

# Predicting Learner Answers Correctness through Brainwaves Assessment and Emotional Dimensions

Alicia HERAZ and Claude FRASSON

*HERON Lab; Computer Science Department; University of Montréal,  
CP 6128 succ. Centre Ville Montréal, QC, H3T-1J4, Canada  
{herazali,frasson}@iro.umontreal.ca*

**Abstract.** We want to explore the relation between affective states, brainwaves and the learner answers during a multi-choice test questions. 24 participants were used in our experiment. While we were measuring their brainwaves, we asked them to answer 35 questions related to the 7 texts they read, for the first time, the day before. During the experiment, the participants can rate, at any time, their emotional dimensions (pleasure, arousal and dominance) on the Self-Assessment Manikin scale (SAM). Measuring the brainwaves determines the learner mental state and the emotional dimensions indicate the learner affective state. When a participant answers, he mentions if he knows the answer or not. Each answer can be either Right or False. The hypothesis of this paper is: "We can predict the learner's answers from his emotional dimensions and his brainwaves". By using some machine learning techniques, we reached 90.49% accuracy. In a future work, these results will be implemented in an agent to improve the pedagogical strategies and the adaptation of the content within an Intelligent Tutoring System (ITS).

**Keywords.** Intelligent Tutoring System, Brainwaves, Learning, Emotional Dimensions.

## Introduction

Many researches in the field of Artificial Intelligence and Cognitive Sciences have contributed to the evolution of the learner model and the tutorial strategies. Thus, to the cognitive model [16] were added other layers like the psychological model [5], the affective model [10], the cultural model [15] and the motivational model [3] with specific tutorial strategies.

To get information about the learner, data collection methods have also evolved from self report [2] to facial expression analysis [13], posture and gestures interpretations [1] to biofeedback measurements [6,7,8]. Recent approaches combine different kinds of information channels to increase the prediction of the emotional and cognitive learner states [9].

In the field of biofeedback measurements, few researches were done and many tracks remain unexplored. Our previous works focused on using the Electroencephalogram

(EEG) and collecting brainwaves during learning tasks [6,7,8]. Results show that the student's affect (Anger, Boredom, Confusion, Contempt, Curious, Disgust, Eureka, and Frustration) can be accurately detected (82%) from brainwaves [7]. We have also conducted an experimentation in which we explored the link between brainwaves and emotional assessment on the SAM scale (pleasure, arousal and dominance). Results were promising, respectively with 73.55%, 74.86% and 75.16% accuracy for pleasure, arousal and dominance [8]. Those results support the claim that all rating classes for the three emotional dimensions (pleasure, arousal and dominance) can be automatically predicted with a good accuracy through the nearest neighbour algorithm. These results suggest that inducing some brainwave states could help learners increase their ability to concentrate and decrease their stress levels.

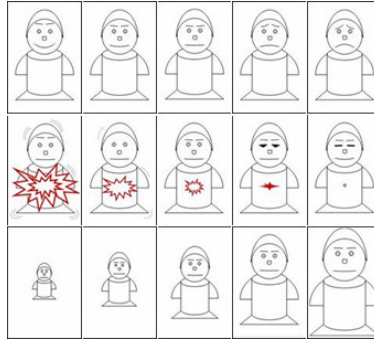
This time we are interested in the test period, when the learner answers questions. We want to track the electrical brain activity of a learner when he is answering questions. We also measure the dimensional emotions that arise during learning. According to Lang (2005), each emotion has three major dimensions: pleasure, arousal, and dominance. The learner indicates his emotional dimensions by rating them on a scale named the Self-Assessment Manikin (SAM), an affective rating system [11].

During the test, a learner can guess the right answer or not and when he answers he indicates if he is sure or not. The question of this paper would be then "Can we predict the learner's answer from his brainwaves and his emotional dimensions?"

We believe that the cognitive state has an impact on brainwaves which inform about the learner's answer. In order to investigate this hypothesis, this paper has reached three objectives. The first one was to conduct an experiment to record the electrical brain activity of 24 participants when they were answering 35 questions. The second objective was to use machine learning technique to predict the learners' answer from his brainwaves. The third objective was to complete our previous work and implement an agent to be added to the architecture of a multi-agent system that measures brainwaves and predicts efficient learner's metrics.

