Visual Mental Workload Assessment from EEG in Manual Assembly Task

Miloš Pušica

School of Food Science & Environmental Health, Technological University Dublin, Ireland. mBrainTrain d.o.o., Belgrade, Serbia. E-mail: milos.pusica@mbraintrain.com

Carlo Caiazzo

Faculty of Engineering, University of Kragujevac, Serbia. E-mail: carlocaiazzo@fink.rs

Marko Djapan

Faculty of Engineering, University of Kragujevac, Serbia. E-mail: djapan@kg.ac.rs

Marija Savković

Faculty of Engineering, University of Kragujevac, Serbia. E-mail: marija.savkovic@kg.ac.rs

Maria Chiara Leva

School of Food Science & Environmental Health, Technological University Dublin, Ireland. E-mail: mariachiara.leva@tudublin.ie

The use of electroencephalography (EEG) to assess mental workload (MWL) has been the subject of many studies. Also, there have been many efforts to achieve task-independent MWL estimation, with the most recent being in the field of machine learning (ML). However, the estimation still remains highly dependent on the specific task used for ML model training. Furthermore, there is a shortage of research that is focused on developing an estimator that would function for multiple different tasks within a specific task domain. The creation of the dataset described in this work is a step towards developing task-independent ML estimator within the scope of visual cognition. An experiment meant for the ML model training is designed to collect EEG signals for different levels of MWL during manual assembly that involves assembly instructions to be visually processed by operators. It includes idle state of an operator, as well as two different complexity levels of the visual instructions. EEG data is collected using wireless EEG-recording cap that can be easily incorporated in everyday assembly line environments.

Keywords: mental workload, visual cognition, electroencephalography (EEG), manual assembly, experiment design, visual instructions, cross-task.

1. Introduction

Mental workload (MWL) assessment is crucial in situations where it is important to understand the cognitive demands placed on an individual, and make necessary adjustments to optimize safety, performance, and well-being. This is relevant in a variety of fields, such as Human Factors and Ergonomics, Transportation, Education and Training, Military, etc. MWL estimation is particularly significant in safety-critical systems (air traffic control, nuclear power plants, industrial control systems, etc.) as high mental workload can lead to fatigue, stress, lack of concentration, thereby causing mistakes, which can have severe implications for safety and reliability (Chen et al., 2016).

MWL is typically estimated through a combination of self-reports and physiological measurements that include EEG, Heart rate variability (HRV), Galvanic skin response (GSR), Eye tracking, etc. It is important to note that the estimation systems, if to be implemented in real-world workplaces, should not interfere with a human work capacity and their well-being. It

Proceedings of the 33rd European Safety and Reliability Conference
Edited by Mário P. Brito, Terje Aven, Piero Baraldi, Marko Čepin and Enrico Zio
 Published by Research Publishing, Singapore
ISBN: 981-973-0000-00-0 :: doi: 10.3850/981-973-0000-00-0 esrel2023-paper

means that the technology must be convenient for everyday use.

EEG is a tool that directly measures brain activity with high temporal and spatial resolution. Additionally, with advancements made in robustness, reliability, precision and convenience of wearable EEG technology, EEG comes on top as the most relevant source of information about human MWL (Hogervorst et al., 2014).

Most common methods for EEG processing for MWL estimation use signal power in specific frequency bands. A widely used spectral power metrics are mental workload index (or cognitive load index), calculated as a ratio of frontal theta and parietal alpha power and engagement index, commonly given with the formula $\beta/(\alpha+\theta)$ at C_z, P_z, P_3, P_4 electrode sites, where α, β, θ refere to signal powers in the alpha, beta and theta bands, respectively (Pope et al., 1995). It was shown that these metrics correlate with some objective metrics of task difficulty (Smith et al., 2001; Berka et al., 2007; Wilson, 2002; Kartali et al., 2019) and subjective self-assessments (Berka et al., 2007). In addition to this, machine learning methods like Linear Discriminant Analysis (LDA), Support Vector Machines (SVM), Bayes-based models, etc. (Zhou et al., 2021), as well as deep neural networks (Craig et al., 2019), are extensively employed. However, ML estimators tend to perform well only for the specific task used in the training dataset. This means that, if to be used in real life, these estimators would have to be trained with the data collected from the same task for which we want to apply the estimator. This is impractical for two reasons. Firstly, the data collection for the training is a challenging and time demanding job, and it is impossible to collect the data for every different task. Secondly, the work performed by humans in real life usually includes multitasking in some form and task type at hand varies throughout time. Hence, it is not clear how we would collect and label the data in this complex scenario. In some experiments in the literature, the estimators were trained for a certain task and then tested for a different one, but the results for cross-task MWL estimation were not satisfactory, even when the results withing the same task were very good (Zhang et al., 2018; Baldwin, Penaranda, 2012).

