Effects of the concentration level, eye fatigue and coffee consumption on the performance of a BCI system based on visual ERP-P300

Cristian Felipe Blanco-Díaz a,b,*, Cristian David Guerrero-Méndez a,b, Teodiano Bastos-Filho a, Sebastián Jaramillo-Isaza b, Andrés Felipe Ruiz-Olaya b

a Postgraduate Program in Electrical Engineering, Federal University of Espírito Santo (UFES), 29075-910 Vitória, Brazil
b Faculty of Mechanical, Electronic and Biomedical Engineering, Antonio Nariño University (UAN), Cra. 3 E No 47A 15 Bogotá, Colombia

1. Introduction

During the last decades, BCI technologies have been used to restore the communication and control skills of people with disabilities. These systems can identify the user’s intention and execute an assigned task with information from Central Nervous System (CNS) without using the Peripheral Nervous System (PNS) (McFarland and Wolpaw, 2011). For this reason, BCI systems have been implemented and researched in education, motor rehabilitation, robotics, and medicine (Chaudary et al., 2016; Abdulkader et al., 2015; Perez Vidal et al., 2016; Rezeika et al., 2018; Picton, 1992). According to the 10–20 international system for EEG electrode positioning, the P300 signal is recorded primarily on the brain’s occipital cortex channels, such as the O2 channel.

Most BCI systems obtain brain information through Electroencephalography (EEG), intending to generate patterns in signals related to specific mental tasks or strategies. One task of selective attention corresponds to the oddball paradigm, which allows the evocation of Event-Related Potential (ERP) P300. The P300 is a positive deviation of microvolts, with a latency (between stimulus and response) of approximately 300 ms, registered through EEG after presenting an infrequent stimulus for the subject (Rezeika et al., 2018; Wolpaw and Wolpaw, 2012). The most used stimulus to evoke P300 is the visual stimulus, used with spellers in studies related to communication rehabilitation (Karimi et al., 2019; Carlson, 2021; Vazquez-Marrufo, 2017; Picton, 1992). According to the 10–20 international system for EEG electrode positioning, the P300 signal is recorded primarily on the parietooccipital cortex of the brain. Visual P300 is also recorded in the brain’s occipital cortex channels, such as the O2 channel.

The main drawback in implementing BCI systems based on P300 lies in the detection of the ERP P300, which has a low amplitude (order of μV) and, with this, a low signal-to-noise ratio (SNR), considering that the P300 signal has components that are involved in psychological and neurophysiological processes. The oddball paradigm is related to concentration tasks because the subject is constantly being stimulated...
with random stimuli. Moreover, when an oddball stimulus is presented, the P300 signal is evoked. For this reason, it has been studied that concentration plays an essential role in the behavior of the P300 signal (Carlson, 2021; Vazquez-Marrufo, 2017), e.g., a decrease in the amplitude of the P300 signal has been demonstrated when the subject performs two simultaneous tasks during an experimental design (Picton, 1992). The concentration factor has also been studied through more interactive BCI systems using a video game, where a positive relationship was reported between the amplitude of the signal and the concentration level of the participants during the experiment (Koike et al., 2019). As well as concentration level, other neurophysiological factors have been of interest in studies of BCI systems (Pitt and Brumberg, 2021), such as motivation (Hammer et al., 2012), emotional stability (Hammer et al., 2018), and working memory (Sprague et al., 2016). Furthermore, moderate but significant correlations have been found between BCI system performance and the psychological characteristics of the individual (Hammer et al., 2012).

Another factor of interest with the BCI using visual stimuli corresponds to the form of stimulation because when the subject is presented with unusual stimuli, such as faces (Lee et al., 2019) or different colors (Guo et al., 2019), the performance of the BCI system is better. However, a problem arises: constant stimulation can generate visual fatigue for an extended period. This factor has been considered in EEG signals involving visual stimulation, where a decrease in accuracy has been reported in individuals who showed eye fatigue (Biever et al., 2010; Collura, 2002; Guerrero-Mendez et al., 2021).

As mentioned in some previous examples, the electrophysiological behavior of the P300 signal can be affected by the user's cognitive, mental, and physical processes during the development of an experimental design in a BCI system. These factors related to individual-specific conditions could generate artifacts that affect the features of the P300 signal. With this, the recognition algorithms in an ERP-based BCI system may reduce or improve their performance whether the subject is focused, unfocused, with eye fatigue, without eye fatigue, or had a consumption of coffee in the hours before experimental development (Carlson, 2021; Picton, 1992; Deslandes et al., 2004; Chen et al., 2019). For this reason, and considering the open challenge of the scientific community for decoding the P300 signal in the presence of noise (Won et al., 2019; Dmitriev et al., 2018; Xiao et al., 2019; Blanco Díaz and Ruiz Olaya, 2020; Philip and George, 2020), the detection of physiological and cognitive variables that decrease the accuracy of the system and affects the performance of BCI are of interest (Hammer et al., 2018, 2012). The goal is to carry out more controlled and accurate experimentation sessions, thus providing greater comfort to the individual who uses the interface and maximizing the system's usability.

