Different effects of using pictures as stimuli in a P300 brain-computer interface under rapid serial visual presentation or row-column paradigm

Álvaro Fernández-Rodríguez[1], María Teresa Medina-Juliá, Francisco Velasco-Álvarez and Ricardo Ron-Angevin

Departamento de Tecnología Electrónica, Universidad de Málaga, Malaga, Spain.

35 Louis Pasteur Boulevard, 29071, Malaga, Spain

[1] Corresponding author. E-mail: afernandezrguez@uma.es.

Abstract

Previous proposals for controlling a P300-based BCI speller have shown an improvement using alternative images instead of letters as target stimuli under a row–column paradigm (RCP). However, the RCP is not suitable for those patients with a lack of gaze control. To solve that, the rapid serial visual presentation (RSVP) paradigm has been proposed in previous studies. The aim of the present work is to assess if a set of alternative pictures that improved performance in RCP could also improve performance in RSVP. Sixteen participants controlled four conditions in calibration and online tasks: letters in RCP, pictures in RCP, letters in RSVP and pictures in RSVP. The effect given by pictures was greater under RCP than under RSVP, both for performance and event-related potential analyses. Indeed, pictures did not show any improvement under RSVP in comparison to letters. In addition, the condition with pictures under RCP was declared the favourite by most users (68.75%), while the condition with pictures under RSVP was not chosen as favourite by any participant. Therefore, this work shows that the improvement related to the use of pictures as alternative flashing stimuli under RCP may not be transferred to RSVP.

Keywords: brain-computer interface (BCI), event-related potential (ERP), speller, paradigm, stimulus.

1. Introduction

The main aim of assistive technology is to improve the quality of life of its potential users. These users are people affected by certain injuries or diseases that cause a deterioration of their motor and/or cognitive abilities. An example of these diseases is amyotrophic lateral sclerosis (ALS), which causes a deterioration of motor skills [1]. In some evolved cases, these patients reach the so-called complete locked-in state (CLIS), in which even the ability to control their gaze is lost [2, 3]. However, even in these severe cases, the cognitive capacity of patients is not necessarily affected [4].

As reported by [5], for most ALS patients, 'the ability to communicate was the most important factor in deciding whether to use (or continue) assisted ventilation. Brain–computer interfaces (BCI) are a type of technology that uses the user's brain activity to establish a communication channel between the user's brain and the device he/she wishes to control [6, 7]. These devices can be, for example, a wheelchair, a domotic system or a virtual keyboard (e.g., [8–10], respectively). As a BCI does not require any motor skills to be handled, it can be an appropriate option for those patients who cannot perform any movements but whose cognitive abilities are kept. Due to its relatively low cost and adequate temporal resolution, the registration method most widely used for recording the brain signal is electroencephalography (EEG).

One of the main applications developed with this technology are virtual keyboards, also known as spellers: they are sets of symbols that can be chosen by the user for communication purposes (see [11] for an extended review of BCI spellers). The most commonly used brain signal for this application is P300 brain potential. P300 is a positive deflection in the voltage of the EEG signal, generally registered from the parietal lobe of the cortex, around 300 ms after the presentation of an uncommon target stimulus [12]. Thus, the user only has to maintain his/her attention over the target stimulus, evoking a P300 signal each time the stimulus appears. Then, the system recognizes the P300 potential and selects the symbol related to that potential; finally, the system executes the command of that stimulus (e.g., type a letter, delete the last character or autocomplete a word).

The presentation paradigm most used by researchers has been the row–column paradigm (RCP), proposed by [8]. In this paradigm, the stimuli are displayed in a matrix of symbols and flashed jointly by rows and columns (i.e., not individually). Due to the distance between stimuli, the RCP requires that the user must have some eye control; otherwise, the performance is severely affected (e.g., [13, 14]). In addition to the RCP presentation mode, there are other alternatives to present stimuli that do not require ocular mobility. The most commonly used paradigm that does not require ocular mobility is rapid serial visual presentation (RSVP) (e.g., [15–18]).

