Abstract
This paper discusses novel research conducted to study the direct impact of learner’s affective changes on the value of a well-established EEG-mental engagement index. An acquisition protocol for recording the electrical activity of the human brain, known as electroencephalography or EEG, was implemented in a learning environment specifically constructed for emotional elicitation. Data was collected from 35 healthy subjects using 8 sensors and two video cameras. A correlation analysis of the engagement index with emotional states was conducted. Results have shown that emotional states are strongly correlated with learners’ mental engagement index and that the later can be used in an educational setting to reliably predict performance.

Keywords: engagement index, affect, intelligent systems.

Introduction
Research on affect modeling and recognition is a well-developed and explored field considering the fact that the precise cognitive mechanisms that underlie and explain emotions are still the subject of much debate. Few will disagree however that emotions are very omnipresent in human life. They influence our behavior and play an important role in our every-day decision making processes (Quartz 2009). Learning activity is also fundamentally related to emotions (Snow, Corno, and Jackson 1996). Cognitive processes, such as problem solving and decision making, not only depend on the individual’s emotional state, but are greatly intertwined with it (Damasio 1994). Moreover, emotions are essential actors for creative thinking, inspiration as well as concentration and motivation (Isen 2000). Hence, learning systems should intelligently adapt their communication and interaction abilities with learners with regards to changes taking place in this important affective dimension.

A growing body of research in the field of artificial intelligence has identified and reproduced emotions using complex models and physiological sensors. Research in emotional AI for example has enabled computers to start recognizing emotion from speech (D’Mello et al. 2005), facial expressions (Pantic and Rothkrantz 2000) or a hybrid of both these methods (Busso et al. 2004). Other researches attempted to recognize affect by analyzing physiological signals like heart rate and skin conductivity (Arroyo et al. 2009). In the majority of these papers, emotional states were associated to various patterns of physiological manifestations with high accuracy and performance.

However, only scarce research has been conducted on studying the impact that affective changes can have on cognitive processes by means of EEG physiological sensors. Nevertheless, in the last few decades, researchers from various scientific communities have made great improvements in methodologies and technologies that give insight into the brain and the learner’s physiological activity. (Pope, Bogart, and Bartolome 1995) at NASA developed an EEG-engagement index based on brainwave band power and applied it in a closed-loop system to modulate task allocation. Performance in a vigilance task improved when this index was used as a criterion for switching between manual and automated piloting mode (Freeman et al. 2000; Pope et al. 1995). We believe that the integration of this index in education could greatly enhance the ability of AI models to continuously detect, adapt and adjust learning to the user’s level of mental engagement.

Nonetheless, research regarding this engagement index has mainly focused on cognitive aspects. It neglected to take into account the impact that the emotional component can have since emotions and cognition are strongly intertwined. Furthermore, this index has barely been used outside the field of automated cognition.

To that end, we propose bringing this valuable index into the field of education by establishing its importance in modeling learner’s engagement and emotional state. In order to do so, we will address in this paper the two following research questions: (1) Can learner’s emotional states have an impact of the evolution of this engagement index? (2) If so, can this new index, combined with the learner’s emotional state, give useful and valid insight regarding the evolution of learner’s performance?
The organization of this paper is as follows: In the first section, we present previous work done in fields similar to our own. In the second section, we lay the grounds on which the core of this paper is based: the computation of the engagement index. In the third section, we detail our experimental methodology. In the fourth section, we present the obtained results and discuss them, in the last section, as well as present future work.

Previous Work

The integration of physio-cognitive data is one of the most important and promising challenges for developing and significantly improving human-technology interaction by enhancing skill acquisition, performance and productivity in educational, military and industrial fields (Parasuraman 2005). Indeed, several physiological sensors were incorporated in various systems for detecting changes in monitored emotional and cognitive states (Arroyo et al. 2009; McQuiggan and Lester 2009; Murugappan et al. 2007). Having said that, the most reliable and accurate physiological signal for monitoring cognitive state changes remains the electroencephalogram (EEG). The EEG data has a high level of time resolution and precision. In fact, EEG information and features extracted from Power spectral distribution (PSD) bands and/or event related potential (ERP) components has served as input for linear and non-linear models to identify and classify cognitive changes such as alertness, attention, workload, executive function, verbal or spatial memory and engagement (Berka et al. 2007; Russell et al. 2005). In one of the most renown studies, (Sterman et al. 1993) used EEG data to assess mental workload in the evaluation of 15 Air Force pilots during refueling and landing exercises performed in an advanced technology aircraft simulator. Most of these models used statistical and artificial intelligence techniques such as discriminant function analysis (DFA), artificial neural networks (ANN), and support vector machines (SVM) to construct their models. One such stunning example is the study by (Wilson 2005) where 38 measures derived from EEG and heart rate were used to classify with high accuracy workload level and verbal/spatial working memory in unmanned combat air vehicle simulation.