## **1. Emotional Dimensions**

To assess the three dimensions of pleasure, arousal or dominance, we use the Self-Assessment Manikin (SAM), an affective rating system devised by Lang [11]. In this system, a graphic figure depicting values along each of the 3 dimensions on a continuously varying scale is used to indicate emotional reactions (Figure 1).



**Figure 1.** Self-Assessment Manikin System

Variance in emotional assessments were accounted for by three major dimensions: affective valence (ranging from pleasant to unpleasant), arousal (ranging from calm to excited) and dominance (or control). For the pleasure dimension, SAM ranges from a smiling, happy figure to a frowning, unhappy figure. For the arousal dimension, SAM ranges from an excited, wide-eyed figure to a relaxed, sleepy figure. For the dominance dimension, SAM ranges from a small figure (dominated) to a large figure (in control). Ratings are scored such that 9 represents a high rating on each dimension and 1 represents a low rating on each dimension.

## 2. Dominant Brainwaves

The neurons in the brain are divided into bands, some slow, some moderate and some fast, measured by cycles per second. Brain activity is characterized by the production of electrical signals reflected in brain waves. These are of very low voltage and are measured in Hertz or cycles per second. Brainwaves are rapid fluctuations of voltage between parts of the brain that are detectable with an EEG. Bioinformation allows us to reorganize the brain's activities through mental training. The brainwaves that we measured were categorized into 6 different frequency bands, or types. According to their frequency, the waves are given the following names: delta, theta, alpha, beta1, beta2 and beta3 waves. Each of these wave type correlates with a particular mental state [14]. Table 1 lists the different frequency bands and their associated mental states. The performance of our mind depends on the predominant type of wave at any given moment:

**Table 1.** Brainwaves and Mental States

Brain Wave	Learner State	Wave EEG
Delta (.05-3 Hz)	Produced in deep, dreamless sleep	
Theta (4-7 Hz)	Drowsiness, inattention, deep meditation.	
Alpha (8-12 Hz)	General relaxation and meditation	
Beta low (12-15 Hz)	Relaxed concentration. Often used for seizure control.	
Beta Midrange (15-20 Hz)	Focused attention	
Beta High (> 20 Hz)	Anxiety	

We call the dominance order of each brainwave the amplitude order of a brainwave. For example: if the amplitudes of the brainwaves (delta, theta, alpha, beta1, beta2, beta3) are respectively (4, 12, 8, 29, 15, 39) then the dominance order is (dom1 = beta3; dom2 = beta1; dom3 = beta2; dom4 = theta; dom5 = alpha; dom6 = delta). We can see that “beta3” is the first dominant brainwave which means that the mental state of the participant at this moment is “Anxiety” (according to table 1).

### 3. Describing the Experiment

Initially, we selected 24 undergraduates from the Computer Science Department at University of Montréal (3 of them were discarded later, see details in section 4). One day before the experiment, participants were asked to read one time and carefully a set of 7 articles. They were selected from old French newspapers. On the day of the experiment, we recorded the brainwaves of each participant when he was answering the 35 questions related to the 7 texts he read before. The EEG used is Pendant EEG [12]. Pendant EEG sends electrical signals to the computer via infrared connection. The electrical signal recorded by the EEG is sampled, digitized and filtered to divide it into different frequency bands. Light and easy to carry, Pendant EEG is not cumbersome and can easily be forgotten within a few minutes. The learner wearing Pendant EEG is completely free of his movements: no cable connects him/her to the computer. Figure 2 shows the overall experiment architecture.

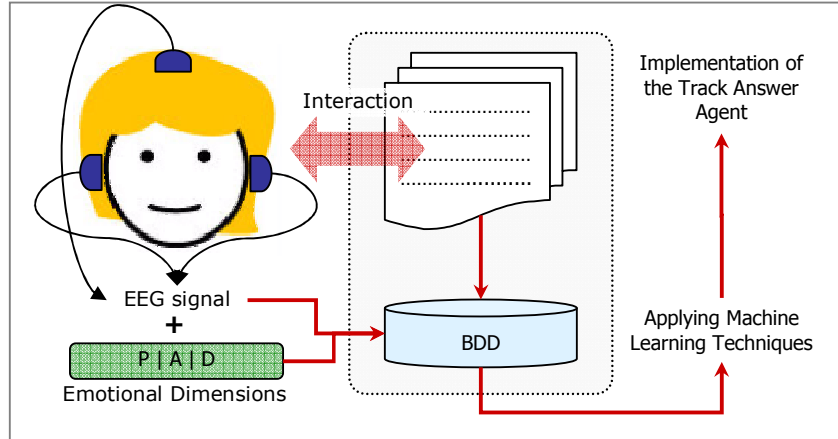


Figure 2. The Overall Experiment Architecture.

