Use of cognitive load measurements to design a new architecture of intelligent learning systems
Amin Zammouri, Abdelaziz Ait Moussa, Sylvain Chevallier

To cite this version:
Amin Zammouri, Abdelaziz Ait Moussa, Sylvain Chevallier. Use of cognitive load measurements to design a new architecture of intelligent learning systems. Expert Systems with Applications, 2023, 237, pp.121253. 10.1016/j.eswa.2023.121253. hal-04350891

HAL Id: hal-04350891
https://universite-paris-saclay.hal.science/hal-04350891
Submitted on 21 Dec 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
A brain-computer interface-based intelligent learning system for learner’s mental efforts recognition

Amin Zammouri\textsuperscript{a,}\textsuperscript{*}, Abdelaziz Ait Moussa\textsuperscript{b} and Sylvain Chevallier\textsuperscript{c}

\textsuperscript{a} EPF Graduate School of Engineering, 94230 Cachan, France  
Email: amin.zammouri@epf.fr

\textsuperscript{b} Faculty of Sciences, Mohammed First University, 60000 Oujda, Morocco  
Email: a_aitmoussa@yahoo.fr

\textsuperscript{c} Université Paris-Saclay, CNRS, Laboratoire interdisciplinaire des sciences du numérique, 91405, Orsay, France  
Email: sylvain.chevallier@uvsq.fr

\textsuperscript{*} Corresponding author.
ABSTRACT

In the context of learning environments, the learner’s attention and mental effort are of primary interest in the process of acquiring knowledge. Due to the skills and abilities of each learner, there is a growing need for generic and adaptive environments. In this work, we introduce a new architecture for such environments to assist the learner with a multi-agent-based approach. Using Electroencephalogram (EEG) data, this architecture generates learning content and interactions adapted for each learner. Moreover, this work presents a new unsupervised approach to estimate and recognize the learner’s cognitive load based on the Standardized Euclidean Distance (SED) and the Power Spectral Density (PSD) of brain rhythms within low frequencies, namely Theta [4–7 Hz] (θ) and Alpha [8–11 Hz] (α). The learner’s outcomes and estimated mental efforts are combined in the evaluation process using a fuzzy logic-based approach. Three experimental protocols are adopted in order to validate our study. These protocols are based on cognitive tasks with different difficulty levels. Experimental results show that PSD in θ and α bands in the occipital lobe accurately describe changes in the learner’s mental efforts and cognitive load according to the cognitive task difficulty level. Based on the Cohen Kappa coefficient, our cognitive load estimation approach, using α, is compared to an existing workload index from the literature. This performance assessment process revealed large values (k ≥ 0.48) in the occipital lobe, which reflects the efficiency of the proposed approach. Results from this study are mainly used in educational engineering and reeducation in order to subjectively assess the approaches and treatments offered in these contexts.

Keywords: Brain-Computer Interface; Cognitive Load; Intelligent Tutoring System; Fuzzy Logic.

1. Introduction

Artificial Intelligence in Education (AIED) and Assistive technologies (AT) aim at developing user specific solutions adapted to the learner’s competencies. A crucial aspect is to take into account specificities of each learner to propose an intelligent learning environment taking advantage of the learner’s interactive behavior. Two main approaches could be distinguished in the context of AIED that are computer-supported learning (Kirschner and Gerjets, 2006) and student-centered learning (Caldier, 2015). In the computer-supported learning, the adaptation of learning content is straightforward, since it provides a suitable background for the implementation of adaptation algorithms (Spüler et al., 2016). Despite the multitude of learning environments, e.g., iWeaver (Wolf, 2003), INSPIRE (Papanikolaou et al., 2002) or Calcularis (Käser et al., 2013), attempts to implement adaptation of the learning process revealed unsatisfactory results. In their interactions with learners, these systems are essentially based on the so-called Pedagogical Agents (PA) which support with great autonomy the learner’s interactions. Regarding the multi-interactions which can take place between learners and PA, these environments can support both the individualized and the collaborative learnings. The common architecture used in these environments is based on four modules (Moreno et al., 2001; Kim and Baylor, 2006; Hooshyar et al., 2015), namely Domain module, Learner module, Pedagogical module and Interface module. In the general case, the domain module represents an expert knowledge on a particular field. It contains not only expertise in acquiring skills, but also offers an internal representation of the competence to build. The domain module must be able to generate solutions to problems in the same context where the learner is placed. This allows the system to determine the differences and correspondences in the learner and tutor actions. The learner module provides the learner knowledge measurement about the problem. It is a tuple of expertise, knowledge, cognitive profile and learner history. The pedagogical module allows to define the mediation to assist the learner in the learning process. It must consider each of educational, pedagogical and psychological principals. The major aim of this module is answering the three questions (1) why interfere? (2) when interfere? and (3) how interfere? The interaction model is the responsible for the system internal representations and the learner interface connection. This module is in a permanent cooperation with the educational system and the learner’s assessing skills. On another hand, it determines the final form used by the system to transfer information.

Compounding the weakness of content adaptation, current learning environments are not generic and specialize in teaching a specific field. For instance, the “Calcularis” environment (Käser et al., 2013) specializes in teaching mathematics to children. In its interactions with children, the domain module of this learning environment adapts the learning content based on the knowledge and skills acquired in the previously addressed levels. In addition, the learner module evaluates the learner’s skills and competencies level by comparing the accuracy of her interactions. However, based only on the accuracy of the learner’s interactions, this solution remains weak in terms of content adaptation. This is because it does not take into account neither the learner’s calculation capacities nor her current internal state and cognitive load. To address these issues, the most sophisticated and suitable approach is to incorporate physiological and brain factors to estimate learners’ mental efforts and cognitive load. Among these factors, we distinguish Heart Rate Variation (HRV), Galvanic Skin Response (GSR), and Electroencephalogram activity (EEG) (Mühl et al., 2014; Ayres et al., 2021).