Due to the difficulty of controlling gaze in some patients, RSVP could be more convenient than RCP for that population. However, most research has been focused on gaze-controlled interfaces (according to [11], 84.4% of papers on P300 BCI spellers have required ocular mobility). The research into P300-based BCI spellers under RCP has been extensive and

productive, as many studies have been carried out regarding the timing, stimulus type, variations in the presentation paradigm or word prediction (e.g., [19–22]). However, it should be questioned if the findings made in RCP can be transferred to RSVP with the stimuli presented in the same central position. For example, in RSVP, the stimuli can compete with each other as they all appear in the same position of the screen; so an increase in the P300 amplitude related to the target stimulus would not be enough to improve the performance if the amplitude is also increased for the non-target stimuli. Otherwise, in RCP, improvement in the target stimuli could increase the P300 amplitude, as deflection of the EEG caused by the non-target stimuli can only occur due to the peripheral vision. Therefore, those stimuli that increase the amplitude of the P300 signal could produce a clear improvement in RCP performance, but not in that of RSVP.

To this day, the stimuli with the best results in RCP are familiar faces (e.g., [20, 23]). However, other recent works have revealed the possibility that a set of varied different pictures (e.g., photographs of objects, people or places) as flashing stimuli – a different one for each symbol – could also improve performance compared to the classical paradigm of letters [24]. Using these pictures, a previous preliminary study has reported a possible interaction effect of the presentation paradigm (RCP and RSVP) and the stimuli (letters and pictures) employed [25], in which a benefit produced by pictures is observed under RCP, but not under RSVP. In comparison to this preliminary work, the present study carried out the following main improvements: a larger sample size (N = 6 vs N = 16), deeper analysis of the results, a higher number of selectable commands (from 6 to 9), picture conditions incorporating the letter they stood for, and inclusion of an online writing task. All these improvements are detailed in the Method section.

In short, the aim of the present study is to test two stimulus sets (letters and pictures) under two different presentation paradigms (RSVP and RCP) in order to assess the differences between conditions.

2. Method

2.1. Participants

The experiment was carried out by 16 participants (aged 22.33 ± 3.98 years, 5 males) who declared normal or corrected-to-normal vision. Five participants had previous experience controlling BCI systems and the rest did not. The study was approved by the Ethics Committee of the University of Malaga and met the ethical standards of the Helsinki Declaration. According to self-reports, none of the participants had any history of neurological or psychiatric illness or were taking any medication regularly. Participants received monetary remuneration of €10 and all of them provided written consent.

2.2. Data acquisition and signal processing

The EEG was recorded at a sample rate of 250 Hz using the electrode positions: Fz, Cz, Pz, Oz, P3, P4, PO7 and PO8, according to the 10/20 international system. All channels were referenced to TP8 and grounded to the position AFz. Signals were amplified by an acti-CHamp amplifier (Brain Products GmbH, Munich, Germany) and band-pass filtered at 0.1–9 Hz. Neither online nor offline artefact detection techniques were employed. All aspects of EEG data collection and processing were controlled by the BCI2000 system [26]. A Stepwise Linear Discriminant Analysis (SWLDA) of the EEG data was performed to obtain the weights for the P300 classifier, calculate the accuracy and enable online spelling.

