

# Design of a Decision-Making Task for a Collaborative Brain-Computer Interface System Based on Emotiv EEG

Ânderson Schuh<sup>(✉)</sup> and Márcia de Borba Campos

Faculty of Informatics (FACIN),  
Pontifical Catholic University of Rio Grande do Sul (PUCRS),  
Porto Alegre, Brazil

anderson.schuh@acad.pucrs.br, marcia.campos@pucrs.br

**Abstract.** This article presents lessons learned in the design, implementation and evaluation of a task of computerized decision-making to be used in a non-invasive collaborative and hybrid brain-computer interface, which used the Emotiv EEG for extract neural feature and response time as behavioral feature. The task developed was based on RSVP and has controlled levels of difficulty that can cause uncertainty. It is believed that the participants' general satisfaction was good, since the majority indicated that they had an easy understanding of the task. The task proved to be efficient for the initial purpose, that is, to generate difficulty to the participants and the experiment can be balanced with respect to the difficulty of executing the task. However, it was not possible to find relationships between the emotions felt by the participants in their subjective answers and in their emotions collected through the Emotiv EEG. It was possible to verify that the participants with less response time tend to answer more correctly, which can indicate their level of confidence, as expected.

**Keywords:** Collaborative BCI · Hybrid BCI · EEG · RSVP · Making-Decision · Response time

## 1 Introduction

A Brain-Computer Interface (BCI) measures directly the brain activity associated with the user's intention, and translates it into control signals which are detected and interpreted by applications [1–3]. The non-invasive BCIs can be based on the electroencephalogram signs (EEG), a device that distributes electrodes by the scalp and, through them, registers the electrophysiological brain activity [4]. A hybrid BCI (hBCI) consists of the combination of two or more types of BCI, two or more signal acquisition techniques, or a combination of a BCI with other interaction techniques not based on BCI [5–8]. But a Collaborative BCI (cBCI) integrates the brain activity of a group of individuals, with the main aim to improve the classification of signs or increase human capacity [9]. Similar to a conventional BCI, a cBCI system has key parts for its operation, and will be applied to two of them: signal acquisition and signal processing. First, the brain signals to a group of users are acquired by various recording devices and, then, are synchronized with environmental events which are common among all

the users. After this, the processing of the collected data takes place in order to decode the intentions of the users, and, finally, they are translated to operating commands [9].

Currently, BCI equipments are easily found on the market, as the case of Emotiv EEG and Neurosky MindWave. In spite of this, it only [10] used resources which are present in this equipment for the prototyping of a cBCI for humorous content classification in images. In this study, we used the Emotiv EEG.

A concern/limitation of the BCI area is the application of these systems in a real environment, especially with regard to the extension of human capacity, because other systems usually present more satisfactory results than a BCI [9]. In this way, it is necessary to find daily tasks that can benefit from its use. The Decision Making (DM) is an issue that has been discussed in the area of Economics and Administration [11]. In addition to being a day-to-day task, it shows, in many cases, complex problems, especially when there is risk or uncertainty [12]. A way to minimize those cases is the DM. However, some research shows that communication is not always an ally of this process [13]. With regard to the cBCIs, where the application domain is the decision making, there are studies that use the techniques of Go/NoGo [15, 16] and Rapid Serial Visual Presentation (RSVP) [17–19]. Through an RSVP, one can simulate adverse conditions to the user. Thus, to decrease the display time of an image at a rate where it is not possible to complete the identification of the elements, there is an uncertainty and the user needs to make a decision.

This article presents lessons learned in the design, implementation and evaluation of a task of computerized decision-making to be used in a non-invasive collaborative and hybrid brain-computer interface, which used the Emotiv EEG.

## 2 Method

We decided to perform an experiment similar to the one presented in [17], which developed a cBCI and used tasks based on RSVP. Despite the record in the RSVP design considerations, it was not considered the information coming from the intentional blink, saccadic movements and memorization after 6 s, which would require a longer time on the data analysis. Some important requirements were raised:

- After performing the task, ask the user for an auto report of his/her experience during the experiment, for possible triangulation of information with the neural and behavioral data collected.
- Some tasks are more emotional and can stimulate various mental processes. With this, the development of a simple task that doesn't involve emotional issues may facilitate the mapping of which processes are important.
- Use simple motor tasks, because, just as the stimuli, can interfere with the mapping that is important.
- Binary responses tasks are the most common and fit well in the original purpose of the work, as visual recognition tasks (search for targets) or planning.

## 2.1 Sample

Unlike what happens in the traditional research, the choice of the participants was intentional (not probabilistic), that is, they were chosen on the basis of the interest issues to the study, profile and availability to perform the tasks.

## 2.2 Data Collection and Instruments

The planning process of the data collection was based on five questions, suggested by [20]. It was analyzed the following characteristics of each participant of the research:

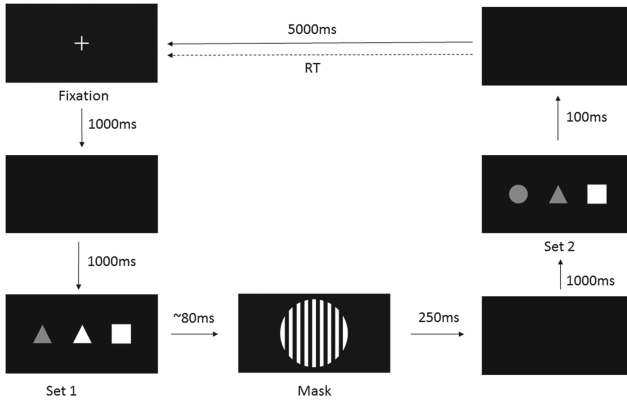
1. Profile by means of a questionnaire;
2. Neural Feature (NF) of the Affective Suite registered by the Emotiv EEG;
3. The Decision Making (DM) registered through the application containing the tasks;
4. Response Time (RT) registered through the application containing the tasks;
5. Perceptions by means of questionnaires.