This paper performs a comparative study to evaluate how specific mental and physical factors could affect the performance of P300-BCI systems. Two methods reported in the literature were used to identify P300: The Mean Amplitude with Linear Discriminant Analysis (MA-LDA) (Hwang et al., 2017; Lee et al., 2019), and the Canonical Correlation Algorithm with Regularized Logistic Regression (CCA-RLR) (Blanco Díaz and Ruiz Olaya, 2020). These methods were implemented using a public EEG database, which was segmented into different subject groups for the following factors: good vs. bad concentration level, high vs. low eye fatigue, and people who consumed coffee before the experiment at different times. In summary, the BCI system based on P300 may be affected by any of the above three conditions.

A maximum average AUC of 85.08% was obtained for the group with high concentration, which was significant concerning those with low concentration (78.74%). Additionally, the group that consumed coffee more than 4 h before the experiment had a significant better performance (88.25%) compared to those who did not consume coffee (84.26%) or consumed coffee less than 4 h before the experiment (81.06%). Finally, the group that did not perceive visual fatigue obtained an increase of 5.08% in AUC compared to the group that perceived visual fatigue. The results allow this work to conclude that concentration factor, visual fatigue, and coffee consumption play an important role in the performance of BCI systems based on the detection of P300 signals. This fact leaves an open door for future studies that allow a more controlled environment of these factors, and maximizing the interface's capacity in detecting the individual's intention.

This article is organized in the following way. The next section describes the experimental methodology that includes the protocol, signal processing, and implementation of algorithms. Then, Section 3 presents the results obtained in the validation. And finally, the discussion and conclusions are presented in the last two sections.

2. Materials and methods

The methodology was segmented into selecting EEG datasets, pre-processing signals, data analysis, P300 identification, performance metrics, and statistical significance analysis. The summary of the methodology is represented in Fig. 1.

2.1. EEG dataset

An open-access dataset of EEG recordings was used for the method implementation, which includes three different paradigms. This dataset was created by Lee et al. (2019). EEG signals were recorded using the BrainAmp device, with a sampling frequency of 1000 Hz that recorded 62 electrodes according to the 10–20 EEG international system. A total of 54 healthy subjects between 24 to 35 years participated in this study and performed two sessions for data acquisition. First, a training session was conducted, where offline data were used to build a classifier. Then, a validation session was conducted, where data were acquired in real-time. The experiment was composed of a speaker based on the oddball paradigm with six rows and six columns with a stimulus interval of 80 ms and an inter-stimulus time of 135 ms. Thirty-three speller characters were used for the training session, and 36 characters were used for the validation session, each repeated 10 times for a total of 330 target stimuli in training (total of 1980) and 360 target stimuli in the validation (total of 2160) (see Fig. 2). At the beginning of each session, the subjects answered a questionnaire to identify their physical and mental factors, i.e., their psychological state during the experiment. For example, the subjects were asked about comfort, motivation, concentration, eye fatigue, drowsiness, physical, and mental conditions, and when coffee was drunk in the last hours (1–24 h) or if not consumed. The questionnaire mentioned above categorizes concentration levels and eye fatigue levels. Values vary from 1 to 5, where 1 represents the highest level (very good), and 5 is the lowest level (very bad). For this study, subjects who manifested high levels (1 and 2) for each situation were grouped and separated from low levels (3, 4, and 5) to form two groups for each factor. Finally, groups were equalized to avoid bias by forming two groups: group 1 for good concentration and no eye fatigue; group 2 for very bad concentration level and eye fatigue, with 25 subjects for level concentration and 14 subjects for eye fatigue in each group.

In contrast, the coffee consumption methodology was different because the subjects answered the questionnaire about the time of the last coffee or no consumption. For this reason, the classification was split into 4 groups: group 1 if the subject did not consume coffee; group 2: if
the subject consumed coffee 1 h before the experiment, group 3: if the subject consumed coffee between 2 h and 4 h before the experiment; and group 4: if the subject consumed coffee between 4 h and 8 h before the experiment. More details on the formed groups is available in Table 1.

2.2. Pre-processing

Considering the temporal and spectral behavior of the ERP P300 (Picton, 1992; Lee et al., 2019), the EEG signals were filtered with a fifth-order Butterworth filter between the frequencies range of 0.5 and 40 Hz and segmented between 0 and 800 ms. The following channels: $F_z$, $C_z$, $P_z$, $P_3$, $P_4$, $O_1$, $O_2$, $PO_1$, and $PO_2$ (see Fig. 3) were used according to the brain lobes where the visual P300 is recorded commonly. These channels also correspond to the brain’s parieto-central and occipital cortex channels (Blanco Díaz and Ruiz Olaya, 2020). Considering that the experimental design consisted of 10 trials for each letter of the speller, for this study 10 trials were averaged per channel for each subject as a noise reduction technique.