However, some tasks engage the same cognitive capacities and are more similar in that sense than the others. For example, tasks from the area of visual cognition activate comparable cognitive processes related to vision and therefore create the same kind of MWL. Having this in mind, we could expect that an estimator trained for a specific task could also perform well on a different task that activates the same brain processes – in this case visual cognition. However, to the best of our knowledge, there has not been much research on cross-task MWL estimation withing a group of tasks engaging similar cognitive processes.

This paper explains EEG and performance-related data collection during manual assembly task, with subjects following visual instructions. The experiment is aimed at evaluating the impact of visual instructions complexity on MWL as evaluated by EEG measurements. It also explains the motivation behind the experiment design and how the data will be used in future studies towards developing a task-domain specific MWL estimator using ML.

2. Metrics for Mental Workload Modelling in Manual Assembly

During the process of manual assembly, mental and physical state of the operator vary depending on the factors such as fatigue, stress, distractions, training and experience, but also task complexity at hand. To have an insight into, possibly in real-time, MWL levels of the operator and model operator performance, certain performance metrics should be monitored. Assuming that the operator on a production line is assembling one item after the other, we can identify the following behavioral metrics that are also measured in our experiment:

- Error rate (item assembled correctly/ incorrectly)
- Assembly time (for one item)
- Self-assessment questionnaire
- Item complexity

An error rate indicates whether the assembly was performed accurately in accordance with the given instructions or if any mistakes were made during the process. It is used as a metric in many studies assessing MWL (Pankok Jr et al., 2017; Kosch at al., 2018; Li et al., 2018). An error made assembling an item could mean that either the task was too difficult, or the operator had lost the focus, meaning that there was a decrease in the MWL level. However, if the task difficulty is kept at the constant level, an error rate is a more reliable indicator of the MWL level.

Assembly time is another measure relevant when assessing MWL. Higher assembly time could mean that either the task was more difficult, or that the operator had lower MWL level. On the other side, if the task difficulty is held constant, higher assembly time is an indicator that the operator invested less mental effort/workload for that time.

A widely used method for obtaining the operator's perspective related to the task experience is through the use of questionaries. There are different kinds of questionaries for MWL assessment: NASA-TLX, SWAT, etc. They can give a good insight into information not otherwise available through real-time physiological measurements. However, filling out questionaries require work interruption and hence it is not applicable for continuous work settings.

Objectively quantifying assembly item complexity is another way towards monitoring MWL. By comparing the complexities of two items, we can infer that if an operator successfully assembles both, they likely put in more effort to assemble the more complex item, indicating a higher MWL level during that task. Many experiments designed for assessing varying MWL levels have employed multiple complexity levels of a certain task (Pankok Jr et al., 2017; Van Acker et al. 2020). The selection of the task type used depends on the specific cognitive demands that we want to evaluate, e.g. memory tasks, attention tasks, problem-solving tasks, etc.

EEG is considered the most relevant physiological measurement for continuous and real-time MWL tracking. Metrics derived from EEG spectral power bands are shown to correlate with the above mentioned MWL metrics. Additionally, EEG is a tool that can be conveniently applied in diverse work environments. Typical metrics used include mental workload index and engagement index. Moreover, a study conducted by Pope et al., 1995 showed that by modifying task demands according to EEG-related indices can lead to the increase in productivity. Although the current metrics are not responsive to the changes in MWL levels that occur frequently over time, this suggests that enhancing EEG-related metrics (e.g. with machine learning) can optimize work environments and increase productivity in various types of workplaces.

3. Experiment

The purpose of the experiment was to create a dataset that could be used to evaluate human mental workload in the domain of visual cognition during manual assembly task as measured by EEG. The idea of the design was to impose variable levels of MWL by changing assembly items complexities. In the following text, we provide an overview of the experimental task and its underlying motivation, as well as a step-by-step description of the experiment procedure.