The questionnaires were made available in digital form and were answered in the presence of the researcher, at specific times, following a given order. The log with the data collected by the BCI was automatically registered in a database.

It was used the Emotiv EEG and its Affectiv Suite for the acquisition of the NF, which monitors in real time the subjective emotions experienced by the user, based on brain waves, being possible to detect: engagement, meditation, frustration, instant excitement and long-term excitement. The detection of the engagement can investigate measures, such as surveillance, alertness, concentration, stimulation and interest. Instant Excitement detection can determine measures, such as excitement, nervousness and restlessness. Finally, the long-term excitement detection performs the same measurements as the previous detection, but its diagnosis is usually more accurate, since it analyzes these measures in a longer period of time (in minutes). The meditation detection represents the relaxation or stress level. The frustration measurement already recognizes what the own name says [21]. For all the detections, the values are represented in the scale of 0 to 1, considering 1 as a strong existence of the emotion, and 0 as the absence. Still, the same detection can point out different emotions, as the meditation detection, in which values close to 1 indicate mental relaxation, and around 0 represents stress or discomfort.

The task was developed using the game engine Unity 3D. For the storage of the collected data, it was used the embedded database SQLite. For the access to the Emotiv EEG Affectiv Suite data, it was used the software Mind your OSCs, which accesses the Emotiv engine, collects the data and makes them available through the OSC Protocol.

Each participant performed a sequence of 2 stages, each one containing 56 tests. Each test begins with the presentation of a white fixation cross centered on the computer screen with a duration of 1 s. This allows the participant to prepare him/herself for the presentation of the stimuli. After this, two screens are presented, each containing a set of images that are geometric shapes. The first set of images appears for about 80 ms and is immediately followed by a mask for 250 ms. The mask is used to erase any remnant of the first set. After 1 s, the second set for 100 ms is shown on.



**Fig. 1.** Developed task RSVP based

As a result, the participant must decide, as soon as possible, if the two sets are identical or different. The answers are given through a conventional QWERTY standard keyboard positioned in front of the user. The F key must be used to indicate the same sets of images while the key J must indicate different sets. These keys were used because they have embossed marking, differentiating them in the keyboard. The response time of each decision is stored. The next test takes 5 s to begin, after the appearance of the second set of elements. Figure 1 illustrates this dynamic.

Each set consists of 3 geometric shapes, which can be any combination of triangles, squares and circles. Each shape can take one of the colors, white (1, 1, 1) or gray (0.65, 0.65, 0.65). With this, it can be said that each of the three elements in a set has two features, color and shape. Therefore, each set has a total of  $(2 \times 3)^3 = 216$  different combination possibilities. If the two sets of one test are considered, we have  $216^2 = 46.656$  possible combinations. In this sense, features may be shared, that is, when they occur in the same position in two sets of an essay.

In this way, each pair of a test set is sorted by the number of these shared features, which is called the Degree of Match (DoM). If all the elements of the first set are distinct from the elements of the second set, we have a  $DoM = 0$ . In the case of an element share a features, for example, the same color, we have a  $DoM = 1$ . In this way, the DoM of each test can range from 0 to 6. To exemplify, in Fig. 1, which has in the first position of the first set a gray triangle. In the second set, the first position has a gray circle. In this case, the two elements share a features, the color. Still, in the second set, the second element is a gray triangle. Although they present the same shape and color, the triangle of the first set does not share features with that one of the second set, since they are not in the same position.

The combination of elements of the first set is randomly generated. In order to produce a proportional experiment at the level of difficulty for the second set, some restrictions were adopted in the randomization of the combinations, using the DoM values to define the difficulty of each test, and, thus, to generate equal proportions of each DoM level in the experiment. In this case, the experiment had in its totality of tests a multiple value of 7, which represents the amount of possible DoM, totaling 96 different

sets and 16 identical sets. After the generation of the elements sets for the assays, they were stored to be used with all the participants, in order to perform the same experiment, increasing its repeatability and reproducibility. No set of stimuli is repeated.

The display time of the first set of elements stays below of what would normally be required for its complete visualization and perception, according to [22]. This can help to identify the confidence level of the participant, who, by failing to identify all the components of the set with high difficulty, may be in doubt, and present a longer response time, as described in the study of [17]. In comparison with [17], the waiting time after the response was increased, so that it was possible to register the features, mainly, of frustration and stress.

### 2.3 Procedures

The experiment was performed in a controlled environment, in a closed room with a table, in which were placed a portable computer that runs the applications, a 19-inch monitor, a mouse and a keyboard, an Emotiv EEG that was connected to the computer through the Wireless adapter, and two chairs. During the experiments, only the participant and the evaluator remained in the room. The laptop screen was used by the evaluator for monitoring, while the 19-inch monitor was used by the participant in the use of the task. All the evaluations were done individually.

For each of the evaluations, the previous configuration of the environment lasted about 40 min. This process took into account the organization of the environment, configuration of the applications and the preparation of the Emotiv. This stage is fundamental to the experiment because the bad hydration and the improper positioning of the electrodes can generate noise in the signals captured by the Affective Suite, or the lack of the Emotiv communication with the application of the data collection or the application of the tasks can generate the lack of some data registration.

With the configured environment, the researcher explained to each participant the research objectives, data collection procedures, the collected data confidentiality, estimated duration of the experiment, discomforts that might be felt during the tasks, among other informations. The participation only occurred after the participant's agreement manifestation, in the informed consent of the research.

The pre-test questionnaire was applied and, afterwards, the researcher positioned the Emotiv on the participant's head, then he tested if all the electrodes were correctly placed and explained the task that was to be performed. The equipment was placed before the actual task was done, so that during the explanation of the task, it was self-calibrated to the participant, as recommended by the equipment manufacturer. Then, the participant positioned himself in front of the monitor, with the index fingers on the keys F and J. After that, the evaluator started the task. After performing the first 56 attempts, the user answered the questionnaire regarding his/her perception about the first stage. After completion, the volunteer performed the second stage of the experiment and answered the second questionnaire. Finally, the participant completed the final questionnaire about his/her overall perception of the experiment. Only at the end is that the equipment was removed from the participant's head. This process took about 40 min for each participant.