2.3. Spelling conditions

The software used to design the interface was the UMA-BCI Speller [27], which serves as a friendly front end for BCI2000. The present work employed four conditions, based on the combination of two paradigms (RCP and RSVP) and two sets of stimuli (letters, and pictures with a letter placed in the middle). The purpose of including the letter in the figure was to facilitate the user not having to memorize the relationship between image and letter for the task in RSVP, where the letters cannot be differentiated by the spatial location. Also, in an article evaluating maintenance of the letter using faces as stimulation showed improved performance compared to the used of faces without a letter [28]. The only difference between the conditions compared was the presentation paradigm and flashing stimuli employed for each condition. Thus, the four conditions presented were: i) letters in RCP (L-RCP), ii) pictures in RCP (P-RCP), iii) letters in RSVP (L-RSVP), and iv) pictures in RSVP (P-RSVP) (Fig. 1). The conditions were displayed on a 15.6 in (39.6 cm) screen at a distance of 60 cm and a refresh rate of 60 Hz. For all conditions, a stimulus onset asynchrony (SOA) of 288 ms and an inter-stimulus interval (ISI) of 96 ms were used, so each stimulus was presented for 192 ms. A 3968 ms pause was established between each selection. On the one hand, the RCP conditions (3 × 3 matrix size) were initially based on the previously mentioned paradigm of [8]. On the other hand, for RSVP, only the flashing stimuli were presented, placed in the middle of the screen. All the stimuli (both letters and pictures with letter) were designed using the image editor software GIMP (www.gimp.org). For all conditions, the font used for letters was Arial in capital letters and they had a size of around 1.8 cm × 2.3 cm, depending on the letter. Regarding the pictures, only used in P-RSVP and P-RCP, their size was 4.7 cm × 3.5 cm (187 px × 140 px). The distance between adjacent stimulus centres, for letters and pictures with letter, were 6.7 cm for left/right stimuli and 5.5 cm

for top/down stimuli. The letters in picture conditions were located in the centre of the pictures with an opacity level of 50%. The purpose of adding the letter to the picture conditions was to facilitate, in RSVP, the participant being able to recognize which image was associated with each letter. However, as mentioned above, the participant had a pause of 3968 ms in which he/she could look at a printed document placed next to the screen that graphically indicated to which figure each letter corresponded. The participant was instructed to rely on the use of pictures for recognition of the target stimulus.

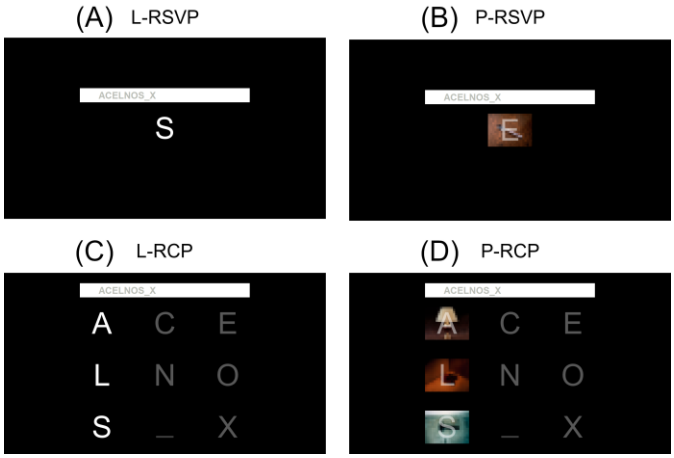

Fig. 1. Conditions tested in the study: letters in row–column presentation (L-RCP), pictures in row–column presentation (P-RCP), letters in rapid serial visual presentation (L-RSVP), and pictures in rapid serial visual presentation (P-RSVP). Due to copyright reasons, the conditions with pictures have been pixelated.

The pictures were obtained from the International Affective Picture System (IAPS) [29]. Those images with the lowest score for arousal level and that filled the aforementioned size (i.e., those pictures that filled all the space and did not have black padding) were chosen. The specific selected pictures according to the IAPS codification for each letter were: 7175 for A, 7010 for C, 7004 for E, 7031 for L, 7020 for N, 7950 for O, 7110 for S, 7080 for _ and 7006 for X. The relation between what is represented in the selected picture and the letter was completely random.

2.4. Procedure

The experiment was carried out in a room where only the participant was present at the time he/she performed the task, in order to concentrate on it without external distractions. It was performed in one session with a duration of approximately 80–100 minutes.