In (Blitz et al., 1970; Jorna, 1992; Mehler et al., 2012; Ayres et al., 2021) it has been demonstrated that heart rate is a sensitive physiological measure for detecting systematic variations in the cognitive load. In a study on the usefulness of HRV, (Aasman et al., 1987) clarified that the measurement of this physiological marker is determined by three different feedback mechanisms related to: respiration, blood pressure and temperature regulation (Ayres et al., 2021). Based on spectral analysis, in the objective to study the periodic behavior of HRV, (Aasman et al., 1987) were able to show that cognitive load is specifically linked to the short-term regulation of blood pressure. The relationship between cognitive load and HRV is indirect (Solhjoo et al., 2019) because an increase in the cognitive load will lead to an increase in blood pressure which will lead to a decrease in HRV. The measurement of HRV is generally accepted as a measure of cognitive load. However, in the work of (Paas et al., 1994) it has been shown that the HRV is mainly used to measure cognitive load when it is a short-term duration cognitive task. In the work of (Paas et al., 1994) it has been demonstrated that for long-duration learning tasks, the validity and sensitivity of HRV spectral analysis is low.
The measurement of GSR is based on the variation of the electrical activity of the eccrine sweat glands on the plantar and palmar surfaces of the hand which are particularly sensitive to psychological stimuli imposing stress. Increased stress will lead to increased sweating, which decreases resistance and increases the electrical conductance of the skin (Dawson et al., 2000; Ayres et al., 2021). The electrodermal signal makes it possible to distinguish, on the one hand, the level of cutaneous conductance which changes over time, which is considered as a measure of psychophysiological activation. On the other hand, it allows to distinguish the response of the cutaneous conductance which changes abruptly. These rapid changes are reflected in spikes in the electrodermal signal and are also referred to as cutaneous galvanic responses (Braithwaite et al., 2013). This component is influenced by stress and excitement (Hoogerheide et al., 2019). In (Setz et al., 2009) and (Larmuseau et al., 2019) it was demonstrated that the increase in skin conductance is a function of cognitive load. However, in the work of (Vanneste et al., 2020) it has been shown that the usability of GSR measures as a cognitive load marker is limited. This is explained by the fact that GSR can only describe a limited proportion of cognitive load variation. (Charles and Nixon 2019) proposed that GSR might be sensitive to sudden but not gradual changes in cognitive load (Ayres et al., 2021).

Given the applicative context of learning environments and strong effects of learners’ cognitive state on EEG, an approach based on EEG seems very appropriate to assist the learner (Wang et al., 2013; Spüler et al., 2016; Zammouri et al., 2017; Zammouri et al., 2018; Lotte et al., 2018). In the context of assistive technology, leveraging the brain activity is often associated with Brain-Computer Interfaces (BCI). These interfaces endow systems of a direct communication relying only on the brain signals (Pfurtscheller et al., 1993; Minguillon et al., 2017; Cirillo et al., 2020; Liu et al., 2020; Zammouri and Zerouali, 2021). While the design of BCI was always focused on developing technical solutions for subjects’ motor capacities rehabilitation (Leeb et al., 2013; Kalunga et al., 2014; Wang et al., 2014; Xie et al., 2022) currently the development of such interfaces converges to the use of spontaneous brain activity in order to derive and recognize the user’s various cognitive states (Blankertz et al., 2010; Anderson et al., 2012; Wang et al., 2011; Hong et al., 2018; Gaume et al., 2019; Mora-Sanchez et al., 2020; Martinez Beltran et al., 2022). Thus, it has given rise to Passive BCIs (pBCIs). Several works from the literature using EEG, functional Near Infrared Response Spectroscopy (fNIRS) (Khalil et al., 2022) or functional Magnetic Resonance Imaging (fMRI) (Baqapuri et al., 2021) have proven their potential to assess users’ cognitive load in the everyday life activities. Good examples are information visualization interfaces (Peck et al., 2013) and Virtual Reality (VR) environments (Mühl et al., 2014). Using pBCIs to estimate the cognitive load aims at following information about the user’s mental activity. Most of the EEG-based approaches used in this context take advantage of the spectral aspect of the brain electrical signal. The electrical brain activity generates different rhythms ( Başar et al., 2001). These rhythms represent a continuum of waves: Delta [0.5-3 Hz] (δ), Theta [4-7 Hz] (θ), Alpha [8-11 Hz] (α), Beta [12-30 Hz] (β) and Gamma (>30 Hz) (γ). For Klimesch (Klimesch, 1999) and Andreason (Andreason, 2000) θ and α oscillations are sensitive to task difficulties. According to Holm (Holm et al., 2009), the increase of activity increases the θ oscillations in the frontal brain area and decreases the α oscillations in the parietal brain area. Apart from brain rhythms, measuring Event Related Potentials (ERPs) represents specific responses to cognitive events. These potentials represent electrical variations measured on the nervous system in response to stimuli. Considered as the commonly used evoked potential, P300 has justified its accuracy to assess the cognitive load through its sensitivity to brain processing competencies (Van Dinteren et al., 2014). However, using ERP signals to measure the cognitive load requires multiplying the number of trials in order to make a one decision. This would be very tedious for learners and not suitable for adapting the content in an online mode.