### 3 Results

Initially, the sample was composed of 11 participants who answered the questionnaires and performed the activities. However, in the analysis of the database, it can be seen that all the Emotiv EEG data coming from a participant had not been recorded. Therefore, this participant was disregarded of the final sample. On the other hand, we chose to keep the data of another participant who did not only have the recorded frustration data. Thus, the final sample consisted of 10 participants.

#### 3.1 Profile of the Participants

The questionnaire regarding the participants' profile contained 8 mandatory questions, besides the name and an identifier, which were filled out by the researcher.

Of the 10 participants of the experiment, 9 were male and 1 was a female. 8 participants are between 18 and 27 years and 2 participants are between 28 and 37 years. As for nationality, 9 were Brazilians and one was Colombian. None of the participants indicated to be photosensitive, thus, do not have sensitivity to light.

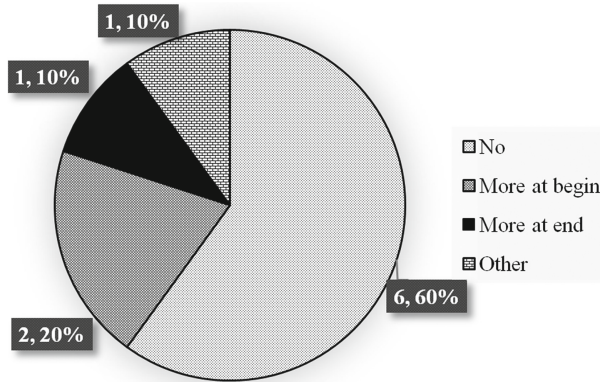
As for the use of the hands, all of them indicated to be right-handed and informed that the right hand is the one with which, preferably, they handle the mouse and the mouse buttons. Regarding the general use of the keyboard, 4 participants reported using both hands and all their fingers and write without looking at the keyboard, 2 reported using their left hand and all their fingers, being that one of them does not look to the keyboard while typing, 2 participants use only their right hand and all their fingers, and 2 others did not report the hand that they usually use to write, but reported typing without looking at the keyboard. To write messages on the smartphone, 8 participants indicated that they use both hands and, predominantly, a finger of each hand, and 2 participants indicated that they use only they right hand and, predominantly, a finger of that hand to write the messages. When asked about the frequency of the perception of games use in the week before the experiment, the majority (60%) reported that they did not use this type of game, 20% played 3-4 days, 10% played every day and 10% played for 1-2 days.

#### 3.2 Questionnaire Stage 1

The explanation of the questions in the Questionnaire Stage 1 was made prior to the activities of the Stage 1. This procedure was adopted in order that the participant had knowledge of what he/she should answer and observe during this stage. At all times, the participant remained with the BCI connected and in operation to prevent it from being recalibrated.

This questionnaire had 6 questions, 4 of which were of multiple choice, 1 with answers based on semantic differential scales, and the other one was open, which are discussed below.

*Have you felt any visual discomfort?* (Fig. 2): participant 3 reported that he had visual discomfort but, even though he did not specify the timing, it is believed to have



**Fig. 2.** Have you felt any visual discomfort?

been during the course of the task because he suggested that the ambient light could be darker. Moreover, comparing the participants' responses with their accuracy and error rates between the stages of the experiment, it cannot be said that such sensations perceived by the participants reflected in their performance.

*In general, about the use of keys in the keyboard to indicate the choices "same" and "different", do you consider that it was: the range of answers had a scale of 1 to 5, representing from the more difficult (1) to the easier one (5).* The average of the answers was 4.1, indicating that the use of the F and J keys were found to be easy to use.

Participant 5 was the only one who considered the use of the keys more difficult than easy, even though he/she indicated that he/she makes use of the keyboard without looking at the keys. It is noteworthy that this participant was the one that obtained the best rate of success in Stage 1, with 91% of success (51 questions). He/she was the participant who obtained the best overall score (88.4%–99 questions). The participants who considered the F keys and J easy to be used had their levels of accuracy rates near average, just above or below.

*Overall, how did you feel during the challenges?* This question was divided in 8 questions, applied with semantic differential scales, as follows, and the answers from 1 to 5:

- Distracted (1) × Tuned (5): the average was 4.1, demonstrating that the participants felt themselves to be more attentive than distracted on the achievement of the challenges. Only the Participant 5, who considered the use of the keys harder to be used, indicated that he/she was more distracted than attentive in performing Stage 1;
- Uncompromised (1) × Engaged (5): the average was 4.4, demonstrating that the participants felt themselves to be more engaged than uncompromised while performing the challenges;
- Frustrated (1) × Satisfied (5): considering that 3 is the central point of the scale and the average was 3.4, it demonstrates that the participants felt themselves a little happier than uncompromised while performing the challenges. Only Participant 2

indicated that he/she was frustrated while performing the activities, however, his/her accuracy rate in Stage 1 remained on the average (82.1% and 46 issues);

- Bored (1) × Excited (5): the average was 4.2, demonstrating that the participants felt themselves more excited than bored while performing the challenges;
- Stressed out (1) × Relaxed (5): considering the central point of the scale and that the average was 3.3, it demonstrates that the participants felt themselves a little more relaxed than stressed while performing the challenges. Participants 2, 5, 11 and 12 indicated that they were more stressed than relaxed while performing Stage 1. Of these, Participant 12 presented the lowest accuracy rate, being of 58.9%, which is equivalent of 33 hits. At the end of the experiment, its average rose to 67.9%, with a total of 76 hits;
- Nervous (1) × Calm (5): the average was 3.7, demonstrating that the participants felt themselves a little calmer than nervous while performing the challenges. Participants 2, 5 and 11 indicated that they were more nervous than calm, and had already indicated they were more stressed than relaxed;
- Angry (1) × Quiet (5): the average was 4.2, demonstrating that the participants felt themselves more calm than angry while performing the challenges. Only Participant 2 indicated to be angrier than calm;
- Uncertain (1) × Confident (5): the average was 3.8, demonstrating that the participants felt themselves confident than uncertain while performing the challenges. Participants 2 and 5 indicated that they were uncertain than confident while performing the activities of Stage 1. On the other hand, if the accuracy rate of the Participant 2 was 82.1% 2 and 46 hits, which was the average of the participants in Stage 1. Participant 1 was the one that got the best accuracy rate in Stage 1, which was of 91.1% and 51 hits.

