

Modeling Mental Workload using EEG Features for Intelligent Systems

Maher CHAOUACHI, Imène JRAIDI, Claude FRASSON

HERON Lab, Computer Science Department
University of Montreal,
2920 chemin de la tour, H3T 1N8, Canada
{chaouacm, jraidiim, frasson}@iro.umontreal.ca

Abstract. Endowing systems with abilities to assess a user's mental state in an operational environment could be useful to improve communication and interaction methods. In this work we seek to model user mental workload using spectral features extracted from electroencephalography (EEG) data. In particular, data were gathered from 17 participants who performed different cognitive tasks. We also explore the application of our model in a non laboratory context by analyzing the behavior of our model in an educational context. Our findings have implications for intelligent tutoring systems seeking to continuously assess and adapt to a learner's state.

Keywords: cognitive workload, EEG, ITS.

1 Introduction

Modeling and developing systems able to assess and monitor users' cognitive states using physiological sensors has been an important research thrust over few past decades [1-5]. These physio-cognitive systems aim to improve technology's adaptability and have shown a significant impact on enhancing users' overall performance, skill acquisition and productivity [6]. Various models tracking shifts in users' alertness, engagement and workload have been successfully used in closed-loop systems or simulation environment [7-9]. By assessing users' internal state, these systems were able to adapt to users' information processing capacity and then to respond accurately to their needs.

Mental workload is of primary interest as it has a direct impact on users' performance in executing tasks [10]. Even though there is no agreement upon its definition, mental workload can be seen in terms of resources or mental energy expended, including memory effort, decision making or alertness. It gives an indication about the amount of effort invested as well as users' involvement level. Hence, endowing systems with workload assessment models can provide intelligent assistance, efficient adaptation and more realistic social communication in the scope of reaching optimal interaction conditions.

Nevertheless, scarce and scattered research has explored these faculties to refine the learning process within educational settings. Intelligent tutoring systems (ITS) are still mainly based on learners' performance in analyzing learners' skill acquisition

process or evaluating their current engagement level and the quality of learning [11-14]. Even though the integration of affective models in ITS added an empathic and social dimension into tutors' behaviors [15-17] there is still a lack of methods helping tutors to drive the learning process and evaluate learners' behavior according to their mental effort. The limited action range offered to learners (menu choice, help, or answers) restricts the ability of forecasting learners' memory capacity and objectively assessing their efforts and implication level [18]. EEG techniques for workload assessment can represent, then, a real opportunity to address these issues. The growing progress in developing non intrusive, convenient and low cost EEG headsets and devices are very promising enabling their use in operational educational environments.

In this paper we model users' workload in a learning environment by developing an EEG workload index and we analyze its behavior in different phases across the learning process. In particular, we performed an experiment with two phases: (1) A cognitive phase, in which users performed different cognitive tasks, was used to derive the workload index. (2) A learning phase during which the developed index was validated and analyzed. We performed data analysis using machine learning techniques and showed that there are identifiable trends in the behavior of the developed index.

The rest of this paper is structured as follows. Section 2 presents previous work on EEG based workload detection approaches. Section 3 presents our experimental methodology. Section 4 discusses our approach and its implications. We conclude in Section 5.

2 Previous Work

Developing EEG indexes for workload assessment is a well developed field especially in laboratory contexts. A variety of linear and non-linear classification and regression methods were used to determine mental workload in different kinds of cognitive tasks such as memorization, language processing, visual, or auditory tasks. These methods use mainly EEG Power Spectral Density (PSD) bands or Event Related Potential (ERP) techniques to extract relevant EEG features [7-9].

Wilson [19] used an Artificial Neural Network (ANN) to classify operators' workload level by taking EEG as well as physiological features as an input. Reported results showed up to 90% of classification accuracy. Gevins and Smith [20] used spectral features to feed neural networks classifying workload of users while performing various memory tasks. In a car driving simulation environment, Kohlmorgen et al. [21] used Linear Discriminant Analysis (LDA) on EEG-features extracted and optimized for each user for workload assessment. Authors showed that decreasing driver workload (induced by a secondary auditory task) improves reaction time. Support Vector Machine (SVM) and ANN were also used to analyze task demand recorded in lecture and meeting scenario as well as in others cognitive tasks (Flanker paradigm and Switching paradigm) using EEG features. Results reached 92% of accuracy in discriminating between high and low workload levels [22, 23].