The signals were decimated by a factor of 10 to reduce the noise in the signal. As a result, the outputs correspond to the target signals

Table 1
Information about the groups formed. $N$ is the sample size. The questionnaire average score corresponds to the mean score per group for the factor of interest (i.e., concentration, eye fatigue, or coffee consumption). Note that for the concentration and eye fatigue factors only two groups were formed.

| Factor           | Group 1   | Group 2   | Group 3 | Group 4 |
|------------------|-----------|-----------|---------|---------|
| $N$              | 25        | 25        | x       | x       |
| average age      | 24.12 ± 12.29 | 24.24 ± 12.37 | x       | x       |
| Sex (male/female)| 14/11     | 14/11     | x       | x       |
| BCI Experience (yes/no) | 9/16      | 14/11     | x       | x       |
| Questionnaire average score | 1.88 ± 0.97 | 3.16 ± 1.61 | x       | x       |
| $N$              | 14        | 14        | x       | x       |
| average age      | 24.5 ± 10.91 | 22.93 ± 10.22 | x       | x       |
| Sex (male/female)| 8/6       | 6/8       | x       | x       |
| BCI Experience (yes/no) | 5/9       | 8/6       | x       | x       |
| Questionnaire average score | 1.93 ± 0.86 | 4.14 ± 1.84 | x       | x       |
| $N$              | 25        | 10        | 7       | 8       |
| average age      | 23.9 ± 12.18 | 24.8 ± 9.80 | 26 ± 7.67 | 29 ± 3.95 |
| Sex (male/female)| 13/12     | 6/4       | 4/3     | 5/3     |
| BCI Experience (yes/no) | 8/17      | 9/1       | 4/3     | 7/1     |
| Questionnaire average score | 0 ± 0.39 | 3.2 ± 0.96 | 7.21 ± 0.68 | |
Logistic regression is a process of modeling the probability of a discrete outcome given for the input variables, which allows the classification of binary outcomes. The correlation coefficients obtained from CCA were classified by the Regularized Logistic Regression (RLR) (Blanco Díaz and Ruiz Olaya, 2020).

### 2.3. Identification of P300 signal

#### 2.3.1. Mean amplitude with linear discriminant analysis (MA-LDA)

The P300 consists of a positive deviation at approximately 300 ms after presenting the stimuli. For this reason, the mean value of the target EEG signals is greater than the mean value of the non-target EEG signals (Hwang et al., 2017; Lee et al., 2019). This study implemented the standard deviation as a feature to improve the classification criteria. In BCI systems, the most implemented classifiers are discriminant classifiers, especially Linear Discriminant Analysis (LDA) (Hwang et al., 2017; Blanco Díaz and Ruiz Olaya, 2020). LDA is based on using hyperplanes to separate the training feature vectors representing two different classes (i.e., P300 and non-P300). The location and orientation of this hyperplane are determined from training data. In this study, a LDA classifier was trained using the training session’s information. Subsequently, the classifier performance was evaluated with the information obtained from the online testing session.

#### 2.3.2. Canonical correlation analysis with regularized logistic regression (CCA-RLR)

Regularized Logistic Regression (CCA-RLR) The Canonical Correlation Analysis (CCA) is a statistical technique that calculates the correlation between two multivariate sets; thus, the canonical correlation coefficient measures the magnitude of association between two canonical variables (Blanco Díaz and Ruiz Olaya, 2020; Lin et al., 2007). The canonical correlation coefficient is defined by:

$$\rho = \max_{w, w_y} \frac{w^T S_{xy} w_y}{\sqrt{w^T S_{xx} w x^T S_{yy} w_y}}.$$  \hspace{1cm} (1)

where $x$ and $y$ correspond to two multidimensional variables, $S$ corresponds to the covariance matrices and $w$ corresponds to the weights that maximize the correlation $\rho$ between $x$ and $y$.

For this study, $x$ is the vector corresponding to the EEG signal (averaged after data analysis), and $y$ corresponds to the Grand Averages (GA) of the training signals for target and non-target (16 reference signals in total, 2 per channel).

The CCA was performed over each channel, generating 16 $\rho$ values per trial, which were used as input features to the classifier. Logistic regression was used as a classifier.