When asked a question, a participant indicates if he knew the answer or if answers randomly. In both cases, the answer could be right or false. Table 2 shows the whole possibilities. The test duration’s varies from 15 to 20 minutes for each participant.

Table 2. Learners Answer’s Types

Code	Answer	Meaning
RG	Right: Guessed	The learner knew the answer and he guessed right
RR	Right: Random	The learner did not know the answer but he guessed right
FG	False: Guessed	The learner thinks he knew the answer but he does not
FR	False: Random	The learner did not knew the answer and he guessed false

#### 4. Data collected: Filtering and Statistics

Over 2 weeks and for a total duration of 30 hours, we used 24 participants and collected 47559 recordings into a database. Before analyzing the data, we did some preliminary treatments and cleanings.

##### 4.1. Data Treatments and Statistics

We discarded 3 participants because they had difficulties to understand the texts due to their French language level of comprehension comparing to the other participants. So the data size was reduced to: 45827 recordings in the database. Figure 3 shows the repartition of the learner's answers types among the 6 frequency bands for dom1. We had 17157 recordings of RG answers, 10803 of FR answers, 5608 of RR answers and 12259 FR answers.

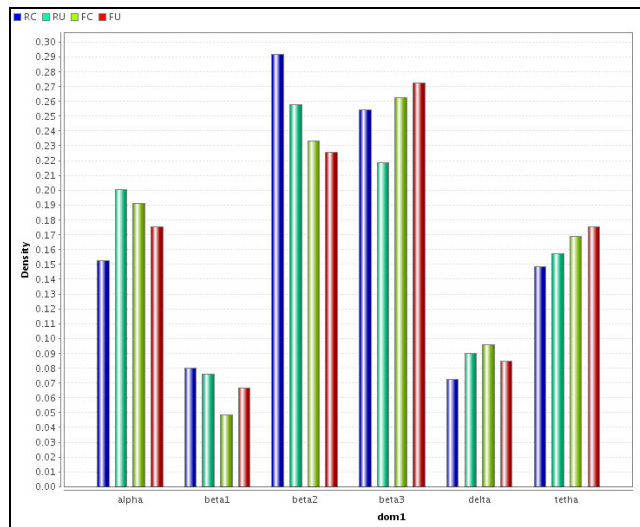


Figure 3. Distribution of Answers among the First Dominant Wave.

Right answers represent 49.68% of the sample. Among of the right answers, 24.63% were a RR answers. We can observe that RG answers occur more frequently when the brainwave beta2 dominates and less frequently when delta dominates.

#### 5. Machine learning results

For classification we used WEKA, a collection of machine learning algorithms recommended for data mining problems implemented in Java and open sourced under the GPL [17]. The problem of determining the learner's answer types from the

emotional dimensions and the brainwaves amplitudes can be represented as the following mapping function.

$$f(p, a, d, dom1, dom2, dom3, dom4, dom5, dom6) \rightarrow response \quad (1)$$

Where  $(p, a, d) \in \{1, 2, 3, \dots, 9\}^3$  are the three emotional dimensions presented in section 1 (respectively: pleasure, arousal and dominance).

And  $(dom1, dom2, \dots, dom6) \in \{\delta, \theta, \alpha, \beta_1, \beta_2, \beta_3\}^6$  are the dominant brainwaves presented in section 2 (first, second, ... and sixth dominance).

Many classification algorithms were tested. The best results were given respectively by the algorithms: Classification via Regression, k-Nearest Neighbour (k=1), Bagging Decision Tree and Random Forest. Table 3 shows the overall classification results using k-fold cross-validation (k = 10). In k-fold cross-validation the data set (N) is divided into k subsets of approximately equal size (N/k). The classifier is trained on (k-1) of the subsets and evaluated on the remaining subset. Accuracy statistics are measured. The process is repeated k times. The overall accuracy is the average of the k training iterations.