3.1. Task paradigm

The goal of the task was to induce different levels of MWL specifically related to visual cognition. Also, the task was meant to simulate manual assembly line workplace. To put together these design goals, the assembly instructions were created to be visually engaging and of distinct complexities and hence require different levels of mental effort to be completed. A subject was given a plexiglass plate with 20 empty holes with switches and asked to connect the holes with wires and toggle the switches according to the visual instructions given on the screen. Basically, participants were asked to assemble a scheme in the same way as it was shown in a picture. As mentioned before, the schemes were of varying complexities. Namely, there were two difficulty types. Easier type schemes were given as shown in Fig. 1 and they were easy to comprehend. It was simple to see which pairs of holes should be connected with wires. On the other side, harder type schemes were designed to be more challenging to grasp. They were presented as photos of assembled items taken from non-ideal angles, with wires tangled up (Fig. 2). Additionally, photos were sometimes rotated to add to the complexity of the task. During the experiment, schemes were presented sequentially on the screen in front of a subject. Also, every scheme had a time limit for its completion: for easier schemes it was given 60 seconds and for harder schemes it was given 90 seconds of assembly time. If the subject would not complete a scheme before the time lapsed, he would leave the current scheme and the corresponding plexiglass plate, take another empty plate and move on with the next scheme appearing on the screen. Otherwise, if the subject completed the scheme in less time than given, he would move on to the next one by clicking on the screen. More time was given for harder schemes as it was anticipated that they would be more demanding.

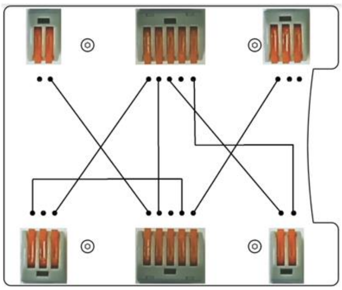

Fig. 1. Low complexity instruction type

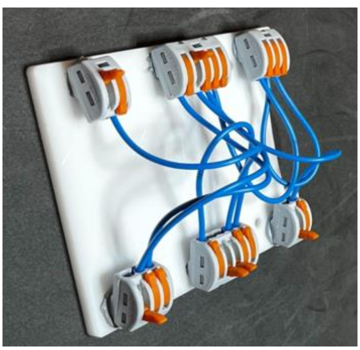

Fig. 2. High complexity instruction type

3.2. Experimental paradigm

The experiment was conducted in a controlled environment, in a modular and adaptive laboratory set-up of the industrial assembly workstation described in Savković et al. 2022. A total of 32 participants were recorded. Every participant was engaged with the task for up to three hours - since the time available per scheme was upper-limited and the next scheme would automatically appear on the screen after the time for the current scheme elapsed, total session time was consequently upper-limited, too. The task was divided into two equal-length sessions to allow for a ten min. break between them. During the break subjects were free to stand up, walk around, and come back for the second session. Every subject has been given a training time to get familiar with the task. Before the task started, baseline measurement was taken for 5 min. while a subject was relaxing. The task consisted of a total of 150 schemes appearing sequentially – 90 easy and 60 hard schemes. Since the time limits for easy and hard schemes were set at 60 and 90 seconds respectively, this resulted in equal maximum time spent on the two scheme types. No scheme was repeated throughout the task. Plates stack was placed on the right side of the subject, and wires pile was located in front, to add to the ergonomic design of the laboratory work station. During the experiment, the subject was seated in a height-adjustable industrial work chair. A touchscreen screen was placed in front of the subject with visual instructions shown one at a time as shown in Fig. 3.

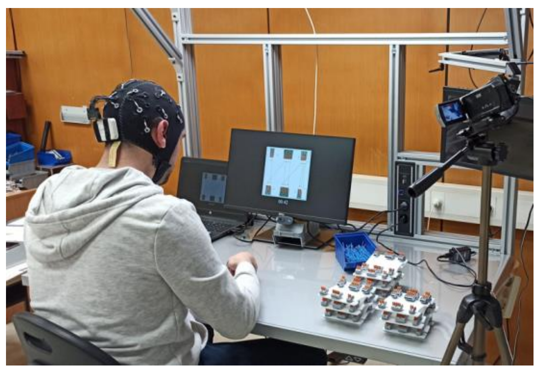

Fig. 3. A subject during the task

Additionally, there was an assistant present during the experiment, who would collect and remove the plates that the subject would place on the left. Also, the assistant would note down for each item whether it was assembled completely or not. For the whole duration of the experiment, the subjects were recorded using wireless EEG cap and frontal and side camera.

3.3. Equipment

To record the brain's activity while performing the task, a gel-based wireless EEG device (mBrainTrain LLC, Belgrade, Serbia) with 24 EEG channels and three accelerometer channels and a sampling rate of 250Hz is employed. The electrodes are integrated into a cap that conforms to the internationally recognized 10-20 system for electrode positioning (Homan, 1988).

The experiment was programmed in the experiment control software *Presentation*, developed by Neurobehavioral Systems (Neurobehavioral Systems, 2021).