An intrasubject, also called repeated measures, design was used, and so all the users went through all the experimental conditions. Each condition consisted of two exercises: a calibration task to adapt the system to the user, and a spelling task – or online task – in which the user actually controlled the interface. The calibration task was composed of three blocks of nine selections, and the spelling task was composed of one block of at least 15 selections (the users could make mistakes and correct them, so the number of selections could be more than 15). The user passed to the next condition when he/she had finished the four blocks of the previous condition. The conditions were presented trying to perform a complete counterbalanced design – that is, a design in which all the conditions are equally distributed between participants – to prevent any unwanted effect, such as learning or fatigue. The symbol selection order was the major-row order for each block in the calibration task (i.e., A, C, E, L, N, O, S, _ and X). On the other hand, the phrase to be typed in the spelling task was 'CENO_EN_LA_CASA' ('I have dinner at home' in Spanish). There was a short break between blocks (variable at the request of the user). A sequence is defined as the minimum number of flashes required per stimulus to locate a target stimulus. Thus, the number of flashes required for one sequence is not the same for RSVP and RCP. While only one flash per stimulus is needed under RSVP to locate the target stimulus, two flashes are required under RCP to locate it due to the system needing to know the corresponding row and column. In the present study, five sequences were used for target stimulus selection, so each symbol flashed five times in RSVP and 10 times in RCP (5 flashes per row and 5 flashes per column). The participant was asked to mentally count these flashes in order to keep his/her attention on the task. The time needed to complete a sequence was 2784 ms in RSVP and 1632 ms in RCP. Offline analysis of the calibration blocks allowed selection of the number of highlighting sequences for the online spelling. The chosen sequence was the one in which the user obtained his/her best accuracy for the first time.

In addition, after the user completed the online task with each condition, he/she had to respond with scores from 0 to 10 (from very low to very high) to the following questions: i) effort needed to perform the task correctly, and ii) level of fatigue produced by control of the interface. Once the user had tested all the conditions, they had to answer the following items: i) preferred presentation paradigm (RSVP or RCP), ii) preferred stimuli (letters or pictures with letter), and iii) favourite

interface (L-RSVP, P-RSVP, L-RCP or P-RCP).

2.5. Evaluation

For calibration and spelling tasks, two parameters were used to evaluate the effect of the presentation paradigm and stimulus type on performance: i) the accuracy (%) of the system in classifying the selections (i.e., the number of correctly predicted selections divided by the total number of predicted selections); and ii) the information transfer rate (ITR, bits/min) based on the formula presented in [30]. Both accuracy and ITR have been used previously to compare the performance of RSVP versus RCP [31]. In addition, the difference in amplitude between target and non-target stimulus signals (μV) in the calibration task was studied in order to observe how the brain signal is affected by the application of different presentation paradigms and sets of stimuli. For analysis of the brain signal, a time interval of 0–800 ms was studied.

To analyse the paradigm and stimulus effects, several analyses were carried out in order to look for significant differences. While the statistical analysis of performance was carried out using SPSS [32], statistical analyses related to the event-related potential (ERP) waveform were executed using EEGLAB [33]. A Bonferroni correction method was applied for multiple comparisons, even for the ERP waveform analyses.

3. Results

3.1. Accuracy in the calibration task

A three-way repeated measures ANOVA (2 × 2 × 5) including the presentation (RSVP and RCP), stimulus (letters and pictures) and sequence (5 sequences) factors was carried out (Fig. 2). The analysis showed a main effect of presentation (F (1, 14) = 36.644, $p < .001$) and sequence (F (4, 56) = 149.715, $p < .001$). Thus, it could be affirmed that the RCP presentation (93.98 ± 10.09%) offered a significantly higher accuracy than RSVP (83.72 ± 16.79%), and that the accuracy was higher as the number of sequences was increased. In addition, the interaction effects presentation × sequence (F (4, 56) = 15.453, $p <$ .001) and paradigm × stimulus × sequence (F (4, 56) = 10.094, $p = .001$) were found. The interaction effect paradigm × sequence suggests that the differences between RCP and RSVP were reduced as the number of sequences was increased. On the other hand, the interaction effect paradigm × stimulus × sequence suggests that the effect of presentation depends on the stimuli used and the number of sequences. Additional analyses considering the four conditions individually (i.e., L-RSVP, P-RSVP, L-RCP and P-RCP) in each sequence were performed. The results of these last analyses are shown in table 1. Moreover, the average number of sequences needed to obtain the first highest accuracy were: 3.93 ± 1.03 for L-RSVP, 3.86 ± 0.92 for P-RSVP, 3.44 ± 0.96 for L-RCP and 2.38 ± 0.96 for P-RCP.

