus interval. *Biological Psychology*, 13, pp. 173-187.

Natani, K., and Gomer, F. E. (1981). Electrocortical activity and operator workload: A comparison of changes in the electroencephalogram and in event-related potentials. (McDonnell Douglas Technical Report E2427). Long Beach, CA: McDonnell Douglas Corporation.

Roy, R. N., Bonnet, S., Charbonnier, S. and Campagne, A. (2013). Mental fatigue and working memory load estimation: Interaction and implications for EEG-based passive BCI. In Conf. Proc. of the IEEE Eng. Med. Biol. Soc., Osaka, Japan, pp. 6607-6610.

Schultheis, H. and Jameson, A. (2004) Assessing Cognitive Load in Adaptive Hypermedia Systems: Physiological and Behavioral Methods. *Lecture Notes in Computer Science*, 313, 225-234.

Smith, M. E., Halgren, E., Sokolik, M., Baudena, P., Musolino, A., Liegeois-Chauvel, C., and Chauvel, P. (1990). The intracranial topography of the P3 event-related potential elicited during auditory oddball. *Electroencephalography and clinical neuro-physiology*, 76, pp. 235-248.

Squires, N. K., Squires, K. C., and Hillyard, S. A. (1975). Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalography and clinical neuro-physiology,* 38, pp. 987-401.

Sternberg, S. (1966). High-speed scanning in human memory. *Science*, vol. 153, pp. 652-654.

Strobel, A., Debener, S., Sorger, B., Peters, J. C., Kranczioch, C., Hoechstetter, K., Engel, A. K., Brocke, B., and Goebel, R. (2008). Novelty and target processing during an auditory novelty oddball: A simultaneous event-related potential and functional magnetic resonance imaging study. *NeuroImage*, 40, 869-883.

Sugimoto, F. and Katayama, J. (2013). Somatosensory P2 reflects resource allocation in a game task: assessment with an irrelevant probe technique using electrical probe stimuli to shoulders. *International Journal of*

Psychophysiology, 87, 200-204.

Ullsperger, V., Freude, G. and Erdmann, U. (2001). Auditory probe sensitivity to mental workload changes - an event-related potential study. *International Journal of Psychophysiology*, 40, pp. 201-209.

Zander, T. O. and Kothe, C. (2011). Towards passive braincomputer interfaces: applying brain-computer interface technology to human-machine systems in general. *Journal of Neural. Enineering.*, 8(2):025005.