The various classification algorithms were successful in detecting the learner answers' types from his emotional dimensions and his brainwaves. Classification accuracy varies from 87% to 90%. Furthermore, Kappa statistic scores is excellent, they vary from 0.83 to 0.87. Kappa statistic measures the proportion of agreement between two rates with correction for chance. Kappa scores ranging from 0.4 – 0.6 are considered to be fair, 0.6 – 0.75 are good, and scores greater than 0.75 are excellent.

**Table 3.** Best Results

Algorithm	Accuracy	Kappa Statistic	Mean Absolute Error
Classification Via Regression	87.70%	0.83	0.12
K-Nearest Neighbor (K=1)	88.16%	0.84	0.06
Bagging	88.35%	0.84	0.09
Decision Tree (J48)	88.57%	0.84	0.07
Random Forest	90.49%	0.87	0.07

For the Random Forest algorithm, table 4 shows the details of classification accuracy among the 4 classes (the 4 types of the learner's answers).

**Table 4.** Detailed accuracy by Class

TP Rate	FP Rate	Precision	Recall	F-Measure	Class
0.914	0.057	0.906	0.914	0.910	RG
0.882	0.032	0.894	0.822	0.888	FG
0.892	0.013	0.907	0.892	0.900	RR
0.918	0.033	0.911	0.918	0.915	FR

For the Random Forest algorithm and according to table 4, we calculated the Youden's J-index to increase the weight to the rating classes with minority instances [18] as the following formula (2):

$$JIndex = Card(RC)^{-1} \sum_{e \in RC} Precision_e \quad (2)$$

With  $Card(RC)$  is the cardinality of rating classes list (4 classes). The JIndex value is 90.45% which is nearly equal to the classification prediction shown in table 3 (90.49%). This result supports the claim that all rating classes for the 4 classes can be automatically detected with good accuracy (90.45%) through the Random Forest algorithm. The highest classification rates appear on the Matrix Diagonal as shown in Figure 4.

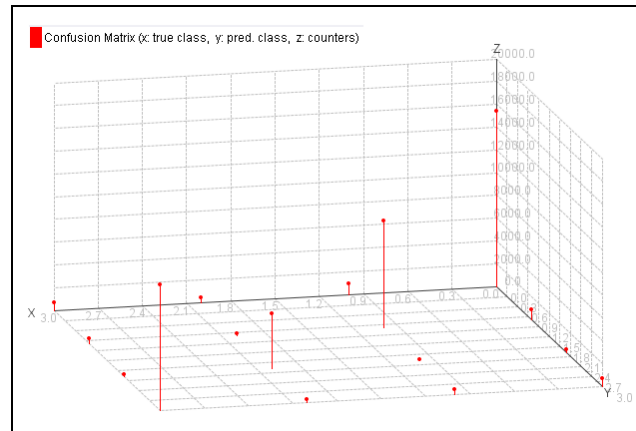


Figure 4. Confusion Matrix.

## 6. Conclusion and Perspective

This study used machine learning techniques to test the hypothesis: “We can predict the guessed and the random learner’s answers from his emotional dimensions and his brainwaves”. 24 participants were recruited, 3 were discarded (because their difficulties to understand the texts due to their bad comprehension level of French language) and 45827 recording were analyzed from the database. We used Pendant EEG to measure the brainwaves while the participants were asked to answer some questions related to the texts they read the day before. The participant wearing Pendant EEG is completely free of his movements: no cable connects them to the machine Pendant EEG sends the electrical signals to the machine via an infrared connection. With this infrastructure, we reduced the possible side effect of the material.

We acknowledge that the use of EEG has some potential limitations. In fact, any movement can cause noise that is detected by the electrodes and interpreted as brain activity by Pendant EEG. However, we gave very strict instructions to our participants. They were asked to remain silent, immobile and calm. We believe that the instructions given to our participants, their number (21) and the database size (45827 records) were able to considerably reduce this eventual noise.