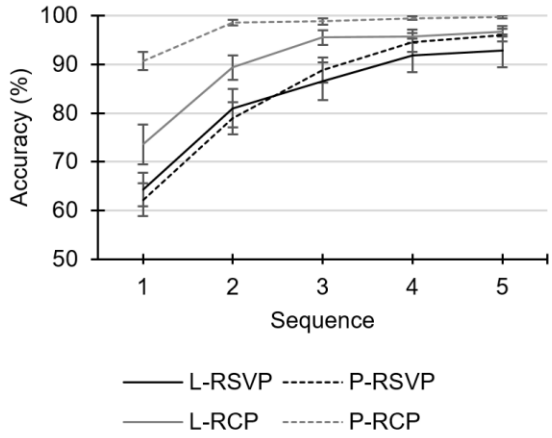

Fig. 2. Accuracy (mean ± standard error) of the different P300-speller conditions as a function of the number of sequences. The conditions tested were: letters under rapid serial visual paradigm (L-RSVP), pictures under rapid serial visual paradigm (P-RSVP), letters under row–column paradigm (L-RCP), and pictures under row–column paradigm (P-RCP).

3.2. ITR in the calibration task

A three-way repeated measures ANOVA (2 × 2 × 5) including the presentation (RSVP and RCP), stimulus (letters and pictures) and sequence (5 sequences) factors was carried out (Fig. 3). The analysis showed a main effect of presentation (F (1, 14) = 579.711, $p < .001$), stimulus (F (1, 14) = 13.678; $p = .002$) and sequence (F (4, 56) = 200.606, $p < .001$). Thus, it could be affirmed that the RCP presentation (42.52 ± 22.85 bits/min) and pictures (33.07 ± 24.21 bits/min) offered a significantly higher ITR than RSVP (17.94 ± 7.85 bits/min) and letters (27.39 ± 16.89 bits/min), respectively. Additionally, the ITR reduced as the number of sequences was increased. In addition, all possible interaction effects were found: presentation × stimulus (F (1, 14) = 10.003, $p = .007$), presentation × sequence (F (4, 56) = 95.208, $p < .001$), stimulus ×

sequence (F (4, 56) = 11.809, p = .002) and presentation × stimulus × sequence (F (4, 56) = 18.918, p < .001). The interaction effect paradigm × stimulus × sequence suggests that the effect of presentation depends on the stimuli used and the number of sequences. As for accuracy, the conditions were considered individually for each sequence; therefore, ANOVA with the condition factor was carried out for each sequence. The results of these last analyses are shown in table 2.

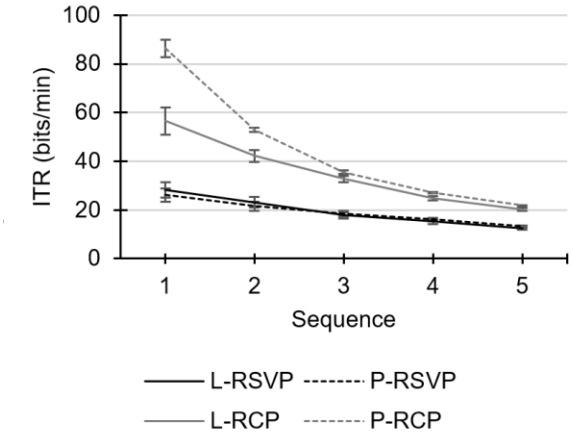

Fig. 3. Information transfer rate (mean ± standard error) of the different P300-speller conditions as a function of the number of sequences. The conditions tested were: letters under rapid serial visual paradigm (L-RSVP), pictures under rapid serial visual paradigm (P-RSVP), letters under row–column paradigm (L-RCP), and pictures under row–column paradigm (P-RCP).