Through the present work we seek to achieve several objectives. First, we seek to identify the brain regions which are solicited while performing a cognitive task. Here we are interested in identifying brain regions which describe changes in the cognitive load levels, and which have not been reported in literature. The illustration of this objective led us to design several experimental protocols and to develop a cognitive load levels classifying algorithm based on θ and α bands. In a second step, and after having illustrated the first objective, we seek to integrate our algorithm in a new architecture of intelligent tutoring systems. To achieve this objective, we designed a dedicated experimental protocol based on the α band. It has been reported that the α band decreases when increasing the cognitive task difficulty level (Sterman and Mann, 1995). The α spectrum also decreases when the duration time to perform the cognitive task puts an additional stress on the subject (Slobounov et al., 2000). This decrease in the α spectrum is mainly identified in the occipital and parietal lobes (Kramer, 1991). This decrease in the α spectrum in these regions is generally attributed to modulation due to attention demand related to the cognitive task. But sometimes, the involvement in a cognitive task can lead to disappearance of the α waves in these regions (Curtis et al., 1993; Puma et al., 2018). In (Zhu et al., 2021), the authors studied variations of θ and α waves when performing scientific problems. According to (Zhu et al., 2021) the increase in the mental effort is accompanied by a decrease in spectrums of θ and α in the occipital lobe. According to (Zhu et al., 2021) these results mean that the mental effort expanded while performing a scientific problem is related to working memory, spatial-visual processing as well as semantic processing.

Since the advent of brain-computer interfaces in the 2000s, works in literature have been interested in exploring and studying cerebral behavior during a cognitive task performance. These works have mainly explored the brain frontal areas because these regions have been associated to the working memory (Baddeley, 2000; Nissim et al., 2017). In works dating from the last decade, links have been established between the cognitive load in general and the brain posterior regions. As an illustration, this link can be found in (Jap et al, 2009) where θ, α and β brain waves measured on the occipital lobe were used to describe levels of mental fatigue and sleepiness in a context of car driving. In this context of studying the cognitive load, our previous work (Zammouri et al., 2018), carried out with the aim of developing a classifier to distinguish two cognitive load levels, demonstrated that the brain waves measured on the occipital lobe describe the cognitive load levels variations. In this way, our present work aims to explore and study the occipital lobe’s EEG signals to describe the brain behavior while using an intelligent tutoring system and while performing a cognitive task with more than two difficulty levels.

This work presents a new architecture of a generic and adaptive computer-supported-based learning environment combining an EEG-based passive BCI to assess the learner’s current mental effort and cognitive load. Such an architecture presents and adapts the learning content in the adequate difficulty level. Our approach consists in designing a new architecture of each of the learner module, the pedagogical module and the interface module. Moreover, we introduce a new unsupervised method for the estimation of cognitive load from the occipital lobe. Regarding the generic character of this new architecture, for instance, we have chosen an implementation in the context of mathematical logic. Learners’ outcomes and the measures of the cognitive load are used as input variables in our fuzzy logic-based evaluation algorithm. Experimental results show, for both θ and α oscillations in the occipital brain area, a decrease in power while
increasing the cognitive task difficulty and vice versa. Based on the Kappa test (Cohen, 1967), the comparison of performances to the rates of Event Related Desynchronization (ERD) and Synchronization (ERS) demonstrates that our proposed cognitive load estimation approach could be of a reliable and optimal alternative solution, especially for developing lightweight systems on devices with limited memory capacities. Moreover, findings presented in this work demonstrate that the developed learning environment could be of considerable help and assessment in the learning process for disabled people.

2. Learning environment architecture

The objective of the learning environment presented in this work aims at evaluating and supervising the learner’s knowledge taking advantage of her cognitive capacities. For this reason, the environment must embed adequate tools to select and sequentially present the learning content to the learner. This requires endowing the environment with a good representational power regarding the domain knowledge and the learner’s one. These information are necessary for the environment to update the learner profile and provide an assessment feedback to the learner according to her profile and current cognitive load. The aspect of Generic Learning Environment (GLE) introduced in our work requires that the domain module should be able to dispose of knowledge on the learning domain. This is modelled in our architecture using a Domain Knowledge Base (DKB). Each domain to learn has its own DKB. A distinction should be made between the domain knowledge and the pedagogical knowledge (Mizoguchi, 2003; Hayashi et al., 2006). Indeed, the pedagogical knowledge represents styles and strategies to interact with the learner. This knowledge must insure the tutoring process and allows to each learner an adequate learning experience. The domain knowledge reflects the subject to be learned, its explanations and all learning activities, i.e., exercises etc. In our architecture we have chosen to associate a Pedagogical Knowledge Base (PKB) to each DKB. The designed architecture to consider these points is presented in Figure 1. The advantage of using a multi-agent architecture is that it exhibits interesting characteristics such as: modularity, ease of evolution, the possibility of repeating agents in the learning environment, which is a distributed environment, taking into account the distributed and asynchronous nature of the problem.

![Fig. 1](image-url) Overview of the developed learning environment architecture. – Brain Signal Measuring Agent: Its role is to measure the EEG signal from the used electrodes. – Brain Load Estimation Agent: It receives the measured EEG signals and process them to estimate the cognitive load level. – Profile Updating Agent: Its role is to update all the information concerning the learner. – Planification Agent: Its role is to select the appropriate learning content for the learner based on her measured cognitive load. – Evaluation Agent: Its role is to assess the learning outcomes of a learner by taking into account the levels of the cognitive load while performing the cognitive task. – Content Presentation Agent: Takes care of presenting, to the learner, the content chosen by the Planification Agent. – Resources Management Agent: It allows to tutor to upload the learning contents. – Assistance Agent: Its role is to suggest help (as indication) when the learner’s cognitive load level is low.