Two cameras recorded the experiment. One was placed on the side and the other was placed in front, below the screen. Frontal camera recorded head movements.

Every subject completed a study-specific questionaries about their experience with the experiment.

4. Results

Following the completion of the experiment, each participant was requested to fill out a questionnaire, which included questions about various aspects of their experience with the experiment. In line with our hypothesis regarding different difficulty levels of the two types of schemes, subjects confirmed that assembling high complexity schemes was more challenging. Also, they did not report any special assembly strategy they could find to use for either type of schemes. Further, it was found that participants worked on the task as fast as they could and mostly did not pay too much attention to the timer displayed on the screen, which is a positive indication, meaning that they were engaged and focused on the task at hand. However, after the time limit for a scheme lapsed, they would immediately leave the scheme and proceed with another one, as instructed by the experiment design. Some participants reported occasional difficulties when handling switches, as they would sometimes close switches they did not intend to.

After the analysis of ten subjects and looking at the times spent on every scheme, it was found that 36.4% of easy schemes were completed in less than 60s (time limit for easy schemes), 49.6% of hard schemes were completed in less than 90s (time limit for hard schemes). However, only 3.6% of hard schemes were completed in time under the time limit for easy schemes, confirming that participants needed more time for harder schemes, as expected by the experiment design.

EEG signals were pre-processed for these ten subjects. Pre-processing was done in Matlab toolbox, EEGLAB (Delorme, Makeig, 2004). The signals were band-pass filtered 1-40 Hz, bad channels were interpolated using pop_interp() function, pop_clean_rawdata() function is executed for artifacts removal and Independent Component Analysis (ICA) was employed for further artifacts removal. All the steps were done through EEGLAB built-in methods. For every subject, MWL index, calculated as a ratio of frontal theta and parietal alpha EEG power was computed for both types of the schemes. Sessions were divided into 2s windows with an overlap of 1s and the index was calculated for every window. Making an average over all the windows belonging to easy schemes assembly time, we got one value for easy schemes for the subject. The same was done for hard schemes. Results can be seen in Fig. 4.

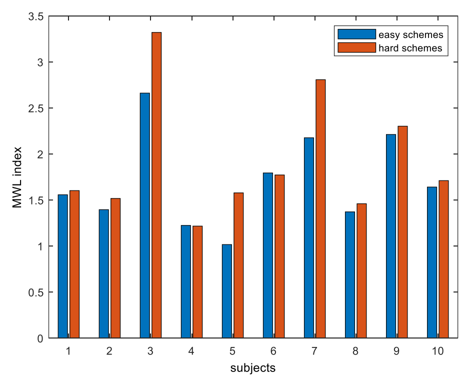

Fig. 4. Bar plot for easy vs hard schemes MWL index

It is clear that MWL index, in majority of the cases, has bigger value for hard schemes, which supports the hypothesis that this schemes type imposes higher cognitive burden on the participants. Only for two subjects the index had nearly the same value for easy and hard schemes type.

5. Conclusions

In the EEG study presented, 30 subjects and around 90h of manual assembly task with two different visual instructions complexities was recorded. The experiment was conducted in a controlled laboratory environment, that simulated a real assembly line workplace. The EEG signals were recorded together with task-related logs (time instances when each instruction scheme appeared on the screen). The experiment was recorded using two cameras. Additionally, data regarding the completion status of each scheme was recorded manually. It is important to note that the only difference between lower and higher complexity instructions was in the visual complexity. This makes the experiment highly appropriate for assessing MWL in the domain of visual cognition.

Ten subjects were processed and analyzed. Findings regarding the times spent on schemes assembly and EEG powerbands-related MWL measure support and justify the design of the two difficulty types of schemes.

6. Future Work

All the data recorded will be analyzed in the follow-up studies. EEG powerbands-related indexes will be extracted and the values for lower and higher complexity schemes will be compared. In addition to the traditional EEG metrics analysis, a neural network will be trained with the experiment data. The network will be trained to classify between EEG segments from lower and higher complexity schemes. To examine whether a neural network trained on the dataset for visual MWL estimation can be used on other experiments from the domain of visual cognition, the network will be evaluated on an external, independent dataset. The outcome of this assessment would provide us with additional insights into the feasibility of estimating mental workload across tasks in the area of visual perception.

Acknowledgement

This research paper was financed from the European Union's H2020 research project under the Marie Skłodowska-Curie Actions Training Network Collaborative Intelligence for Safety-Critical Systems (Grant Agreement ID: 955901).

References

Pope, Alan T., Edward H. Bogart, and Debbie S. Bartolome. "Biocybernetic system evaluates indices of operator engagement in automated task." Biological psychology 40, no. 1-2 (1995): 187-195.