### 2.4. Performance metrics

According to the confusion matrices, the most commonly used metrics reported in the literature for P300 classification are: Sensitivity (Sen), Specificity (Spe), Accuracy (Acc), and the Area Under the Curve (AUC) of the Received Operation Curve (ROC) (Hwang et al., 2017; Won et al., 2019; Colwell et al., 2014; Dmitriev et al., 2018; Xiao et al., 2019; Blanco Díaz and Ruiz Olaya, 2020; Philip and George, 2020). The following equations define these metrics:

- Sensitivity (Sen): $\text{Sen} = \frac{TP}{TP + FN}$ \hspace{1cm} (2)
- Specificity (Spe): $\text{Spe} = \frac{TN}{TN + FP}$ \hspace{1cm} (3)
- Accuracy (Acc): $\text{Acc} = \frac{TP + TN}{TP + TN + FP + FN}$ \hspace{1cm} (4)
- AUC: $\text{AUC} = \int TPR \, d(FPR)$ \hspace{1cm} (5)

where $TP$, $TN$, $FP$ and $FN$ correspond to the true positives, true negatives, false positives and false negatives, respectively. $TPR$ and $FPR$ are defined by:

- $TPR = \frac{TP}{TP + FN}$
- $FPR = \frac{FP}{FP + TN}$

### 2.5. Statistical analysis

An unpaired t-test was performed to compare the means and standard deviations of the BCI performance and to identify significant differences between the conditions of each physical or mental factor according to the four evaluation metrics (Sen, Spe, Acc, and AUC). For this, a Kolmogorov–Smirnov analysis was initially implemented to verify whether the data have a normal distribution. Then, the unpaired t-test was implemented with a $p$-value of 0.05, where the alternative hypotheses were that the best algorithmic performance is obtained with a good state of eye fatigue, high concentration, and coffee consumption. However, the null hypothesis was the opposite. Additionally, Spearman’s correlation was obtained between the four evaluation metrics and the scores obtained from the questionnaires for the indices of ocular state, concentration, and coffee consumption (ordinal vector) (Sprague et al., 2016; Hammer et al., 2018). The latter test was also performed with a significance of 0.05. This value allows searching for linear relationships between the psychological factors and the performance of the BCI system (Hammer et al., 2018; Sprague et al., 2016).
Table 2

| Factor            | Metric | MA-LDA | CCA-RLR | MA+CCA-RLR |
|-------------------|--------|--------|---------|------------|
|                   | Sen    | Spe    | Acc     | AUC        | Sen    | Spe    | Acc     | AUC        | Sen    | Spe    | Acc     | AUC        |
| Concentration     | 0.12   | 0.12   | 0.13    | 0.15       | -0.20  | -0.38*  | -0.34*  | -0.28      | -0.24*  | -0.38*  | -0.37*  | -0.34*     |
| Eye Fatigue       | 0.17   | 0.16   | 0.17    | 0.09       | 0.24   | 0.36*   | 0.35*   | 0.33*      | 0.13    | 0.18    | 0.17    | 0.19       |
| Coffee consumption| -0.02  | -0.16  | -0.17   | -0.09      | 0.24   | 0.36*   | 0.35*   | 0.33*      | 0.13    | 0.18    | 0.17    | 0.19       |

* is significant $p < 0.05$.

3. Results

The Grand Average (GA) is the reference EEG signal for the CCA method, obtained from each channel's grand averaging of the target and non-target signals. The reference signals are calculated for a test subject and shown in Fig. 4. The results of metrics of P300 identification, calculated with the Eqs. (2)–(5), using the MA-LDA, CCA-RLR, and MA+CCA-RLR methods for all subjects, are represented in Fig. 5, using violin plots.

The results of the Spearman correlation analysis between analyzed factors and the performance metrics are presented in Table 2. In addition, correlations of all methods are presented in Supplementary Material Fig. 1 for the metric results considering the Concentration level factor. Eye state in Supplementary Material Fig. 2, and Effect of Coffee Consumption in Supplementary Material Fig. 3.

According to Fig. 1 and Table 1 the groups were segmented into groups by each factor, considering the concentration level, eye fatigue, and coffee consumption. The results of these metrics for the concentration level factor are shown in Fig. 6. The metrics for the eye fatigue factor are shown in Fig. 7. Finally, the results for the metrics for coffee consumption are shown in Fig. 8.

A high standard distribution can be observed in Figs. 5, 6, 7, and 8. This behavior indicates that in some cases, the last inter quartile range can be below 0 and above 1. However, this represents that the data have a low probability of having strictly lower or higher performance value for a BCI system. In contrast, the probability density is centered in the Inter quartile range where the average and median are located, which allows interpreting that the data have a high probability of having a normal distribution. These results support the Kolmogorov–Smirnov analysis, where it was obtained that the data have a high probability of having a normal distribution with a $p$-value of 0.05.