*Overall, do you think that your accuracy rate was better at identifying the “same” or “different”?* As previously described, this stage had 56 pairs of screens (stimulus set), 8 of which represented the same screens and 48 different screens. Most participants (80%) (Participants 2, 3, 4, 5, 6, 7, 9, 10, 11) consider that their accuracy rate was better in identifying the different, one participant (10%) (Participant 12) considers that his/her rate was better in identifying the same and the other (10%) (Participant 6) considers that there is no significant difference in identifying the same or the different. These results demonstrate that the participants are aware of the difficulty to observe small details in stimuli, demonstrating the efficiency of the proposed task.

*Do you think that your accuracy rate was better in that phase of Stage 1?* The majority (80%) (Participants 2, 3, 4, 6, 9, 10, 11 and 12) consider that their accuracy rate was better in the final phase of the stage, in keeping with the already informed statements. Of these, only Participant 3 did not have his/her answer confirmed, being that in the first half of Stage 1, he/she made 23 hits, and in the second half he/she hit 21 times. One (10%) (Participant 5) found that there was no significant difference in his/her accuracy rate, confirmed by the correct answers, being that 25 were in each half of Stage 1. Another (10%) (Participant 7) could not state, being that there was an increase of 6 hits in the final half of Stage 1.

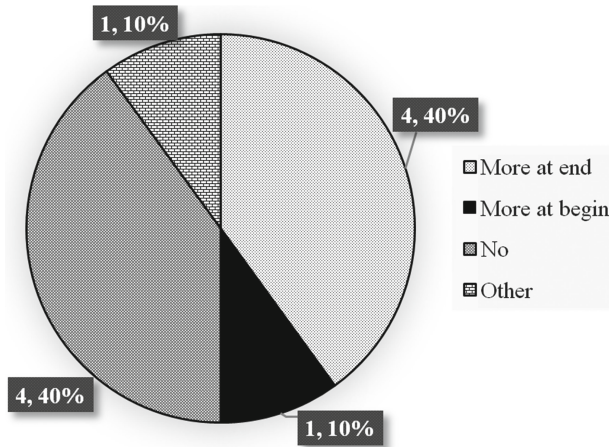


### 3.3 Questionnaire Stage 2

The procedure adopted in the second stage was the same from the previous stage, that is, before the battery of tests was started, the questionnaire was read and explained, which contained 6 identical questions to the previous questionnaire, which are discussed below.

#### *Have you felt any visual discomfort?*

Participants were asked if they felt uncomfortable sensations during the visual task. The answers are illustrated in Fig. 3.



**Fig. 3.** Have you felt any visual discomfort?

4 participants (40%) (Participants 2, 4, 6 and 7) reported that they had more visual nuisance at the end of the stage. Of these, Participant 2 had also indicated that he/she had it only at the end of Stage 1, Participants 4 and 7 indicated that they had no nuisance in Stage 1, and Participant 6 had indicated that he/she had at the beginning of Stage 1.

Four participants (40%) (Participants 5, 9, 10 and 11) had no nuisance. Of these, Participants 5, 10 and 11 also did not indicate nuisances in Stage 1. And Participant 9 had indicated nuisance at the beginning of Stage 1.

Participant 1 (10%) (Participant 12) reported that he/she had nuisance at the beginning of the stage, being that he/she indicated no nuisance during Stage 1.

The other one (10%) (Participant 3) did not specify in what time he/she had visual nuisance, but continued recording suggestion so that the test environment was less bright.

Overall, about the use of keys on the keyboard to display the choices “same” and “different”, do you consider that it was: The answers range had a scale of 1 to 5, representing from the more difficult (1) to the easier one (5). The answers average was of 4.4, being that most of the participants kept their assessments. Participants 5 and 12 considered that the use of these keys got easier in Stage 2. Thus, considering that the average of Stage 1 was of 4.1, it can be inferred that the use of the keys throughout the experiment facilitated their learning and use.

*Overall, how did you feel during the challenges?* This question was divided in 8 questions, applied with semantic differential scales, as follows, and the answers were registered from of 1 to 5:

Distracted (1)  $\times$  Tuned (5): the average was 4.3, demonstrating that the participants felt themselves to be more attentive than distracted on the achievement of the challenges. Only the Participants 4 and 6 indicated having diminished their attention, from 4 to 3 and from 5 to 3, respectively. The other ones indicated or an increase (Participants 3, 5, 9 and 11) or kept the same level from Stage 1 (Participants 2, 7, 10, 12);

- Uncompromised (1)  $\times$  Engaged (5): the average was 4.6, demonstrating that the participants felt themselves to be more engaged than uncompromised while performing the challenges;
- Frustrated (1)  $\times$  Satisfied (5): considering that 3 is the central point of the scale and the average was 3.7, it demonstrates that the participants felt themselves a little happier than uncompromised while performing the challenges;
- Bored (1)  $\times$  Excited (5): the average was 4.2, demonstrating that the participants felt themselves more excited than bored while performing the challenges;
- Stressed out (1)  $\times$  Relaxed (5): considering the central point of the scale and that the average was 3.6, it demonstrates that the participants felt themselves a little more relaxed than stressed while performing the challenges;
- Nervous (1)  $\times$  Calm (5): the average was 3.6, demonstrating that the participants felt themselves a little calmer than nervous while performing the challenges;
- Angry (1)  $\times$  Quiet (5): the average was 4, demonstrating that the participants felt themselves more calm than angry while performing the challenges;
- Uncertain (1)  $\times$  Confident (5): the average was 4, demonstrating that the participants felt themselves confident than uncertain while performing the challenges.