A learning scenario comprises a set of activities. These activities are all initiated by the learner and differ among reading a course, doing exercises, answering quiz and interacting with forums of the learning environment. Choosing these educational activities allows the system to present concepts in different ways in order to reach a best possible presentation. Each learning scenario has a specific instance of Learning Session in order to accurately modelling the learning activities and building a learner’s specific history. In the Pedagogical Module, planification rules define “learning levels” associated to each learning domain. Each learning level consists of a set of learning activities. The induced learning activities can be of three different levels of difficulty (Low: l₁, Medium: l₂ and Difficult: l₃). The adopted methodology advocates relying on modelling the learner profile through (1) the observation of the learner’s learning activity, (2) estimation of the learner’s brain activity and (3) the construction of the observables in the learner profile. All information are organized to allow the environment to clearly define the learner’s situation and guide her to the best in the learning process. During her learning process, the learner performs an effective trajectory, identified as her personal path. This trajectory probably differs from that one which was planned by the tutor. This path allows tracking the learner’s learning process.

Taking advantage of the ontology engineering approach (Hayashi et al., 2006), we developed our simple ontology to represent activities of a learning scenario. We propose the concept of Learning Activity Objectives (LAO). For each learning level, a set of objectives are assigned to the learner in order to decide her transition to the next learning level. This allows the learner module to determine the learner’s learning process by searching the existing correlation between expended cognitive efforts, learning outcomes and the learning taken time. Hence, the environment can distinguish when learning results from good acquisition of the learning content. Also, using the learner module information, the environment selects and presents appropriately the content to learn in order to assist and/or improve the learning process. These interactions with the learning environment are analyzed and designed based on a multi-agent-oriented software engineering.
method. The behavior of each agent comprising the learning environment depends on the learner’s brain activity and the resulting actions. The use of a multi-agent-based architecture offers a high flexibility, for modelling, the learner profile, her interactions as well as the creation of the content adaptation strategies.

In the architecture shown in Figure 1, a learning session begins with an interaction between the learner and the Interface Module. This interaction solicits the Content Presentation Agent and the Brain Signal Measuring Agent. These two agents are executed in the background since they operate permanently throughout the learning session. The Content Presentation Agent communicates with the Planification Agent in order to select the learning content for the learner. To do this, the Planification Agent requests the Profile Updating Agent which provides the necessary information on the learner. This content is used by the Content Presentation Agent. On the other hand, the Brain Load Estimation Agent runs permanently and in the background while communicating with the Brain Signal Measuring Agent. The Brain Load Estimation Agent controls the operations of the Assistance Agent. In fact, if the Brain Load Estimation Agent gives an estimation on mental fatigue in the learner, the Assistance Agent triggers the rules entered by the tutor in the PKB. For example, the PKB could contain response indications to be communicated to the learner in the case of difficulty. At the end of a learning activity, the Evaluation Agent is triggered to give a decision on the learner’s learning experience. The result of the Evaluation Agent is received by the Profile Updating Agent which saves it in the “Learner Base” (LB). Since in our experiments we use the OpenViBE software (Renard et al., 2010) for the EEG signals measurements, these are routed from the acquisition software to the learning environment based on a virtual-reality peripheral network (VRPN) server. The arrival and management of the EEG signals on the learning environment is ensured by the Brain Signal Measurement Agent.

The Assistance Agent, Evaluation Agent and Planification Agent represent the kernel of our learning environment. They primarily model the nature of the subject to be taught and the difficulty levels of the considered learning activities. Thus, the pedagogical module performs the measurement and the evaluation of the learner’s knowledge in relation to the considered domain knowledge. This is done based on her learning activities in the environment and on measures of her expended mental efforts. Indeed, the Evaluation Agent analyses the learner’s interactions based on checking results achieved during the learning activities. We incorporated in this agent a fuzzy logic-based evaluation algorithm. This algorithm takes advantage of qualitative measures, i.e., correct outcomes during the learning activities, and measures of the cognitive load and mental efforts. This combination provides more qualitative characteristics of the learner’s learning process.

3. Cognitive load estimation approach

3.1. Experimental setups

In this study three experimental protocols are adopted. In order to evaluate our cognitive load estimation approach, we have chosen two experimental protocols. The first experimental protocol is based on the Raven’s matrix test. Considered as an intelligence test in competitive recruitment, Raven’s matrix test consists of either completing a figure with a missing part or finding the law of progression of graphic series of matrices (2,2), (3,3) etc. In our protocol we used matrices of size (3,3). As a first step of validation of our classification algorithm, we studied the θ and α modulations while going from a rest state to a state of Raven’s matrix resolution. In the second experimental protocol we used the N-Back test as a cognitive task. This test consists of presenting the participant with a series of consecutive letters. The participant has 5 seconds to memorize them and then correctly answer the questions asked: letters are suggested to her and she must indicate whether or not they exist in the series supposed memorized. We chose to introduce this experimental protocol in order to identify brain regions which undergo significant modulations during an auditory cognitive task. The interest behind this is to study the modulations of the θ and α waves in the occipital lobe during an auditory cognitive task. This will allow us to illustrate the involvement and modulation of the θ and α rhythms while performing a cognitive task which does not require visual attention.