The unpaired t-test statistical analysis yielded significant results ($p < 0.05$) or all performance metrics between group 1 and group 2 concerning the concentration factor, using the P300 CCA-RLR decoding method and the hybrid method. Thus, the alternative hypothesis that the low concentration group had a worse performance than the high concentration group scores is accepted. Additionally, this statistical test showed significant results for the CCA method between groups 1, 3, and 4, considering the coffee consumption factor for the Spe, Acc, and AUC metrics. The other results show an average difference in performance between the groups. However, this difference is not statistically significant for a value of $p = 0.05$. 
Fig. 6. Comparison of metrics for group 1 and group 2 of concentration level factor with the three P300 identification methods: (A) Sen (B), Spe (C), Acc (D) AUC. Purple represents the distributions of group 1; Pink represents the distributions of group 2. The black line represents the mean. The red line represents the median. * represents a statistically significant difference according to the unpaired t-test ($p < 0.05$). ** represents significant Spearman’s correlation coefficient ($p < 0.05$).

Fig. 7. Comparison of metrics for group 1 and group 2 of eye fatigue factor with the three P300 identification methods: (A) Sen (B), Spe (C), Acc (D) AUC. Purple represents the distributions of group 1; Pink represents the distributions of group 2. The black line represents the mean. The red line represents the median. * represents a statistically significant difference according to the unpaired t-test ($p < 0.05$). ** represents significant Spearman’s correlation coefficient ($p < 0.05$).

Finally, the analysis based on the Spearman correlation coefficients obtained in Table 2, represents statistical significance for the variables Sen, Spe, Acc, and AUC, in the three factors analyzed using the CCA RLR method and the hybrid method, obtaining a maximum negative correlation of $-0.38$. Additionally, this correlation is negative for the factors of concentration and visual state, i.e., as soon as the subjects perceived a more significant state of ocular fatigue and low concentration, the performance was moderately lower. On the other hand, in most cases, the obtained correlation was positive for coffee consumption. Although the absolute value of the correlation coefficients
is low, these results support works reported in the literature related to BCI systems and psychological and cognitive factors, which will be compared below (Hammer et al., 2012, 2018). Additionally, when implementing Spearman's correlation to evaluate the existing correlations between the three different factors, it was possible to observe a correlation between concentration levels and eye fatigue of 0.5644. On the other hand, the other two factors obtained correlation values close to zero, which leads to the conclusion that there is no correlation between coffee consumption and eye fatigue or concentration levels.

4. Discussion

4.1. General discussion

Regarding the neurophysiological behavior of the P300 signal and its latency and amplitude alterations, this is probably a consequence of differences in the physiological and cognitive factors of the test subject during the execution of the experimental design sessions with BCI. The literature has reported that training the extraction and classification algorithms in subjects under the factors mentioned above affects the decoding of the patient's intention. Thus, a variation in the system’s performance is produced (Hammer et al., 2012; Pitt and Brumberg, 2021). Therefore, this study aimed to analyze these differences through different metrics that quantify the ability of the classifiers to identify the ERP for three main factors: level of attention during the experiment, eye fatigue during the experimental design, and coffee consumption in the hours before the experiment. In addition, this study used recognition methods based on the signal’s shape.

Before making comparisons between factors, it is essential to validate the results of the implemented methods. Consequently, from the results, it is possible to observe that the method that implements hybrid characteristics of the signal mean amplitude (Hwang et al., 2017; Lee et al., 2019) combined with the CCA (Blanco Díaz and Ruiz Olaya, 2020) is the one that obtains the best average performance for all subjects. Additionally, that method exhibits an average maximum performance of 84.17%, Spe of 85.99%, Acc of 85.68%, and AUC of 85.08%.

Concerning the concentration level factor, it is possible to observe in Fig. 6 that the BCI system performance metrics for group 1 (high concentration) show an improvement regarding group 2 (low concentration). According to Sen, group 1 reflects an average percentage improvement of 6.11%, Spe of 5.49%, Acc of 5.59%, and AUC of 5.8%, with statistically significant differences for the unpaired t-test ($p < 0.05$) by using the methods CCA-RLR and MA+CCA-RLR. These results support that concentration tasks directly influence the P300 signal evoked through the oddball paradigm. For that reason, it is affected by the subject’s level of attention. In other words, the user’s concentration level factor can improve the performance of a real-time BCI system based on event-related potentials (Carlson, 2021; Sprague et al., 2016).

Regarding the eye fatigue factor, through Fig. 7 it is possible to observe a positive difference between group 1 (low eye fatigue) and group 2 (high eye fatigue) metrics. An average Sen was obtained, increases by an amount of 6.35%, a Spe increases by 3.81%, Acc increases by 4.23%, and AUC increases by 5.08%, although it is not significant for the unpaired t-test. Different types of visual stimulation have been shown to affect the P300 signal (Guo et al., 2019). This response implies a similar behavior to other visual systems. Studies with Steady-State Evoked Potentials (SSVEP) indicate that ocular fatigue generates a percentage decrease in variables such as Acc in BCI systems (Bieger et al., 2010; Guerrero-Mendez et al., 2021). The stimulus method based on the oddball paradigm differs from the one used for SSVEP. Despite that fact, the results of this study are related to those reported in the literature. Hence, it is possible to observe that the group that did not perceive visual fatigue during the experimental development obtained better Sen, Spe, Acc, and AUC percentages than the group that perceived visual fatigue (Collura, 2002; Guerrero-Mendez et al., 2021).