Berka and colleagues [1, 24] developed a workload index using Discriminant Function Analysis (DFA) for monitoring alertness and cognitive workload in different environments. Several cognitive tasks such as grid location task, arithmetic computing, and image memorization were analyzed to validate the proposed index. The same index was used in an educational context to analyze students' behavior while acquiring skills in a problem solving context [18, 25, 26].

In this paper, we propose to model users' workload from EEG extracted features through a cognitive task activity. The major contribution of this study is to validate the model within a learning activity as opposed to the work of Berka and colleagues [1, 24] where the proposed index was validated only according to purely cognitive tasks. Our study uses Gaussian Process Regression to train workload models in a first phase especially designed to elicit different levels of workload and applied in a second phase, within a learning environment to detect different trends in learners' mental workload behavior. We will now describe our experiment.

3 Methodology

The aim of this study was to model and evaluate mental workload induced during human-computer interaction using features extracted from EEG signals. The experimental process was divided into two phases: a cognitive activity phase including three cognitive tasks designed with incrementally increasing levels of difficulty to elicit increasing levels of required mental workload and a learning activity phase about trigonometry consisting of three main steps. Data gathered from the first phase were used to thoroughly derive a workload index whereas data from the second phase were used to validate the computed index in a "non-laboratory" context. Our experimental setup consists of a 6-channel EEG sensor headset and two video feeds. All recorded sessions were replayed and analyzed to accurately synchronize the data using necessary time markers.

17 participants were recruited for this research. All participants were briefed about the experimental process and objectives and signed a consent form. Participation was compensated with 20 dollars. Upon their arrivals, participants were equipped with the EEG-cap and familiarized with the materials and the environment. All subjects completed a five minutes eyes open baseline followed by another five minutes eyes closed baseline. During this period, participants were instructed neither to be active nor to be relaxed. This widely used technique enabled us to establish a neutral reference for the workload assessment. Then, participants completed the cognitive activity phase. This phase consists of three successive tasks: (1) Forward Digit Span (FDS) (2) Backward Digit Span (BDS) and (3) Logical Tasks (LT). Each task has between three and six difficulty levels. All participants performed these tasks in the same order and were allowed to self-pace with respect to the time required to complete each task.

Forward Digit Span (FDS). This test involves attention and working memory abilities. In this task, a series of single digits were successively presented on the screen. Participants were asked to memorize the whole sequence, then prompted to

enter the digits in the presented order. This task included 6 difficulty levels incrementally increasing by increasing the length of the sequence that participants have to retain. Level one consisted of a series of 20 sets of 3 digits, level two: 12 sets of 4 digits, level three: 8 sets of 5 digits, level four: 6 digits and 6 sets, level five: 7 digits and 4 sets and level six: 4 sets of 8 digits.

Backward Digit Span (BDS). The principle of this test is similar to the FDS task. Participants had to memorize a sequence of single digits presented on the screen. The difference was that they were instructed to enter digits in the reverse order from the one presented. Five levels of difficulty were considered by increasing the number of digits in the sequence. The first level consisted of a series of 12 sets of 3 digits; the second level: 12 sets of 4 digits, the third level: 5 digits and 8 series, the fourth level: 6 sets of 6 digits and the fifth level: 4 sets of 8 digits. No time limit was fixed for FDS and BDS tasks.

Logical Tasks (LT). These tasks require inferential skills on information series and are typically used in brain training exercises or in tests of reasoning. In these tasks, participants were instructed to deduce a missing number according to a logical rule that they had to infer from a series of numbers displayed on the screen, within a fixed time limit of 30 seconds. An example of such series is “38 - 2 - 19 - 9 - 3 - 3 - 40 - 4 - ?” and one should deduce that the missing element (“?”) would be “10”. That is, the logical rule that one should guess is that for each group of three numbers the last number is the result of the division of the first by the second. The logical tasks involved three difficulty levels. Each level consisted of a series of 5 questions and the difficulty level was manipulated by enhancing the difficulty of the logical rule between the data.