However, to the best of our knowledge, the closest work we could find relevant to our present research is in biocybernetic systems. An EEG-based engagement index was proposed by (Freeman et al. 1999; Pope et al. 1995) and used on a closed-loop method to adjust modes of automation according to operator’s level of engagement. Performance improvement was reported using this engagement index for task allocation mode (manual or automated). Getting closer to education and problem solving, (Stevens, Galloway, and Berka 2007) used EEG-based engagement index to relate mental changes during problem solving task.

This mental engagement index however has always been established without taking into consideration the emotional state of the learner. We propose to extend this index into the educational field and refine it by adding an affective analysis. We will start by explaining how to compute the mental engagement index.

Computing Engagement Index

Before going further, let us summarize in a brief, but concise way, the nature of an EEG signal. EEG, like any electrical signal, is composed of frequencies resulting from electrical neural activity in the brain. These frequencies are often grouped in sequence and are known as bands. Theta band, for example, is the name given to frequencies ranging from 4 to 8 Hz. These bands reflect specific and different cognitive processing abilities in specific areas of the brain (Lubar et al. 1995). Thus, the computation and analysis of frequency bands within power spectral density (PSD) combined with numerous research on alertness and attention provides a powerful tool for monitoring and mapping mental engagement (Lubar et al. 1995). As previously mentioned, (Pope et al. 1995) developed an engagement index using three EEG bands: Theta (4–8 Hz), Alpha (8–13 Hz) and Beta (13–22 Hz). The ratio used was: Beta / (Alpha + Theta). This ratio was also found as being the most effective when validated and compared to many other indices (Freeman et al. 1999).

In our study, we computed the engagement index by applying a Fast Fourier transformation to convert the EEG signal from each active site into a power spectrum. Bin powers (the estimated power over 1Hz) were summed together with respect to each band in order to compute total power and produce the EEG band ratio. By combined power, we mean the sum of band power computed from each measured scalp site. The EEG engagement index at instant T is computed by averaging each engagement ratio within a 40s sliding window preceding instant T. This procedure was repeated every 2s and a new 40s sliding window is used to update the index. Two main methods exist to help interpret this index:

The slope method: the slope of successively derived engagement indexes (every 2s) is computed. More importantly the sign, negative (Low engagement tendency) or positive (high engagement tendency) is also considered.

The absolute method: sets an engagement threshold by averaging engagement index values over a period of time prior to testing (baseline). During task performance, engagement index exceeding the threshold is considered positive and values below the threshold negative.

We opted to follow the absolute method when evaluating changes in the engagement index. Henceforth, we designed and integrated this method in an experiment that will be discussed in the following section.

Experimental Methodology

In order to assess any relation between emotions and the evolution of the engagement index, an experiment designed to provoke specific emotional reactions was conducted. In this experiment, two video feeds, an EEG
headset and physiological sensors were used to monitor and record the user’s reactions throughout the learning process. All the data was synchronized. This setup is important for our investigation. It permits us to compute, offline, the engagement index from the EEG headset sensors. Furthermore, each computed index can easily be paired with the emotional state of the learner obtained from two separate physiological sensors called BVP for blood volume pressure and SC for skin conductance. This pairing (EEG & BVP-GSR), as discussed in the results section, allows us to map and follow the evolution of both these important cognitive and emotional parameters involved in learning. Further widely used parameters were also recorded for the offline analysis, namely question response time (RT) as well as a pre and post questionnaires for general evaluation purposes.

Before the beginning of the experiment, a one minute baseline was recorded for each participant. This widely used technique establishes a neutral emotional and engagement state for future comparisons. During this time, learners were instructed neither to be engaged nor too relaxed. Learners were then told to respond to three series of ten successive true/false questions.