*Overall, do you think your hit rate was better in identifying the "same" or "different"?* Most of the participants (80%) (Participants 2, 4, 5, 7, 9, 10, 11 and 12) consider that its accuracy rate was better in identifying the different, one participant (10%) (Participant 3) considers that its rate was better in identifying the same and other participant (10%) (Participant 6) considers that there is no significant difference in identifying the same or the different. As in Stage 1, the participants demonstrate awareness of the difficulty to observe small details in stimuli, demonstrating the efficiency of the proposed task.

*Do you think that your accuracy rate was better in which phase of the Stage 1?* Three participants (30%) (Participants 9, 11 and 12) believe that their accuracy rate was better in the early phase of the stage, however, they had in average 4.6 hits in the final half of Stage 2. Three participants (30%) (Participants 4, 6 and 10) believe that there was no significant difference in their accuracy rate, in spite of this, the participants had an increase at the end of Stages 5 and 8 and 1 hit, respectively. One participant (10%) (Participant 3) thinks that he/she was better between the phases, which was validated, and between the Test 70 and the 88, he/she obtained 25 hits, against 23 hits in the other tests of Stage 2. One participant (10%) (Participant 2) believes that his/her accuracy rate was better in the final phase and other participant (10%) (Participant 7) didn't know what to inform.

### 3.4 Final Questionnaire

Upon completion of Stage 2, and completing the questionnaire relating to that stage, a final questionnaire was applied.

*Do you consider that the more challenges similar to those that have been presented have to solve, the better your attention to observe visual sequences?* Eight participants (80%) (Participants 2, 3, 5, 7, 9, 10, 11 and 12) consider that the greater the experience with the task, the higher their attention. One participant (10%) (Participant 4) believes that there would be no improvement in attention and other one (10%) (Participant 6) believed to be indifferent. However, one can't see this performance improvement during the experiment.

*Do you consider that the more similar challenges to those which were presented you have to solve, the better will be your agility to use the keys of the keyboard to indicate "same" and "different"?* Six participants (60%) (Participants 2, 4, 5, 7, 9, and 12) consider that the greater the experience with the task, the higher will be their agility with the keys. Three participants (30%) (Participants 3, 6 and 10) positioned themselves as indifferent to the improvement. Finally, one participant (10%) (Participant 11) considers that it wouldn't make any difference. It was found that there is a small improvement in the average answer time of the participants between the stages, being 923 ms in stage 1 and 870 ms in stage 2.

*Considering that it is a visual attention challenge, do you consider that the time that separates one challenge from another (marked by the screen +) was:* Eight participants (80%) (Participants 2, 3, 4, 5, 6, 7, 9 and 12) consider that time is suitable and 2 participants (20%) (Participants 10 and 11) indicated that it could be higher.

*Considering that it is a challenge of visual attention, do you consider that the time to identify the sequences same and different was:* The majority of the participants (70%) (Participants 2, 3, 4, 5, 6, 10 and 12) consider that the time could be higher, while 3 participants (30%) (Participants 7, 9 and 11) believe that the time was appropriate. This confirms that the participant's perception goes against the difficulty proposed in the task.

*Considering that it is a challenge of visual attention, do you consider that the time to identify the sequences same and different should vary along the Stage? For example, start with a longer duration and reduce it until it stabilizes.* Nine participants (90%) (Participants 2, 3, 4, 5, 6, 9, 10, 11 and 12) considered that time could be reduced according to the experience with the task, but 1 participant (10%) (Participant 7) felt that it could not be reduced.

*Considering that it is a challenge of visual attention, do you consider that the background colors and objects are clearly identified?* Five participants (50%) (Participants 6, 7, 9, 11 and 12) consider that the colors used are clearly identified. Four participants (40%) (Participants 2, 3, 5 and 10) consider that the colors are not easily identified and, finally, one participant (10%) (Participant 4) considers it as indifferent. The difficulty in identifying the colors is related to the proposed time of exhibition, however, the majority of the participants were able to notice these differences.

### 3.5 Relationship Between Measures and Responses

In the previous sections, we sought to present the participants' perception to perform the experiment, data resulting from the data collection by means of a questionnaire. This section seeks to present the data collected through the BCI and discuss possible relations between neural and behavioral measures, as well as the participants' perceptions. It also includes discussion of error rates and relation to the difficulties of the proposed task.

As for the data collected regarding to the neural measurements, each test generated, on average, 70 records of each of the 5 measures extracted by the Emotiv EEG, which are engagement, frustration, meditation, excitement, and long-term excitement. Also, the response times of each test were stored.

Due to the generated data volume, it was performed summarization of the neural feature data per assay. The summarization is a descriptive statistics area technique, which consists of the synthesis of the data collected and, despite the loss of informations, it is still a minor factor compared to the gain provided in the interpretations [23]. The average will be used as a summarization method, since it is tried to understand if the increase or decrease of some of the neural features can influence in the decision making.

In Table 1 are presented the average of each of the neural and behavioral measures, along with its standard deviation, minimum and maximum value. Due to a read error, the frustration data of Participant 7 were not recorded correctly and not to interfere with the general analysis, the frustration data of this participant were not considered in the analysis. Therefore, the other features are based on data of 10 participants while the frustration is based on the data of 9 participants. It turns out that, considering the central measures and the standard deviation, the analyzed periods do not have considerable differences.