After completing the cognitive activity phase, participants took a little break, and then were invited to perform the learning phase during which a trigonometry course was given. This phase consists of three successive steps (1) a pretest session, (2) a learning session and (3) a problem solving session. Before starting these tasks, participants were asked to report their self-estimated skill level in trigonometry (low, moderate or expert).

Pretest. This task involved 10 (yes/no/no-response) questions that covered some basic aspects of trigonometry (for instance: “is the tangent of an angle equal to the ratio of the length of the opposite over the length of the adjacent?”). In this part, participants had to answer to the questions without any interruption, help or time limit.

Learning Session. In this task, participants were instructed to use a learning environment covering the theme of trigonometry and specially designed for the experiment. Two lessons were developed explaining several fundamental trigonometric properties and relationships. The environment provides basic

definitions as well as their mathematical demonstrations. Schemas and examples are also given for each presented concept.¹

Problem Solving. Problems presented during this task are based on participants' ability to apply, generalize and reason about the concepts seen during the learning session. No further prerequisites were required to successfully resolve the problem except the lessons' concepts. However a good level of reasoning and concentration is needed to solve the problems. A total of 6 problems with a gradually increasing difficulty level were selected and presented in the same order for all the participants. Each problem is a multiple-choice question illustrated by a geometrical figure. A fixed time limit is imposed for each problem varying according to its difficulty level. Each problem requires some intermediate steps to reach the final solution and the difficulty level was increased by increasing the number of required intermediate steps. For example, to compute the sinus of an angle, learners had first to compute the cosines in the first step. Then, they had to square the result and to subtract it from 1 in the second step. Finally the third step consisted of computing the square root.

The problem solving environment provided a limited number of hints for each problem as well as a calculator to perform the needed computations. All the problems were independent in terms of learned concepts except for problem 4 and 6 that shared the same geometrical rule required to solve the problem (i.e. "the sum of the angles of a triangle is equal to 180 degrees").

3.1 Subjective and Objective Measures of Workload

After completing each task level, participants were asked to evaluate their mental workload both in the cognitive activity phase and the learning activity phase. We used the *NASA-Task Load Index* (NASA-TLX) technique [27]. As for other subjective measures of workload, NASA-TLX relies on subjects' conscious perceived experience with regards to the effort produced and the difficulty of task. NASA_TLX has the advantage of being quick and simple to administer.

In addition to the subjective ratings, other objective factors that may be used for assessing workload were considered in this study, such as performance (i.e. proportion of correct answers in cognitive tasks, pretest and problem solving) and response time.

3.2 EEG Recording

EEG is a measurement of the brain electrical activity produced by synaptic excitations of neurons. In this experiment, EEG was recorded using a stretch electro-cap. EEG signals were received from sites P3, C3, Pz and Fz as defined by the international 10-20 electrode placement system (Jasper 1958). Each site was referenced to Cz and grounded at Fpz. Two more active sites were used namely A1 and A2 typically known respectively as the left and right earlobe. This setup is known as "referential

¹ At the end of the experiment, all participants reported that they were satisfied with the quality of the environment as well as the pedagogical strategy used for presenting the materials.

linked ear montage” and is illustrated in figure 1. Roughly speaking, in this montage the EEG signal is equally amplified throughout both hemispheres. Moreover, the “linked ears” setup yields a more precise and cleaner EEG signal by calibrating each scalp signal to the average of the left and right earlobe sites (A1 and A2). For example, the calibrated C3 signal is given by $(C3-(A1+A2)/2)$. Each scalp site was filled with a non-sticky proprietary gel from Electro-Cap and impedance was maintained below 5 Kilo Ohms. Any impedance problems were corrected by rotating a blunted needle gently inside the electrode until an adequate signal was obtained. The recorded sampling rate was at 256 Hz.