With coffee consumption, it is possible to observe in Fig. 8 that the best performance for all metrics was obtained by group 4 (the group...
that consumed coffee in a lapse of more than 4 h), with an average Sen of 88.10%, Spe of 88.41%, Acc of 88.36% and AUC of 88.25% using the MA+CCARLR method. Furthermore, this outstanding performance was statistically significant versus groups 1 (did not consume coffee) and 3 (consumed coffee between 2 and 4 h) using the CCA-RLR method with the Spe, Acc, and AUC evaluation metrics ($p < 0.05$) according to the unpaired t-test.

4.2. Related works

Violin plots are frequently used to visualize numerical data's full distribution and compare distributions between multiple groups (Hintze and Nelson, 1998). In this study, Fig. 5 shows that the P300 extraction methods for all subjects have average Sen values of 73.67%, 79.56%, and 84%, average Spe values of 73.54%, 84.04%, and 85.99%, and average Acc values of 73.56%, 83.30% and 85.69%, and average AUC of 73.61%, 81.80%, and 85.08%, respectively. These preliminary results are in good agreement with those reported in the literature. Firstly, Hwang et al. reported Acc values close to 75% with four subjects using multiple trials (Hwang et al., 2017). Then, Won et al. reported AUC values of approximately 84% using multiple trials and preprocessing techniques such as Independent Component Analysis (ICA) (Won et al., 2019). Finally, Collwell reported AUC values of 75% and 86% using a derivation of LDA and multiple trials with 18 subjects (Collwell et al., 2014).

It is worth mentioning that studying the influence of psychological and cognitive factors that affect the performance of BCI systems is of great interest to the scientific community. This interest is because it allows the recognition of factors that can reduce performance in experimental sessions and is helpful in neuroscience for exploring brain behavior. Hammer et al. have been part of this challenge since they performed correlation studies between psychological factors and motivation in the performance of a BCI system based on motor imagery (Hammer et al., 2012). However, this study reported no significant correlations around 0.5 or significant correlations for specific tasks involving concentrating on neuromotor movements. These results demonstrate that the psychological predictors studied had a non-direct influence on the performance of BCI systems. Subsequently, Hammer et al. attempted in Hammer et al. (2018) to replicate a similar study with a BCI P300 system. That study reported negative Spearman correlations of approximately -0.41 between the performance of BCI systems with the psychological factor of emotional stability. That study concluded that people with high emotional stability scores had worse performance than those with low emotional stability. Furthermore, Sprague et al., in a study conducted in Sprague et al. (2016) demonstrated that executive function, general intelligence, and working memory might cause individual differences in P300-based BCI performance. Finally, it should be highlighted that a correlation of near 0.56 was found between the physiological factors of concentration levels and eye fatigue; however, the correlation value between these two factors and coffee consumption are close to zero. Therefore, this behavior is possibly linked to coffee consumption being considered an external stimulant (Chen et al., 2020).

4.3. Final discussion

The results obtained in this study support the studies reported in the literature where neurophysiological factors affect the user's intention decoding algorithms, where a better performance for all metrics with statistically significant differences between the group that reported high concentration versus low concentration is noteworthy. Additionally, significant Spearman correlations were obtained for these metrics, which despite not being high, demonstrate a negative behavior that may correlate with an indirect influence between the concentration factor and Speller BCI performance (Sprague et al., 2016; Hammer et al., 2018). Similar occurred with the visual fatigue factor, where the unpaired t-test showed no significant differences between the two groups; however, a worse percentage performance can be observed among the group that showed greater visual fatigue. These metrics obtained significant Spearman correlations for the Spe and AUC metrics using the decoding methods influenced by the CCA. In light of the results presented in Figs. 6 and 8, it can be stated that the CCA classifier may be more sensitive to changes in the responses recorded in the evoked potentials concerning variables such as amplitude and latency. This behavior is possibly induced by physiological and cognitive differences derived from the stimuli presented to the subjects, thus altering the shape of the P300 signal and leading to a decrease in the performance of this classifier. These metrics also obtained negative significant correlation coefficients for the two metrics mentioned above. These results suggest that visual fatigue is indirectly related to a decrease in BCI system performance, especially in the differentiation between the negative variable (non-P300) and the positive variable (P300). On the other hand, for the coffee consumption factor, the best performance in the Spe, Acc, and AUC metrics was obtained by the subjects who consumed coffee for a period of approximately 4 h compared to those who did not consume coffee or consumed it in a period between 2 to 4 h. Additionally, these metrics also obtained significant positive correlations. These times of 4 and 8 h align with the research conducted by Chen et al. (2020). The caffeine/placebo administration times were vital in demonstrating variations in the latencies and amplitudes of the ERPs generated. This fact allows concluding that coffee consumption could benefit cognitive brain functions.