**Table 1.** General information about the recorded measures, \* 9 data users.

Feature	Average	Median	STD	Minimum	Maximum
Engagement	0.5712	0.5750	0.1187	0.0984	1.0000
Excitement	0.3360	0.3010	0.1552	0.0214	0.9885
Meditation	0.3811	0.3659	0.0636	0.2149	0.6667
Frustration*	0.4599	0.4177	0.1734	0.1203	1.0000
Response time	0.8967	0.8000	0.4155	0.0800	3.8600

Table 2 summarizes the amounts of errors and hits by difficulty level of all users. There are 112 trials per participant, distributed equally in the DoM 7 levels, from 0 to 6. It is verified that the greater the DoM, the greater the difficulty of the user to perceive the details of the image, and, in this way, greater the error rate. On the other hand, it is verified that DoM 6 has more hits than DoM 5. This occurs, since the user, when he/she cannot identify a different features with a higher incidence in DoMs 5 and 6, ended up voting as if it were the same. In the case of DoM 6, the answer is considered correct. In this way, it is believed that the applied theory of RSVP in the proposed task, fulfills its role, generating the difficulty in recognizing the more complex images.

**Table 2.** Accuracy rate and error by levels of DoM

Difficulty	Accuracy	Errors	Not rated
0	156	4	0
1	150	10	0
2	153	6	1
3	131	28	1
4	126	34	0
5	89	71	0
6	109	51	0

Table 3 presents the accuracy and error rates of the participants in stages 1 and 2. The data has been leased to a decimal place. It can be verified that 2 participants (2 and 9) had a small increase in the accuracy rate (both from 82.1% to 83.9%), being only one question. Three participants (7, 11 and 12) had greater increases in the accuracy rate (from 78.6% to 87.5%, from 73.2% to 85.5%, and from 58.9% to 76.8%, respectively), with 5, 7 and 10 questions, respectively. Three participants (3, 4, 10) had a slight increase in the error rate (from 21.4% to 23.2%, from 14.3% to 16.1% and from 16.1% to 17.9%, respectively), being only of one question, while one participant (5) had an increase of 3 wrong questions (from 8.9% to 14.3%). One participant (6) maintained his/her score in two stages, with 85.7% of accuracy.

**Table 3.** Accuracy rate per participant between stages

	Stage 1				Stage 2				Total			
	Accuracy		Error		Accuracy		Error		Accuracy		Error	
ID	Qty.	%	Qty.	%	Qty.	%	Qty.	%	Qty.	%	Qty.	%
2	46	82.1	10	17.9	47	83.9	9	16.1	93	83.0	19	17.0
3	44	78.6	12	21.4	43	76.8	13	23.2	87	77.7	25	22.3
4	48	85.7	8	14.3	47	83.9	9	16.1	95	84.8	17	15.2
5	51	91.1	5	8.9	48	85.7	8	14.3	99	88.4	13	11.6
6	48	85.7	8	14.3	48	85.7	8	14.3	96	85.7	16	14.3
7	44	78.6	12	21.4	49	87.5	7	12.5	93	83.0	19	17.0
9	46	82.1	10	17.9	47	83.9	9	16.1	93	83.0	19	17.0
10	47	83.9	9	16.1	46	82.1	10	17.9	93	83.0	19	17.0
11	41	73.2	14	25.0	48	85.7	8	14.3	89	79.5	22	19.6
12	33	58.9	22	39.3	43	76.8	13	23.2	76	67.9	35	31.3
avg	46	82.1	10	17.9	47	83.9	9	16.1	93	83.0	19	17.0
std	5.0	8.9	4.6	8.3	2.1	3.7	2.1	3.7	6.4	5.7	6.1	5.4

Still, the next questions addressed the user’s perception during the execution of the experiment. Table 4 present a summary of the data collected.

**Table 4.** Subjective answers per participant (ID) ans per Stage 1 (S1) and Stage 2 (S2).

ID	Dis. x Tun.		Unc. x Eng.		Frustr. x Sat.		Bor. x Exc.		Str. x Rel.		Ner. x Cal.		Ang. x Qui.		Uns. X Con.	
	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
2	4	4	5	5	2	2	4	5	2	2	2	2	2	2	2	2
3	4	5	5	5	5	4	5	5	5	5	5	5	5	5	5	5
4	4	3	5	5	4	5	5	5	5	5	5	5	5	5	4	5
5	2	4	3	4	3	3	4	3	2	2	2	2	3	3	2	4
6	5	3	3	4	3	4	4	4	4	4	3	4	5	4	5	4
7	4	4	5	5	3	4	4	4	3	5	5	4	4	4	4	4
9	4	5	4	4	3	3	4	4	4	3	4	3	4	3	4	4
10	5	5	5	5	3	4	5	5	4	4	4	5	5	5	4	4
11	4	5	5	5	5	5	4	4	2	2	2	2	5	5	5	5
12	5	5	4	4	3	3	3	3	2	4	5	4	4	4	3	3
avg	4.1	4.3	4.4	4.6	3.4	3.7	4.2	4.2	3.3	3.6	3.7	3.6	4.2	4	3.8	4

In order to compare the neurophysiological measures with the participants’ subjective answers, a table of the same scale of the collected neurophysiological measures was created. Initially, for each participant, the maximum and minimum values of each of the neural variables were found. After that, the maximum by the minimum was subtracted, and the result was divided by 5, in order to obtain the extreme values of each of the 5 levels of the scale. After that, the average of each of the neural variables was extracted for each of the stages. Afterwards, these values were classified in a scale of 1 to 5, based on the values obtained in the first operation. In the case of the neural measure Frustration, the values were inverted, being the lowest value 5 and the higher 1. This was necessary to adjust the scale used in the questionnaire that varied from Frustrated (1) to Satisfied (5). The obtained values are shown in Table 5.