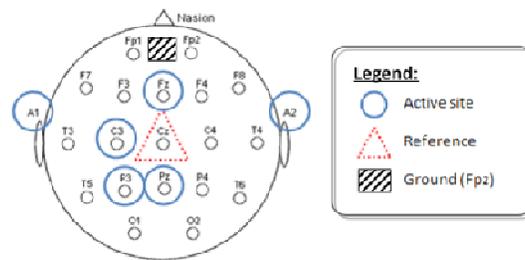


Figure 1. EEG channel electrode placement

Data Preprocessing and Features Extraction. Due to its weakness (on the order of micro volts: 10^{-6} volts), the EEG signal needs to be amplified and filtered. The brain electrical activity signal is usually contaminated by external noise such as environmental interference caused by surrounding devices. Such artifacts alter clearly the quality of the signal. Thus a 60-Hz notch filter was applied during the data acquisition to remove these artifacts. In addition, the acquired EEG signal easily suffers from noise caused by user body movements or frequent eye blinks or movement. Therefore, an artifact rejection heuristic was applied to the recorded data using a threshold on the signal power with regards to the eyes open and eyes closed baseline. If the amplitude of any epoch in any channel exceeded the threshold by more than 25%, the epoch was considered as contaminated and was excluded from subsequent analysis.

For each participant, EEG data recorded from each channel were transformed into a power spectral density using a Fast-Fourier Transform (FFT) applied to each 1-second epoch with a 50 % overlapping window multiplied by the Hamming function to reduce spectral leakage. Bin powers (the estimated power over 1 Hz) of each channel ranging from 4 Hz to 48 Hz were concatenated to constitute the feature vector.

To sum up, 176 features (4 channels x 48 bins) were extracted from the signal each second. To reduce the data input dimensionality and improve the workload model, a Principal Component Analysis (PCA) was applied on the EEG data of each participant. The number of features was reduced to 25 (78.63% reduction rate) explaining 85.98% to 94.71% of the variability ($M = 90.42\%$, $SD = 3.30\%$). PCA scores were then z-normalized and the resulting vector used as an input for the model.

4 Results and Discussions

The experimental results are presented in the following subsections. The first part is concerned with the cognitive phase data analysis. The second part describes mental workload modeling from the EEG. The third part deals with the validation and evaluation of the model within the learning activity.

4.1 Cognitive Activity Results

In order to evaluate NASA_TLX subjective estimates of mental workload, correlational analysis was performed with regards to the task design and performance and response time objective variables.

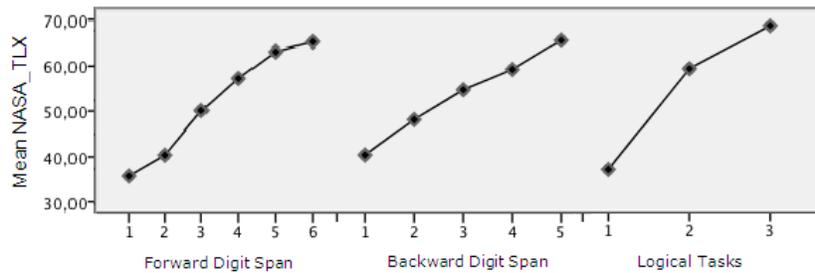


Figure 2. Mean NASA_TLX workload estimate for each difficulty level for the forward digit span, backward digit span and logical tasks

Repeated measures one-way ANOVA ($N = 17$) was performed in the FDS, LT, and BDS cognitive activities across their associated difficulty levels. Degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity for FDS and BDS (epsilon = 0.35 and 0.54 respectively) as the assumption of sphericity had been violated (chi-square = 60.11 and 54.40 respectively, $p < .05$) while Mauchly's test was non significant for the LT. Results revealed significant changes between NASA_TLX scores with regards to the task demand of each level in the three tasks ($F(1.73, 27.65) = 25.65$, $p < 0.001$ for FDS, $F(2.18, 34.89) = 18.25$, $p < 0.001$ for BDS and $F(2, 32) = 43.51$, $p < 0.01$ for LT), showing a significant linear increase of subjective workload estimates as the level of difficulty increased (figure 2).