In order to evaluate and validate the usability of the new architecture of ITS presented in this study, we introduced a third experimental protocol in which we have chosen as a test scenario conducting a positioning test. The chosen learning activity consists of 20 multiple-choice questions covering mathematical and arithmetic logic. The 20 questions induce three different difficulty levels. In the first level, denoted by l₁, subjects were asked to complete the missing element in a numerical series. In the second level, denoted by l₂, subjects were asked to solve problems such as conversion of time and temperature units. Finally, in the third level, denoted by l₃, we presented to subjects questions which require problem solving using different arithmetic operations. These questions are presented randomly to the learner. This presentation mode is used to assess the learner’s brain reactions to the difficulty level changes. Each question is presented separately with four answer proposals. The learner selects her answer using the computer mouse. The learner is asked to restrict her movements and move only her hand in order to select the answer.

Experiments from each protocol include ten subjects (i.e., a total of 30 participants) who have voluntarily accepted to participate in this study. Subjects are all male university students (from different countries) aged from 20 to 30 years old. Informed consent is obtained from each subject. The participants do not present any mental or ocular trouble that may affect the experimental results. In all our experimental protocols, a 5-minutes relaxation session (with closed eyes) is required to each subject before starting the experiment. This relaxation session aims to reduce the effects of the learner’s activities before starting the experiment. For example, walking to the place of experimentation could influence brain rhythms associated with motor activity.

3.2. EEG data recording

For experimentations from the first and second protocols, the EEG data acquisition system uses 14 channels. Electrodes are placed at the occipital, parietal, frontal and temporal lobes following the extended 10-20 international system. The used electrodes are: AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8 and AF4 as presented in Figure 2. EEG data are measured continuously based on the epoc emotiv acquisition device using a sampling frequency of 128Hz.

In the third experimental protocol, the EEG data acquisition system uses eight channels with a reference to the left earlobe. Electrodes are placed at the occipital and parietal lobes following the extended 10-20 international system. The used electrodes are Pz, PO3, POz, PO4, O1, Oz, O2 and Iz as presented in Figure 2. EEG data are measured continuously based on the g.Mobilab+ acquisition device using
a sampling frequency at 256Hz. Since the learning environment is destined for a real-time use, no sophisticated method for noise elimination is applied.

In all protocols, EEG data are acquired using the OpenViBE platform (Renard et al., 2010) and are filtered using a pass band filter for 1-30 Hz.

![Fig. 2 Placements of the used electrodes according to the extended 10-20 international system. In orange color are presented electrodes used in the first and second protocols. The blue color represents electrodes used in the third experimental protocol. The green color represents common electrodes to all the experimental protocols.](image)

3.3. Cognitive load estimation approach

In order to control and supervise the learner’s different levels of cognitive load during her interaction with the learning environment, we introduce an unsupervised approach to estimate such levels of cognitive load from the EEG signals. On the one hand, our method consists in calculating the EEG power spectrum using a Short Time Fourier Transform (STFT). The cognitive load is estimated on a 5-seconds moving window and using an averaged periodogram based on the Hanning-window. According to literature works, developed approaches in the context of cognitive load estimation in general are based on analyzing θ and α waves. In this way, we have chosen to focus our study on these two brain rhythms, thus, through this approach two Cognitive Models (CM) are designed. These CM describe changes in the cognitive load levels based on the power spectrum of each of θ and α rhythms. These rhythms are the most likely to reflect changes in the cognitive load. For each of the two models we propose to compute a Cognitive Reference Model (CRM). These CRM are denoted by $CRM^i (i \in \{\theta, \alpha\})$. A $CRM^i$ is an epoch which represents the baseline power spectrum of the considered rhythm. It is computed using a 5-seconds time window. This time window is chosen before starting the experimentation and after having finished the relaxation session and so that the power spectrum of the corresponding rhythm is a normal distribution. We assume that before starting the experimentation, the learner is in a state of quasi-rest. The $CRM^i$ allow evaluating and estimating the level of the learner’s cognitive load by computing the distance from these models. We assume that a large distance represents a great expenditure of mental effort and cognitive load.

In our investigation we used the Standardized Euclidean Distance (SED). The Euclidean distance calculates the distance between two real-valued vectors. If vectors have values with different scales, it is common to normalize the numeric values across all columns before calculating the Euclidean distance. Otherwise, columns with large values will dominate the distance measurement. The obtained scores are used to find the most similar examples. In our study we apply the SED to compare if the vector $P^i$ is from the same distribution of each 5-seconds time window of the power spectrum of the considered rhythm, i.e., θ or α. Let us consider the case of a single channel; each one of the cognitive reference models is represented by $CRM^i(\mu^i,(\sigma^i)^2)$, where $\mu^i$ and $(\sigma^i)^2$ represent respectively the mean and the variance of the $CRM^i$. Let $P^i$ be the power spectrum vector during a given instant of the moving-window. The standardized Euclidean distance from each cognitive reference model is obtained as follows:

$$SED^i(P^i, CRM^i) = \sqrt{\frac{\sum_{n=1}^{N} w_n (P^i_n - CRM^i_n)^2}{\sum_{n=1}^{N} w_n \sigma_n^2}}$$

(1)

where $w_n$ is the inverse of the variance, $P^i_n$ is the power spectrum vector and $N$ is the number of observations during the considered epoch. This distance is then used by the Evaluation Agent in order to determine the current learner’s cognitive load. For each 5-seconds epoch, the classifier makes its decision by computing the $SED^i$ distances. The final decision on the expended cognitive load when following a learning activity is made by computing the average of $SED^i$ distances over the learning activity epochs. This average is denoted by $ASED^i$. The classifier performance is assessed based on the $ASED^i$ averages.