3.3. Event-related potential in the calibration task

shows the grand average ERP waveform for the amplitude difference (μV) between target and non-target stimuli as a function of the stimulus used (letters and pictures) for each presentation type (RCP and RSVP) and EEG channel. In addition, this figure presents a statistical analysis between letters and pictures for each presentation paradigm and channel; the significant time intervals (p < .05) are pointed out on the time axis of each plot. As can be observed, the RCP presentation mode showed more significant differences – between letters and pictures – than the RSVP mode. These differences under RCP were mainly found for those brain components around 220–500 ms (such as N2, P300, N400) for midline (Fz, Cz, Pz and Oz), PO7 and PO8. Otherwise, under the RSVP conditions, the differences between letters and pictures were only found for a short time in Oz (392–420 ms) and PO7 (228–248 ms). Therefore, these results show that the difference between target and non-target stimuli is differentially modulated by the use of different stimuli depending on the presentation paradigm.

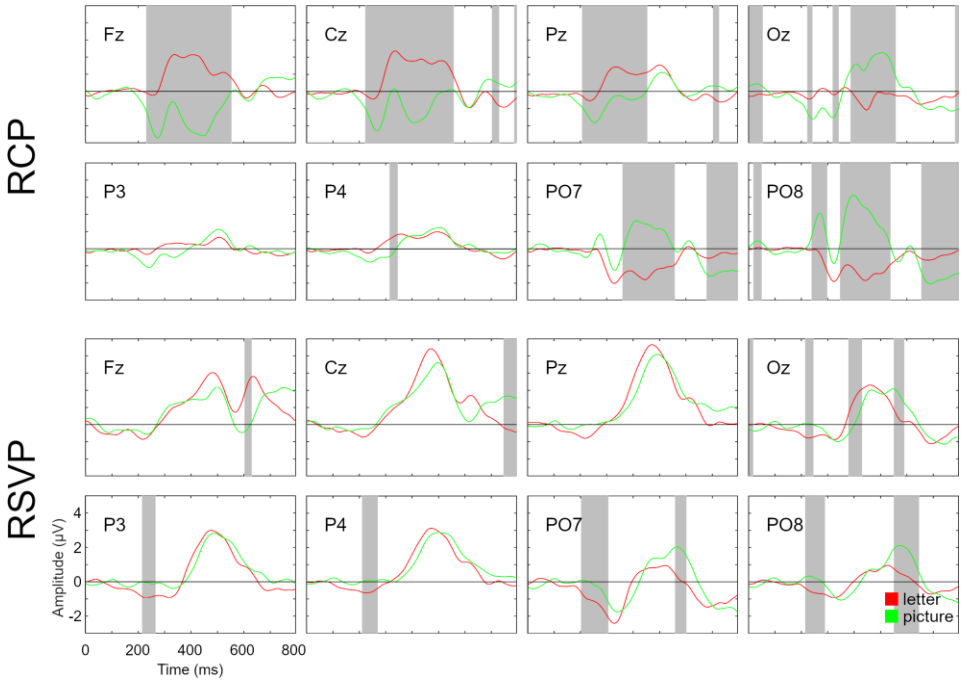

Fig. 4. Grand-average event-related potential (ERP) waveforms of amplitude differences (μV) between target and non-target stimuli for each presentation mode (row–column paradigm (RCP) and rapid serial visual presentation (RSVP)) and type of stimulus (letters and pictures), separated by channel (Fz, Cz, Pz, Oz, P3, P4, PO7 and PO8) and represented in the time interval 0–800 ms. The significant differences (p < .05) in amplitude difference (μV) between type of stimulus are denoted with a bold black line over the time axis.

Table 1. Accuracy (%, mean ± standard deviation) comparisons between the four different spellers for sequences 1–5.