Bi-variate correlations between NASA_TLX and objective measures of task performance and response time are illustrated in figure 3 (a). Correlations were calculated for each individual across the 14 levels (FDS, BDS and LT). The median correlation between NASA_TLX and task performance was -0.58 with a range of -0.89 to 0.15 and the median correlation between NASA_TLX and response time was 0.42 with a range of -0.16 to 0.72. Correlations were also computed for each of the 3 activities across the 17 participants. Performance decreased linearly as the workload increased in FDS ($r = -0.59$, $p < 0.001$) and LT ($r = -0.50$, $p < 0.001$) while the relationship was not linearly significant in the BDS ($r = -0.19$, $p = \text{n.s.}$). Response

time increased linearly as a function of workload in FDS ($r = 0.30$ $p < 0.05$) and LT ($r = 0.35$, $p < 0.05$) and did not linearly correlate in BDS ($r = 0.014$, $p = \text{n.s.}$)

To summarize, NASA_TLX workload ratings showed linear relationship with the objective measures except for the BDS tasks. Besides, NASA_TLX accurately tracked the intended pattern of the task design that used incrementally increasing levels of difficulty to elicit increasing levels of mental workload required for the task. This manifest trend suggests that NASA_TLX subjective ratings can be a reliable indicator of mental workload for training the predictive models.

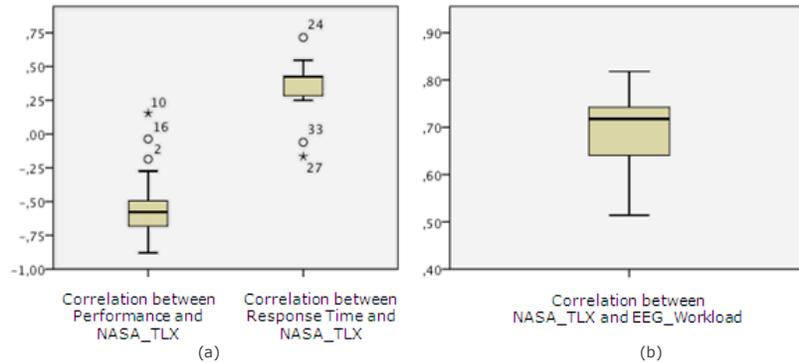


Figure 3. Bivariate correlations between: (a) NASA_TLX and objective measures (b) NASA_TLX and EEG_Workload

4.2 EEG Mental Workload Index

Our aim was to quantitatively predict the mental workload using EEG extracted features. Specially, we were interested in using workload patterns detected in the cognitive activity phase to analyze the learning activity. An individual model was developed for each participant by training a Gaussian Process Regression (GPR) function [28] with a squared exponential kernel and noise parameter $\sigma^2 = 0.1$. NASA_TLX subjective ratings were introduced as a target variable of the training data in all models.

NASA_TLX workload values were classified into three classes namely low, medium and high (low < 30 , $30 \leq \text{medium} < 70$ and high ≥ 70). The same classification was done for the predicted GPR values. A mean accuracy rate of 91% across all the participants was reached with models trained on a split of 90% of the cognitive task data and tested on the other 10%. The mean EEG_Workload indices derived by the model were computed across each task of the learning activity phase.

4.3 Learning Activity Results

Our next objective was to validate the computed EEG_Workload model within the learning context. The box plot in figure 3 (b) illustrates the results of the bi-variate

correlations between EEG_Workload and NASA_TLX subjective metrics. Correlations were computed across participants in the pretest, the learning session and each of the six problems revealing significant relationships in the eight activities. The median correlation was 0.72 with a range of 0.52 ($p < 0.001$) for problem 2 to 0.82, ($p < 0.05$) for problem 1. These results provide confirmation of the validity of the computed EEG model of mental workload.