In our experimental protocols, EEG signals are measured based on multiple electrodes, but it is possible to reduce the number of electrodes while keeping a good accuracy. A subset of electrodes could be selected to obtain a very lightweight system as detailed in the next sections.

3.4. Fuzzy-logic-based knowledge evaluation algorithm

Fuzzy logic or fuzzy sets theory offers tools to deal with uncertainties in tasks associated with data mining contexts, identification and control of systems, signal and image processing and patterns classification. One of the important directions, in theory and practice, of
machine learning methods is kernel-based methods, especially Support Vector Machine (SVM) (Suthaharan, 2016). According to Chen (Chen, 1998), under a general hypothesis made on membership functions, a fuzzy logic classifier is equivalent to a kernel-based classifier in terms of decision boundaries. Our fuzzy logic-based evaluation algorithm consists of three steps, namely (1) the definition of the input and output variables as well as the definition of their matching membership functions (2) the design of a fuzzy rules base (inference process) (3) the projection of the outputs of the inference process in the real domain. In the first step, we define variables which compose inputs of the evaluation algorithm. We are interested in the learner’s outcomes and cognitive load estimation. Our evaluation algorithm tracks the accuracy of the learner’s response (\(Corr_{out}\)) and the cognitive load estimation (\(ASED^i, i \in \{\theta, \alpha\}\)). Setting these variables allows designing the membership functions in order to decide on the learner’s knowledge. These functions place decisions on learner’s knowledge in the interval of values ranging from 0 to 1. A value of 0 indicates that the decision is not included, while a value of 1 corresponds to a fully included one. The membership function acts on the set of values taken by an input variable. This set is called the universe of discourse. Let \(I\) be this set; in our case \(I\) consists of values of the variables \(Corr_{out}\) and \(ASED^i\). We give the formal definition of the membership function as follows:

\[
\mathcal{M}: \ I \rightarrow [0,1].
\]  

(2)

In the second step of the evaluation algorithm, we define the fuzzy rules base. This base consists of rules that are used in parallel. Regarding the generic aspect induced by our proposed ITS architecture, each learning activity has its own inference process. The latter is defined by the expert of the considered domain. Thus, the fuzzy rules operate based on the knowledge base from the expertise of the tutor. One should note that the fuzzy rules are stored in the PKB. The last step, for an operational evaluation process, is the defuzzification. Indeed, in the second step, a set of commands is generated upon the execution of the fuzzy rules. The aim of the defuzzification is to transform the resulting parameters in digital format. In Figure 3 we give an example of this evaluation process.

![Figure 3](image)

**Fig. 3** Example of fuzzification and defuzzification of our knowledge evaluation algorithm while using the \(\alpha\) rhythm. -\(a\)- fuzzy rules applied on the \(Corr_{out}\) input, -\(b\)- fuzzy rules applied on the \(ASED^i\), -\(c\)- fuzzy rules related to the decision on the learner’s knowledge.

4. **Experimental results**

To begin with, our study is conducted with the aim to identify the brain regions which describe the variations of the cognitive load, and this through the development of a classifier of cognitive load levels. In this first step, two experimental protocols are adopted. We designed these protocols to demonstrate the sensitivity of our classifier when it comes to evaluating the level of cognitive load when moving from a state of rest to a state of performing a cognitive task. In order to demonstrate the implication of the \(\alpha\) wave in any cognitive task, including those which do not require visual attention, we chose Raven’s matrices and the N-Back test as cognitive tasks of test. In Figure 4 and Figure 5 we present the obtained results. In these figures, the origin of the y-axis actually represents the \(CRM^i\) vector. The value of a bar, meanwhile, represents, on average, the distance separating \(P_{i}^\theta\) from \(CRM^i\). These results represent, for each used electrode, the averaged \(ASED^i\) for all participants in each experimental protocol of this first step. In general, and in the case of both \(\theta\) and \(\alpha\) waves, these results show large \(ASED^i\) values on the frontal lobe electrodes. In the case of the Raven matrix test we observe large values on the occipital lobe's electrodes. Regarding the N-Back test, in addition to the large values from the frontal electrodes, results show large values on the temporal and occipital lobes.

In a second step, and after validating our cognitive load classifier, we integrated it into our new ITS architecture. To do this, we designed a third experimental protocol. In this protocol the cognitive load classifier is applied to all EEG recordings measured on ten subjects. The durations of experimentations range from 5 to 20 minutes. This difference is due to the various strategies of each subject to analyze and answer questions. We chose not to fix the time to answer questions in order not to stress the subject and therefore influence the outcome of this study. The 20 questions used in this experimental protocol induce three difficulty levels. However, the subject’s analysis of the question to build her answer may involve different cognitive load sublevels associated to each difficulty level. In order to select a subset of electrodes, we were interested to the correlation coefficients between time on task, \(ASED^i\) and scores.
The averaged correlation coefficients obtained on all the ten subjects show a strong positive correlation on electrodes Oz, PO3, POz, O2, Iz, and O1 as presented in Figure 6. We choose to select three electrodes with the greatest correlation coefficients, i.e., electrodes O1, O2 and Oz.