Sequence	Speller			
	(1) L-RSVP	(2) P-RSVP	(3) L-RCP	(4) P-RCP
1st	64.36 ± 13.85 [4]	62.22 ± 13 [4]	73.59 ± 15.77 [4]	90.75 ± 7.52 [1, 2, 3]
2nd	81.02 ± 15.88 [4]	78.95 ± 12.85 [3, 4]	89.39 ± 9.79 [2, 4]	98.61 ± 2.29 [1, 2, 3]
3th	86.57 ± 15.41	88.88 ± 10.07 [4]	95.56 ± 5.97	98.84 ± 2.61 [2]
4th	91.9 ± 13.74	94.57 ± 7.52	95.8 ± 5.21	99.54 ± 1.26
5th	92.83 ± 13.62	96.05 ± 4.94	96.79 ± 4.17	99.77 ± 0.93

Note: letters under rapid serial visual paradigm (L-RSVP), pictures under rapid serial visual paradigm (P-RSVP), letters under row–column paradigm (L-RCP), and pictures under row–column paradigm (P-RCP). Significant differences between spellers ($p < .05$) are denoted with a superindex to show which speller average they are different to (1 for L-RSVP, 2 for P-RSVP, 3 for L-RCP, and 4 for P-RCP). The Bonferroni correction was applied.

Table 2. Information transfer rate (bits/min, mean ± standard deviation) Accuracy (%, mean ± standard deviation) comparisons between the four different spellers for sequences 1–5.

Sequence	Speller			
	(1) L-RSVP	(2) P-RSVP	(3) L-RCP	(4) P-RCP
1st	26.28 ± 11.63 [3, 4]	24.33 ± 9.89 [3, 4]	59.84 ± 23.15 [1, 2, 4]	91.43 ± 15.6 [1, 2, 3]
2nd	21.58 ± 7.7 [3, 4]	20.09 ± 6.73 [3, 4]	44.67 ± 10.26 [1, 2, 4]	56 ± 3.61 [1, 2, 3]
3th	16.55 ± 4.76 [3, 4]	17.3 ± 4.28 [3, 4]	34.61 ± 4.92 [1, 2]	37.65 ± 2.63 [1, 2]
4th	14.25 ± 3.69 [3, 4]	14.96 ± 2.73 [3, 4]	26.14 ± 3.57 [1, 2]	28.73 ± 1.06 [1, 2]
5th	11.67 ± 2.9 [3, 4]	12.31 ± 1.57 [3, 4]	21.39 ± 2.37 [1, 2]	23.14 ± 0.62 [1, 2]

Note: letters under rapid serial visual paradigm (L-RSVP), pictures under rapid serial visual paradigm (P-RSVP), letters under row–column paradigm (L-RCP), and pictures under row–column paradigm (P-RCP). Significant differences between spellers ($p < .05$) are denoted with a superindex to show which speller average they are different to (1 for L-RSVP, 2 for P-RSVP, 3 for L-RCP, and 4 for P-RCP). The Bonferroni correction was applied.

3.4. Accuracy in the online task

A repeated two-way measures ANOVA (2 × 2) including the paradigm (RSVP and RCP) and stimulus (letters and pictures) factors was carried out (Fig. 5). The analysis showed no main or interaction effect for any factor. However, it should be noted that the interaction effect presentation × stimulus was close to offering significant differences (F (1, 14) = 4.295; p = .057).

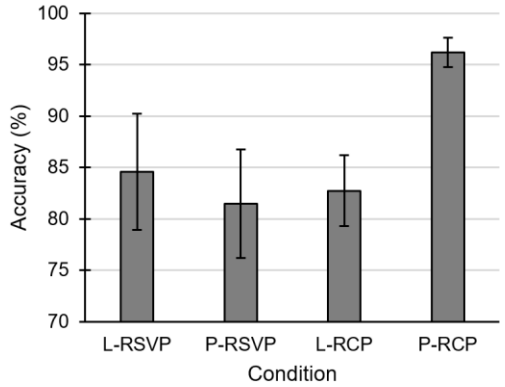

Fig. 5. Accuracy (mean ± standard error) of the different P300-speller conditions: letters under rapid serial visual paradigm (L-RSVP), pictures under rapid serial visual paradigm (P-RSVP), letters under row–column paradigm (L-RCP), and pictures under row–column paradigm (P-RCP).