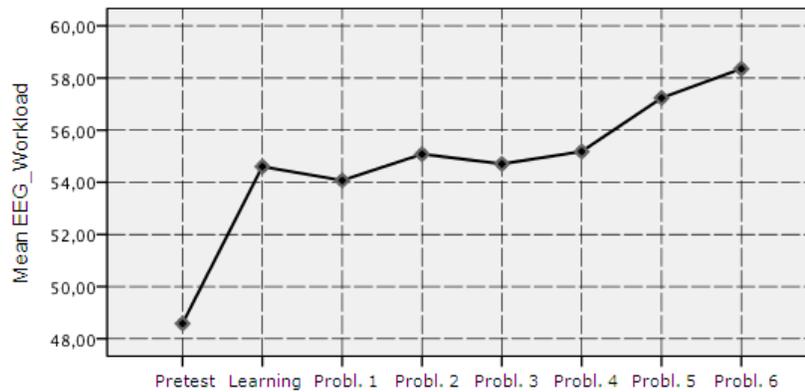


Figure 4. Mean EEG_Workload for each activity in the learning environment

Our next investigation was to evaluate the progression of the workload level across the learning tasks. A repeated measures ANOVA revealed that there were significant changes in the EEG_Workload between the learning activities $F(3.23, 51.61 = 2.76, p < 0.05)$. Degrees of freedom were corrected using the Greenhouse-Geisser estimates of sphericity ($\epsilon = 0.46$). Post hoc results showed that the EEG_Workload measures significantly increased during the learning session when compared with the pretest ($p < 0.05$). This increase can be explained by the effort produced by learners in understanding concepts and acquiring skills in the learning phase compared to the pretest session where learners responded to questions that did not require particular engagement and concentration levels. In fact, during the pretest no pressure was put on learners who had simply to situate their knowledge in trigonometry.

Significant increases were also registered between the pretest and problems 4, 5 and 6 with the highest workload level (see figure 4) revealing that mental workload significantly increases from the beginning to the end of the learning interaction. A significant increase was also found during problem 5 when compared with problems 1, 3 and 4 suggesting that learning tasks varied in terms of the cognitive workload demand required from the participants.

To summarize, EEG_Workload was validated according to the learners' subjective ratings. The index value increased clearly as learners acquired skills in trigonometry. This trend was significant when we compared the pretest with the learning session and the problem solving tasks.

Our next concern was to evaluate the EEG_Workload model with regards to the objective metrics in the learning environment. A bi-variate correlation was computed between EEG_Workload and response time across participants in the six problems (N

= 17 x 6) showing a rather small but significant relationship ($r = 0.21$, $p < 0.05$). Besides, correlational analysis of each problem apart revealed a significant correlation only in problem 5 ($r = 0.56$, $p < 0.05$). Indeed, unlike pure cognitive tasks with strict laboratory conditions and imposed time limits, in a more complex learning environment, learners are less restricted and a longer response time does not necessarily imply higher mental workload which can also be distilled into other mental processes.

Bi-variate correlations were computed between EEG_Workload and performance in the pretest and each problem. Results showed a significant linear relationship between EEG_Workload and performance in the pretest ($r = -1.88$, $p < 0.05$) while no significant correlation was found between EEG_workload and performance in any of the problems. Again, these data suggest that changes in workload are not related to performance in complex learning tasks and that a more complex relationship may exist between mental workload and performance.

Looking at participants' self reported skill levels in trigonometry, ANOVA tests revealed a reliable effect of the skill level on the mean EEG_Workload in the problems ($F(2, 14) = 11.93$, $p < 0.05$). We found that participants with a moderate skill level have had the highest workload values ($M = 68.97$; $SD = 14.16$) as compared to participants with low skill level ($M = 57.25$; $SD = 4.06$) and to expert participants who showed the lowest workload values ($M = 41.35$, $SD = 7.52$). One can explain this result by the fact that learners with moderate skills tend to produce more effort than experts who might be more at ease with trigonometry and do not produce a lot of effort. On the other side, learners with the lowest skill level tend to be rather disengaged in the task compared to learners with moderate to high skills.

Our last investigation dealt with the overall impact of learners' workload, response time and skill level on their performance in the problems. A multiple regression analysis was conducted to measure the influence of each of these parameters. The overall model was significant ($F(3, 98) = 8.48$, $p < 0.01$, $R = 0.41$). Conditional main effect analyzes revealed a positive effect of workload ($\beta = 0.24$, $p < .05$) and skill level ($\beta = .42$, $p < .001$) and a negative effect of response time ($\beta = -0.27$, $p < .05$) suggesting that a combination of these variables can be used to predict learning performance. This result suggests that a combination of several variables can improve the accuracy of systems in assessing learners' skill acquisition process.