One should note that in the $l_1$ level subjects were asked to complete the missing element in a numerical series. In the $l_2$ level we asked subjects to solve problems such as conversion of time units. Finally, in the $l_3$ level, questions that require problem solving using different arithmetic operations are induced. Figure 7 displays the averaged value, across the ten subjects, of the $ASED^\theta$ in each difficulty level. These results show that for the $\theta$ band, maximal distances are obtained while answering level $l_3$ questions. The question of level $l_2$ displays intermediate $ASED^\theta$ values, while $l_1$ level show lower values. These findings are similar for all the three electrodes except O2. This exception is due to the fact that the subject 2 has produced more mental effort to answer the question of $l_2$ level. Thus, these results...
show that the $ASED^\theta$ moves away from the $CRM^\theta$ according to the difficulty level. On another hand, for the $\alpha$ band, results displayed in Figure 7 illustrate maximum distance in the case of questions from the $l_3$ difficulty level. These distances are immediately followed by the distance value corresponding to $l_2$ level. Finally, we find the matching value to the $l_1$ difficulty level. Thus, and similar to the $\theta$ band, our results illustrate that the $ASED^\alpha$ moves away from the $CRM^\alpha$ according to the cognitive task difficulty level.

In both cases, i.e., $\theta$ and $\alpha$, increasing the difficulty level of the cognitive task involves a large $SED^i$ (consequently $ASED^i$) from the corresponding $CRM^i$ and vice versa. To evaluate the performance of our cognitive load estimation method, we compare the results of our method in the $\alpha$ band with the reference neuro-index of the literature that are the ERD and ERS rates. According to (Fink et al., 2005), (Grabner et al., 2004) these rates well measure the oscillatory dynamics. Indeed, it reflects the perceptual decrease (ERD) or increase (ERS) in band power of a brain wave during a test epoch compared to a baseline epoch (Fink et al., 2005). Let $e_b$ be the baseline time epoch. We denote by $P_{e_b}^\alpha$ the band powers during the baseline time epoch. $P_{e_b}^\alpha$ is nothing other than the $CRM^\alpha$. Thus, we define the index $z = ERD/ERS$ as follows:

$$z = \frac{P_{e_b}^\alpha - P_b^\alpha}{P_b^\alpha} \times 100$$  \hspace{1cm} (3)

According to Antonenko (Antonenko et al., 2010), a positive value of $z$ indicates a decrease in the band power i.e., ERD, whereas a negative value indicates an increase in the band power i.e., ERS. For Grabner (Grabner et al., 2004), (Brouwer et al., 2012) performing a difficult cognitive task increases the ERD in the $\alpha$ which implies a decrease in the power spectrum. The performance evaluation using the $z$ index is based on the computation of the Cohen Kappa coefficient (Cohen, 1967). This test represents a statistical metric of agreement between the classifiers as defined in equation (4):

$$k = \frac{P_b(\alpha) - P_b(\varepsilon)}{1 - P_b(\varepsilon)}$$  \hspace{1cm} (4)
where $P_z(\varepsilon)$ represents the agreement which exists between decisions made by our approach and those reflected by the $z$ index. $P_z(\varepsilon)$ is the probability of a random agreement. We use this coefficient to interpret the agreement rate between decisions made by our approach and those of the $z$ index when changing the task difficulty level in $l_1$ and $l_2$. The obtained results of the agreement, in the case of $\alpha$, show a strong agreement, $k = 0.63$, between the $z$ index and our SED-based classifier while using electrode O2. For O1 and Oz electrodes, the Kappa test gives, respectively, values $k = 0.51$ and $k = 0.48$ which reflects a moderated agreement.

We present on Table 1 results from the evaluation process based on the outcomes of each participant from the third protocol and using workload measures from electrode O2. For our experimental test, we have two inputs, namely Correct and the mean of the $ASED^\ast$ values from O2 all during the experimental test. The evaluation represents an estimation of the learner’s knowledge. These evaluations are obtained by using trapezoid membership functions. The results of Table 1 are compared to the traditional evaluation which takes into consideration only the rate of correct outcomes.

5. Discussion

In this work, we present a new learning environment architecture. This multi-agent-based environment exploits brain signal measurements for the learner’s knowledge evaluation and content adaptation. Our estimation approach of the learner’s cognitive load analyzes the evolution of these signals from an initial state in which the learner is not considered tired. The evolution or the moving distance is estimated by computing the standardized Euclidean distance on the power spectral densities of $\theta$ and $\alpha$ measured from different brain regions. For this classifier, a large distance from the baseline reflects an important expanded cognitive load. This is demonstrated with all experimental protocols where $ASED^\ast$ increased according to the increase in the cognitive task difficulty level and vice versa. This result is well demonstrated on Figure 8 in which, variations of $ASED^\ast$ regarding each question are given. This assumption is justified and proved by comparing our decisions to the ERD and ERS rates. Indeed, using the Cohen Kappa test we found (in averaged value) a strong agreement between decisions from our classifier and those using the $z$ index. Through the EEG data measured on 30 participants, we were able to demonstrate the involvement of $\theta$ and $\alpha$ waves in the description of the cognitive load levels variations during the performance of a cognitive task of long duration. In addition to findings already known in the literature, and which concern the involvement of these two brain rhythms from the frontal lobe, our results have highlighted the modulations of the occipital $\theta$ and $\alpha$ waves when it comes to perform a long-lasting cognitive task. Furthermore, our results show significant modulations of $\theta$ and $\alpha$ in the occipital lobe even when it comes to performing a cognitive task which does not require any visual attention, and which could influence posterior brain regions.