General Engagement was compared to Distracted vs. Tuned and Uncompromised × Engaged, general Excitement was compared to Bored vs. Excited, general Meditation was compared to Stressed vs. Relaxed, Nervous vs. Calm and Irritated vs. Quiet, and lastly, general Frustration with Frustrated vs. Satisfied.

From 160 possible answers (8 subjective questions in 2 stages to the 10 participants), 4 participants had no equal value between the subjective and the neural answers. Only 15 answers of all participants obtained the same values between the subjective responses and the neural variables, however, these were not repeated between the stages. As an example, Participant 2, for the subjective answer of the Stage Distracted vs. Tuned, got level 4, the same level of his/her neural feature Engagement for Stage 1. When analyzing the increase or decrease in the levels between stages, we noticed 22 times of the 84 possible changes, being that from the 22 times, 12 times remained the

**Table 5.** General levels for neural features per participant (ID) ans per Stage 1 (S1) ans Stage 2 (S2).

ID	Gen. engagement		Gen. excitement		Gen. meditation		Gen. frustration	
	S1	S2	S1	S2	S1	S2	S1	S2
2	4	3	2	2	1	2	1	2
3	3	2	2	2	3	3	3	1
4	3	2	2	2	3	3	2	3
5	2	3	3	2	1	2	3	2
6	3	2	3	3	2	2	3	3
7	4	2	1	2	2	4		
9	4	2	2	2	2	2	3	2
10	3	2	2	2	2	3	2	2
11	3	3	2	3	3	3	2	2
12	3	1	2	1	4	3	1	1
avg	3.20	2.20	2.10	2.10	2.30	2.70	2.22	2.00

same values for stage 1 and stage 2. As an example, participant 4, for the subjective answer Distracted vs. Tuned, had a variation of 1 level between stage 1 (level 4) and stage 2 (level 3), as to the levels of the neural feature Engagement, also occurred a decrease of 1 level between stage 1 (level 3) and stage 2 (level 2), i.e., the same behavior.

When analyzing the participants' average, a similar behavior between the subjective answers and the neurophysiological measures can be noticed in two cases. As for the General Excitement, 2.1 was obtained in both stages, being that in the "Bored  $\times$  Excited" questionnaire it was obtained 4.2 in both stages. Still, in the case of general Meditation regarding the question "Stressed  $\times$  Relaxed", the two participants had an increase between 0.3 and 0.4 points. However, for both cases, physiological measures presented medium-low levels and the subjective answers presented medium-high levels.

In this sense, we sought to analyze the relationship between the subjective answer for "Uncertain  $\times$  Confident" with the participant's accuracy rate. Participant 5 got the best accuracy rate, reaching 88.4%, however, it was not the participant who said to be confident. In the first stage, his/her accuracy rate was 91.1%, and his/her subjective answer was 2, the lowest recorded. The same participant also indicated an improvement in his/her security, raising from 2 to 4 in the second stage, however, his/her accuracy rate decreased to 85.7%. Participant 4 indicated an increase of his/her security from 4 to 5 between stage 1 and stage 2. On the other hand, his/her security rate fell from 85.7 to 83.9, i.e. a hit less. Participant 6 indicated a drop from 5 to 4 points in his/her security between the stages, but his/her accuracy rate remained at 85.7% in both stages. Participants 7, 11 and 12 obtained the highest accuracy rate variations between the both stages, however, their perceptions among the stages remained the same. In this way, no evidence was found that could satisfactorily relate the participants' subjective answers

that were collected through questionnaires with the neural measurements, which were recorded by the BCI through the adopted formulas.

The RT was analyzed using boxplot graphs (Fig. 4), where it was found that the values for correct answers are concentrated in lower values than in the wrong answers. This behavior was expected, as previously described, since participants with greater certainty make decisions more quickly. However, we can see that most of the values still share the same distribution. This may indicate that the participants are right even in doubt.

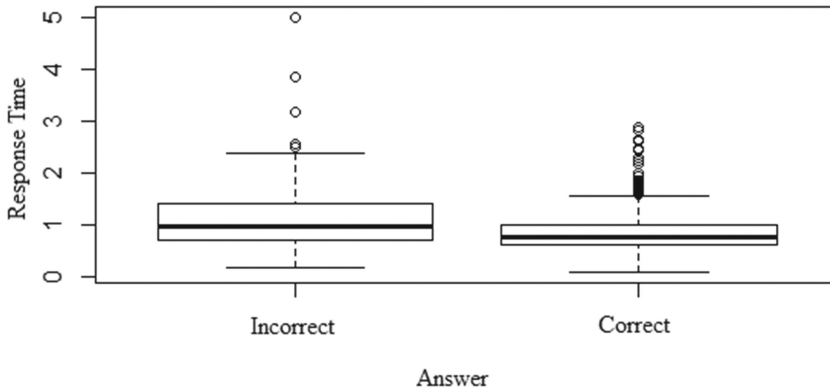


Fig. 4. Response times boxplot graph for all participants

## 4 Conclusion

This article presented lessons learned from the design, implementation and evaluation of a computerized decision-making task for use in a non-invasive, collaborative and hybrid brain-computer interface using Emotiv EEG. The BCI area has advanced in the last few years, especially in new approaches such as hybrid BCIs and collaborative BCIs. End-user equipment, such as Emotiv EEG, facilitates the access to the signal acquisition equipment.

It is believed that the participants' general satisfaction was good, since the majority indicated that they had an easy understanding of the task. As for the visualization time of the stimuli, the task proved to be efficient for the initial purpose, that is, to generate difficulty to the participants, along with the DoMs. In this way, the experiment can be balanced with respect to the difficulty of executing the task.

However, it was not possible to find relationships between the emotions felt by the participants in their subjective answers and in their emotions collected through the Emotiv EEG's Affective Suite. Moreover, it was possible to verify that, in an empirical way, the participants with less response time tend to answer more correctly, which can indicate their level of confidence, as expected.