5 Conclusion

In this paper we have presented a workload index based on features extracted from EEG signals. 17 participants were recruited for this experiment and were equipped with a 6-channel EEG headset. The developed workload index uses a Gaussian Process Regression model trained on data gathered from purely cognitive tasks with incrementally increasing levels of difficulty to elicit increasing levels of required mental workload. Our model was validated in a learning activity during which learners interacted with an educational environment about trigonometry including a pretest, a learning session and six problem solving tasks and self-reported their workload level. Results showed that our index was significantly correlated to

learners' subjective ratings and gradually increased from the pretest to the end of the session. Correlational analysis showed that mental workload was not necessarily, linearly associated to performance and response time objective variables in the learning context, as opposed to the strict laboratory conditions of the cognitive task activity.

Future work involves developing a generalized model and incorporating it within a real time interaction based tutoring system. Further variables from the learners' profile will be considered in the development of the system that will be centered in optimally adapting content, problem level and interactions to learners' mental states.

Acknowledgments. We acknowledge the CRSNG (Conseil de Recherches en Sciences Naturelles et en Génie du Canada) for funding this research.

References

1. Berka, C., Levendowski, D.J., Cvetinovic, M.M. et al.: Real-Time Analysis of EEG Indexes of Alertness, Cognition, and Memory Acquired With a Wireless EEG Headset. *International Journal of Human-Computer Interaction* 17, 151 - 170 (2004)
2. Gevins, A., Smith, M.E.: Assessing fitness-for-duty and predicting performance with cognitive neurophysiologic measures. In: Caldwell, J.A., Wesensten, N.J. (eds.) *Biomonitoring for Physiological and Cognitive Performance during Military Operations*. Proceedings of SPIE, vol. 5797-18. pp. 127-138 (2005)
3. Murata, A.: An Attempt to Evaluate Mental Workload Using Wavelet Transform of EEG. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 47, 498-508 (2005)
4. Smith, M.E., Gevins, A.: Neurophysiologic monitoring of mental workload and fatigue during operation of a flight simulator. In: Caldwell, J.A., Wesensten, N.J. (eds.) *Proceedings of SPIE Defense and Security Symposium, Biomonitoring for Physiological and Cognitive Performance during Military Operations*, pp. 116–126. SPIE, Orlando, FL (2005)
5. Wilson, G.: Operator functional state assessment for adaptive automation implementation. In: Caldwell, J.A., Wesensten, N.J. (eds.) *Proceedings of SPIE Defense and Security Symposium, Biomonitoring for Physiological and Cognitive Performance during Military Operations*, pp. 100–104. SPIE, Orlando, FL (2005)
6. Parasuraman, R.: *Neuroergonomics: the brain at work*. Oxford University Press, New York (2005)
7. Pope, A.T., Bogart, E.H., Bartolome, D.S.: Biocybernetic system evaluates indices of operator engagement in automated task. *Biological Psychology* 40, 187-195 (1995)
8. Prinzel, L.J., Freeman, F.G., Scerbo, M.W.: A Closed-Loop System for Examining Psychophysiological Measures for Adaptive Task Allocation. *International Journal of Aviation Psychology* 10, 393-410 (2000)
9. Serman, M.B., Kaiser, D.A., Mann, C.A. et al.: Application of Quantitative EEG Analysis to Workload Assessment in an Advanced Aircraft Simulator. In: *Human Factors and Ergonomics*, Seattle, Washington, USA, 118-121 (1993)
10. Hancock, P.A., Warm, J.S.: A Dynamic Model of Stress and Sustained Attention. *Human Factors and Ergonomics Society* 31, 519-537 (1989)
11. Arroyo, I., Woolf, B.: Inferring learning and attitudes from a Bayesian Network of log file data. In: *Proceeding of the 2005 conference on Artificial Intelligence in Education*:

- Supporting Learning through Intelligent and Socially Informed Technology, IOS Press, 33-40 (2005)
12. Baker, R.S., Corbett, A.T., Koedinger, K.R.: Detecting Student Misuse of Intelligent Tutoring Systems. In: Intelligent Tutoring Systems, Springer Berlin / Heidelberg, 54-76 (2004)
 13. Beck, J.E.: Engagement tracing: using response times to model student disengagement. In: Proceeding of the 2005 conference on Artificial Intelligence in Education: Supporting Learning through Intelligent and Socially Informed Technology, IOS Press, 88-95 (2005)
 14. Johns, J., Woolf, B.: A dynamic mixture model to detect student motivation and proficiency. In: Proceedings of the 21st national conference on Artificial intelligence - Volume 1, AAAI Press, Boston, Massachusetts, 163-168 (2006)
 15. Arroyo, I., Cooper, D.G., Bursleson, W. et al.: Emotion Sensors Go To School. In: Proceeding of the 2009 conference on Artificial Intelligence in Education: Building Learning Systems that Care: From Knowledge Representation to Affective Modelling, IOS Press, 17-24 (2009)
 16. D'Mello, S., Craig, S., Witherspoon, A. et al.: Automatic detection of learner's affect from conversational cues. *User Modeling and User-Adapted Interaction* 18, 45-80 (2008)
 17. Forbes-Riley, K., Rotaru, M., Litman, D.J.: The relative impact of student affect on performance models in a spoken dialogue tutoring system. *User Modeling and User-Adapted Interaction* 18, 11-43 (2008)
 18. Stevens, R., Galloway, T., Berka, C.: Integrating EEG Models of Cognitive Load with Machine Learning Models of Scientific Problem Solving. In: Schmorow, D., Stanney, K., Reeves, L. (eds.) *Augmented Cognition: Past, Present and Future*, Strategic Analysis, Inc., Arlington, 55-65 (2006)
 19. Wilson, G.F.: An analysis of mental workload in pilots during flight using multiple sychophysiological measures. *Int J Aviat Psychol* 12, 3-18 (2004)
 20. Gevins, A., Smith, M.E.: Neurophysiological measures of cognitive workload during human-computer interaction. *Theoretical Issues in Ergonomics Science* 4, 113-131 (2003)
 21. Kohlmorgen, J., Dornhege, G., Braun, M. et al.: Improving human performance in a real operating environment through real-time mental workload detection. In: *Toward Brain-Computer Interfacing*, pp. 409-422. MIT Press, Cambridge, MA (2007)
 22. Heger, D., Putze, F., Schultz, T.: Online Workload Recognition from EEG Data during Cognitive Tests and Human-Machine Interaction. In: Dillmann, R., Beyerer, J., Hanebeck, U., Schultz, T. (eds.) *KI 2010: Advances in Artificial Intelligence*, vol. 6359. pp. 410-417. Springer Berlin / Heidelberg (2010)
 23. Honal, M., Schultz., T.: Determine task demand from brain activity. In: *3rd International Conference on Bio-inspired Systems and Signal Processing* (2008)
 24. Berka, C., Levendowski, D.J., Lumicao, M.N. et al.: EEG Correlates of Task Engagement and Mental Workload in Vigilance, Learning, and Memory Tasks. *Aviation, Space, and Environmental Medicine* 78, B231-B244 (2007)
 25. Stevens, R., Galloway, T., Berka, C.: EEG-Related Changes in Cognitive Workload, Engagement and Distraction as Students Acquire Problem Solving Skills. In: Conati, C., McCoy, K., Paliouras, G. (eds.) *User Modeling 2007*, vol. 4511. pp. 187-196. Springer Berlin / Heidelberg (2007)
 26. Stevens, R.H., Galloway, T., Berka, C.: Integrating innovative neuro-educational technologies (I-Net) into K-12 science classrooms. In: *Proceedings of the 3rd international conference on Foundations of augmented cognition*, Beijing, China, 47-56 (2007)
 27. Hart, S.G., Staveland, L.E.: Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In: Hancock, P.A., Meshkati, N. (eds.) *Human mental workload*, pp. 139-183. North-Holland, Amsterdam (1988)
 28. Rasmussen, C.E., Williams, C.K.I.: *Gaussian Processes for Machine Learning*. MIT Press;, (2006)