Findings from our work corroborate those previously found in the literature. Brouwer (Brouwer et al., 2012) demonstrated, through an experimental protocol based on the N-Back test, that mental efforts increase with respect to the increase in N-Back difficulty level. Using an SVM-based approach, Brouwer demonstrated 80% of classification while using the $\alpha$ band, and 72% while using the $\theta$ band. In a work of Röy (Röy et al., 2016) authors studied features of mental fatigue. Also based on the N-Back test, this study demonstrated that the $\alpha$ band in the parieto-occipital lobe well describes changes in individuals’ cognitive load and mental fatigue. In another work (Charbonnier et al., 2016), the same authors demonstrated that using ERPs describes the workload changes better than spectral features, i.e., brain rhythms. One can easily establish similarities which exist between these works and findings presented in our work. However, the robustness aspect from our work lies in the optimal required processes contrary to these approaches. Indeed, in the former work of Röy (Röy et al., 2016) the designed mental fatigue marker makes use of the Frobenius distance of the EEG spatial covariance matrices of different brain regions calculated on 20 seconds epochs to a mean covariance matrix learned during an initial reference state. Similar to our approach Pal (Pal et al., 2008) have shown, based on the Mahalanobis distance and using only the electrode Oz, that changes in the $\theta$ and $\alpha$ powers perfectly describe drivers’ drowsiness and depart from alertness. Despite the robustness of these approaches, their uses in the context of learning environments are limited by some constraints. Indeed, in such environments, some interactions like answering multi choice questions can take a duration much less than 20 seconds. The 5 seconds moving time-window used in our investigation seems suitable for the necessary process of estimating spectra. Moreover, it takes into account the cognitive load estimation in short interactions with the environment. However, the use of the Mahalanobis distance on this time window is not possible since it requires an input matrix with certain rows and columns. Differently to these approaches as well as to the investigation of Spüler (Spüler et al., 2016), using standardized Euclidean distance allowed distinguishing more than two cognitive workload levels in learners. This could be explained by the fact that experimental protocols used in these works generate only two levels of brain workload which well matches to the study of drowsiness.

![Fig. 8](image_url)

Fig. 8 $ASED^\ast$ distance variations according to questions of the experimental test on each electrode (averaged on all subjects). Questions 4, 10 and 14 correspond to the $l_3$ difficulty level. Question 19 corresponds to the $l_2$ difficulty level. Other questions correspond to the $l_1$ difficulty level.
Results, in terms of evaluation of learners’ knowledge levels, show that estimates made by the evaluation algorithm correspond at 90% to the traditional evaluation which makes use of correct outcomes rates. This agreement allows postulating that the Evaluation Agent incorporated in our environment returns a reliable estimation of the learner’s knowledge level. Unlike traditional evaluation, using an approach that analyzes the learner’s cognitive load allows to decide when answering a learning activity results from a sustained mental effort and not to chance. This is illustrated in results of Table 1. Indeed, using only the correct outcomes of the subject S7, the traditional evaluation process returns “Excellent” as level of knowledge, whereas using our evaluation algorithm, the decision gives the level “Good” with the percentage 0.4. On another hand, for subjects S4 and S5, the two evaluation processes returned “Medium” as decision on the knowledge level. Yet, our evaluation algorithm gives this decision with a low percentage of membership (0.26). According to the inference rules of the pedagogical agent, this information could be used to decide and design the adequate remedial learning activities. Apart from this, it illustrates that using a fuzzy logic-based approach, which combines information on the learner’s cognitive load, could address some learning problems such as uncertainty of learners in answering questions.

Conclusion

This work presents a new architecture of a learning environment based on brain signals in order to evaluate learners’ knowledge and adapt the learning contents according to their current mental states and cognitive load. The presented architecture makes use of a multi-agent concept to design the different modules composing the learning environment. Results obtained from the used experimental protocol illustrate that increasing the difficulty level in the environment learning activities implies modulations of $\theta$ and $\alpha$ bands in the occipital lobe. The novelty aspect and the robustness of this architecture lie in the incorporation of various agents which communicate and share functionalities. The cognitive load classifier incorporated in this environment is based on the power spectral density analysis, using an averaged periodogram, and the standardized Euclidean distance. EEG powers are computed in $\theta$ and $\alpha$ PSD in the occipital brain area. These powers and the learner’s outcomes are then used as input variables of our fuzzy logic-based evaluation algorithm. The Brain Load Estimation Agent and Evaluation Agent represent the kernel of our new learning environment architecture. In the Brain Load Estimation Agent, we incorporated a new cognitive load classifier which makes use of the $\theta$ and $\alpha$ PSD based on an averaged periodogram and the standardized Euclidean Distance. EEG PSD are computed in the occipital brain area. Based on the $\alpha$ rhythm, our classifier is compared to the ERD/ERS ratio based on the Kappa coefficient. This comparison revealed a strong agreement ($k = 0.63$). This $\alpha$ PSD-based classifier is used in the Evaluation Agent. Indeed, the ASED$^\alpha$ and the learner’s Corr$_{out}$ are used as input variables in our fuzzy-logic based evaluation algorithm.

In this study we were interested in developing an approach which makes it possible to estimate the cognitive load while answering a cognitive task in general. It is limited to healthy individuals who have participated in the experiments under conditions which no longer disturb their cognitive activities. As a perspective of our work, we aim to study the effect of certain phenomena, such as stress, on the performance of our proposed approach. On the other hand, we seek to apply our approach to patients suffering from certain diseases such as cerebral palsy. The development and adaptation of our algorithm could make it possible to study and analyze the learning process in children suffering from certain difficulties such as dyslexia, dyscalculia, etc. This could be achieved by studying the variations in their cognitive load levels while performing cognitive tasks on our learning environment. Such an adaptation would make our intelligent tutoring system a tool for subjective evaluating rehabilitation methods.

References