Most of the questions in this experiment were relatively simple and did not require any prerequisite knowledge or specific skills. However, a good level of attention and alertness is required to avoid making easy mistakes. Question response time was limited to 20 seconds. Participants were informed that a correct answer was rewarded one point whereas no points were given for a bad answer or no answer at all. The goal set for all participants was to obtain the highest possible score within the imposed time limit. Apparently obvious questions were designed to mislead learners on purpose. Learners were sometimes so confident of their answer that they would get surprised, frustrated or upset after discovering that they were wrong.

The first series presented general knowledge questions. One sample true/false question would be “Is Rio De Janero the capital of Brazil?” The second series pertained to spell checking. Carefully chosen words were presented one by one on the screen. The task is to determine whether the presented word is properly spelled or not. Finally, participants were asked to respond by true or false to a series of logical statements: “If X < Y-2 then X < Y.”

After each given answer, the system interacted with learners by sending different textual emotional messages to inform them about the correctness of their response. When a good answer is given, the message was encouraging, for example, “Excellent answer! You seem to be very concentrated.” Conversely, in case of a wrong answer, the message could be empathic or may contain an advice, for example, “I’m sure that you know the correct answer” or “Wrong! You need to be more concentrated”.

**EEG recordings.** During data acquisition, learners wore an electro-cap and data was recorded from six active sites, four located on the scalp at locations P3, C3, Pz, Fz as defined by the international 10-20 system (Jasper 1958) and referenced to Cz. The last two active sites are A1 and A2 and are more typically known respectively as the left and right ear. This specific setup, also called a montage, is technically referred to as a “referential linked ears montage” and is illustrated on figure 1. The details and specifics of this montage being out of the scope of the present paper, suffice it to say that the distinct advantage of a referential montage over other setups is that the EEG signal is equally amplified throughout both hemispheres. Furthermore, the “linked-ears” aspect allows us to mathematically obtain a much more precise and cleaner EEG signal by correcting each scalp location signal to that of the middle of the brain. For example, the corrected C3 would become \(C3 = C3 - (A1 + A2)/2\). Overall, one can say that we obtain a “centrally calibrated equally amplified” EEG signal. Electrode impedance was kept below 5 kilo Ohms. A non sticky proprietary gel from Electro-Cap was also used. The recorded sampling rate was 256 Hz.

![Figure 1: Channel electrode placement.](image)

Furthermore, the brain electrical activity is very weak (in the order of micro volts: \(10^{-6}\) volts) and usually contains a lot of noise. Thus all EEG signals were amplified and filtered. Sources of noise are static electricity or electromagnetic fields produced by surrounding devices. A 60-Hz notch filter was applied to remove such environmental interference during the data acquisition phase. It is important to specify that a 50-Hz notch filter should be used in Europe where the power distribution (110 volts) differs from that of North America (120 volts). In addition to external noise, the EEG signal can be heavily contaminated by artifacts that originate from body movement or very frequent eye blinks. Therefore, a 48-Hz high pass and 1-Hz low pass filters were applied for artifact rejection.

**Affect detection with physiological sensors.** To detect affective states, as previously mentioned, learners were equipped with both a blood volume pressure sensor and skin conductance sensor. BVP signals were used to derive the heart rate (HR) whereas SC sensors computed galvanic skin response (GSR). Affective data was recorded at 1024 Hz of sampling rate. These sensors are known to reliably measure specific emotional activations and are widely used for emotional detection. Indeed, as emotions can be characterized in terms of judged valence (pleasant or unpleasant) and arousal (calm or aroused), collected physiological signals were analyzed according to the arousal/dominance emotional space. GSR increases linearly with a person’s level of arousal, while HR has been shown to correlate with valence (Lang 1995). We established four quadrants, labeled Q1 to Q4, with regards
to signal variations in both HR and GSR (figure 2). Thus, learner’s affective state is determined by normalizing HR and GSR variations with regards to the baseline. For example, a positive HR signal and a positive GSR signal will be considered as a signal located in Q1. Normalization is done by mean-shifting (subtracting current values from the baseline and dividing the difference with the standard deviation). For readability purposes, we will refer to the mean-shifted normalized values simply as mean HR and mean GSR from now on.

Figure 2: Lang’s 2D affective space labeled by quadrant

**Participants.** Thirty-five learners (13 women) with a mean age of 27.2 ± 6.91 years, ranging from 19 to 46 years, took part in the experiment. Participation was compensated with 10 Canadian dollars. All participants signed a written consent form, were French speakers with normal or corrected-to-normal vision and without any neuropsychological disorder according to self report.