Results are encouraging. In fact, the Random Forest analyses resulted in accurate predictions 90.49% and the Yuden’s J-Index is 90.45%. These results show that we could implement in the future an agent that communicates information about the

learner state to an ITS. This will increase the accuracy of the learner model. An ITS will choose an adequate pedagogical strategy that adapt to certain learner's mental states correlated to the brainwaves frequency bands, his emotional dimensions and his emotional states.

## 7. Acknowledgments

We acknowledge the support of the FQRSC (Fonds Québécois de la Recherche sur la Société et la Culture) and NSERC (National Science and Engineering Research Council) for this work.

## References

- [1] Ahn, H.I., Teeters, A., Wang, A., Breazeal, C., and Picard, R.W., (2007). Stoop to Conquer: Posture and affect interact to influence computer users' persistence, The 2nd International Conference on Affective Computing and Intelligent Interaction, September 12-14, Lisbon, Portugal.
- [2] Anderson, J. R. (2001). Tailoring Assessment to Study Student Learning Styles. In American Association for Higher Education, (53) 7.
- [3] Boyer K.E. , Phillips R., Wallis M., Vouk M., and Lester J., Balancing Cognitive and Motivational Scaffolding in Tutorial Dialogue. (2008) In Proceedings of the Ninth International Conference on Intelligent Tutoring Systems, Montreal, Canada.
- [4] DMello, S., Jackson, T., Craig, S., Morgan, B., Chipman, P., White, H., Person, N., Kort, B., el Kaliouby, R., Picard, R.W. and Graesser, A. (2008), AutoTutor Detects and Responds to Learners Affective and Cognitive States, Workshop on Emotional and Cognitive Issues at the International Conference of Intelligent Tutoring Systems, June 23-27, Montreal, Canada.
- [5] Grandbastien M. et J.M. Labat. (2006). Environnements informatiques pour l'apprentissage humain. Éditions Lavoisier, Paris, France.
- [6] Heraz A., Daouda T. and Frasson C., (2008) Decision Tree for Tracking Learner's Emotional State predicted from his electrical brain activity. ITS2008. Montréal, Canada.
- [7] Heraz A., Razaki R. and Frasson C., (2007). Using machine learning to predict learner emotional state from brainwaves. ICALT. Niigata, Japan.
- [8] Heraz A. and Frasson C., (2007). Predicting the Three Major Dimensions of the Learner's Emotions from Brainwaves. International Journal of Computer Science.
- [9] Kapoor, A., Ahn, H.I., Picard, R.W. Mixture of Gaussian Processes for Combining Multiple Modalities, in Proceedings of Multiple Classifier Systems, Eds. N. C. Oza, R. Polikar, J. Kittler, and F. Roli, 6th International Workshop, MCS 2005, June 2005, Seaside, CA, pp. 86-96.
- [10] Kort B., R. Reilly and R.W. Picard (2001), An Affective Model of Interplay Between Emotions and Learning: Reengineering Educational Pedagogy-Building a Learning Companion, In Proceedings of International Conference on Advanced Learning Technologies (ICALT 2001), August 2001, Madison.
- [11] Lang, P.J., Bradley, M.M., et Cuthbert, B.N. (2005). International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Technical Report A-6. University of Florida.
- [12] McMillan Bruce (2006). <http://www.pocket-neurobics.com>
- [13] Nkambou R. et Héritier V. (2004). Facial expression analysis for emotion recognition in ITS. In: ITS'2004 workshop on Emotional Intelligence proceedings.
- [14] Norris, S. L., and Currier, M. Performance enhancement training through neurofeedback. In Introduction to Quantitative EEG and Neurofeedback, (1999) J. R. Evans and A. Abarbanel, Eds.
- [15] Ogan, A., Alevin, V., & Jones, C. (2008). Pause, predict, and ponder: Use of narrative videos to improve cultural discussion and learning. In Proceedings of the Twenty-Sixth Annual SIGCHI Conference on Human Factors in Computing Systems (pp. 155-162). New York, NY: ACM.
- [16] Wenger, E. (1987). Artificial Intelligence and Tutoring Systems. Los Altos, CA: Morgan Kaufmann.
- [17] Witten Ian H. and Eibe Frank (2005) Data Mining: Practical machine learning tools and techniques, 2nd Edition, Morgan Kaufmann, San Francisco, 2005.
- [18] Youden. W. J. (1961). How to evaluate accuracy. Materials Research and Standards, ASTM.