## References

1. Wolpaw, J.R.: Brain-computer interfaces as new brain output pathways. *J. Physiol. Online* (2007). doi:[10.1113/jphysiol.2006.125948](https://doi.org/10.1113/jphysiol.2006.125948)
2. Leuthardt, E.C., Schalk, G., Roland, J., Rouse, A., Moran, D.W.: *Neurosurgical Focus* (2009). doi:[10.3171/2009.4.FOCUS0979](https://doi.org/10.3171/2009.4.FOCUS0979)
3. Graimann, B., Allison, B., Pfurtscheller, G.: Brain-computer interfaces: a gentle introduction. *Brain-Comput. Interfaces* (2010). doi:[10.1007/978-3-642-02091-9\\_1](https://doi.org/10.1007/978-3-642-02091-9_1)
4. Lebedev, M.A., Nicolelis, M.A.: Brain-machine interfaces: past, present and future. *Trends Neurosci.* (2006). doi:[10.1016/j.tins.2006.07.004](https://doi.org/10.1016/j.tins.2006.07.004)
5. Wolpaw, J.R.: Brain-computer interfaces as new brain output pathways. *J. Physiol.* (2007). doi:[10.1113/jphysiol.2006.125948](https://doi.org/10.1113/jphysiol.2006.125948)
6. Nicolas-Alonso, L.F., Gomez-Gil, J.: Brain computer interfaces, a review. *Sensors* (2012). doi:[10.3390/s120201211](https://doi.org/10.3390/s120201211)
7. Amiri, S., Fazel-Rezai, R., Asadpour, V.: A review of hybrid brain-computer interface systems. *J. Adv. Hum.-Comput. Interact. Spec. Issue Using Brain Waves Control Comput. Mach.* (2013). doi:[10.1155/2013/187024](https://doi.org/10.1155/2013/187024)
8. Pfurtscheller, G., et al.: The Hybrid BCI. *Front. Neurosci.* (2010). doi:[10.3389/fnpro.2010.00003](https://doi.org/10.3389/fnpro.2010.00003)
9. Wang, Y., Wang, Y.-T., Jung, T.-P., Gao, X., Gao, S.: A collaborative brain-computer interface. In: *Proceedings of 2011 4th International Conference on Biomedical Engineering and Informatics* (2011). doi:[10.1109/BMEI.2011.6098286](https://doi.org/10.1109/BMEI.2011.6098286)
10. Stoica, A.: MultiMind: multi-brain signal fusion to exceed the power of a single brain. In: *Proceedings of the 3rd International Conference on Emerging Security Technologies* (2012). doi:[10.1109/EST.2012.47](https://doi.org/10.1109/EST.2012.47)
11. Chiavenato, I.: *Administração: Teoria, Processo e Prática*. Makron Books (2000). ISBN:85.346.1078-9
12. Bahrami, B., Olsen, K., Latham, P.E., Roepstorff, A., Rees, G., Frith, C.D.: Optimally interacting minds. *Science* (2010). doi:[10.1126/science.1185718](https://doi.org/10.1126/science.1185718)
13. Manes, F., Sahakian, B., Clark, L., Rogers, R., Antoun, N., Aitken, M., Robbins, T.: Decision-making processes following damage to the prefrontal cortex. *Brain J. Neurol.* (2002). doi:[10.1093/brain/awf049](https://doi.org/10.1093/brain/awf049)
14. Hagen, G.F., Gatherwright, J.R., Lopez, B.A., Polich, J.: P3a from visual stimuli: task difficulty effects. *Int. J. Psychophysiol.* (2006). doi:[10.1016/j.ijpsycho.2005.08.003](https://doi.org/10.1016/j.ijpsycho.2005.08.003)
15. Valeriani, D., Poli, R., Cinel, C.: A collaborative brain-computer interface to improve human performance in a visual search task. In: *International IEEE/EMBS Conference on Neural Engineering* (2015). doi:[10.1109/NER.2015.7146599](https://doi.org/10.1109/NER.2015.7146599)
16. Yuan, P., Wang, Y., Gao, X., Jung, T.-P., Gao, S.: A collaborative brain-computer interface for accelerating human decision making. In: Stephanidis, C., Antona, M. (eds.) *UAHCI 2013. LNCS*, vol. 8009, pp. 672–681. Springer, Heidelberg (2013). doi:[10.1007/978-3-642-39188-0\\_72](https://doi.org/10.1007/978-3-642-39188-0_72)
17. Poli, R., Valeriani, D., Cinel, C.: Collaborative brain-computer interface for aiding decision-making. *PLoS one* (2014). doi:[10.1371/journal.pone.0102693](https://doi.org/10.1371/journal.pone.0102693)
18. Poli, R., Cinel, C., Sepulveda, F., Stoica, A.: Improving decision-making based on visual perception via a collaborative brain-computer interface. In: *IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support, CogSIMA 2013* (2013). doi:[10.1109/CogSIMA.2013.6523816](https://doi.org/10.1109/CogSIMA.2013.6523816)

19. Valeriani, D., Poli, R., Cinel, C.: A collaborative brain-computer interface for improving group detection of visual targets in complex natural environments. International IEEE/EMBS Conference on Neural Engineering (2015). doi:[10.1109/NER.2015.7146551](https://doi.org/10.1109/NER.2015.7146551)
20. Rogers, Y., Sharp, H., Preece, J.: Design de interação: além da interação humano-computador, 3rd edn. Bookman, Porto Alegre (2013)
21. Emotiv, July 2016. [www.emotiv.com](http://www.emotiv.com)
22. Spence, R., Witkowski, M.: Rapid Serial Visual Presentation: Design for Cognition. Springer, London (2013). doi:[10.1007/978-1-4471-5085-5](https://doi.org/10.1007/978-1-4471-5085-5)
23. Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E.: Multivariate Data Analysis. Prentice Hall, New Jersey (2005)