Smith, Michael E., Alan Gevins, Halle Brown, Arati Karnik, and Robert Du. "Monitoring task loading with multivariate EEG measures during complex forms of human-computer interaction." Human Factors 43, no. 3 (2001): 366-380.

Kartali, Aneta, Milica M. Janković, Ivan Gligorijević, Pavle Mijović, Bogdan Mijović, and Maria Chiara Leva. "Real-time mental workload estimation using eeg." In Human Mental Workload: Models and Applications: Third International Symposium, H-WORKLOAD 2019, Rome, Italy, November 14–15, 2019, Proceedings 3, pp. 20-34. Springer International Publishing, 2019.

Berka, Chris, Daniel J. Levendowski, Michelle N. Lumicao, Alan Yau, Gene Davis, Vladimir T. Zivkovic, Richard E. Olmstead, Patrice D. Tremoulet, and Patrick L. Craven. "EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks." Aviation, space, and environmental medicine 78, no. 5 (2007): B231-B244.

Zhou, Yueying, Shuo Huang, Ziming Xu, Pengpai Wang, Xia Wu, and Daoqiang Zhang. "Cognitive workload recognition using EEG signals and machine learning: a review." IEEE Transactions on Cognitive and Developmental Systems (2021).

Craik, Alexander, Yongtian He, and Jose L. Contreras-Vidal. "Deep learning for electroencephalogram (EEG) classification tasks: a review." Journal of neural engineering 16, no. 3 (2019): 031001.

Zhang, Xixie, Laurens R. Krol, and Thorsten O. Zander. "Towards task-independent workload classification: shifting from binary to continuous classification." In 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 556-561. IEEE, 2018.

Savković, Marija, Carlo Caiazzo, Marko Djapan, Arso M. Vukićević, Miloš Pušica, and Ivan Mačužić. "Development of modular and adaptive laboratory set-up for neuroergonomic and human-robot interaction research." Front. Neurorobotics (2022).

Homan, Richard W. "The 10-20 electrode system and cerebral location." American Journal of EEG Technology 28, no. 4 (1988): 269-279.

Neurobehavioral Systems, Inc., Berkeley, CA, www.neurobs.com, Presentation

mBrainTrain, https://mbraintrain.com/smarting-mobi/

Pankok Jr, Carl, Maryam Zahabi, Wenjuan Zhang, Inchul Choi, Yi-Fan Liao, Chang S. Nam, and David Kaber. "The effects of interruption similarity and complexity on performance in a simulated visual-manual assembly operation." Applied ergonomics 59 (2017): 94-103.

Van Acker, Bram B., Klaas Bombeke, Wouter Durnez, Davy D. Parmentier, Joao Costa Mateus, Alessandro Biondi, Jelle Saldien, and Peter Vlerick. "Mobile pupillometry in manual assembly: A pilot study exploring the wearability and external validity of a renowned mental workload lab measure." International Journal of Industrial Ergonomics 75 (2020): 102891.

Chen, Jiayu, Xinyi Song, and Zhenghang Lin. "Revealing the “Invisible Gorilla” in construction: Estimating construction safety through mental workload assessment." Automation in Construction 63 (2016): 173-183.

Hogervorst, Maarten A., Anne-Marie Brouwer, and Jan BF Van Erp. "Combining and comparing EEG, peripheral physiology and eye-related measures for the assessment of mental workload." Frontiers in neuroscience 8 (2014): 322.

Wilson, Glenn F. "An analysis of mental workload in pilots during flight using multiple psychophysiological measures." The International Journal of Aviation Psychology 12, no. 1 (2002): 3-18.

Baldwin, Carryl L., and B. N. Penaranda. "Adaptive training using an artificial neural network and EEG metrics for within-and cross-task workload classification." NeuroImage 59, no. 1 (2012): 48-56.

Kosch, Thomas, Markus Funk, Albrecht Schmidt, and Lewis L. Chuang. "Identifying cognitive assistance with mobile electroencephalography: A case study with in-situ projections for manual assembly." Proceedings of the ACM on Human-Computer Interaction 2, no. EICS (2018): 1-20.

Li, Dan, Sandra Mattsson, Omkar Salunkhe, Åsa Fast-Berglund, Anders Skoogh, and Jesper Broberg. "Effects of information content in work instructions for operator performance." Procedia Manufacturing 25 (2018): 628-635.

Delorme, Arnaud, and Scott Makeig. "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis." Journal of neuroscience methods 134, no. 1 (2004): 9-21.