**Experimental Results and Discussion**

During the session, computed engagement index was associated with one of the four quadrants of the 2D affective space for all learners. The obtained results will be presented in following four sub-sections.

**Emotional states impact the index.** Our first research question was to investigate the possible impact that learner’s emotional states can have on the evolution of the engagement index. The obtained results in this sub-section strongly suggest such impact. In order to analyze the effect of the learner’s emotional state on the engagement index, a one-way ANOVA was performed. The result shows that there is a significant main effect of the emotional state on the index value, \( F(3, 6064) = 115.749, p < 0.01 \) for all participants. Specifically, the analysis of this result revealed that mean engagement index values were significantly higher when learner’s emotional state was in Q1 (Positive valence and high arousal: \( M = 0.769, SD = 0.085 \)) compared to the other quadrants (see figure 3). Giving this result, we can state that positive emotions arising in Q1 (such as joy or excitation) seem to lead to the highest level of learner engagement.

However, the second highest mean engagement index was found in Q3 (Negative valence and high Arousal: \( M = 0.743, SD = 0.074 \)) indicating that emotions in this state (ex: confusion or frustration) might also elicit high engagement levels. An obvious example is the following: the learner answers a question being 100% sure of the correctness of his response. The system then reveals to the learner that he was wrong, placing him in a state of confusion. Thus, learner’s engagement increases to recoup and to better perform on the rest of the test. Finally, the two lowest mean engagement indexes were registered in Q2 (Positive valence and negative arousal: \( M = 0.72, SD = 0.086 \)) and Q4 (Negative valence and negative arousal: \( M = 0.712, SD = 0.105 \)) suggesting that better engagement levels can be attained when the learner is relaxed rather than bored.

**Engagement index correlates with GSR and HR.** The significant impact of the emotional state on learner’s engagement index is a promising result. Nonetheless, we wished to examine in more depth the nature of the relation between the engagement index and the mean GSR and HR signals across all learners. To this end, a bi-variate correlation was computed. Results show a significant linear relation between the engagement index and the GSR signal \( r=0.68, p=0.032 \). However, a non significant linear relation between the engagement index and the HR was found \( r=-0.042, p=0.197 \). Table 1 sums these results. We believe that mentioning this relation with HR is important even though it seems non-significant at first glance because it may hide a subtle, but revealing, significance. Indeed, we have observed that emotional state shifts from Q1 to Q3 (joyful to frustrated) or from Q2 to Q4 (relaxed to bored) explained by a change in the valence dimension (computed from HR) do have a visible but small effect on engagement index. The obtained results in this sub-section showed that better engagement levels can be attained when the learner is relaxed rather than bored.

![Table 1: Results for the bi-variate correlation](image)

<table>
<thead>
<tr>
<th>Index correlation with arousal (GSR)</th>
<th>0.680</th>
<th>0.032</th>
</tr>
</thead>
<tbody>
<tr>
<td>Index correlation with valence (HR)</td>
<td>-0.042</td>
<td>0.197</td>
</tr>
</tbody>
</table>

Table 1: Results for the bi-variate correlation

Therefore, as shown in the first sub-section, higher mean engagement index values are located within Q1 and Q3 and lower mean engagement index values can be traced to
Q2 and Q4. It is important to mention that Q1 and Q3 explain an increase in the engagement index while Q2 and Q4 explain a decrease. In the light of those results, it seems clear that intelligent systems wishing to influence the engagement index, either by increasing or decreasing it, should clearly be aware of the strong influence exerted by the emotional state of the learner. Any AI model wishing to implement this index should take into account learner’s emotional state when planning an intervention or interaction during any given problem solving task. This will strongly attune the intervention, thus sustaining a stronger motivational state and consequently enhance performance. Figure 4 shows an example of one learner’s visible engagement index increase following a positive intervention. We can clearly observe that the positive message was followed by an emotional state shift from Q2 to Q1 resulting in an increase in mental engagement.

**Engagement index and task performance.** We have answered, in the two previous two sub-sections, our first research question pertaining to the importance of the impact of learner’s emotional states on the engagement index. In light of those results, we took a step further by answering the second research question: can this new index, combined with learner’s emotional state, give us useful and valid insight regarding the evolution of learner’s performance? Results in this sub-section will point to the affirmative. In order to analyze the impact of mental engagement on task performance, measured by the obtained score, two groups were considered: (G1) Learners whose mean engagement level was lower than the neutral baseline throughout the experiment (they were less engaged in the task) and (G2) learners whose level engagement was higher than the baseline.