The values of the correlations obtained in this study are close to those reported in the literature. This fact may be due to, as shown in Table 1, the population used in the study being more homogeneous than the general population, which is related to a reduction in the variance of the sample. The results of the statistical tests allow observing a direct influence on the concentration factor and coffee consumption in specific cases. The results of this study leave an open door for further study of factors such as concentration, visual fatigue, and the consumption of substances such as coffee in more controlled environments since, as observed, they can quantitatively affect the performance of BCI speller systems.

5. Conclusion

Physical and cognitive parameters such as concentration, visual fatigue, and coffee consumption play an important role in P300 speller-based BCI performance. Identifying these factors would allow for an improved capability of BCI systems in detecting P300 signals, leading to better performance and, therefore, may aid in decoding the individual’s intent. Additionally, this study had some limitations, especially related to the experimental conditions used in the public dataset and the limited number of subjects involved in the experiment.

Future studies will focus on studying more physical and mental factors that may affect the neurophysiological behavior of the P300 signal to maximize the performance of a BCI system based on sensing processes, such as those related to anxiety, stress, or other substance consumption.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.
Acknowledgments

The authors would like to thank Antonio Nariño University (UAN/Colombia) under the project number 2021020 “Model based on multimodal EEG-EMG information to improve motion intention decoding for the control of a BCI system” and the Federal University of Espírito Santo (UFES/Brazil) and FAPES/I2CA (Resolution N° 285/2021) by the MSc scholarships awarded to the first two authors.

Funding

This work was supported by Antonio Nariño University (UAN/Colombia) under the project number 2021020 “Model based on multimodal EEG-EMG information to improve motion intention decoding for the control of a BCI system”, Federal University of Espírito Santo (UFES/Brazil) and FAPES/I2CA (Resolution N° 285/2021) by the MSc. scholarships awarded to the first two authors.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jneumeth.2020.109722.

References

Abdulkader, S., Atia, A., Mostafa, S., Mostafa, M., 2015. Brain computer interfacing: Applications and challenges. Egypt. Inform. J. 16, 213–230. http://dx.doi.org/10.1016/j.egyij.2015.06.002.

Biegler, G., Molina, G.G., Zha, D., 2010. Effects of stimulation properties in steady-state visual evoked potential based brain-computer interfaces. In: Proceedings of 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society. http://dx.doi.org/10.1109/EMBS.2010.5596950.

Blanco Díaz, C.F., Ruiz Olaya, A.F., 2020. A novel method based on regularized logistic regression and CCA for P300 detection using a reduced number of EEG trials. IEEE Lat. Am. Trans. 18 (12), 2147–2154. http://dx.doi.org/10.1109/TLA.2020.9400443.

Carlson, J., 2021. A systematic review of event-related potentials as outcome measures of attention bias modification. Psychophysiology 58, http://dx.doi.org/10.1111/psyp.13801.

Chaudary, U., Birbaumer, N., Ramos, A., 2016. Brain-computer interfaces for communication and rehabilitation. Nat Rev. Neurol. 12, 513–525. http://dx.doi.org/10.1038/nrneurol.2016.113.

Chen, X., Zhang, L., An, G., Yang, D., Li, C., Shao, Y., Wang, J., Fan, R., Ma, Q., 2020. Effects of caffeine on event-related potentials and neuropsychological indices after sleep deprivation. Front. Behav. Neurosci. 14, http://dx.doi.org/10.3389/fnbeh.2020.00108.

Collura, T.F., 2002. Application of repetitive visual stimulation to EEG neurofeedback protocols. J. Neurother. 6 (2), 47–70. http://dx.doi.org/10.1081/JNN-114006602.

Colwell, K.A., Ryan, D.B., Throckmorton, C.S., Sellers, E.W., Collins, L.M., 2014. Channel selection methods for the P300 speller. J. Neurosci. Methods 232, 6–15. http://dx.doi.org/10.1016/j.jneumeth.2014.04.009.

Deslandes, A., Alves, H., Cagy, M., Piedade, R., Pompeu, F., Ribeiro, P., 2004. Effects of caffeine on visual evoked potential (P300) and neuromotor performance. Arq. Neuropsiquiatr. 62, http://dx.doi.org/10.1590/S0004-282X2004000300002.

Dmitriev, A., Al-barodsh, M., Igór, S., Nikolaev, A., 2018. The optimal stimulation mode and the number of averaging epochs selection for P300 detection. In: 2018 1st Symposium on Biomedical Engineering, Radioelectronics and Information Technology. USBEREIT, pp. 91–94. http://dx.doi.org/10.1109/USBEREIT.2018.8384558.

Guo, M., Jin, J., Yao, W., Xu, C., 2019. Investigation of visual stimulus with various colors and the layout for the oddball paradigm in evoked related potential-based brain-computer interface. Front. Comput. Neurosci. 13, http://dx.doi.org/10.3389/fncom.2019.00624.

Hammer, E.M., Halder, S., Blankertz, B., Sannelli, C., Dickhaus, T., Klieh, S.C., Müller, K.-R., Kübler, A., 2012. Psychological predictors of SMR-BCI performance. Biol. Psychol. 89, 80–86. http://dx.doi.org/10.1016/j.biopsycho.2011.09.006.

Hammer, E.M., Halder, S., Klieh, S.C., Kübler, A., 2018. Psychological predictors of visual and auditory P300 brain-computer interface performance. Front. Neurosci. 12, http://dx.doi.org/10.3389/fnins.2018.00307.

Hinze, J.L., Nelson, R.D., 1998. Violin plots: A box plot-density trace synergism. Amer. Statist. 52 (2), 181–184. http://dx.doi.org/10.1080/00031305.1998.10480559.

Hwang, J., Lee, M., Lee, S., 2017. A brain-computer interface speller using peripheral stimulus-based SSEVP and P300. In: The 5th International Winter Conference on Brain-Computer Interface, BCI, Sabuk, South Korea, http://dx.doi.org/10.1109/IIW-BCI.2017.7858164.

Karimi, S., Mirzakuchaki, S., 2019. Comparison of the P300 detection accuracy related to the BCI speller and image recognition scenario. http://dx.doi.org/10.1109/斠iscience.gpit0902.

Lin, Z., Zhang, C., Wu, W., Gao, X., 2007. Frequency recognition based on canonical correlation analysis for SSEVP-based BCIs. IEEE Trans. Biomed. Eng. 54 (6), 1172–1176. http://dx.doi.org/10.1109/TBME.2006.889197.

McFarland, D.J., Wolpaw, J.R., 2011. Brain-computer interfaces for communication and control. Commun. ACM 5, 60–66. http://dx.doi.org/10.1145/1941487.1941506.

Patel, V., Patel, M., 2019. Brain computer interface: Applications and P300 overview. In: The 10th ICCCNT International Conference on Computing, Communication and Networking Technologies. Kanpur, India, http://dx.doi.org/10.1109/ICCTN74560.2019.8944461.

Perez Vidal, A.F., Oliver Salazar, M.A., Salas Lopez, G., 2016. Development of a brain-computer interface based on visual stimuli for the movement of a robot joints. IEEE Lat. Am. Trans. 14 (2), 477–484. http://dx.doi.org/10.1109/TLA.2016.7437182.

Philip, J.T., George, S.T., 2020. Visual P300 mind-speller brain-computer interfaces: A walk through the recent developments with special focus on classification algorithms. Clin. EEG Neurosci. 51 (1), 19–33. http://dx.doi.org/10.1177/1550059419842753.

Pitton, T., 1992. The P300 wave of the human event-related potential. J. Clin. Neurophysiol. 9, 456–479. http://dx.doi.org/10.1097/00004691-199210000-00002.

Pitt, K.M., Brumberg, J.S., 2021. Evaluating person-centered factors associated with brain–computer interface access to a commercial augmentative and alternative communication paradigm. Assist. Technol. 1–10. http://dx.doi.org/10.1002/10400435.2021.1872737, PMID: 33667154.

Rezeika, A., Benda, M., Statwicki, P., Ermens, E., Boaboo, A., Volosyak, I., 2018. Brain-computer interface spellers: A review. Brain Sci. 8, 1–38. http://dx.doi.org/10.3390/brainsci800002.

Sprague, S.A., McIe, M.T., Sellers, E.W., 2016. The effects of working memory on brain-computer interface performance. Clin. Neurophysiol. 127 (2), 1331–1341. http://dx.doi.org/10.1016/j.clinph.2015.10.038, URL: https://www.sciencedirect.com/science/article/pii/S1382815015010160.

Vazquez-Marrufo, M., 2017. Event-related potentials for the study of cognition. In: Event-Related Potentials and Evoked Potentials. IntechOpen, http://dx.doi.org/10.5772/intechopen.69308.

Wolpaw, J., Wolpaw, E., 2012. Brain-Computer Interfaces: Principles and Practice. Oxford University Press, New York (NY).

Wong, K., Kwon, M., Jung, S., Ahn, M., Jun, S.C., 2019. P300 speller performance predictor based on RSVP multi-feature. Front. Hum. Neurosci. 13 (261), 1–14. http://dx.doi.org/10.3389/fnhum.2019.00261.

Xiao, X., Xu, M., Jin, J., Wang, Y., Jung, T., Ming, D., 2019. Discriminative canonical pattern matching for single-trial classification of ERP components. IEEE Trans. Biomed. Eng. 1. http://dx.doi.org/10.1109/TBME.2019.2958641.