<table>
<thead>
<tr>
<th>Overall performance comparison between G1 and G2</th>
<th>Mean score</th>
<th>Mean score SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1: mentally less engaged</td>
<td>16.00</td>
<td>3.120</td>
</tr>
<tr>
<td>G2: mentally engaged</td>
<td>19.78</td>
<td>2.025</td>
</tr>
</tbody>
</table>

Table 2: Performance comparison between G1 and G2

Results from a one-way ANOVA showed that the performance in G2 was significantly higher than in G1: $F(1, 33) = 19.782, p < 0.01$. Table 2 shows that learners who stayed engaged performed statistically better ($M = 19.8, SD = 2.025$) on average than those who were less engaged in the task ($M = 16, SD = 3.120$).

**Linear regression analysis.** Knowing that the engagement index can serve as an indicator for learner’s performance, we computed a linear regression analysis between emotional indicators and this index. The chosen dependant variable was the mean engagement index on each question. Four predictors were introduced in the model: (1) mean HR, (2) mean GSR, (3) response time (RT) for answering the current question and (4) learner’s current question result, coded +1 for a good answer and -1 otherwise.

The overall model was statistically significant $F(4, 947) = 121.45, p < 0.01; R^2 = 0.289$. Furthermore, conditional main effect analysis showed an effect of the mean HR (beta = -0.001, $p < 0.05$), the mean GSR (beta = 0.42, $p < 0.05$) as well as the mean RT on the current question (beta = 0.21, $p < 0.05$), but a non-significant effect of the answer from the current question (beta = 0.03, $p = 0.46$). At first glance, one might find such result counterintuitive. After all, how can a given answer, either wrong or right, not significantly explain, impact or influence learner’s engagement level? Is it possible that the contribution of the fourth predictor simply got drowned in the averaging process throughout the 35 learners? The answer to both of these questions, after careful observation, resides in the participant himself.

Indeed, we have observed different trends in different learners. For some, a consecutive series of wrong answers lowered their mental engagement. For other, the same effect was observed however for a consecutive series of good answers. This pattern of low engagement index after successive responses was observed on a large portion of the participants. Indeed, learners tend to relax after getting a few good answers and consequently become less engaged. This is where the contribution of the emotional factor is crucial. Intelligent systems should integrate into their model the values from the two emotional dimensions (arousal and valence) to guide the system towards choosing the best intervention strategy for a specific learner. In the case of learners relaxing and becoming mentally disengaged, the system would intelligently recognize this situation and intervene in order to increase their engagement levels. Hence, we propose the construction of an individualized computational model for each learner.

Furthermore, we believe that multiple factors, besides emotions, can influence the engagement index, notably the objectives of the learner (does he want to achieve a medium score or be the best?) and his personality. We believe the obtained results thus far are very encouraging and a deeper analysis is called for in order to successfully construct such a personalized model.

**Conclusion and Future Works**

We have presented in this paper evidence of a direct link between the learner’s emotional state and the engagement index in a problem solving environment. The two
hypotheses in this paper were (1) does the emotional state influence the engagement index and if so (2) what insight can it give us with regards to the learner’s performance? We have successfully answered those questions with an experiment in which learners were asked to answer three series of problem solving questions. Throughout that experiment, physiological and EEG sensors were used to monitor and record the learner’s activity followed by an offline analysis to compute and link engagement index with emotional state changes. Performance was also analyzed. The presented results have clearly shown that (1) emotions impact the engagement index, (2) the engagement index is more correlated with arousal than valence, (3) the engagement index is a valid indicator of learners’ performance and (4) construction of a personalized model to predict the variation of that engagement index is not only possible but highly recommended.

The use of such an index varies from pure AI modeling to intelligent system design. For example, given a required level of engagement for a specific task, a system using this index could calm learners in order to decrease their engagement levels or push and even encourage them to get the opposite effect.

As future work, we propose the elaboration and development of a pedagogical intervention strategy aimed at optimizing learning. This strategy has to take into account the mental engagement index and the learner’s emotional state. We also propose that this strategy considers the workload as well as personal attributes such as personality and objectives in order to intelligently adapt the system’s interventions.

Acknowledgement
We would like to acknowledge the FQRSC and the CRSNG for funding this work.

References