3.5. ITR in the online task

A repeated two-way measures ANOVA (2 × 2) including the paradigm (RSVP and RCP) and stimulus (letters and pictures) factors was carried out. The analysis showed a main effect of presentation (F (1, 14) = 107.809, p < .001) and stimulus (F (1, 14) = 90.215; p < .001). Thus, it could be affirmed that the RCP presentation (35.08 ± 18.37 bits/min) and pictures (31.55 ± 21.27 bits/min) offered a significantly higher ITR than RSVP (12.58 ± 5.77 bits/min) and letters (16.83 ± 8.81 bits/min), respectively. In addition, the interaction effect presentation × stimulus offered significant results (F (1, 14) = 42.468, p < .001), suggesting that the effect of stimulus depends on the presentation used. Specifically, Fig. 6 shows how the effect of pictures is larger in RCP than in RSVP.

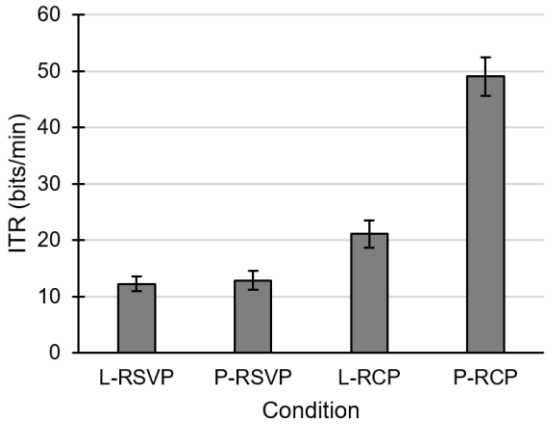

Fig. 6. Accuracy (mean ± standard error) of the different P300-speller conditions: letters under rapid serial visual paradigm (L-RSVP), pictures under rapid serial visual paradigm (P-RSVP), letters under row–column paradigm (L-RCP), and pictures under row–column paradigm (P-RCP).

3.6. Subjective questionnaire

Fig. 7 shows the average scores given by users for the effort required to control each condition and the fatigue resulting from controlling each condition. In reference to the effort, the analysis shows a main effect of presentation (F (1, 15) = 15.244; p = .001), indicating that RCP (4.03 ± 2.56) requires less effort than RSVP (5.59 ± 2.51). Also, the interaction effect presentation × stimulus (F (1, 15) = 5.714, p = .03) was also found. Likewise, regarding fatigue, the analysis showed a main effect of presentation (F (1, 15) = 35.704, p < .001), indicating that RCP (3.31 ± 2.35) causes less fatigue than RSVP (5.34 ± 2.38); however, the interaction effect presentation × stimulus factor was not found.

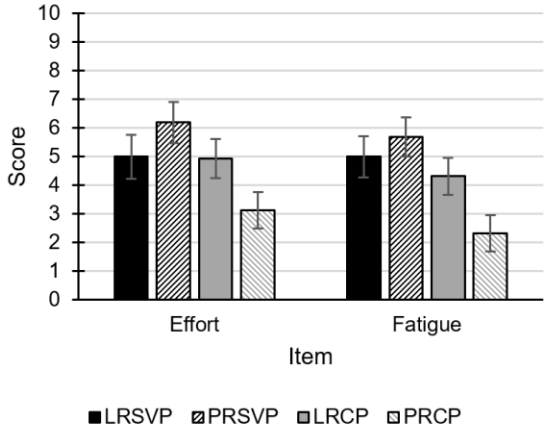

Fig. 7. Score (mean ± standard error) of the different P300-speller conditions denoted by users in the subjective questionnaire for effort (left) and fatigue (right). The conditions tested were: letters under rapid serial visual paradigm (L-RSVP), pictures under rapid serial visual paradigm (P-RSVP), letters under row–column paradigm (L-RCP), and pictures under row–column paradigm (P-RCP).

Fig. 8 shows the preferences for presentation paradigm, stimulus type and condition. First, users significantly preferred RCP (87.5% of users) versus RSVP (12.5% of users) (Z = 2.75, p = .004). Second, no significant results were found regarding the preferred stimuli. Finally, the users showed significant differences in preference between the different

conditions (χ^2 (2) = 9.125; p = .01), P-RCP being the most chosen preferred condition. It should be noted that no user selected the P-RSVP condition as their favourite.

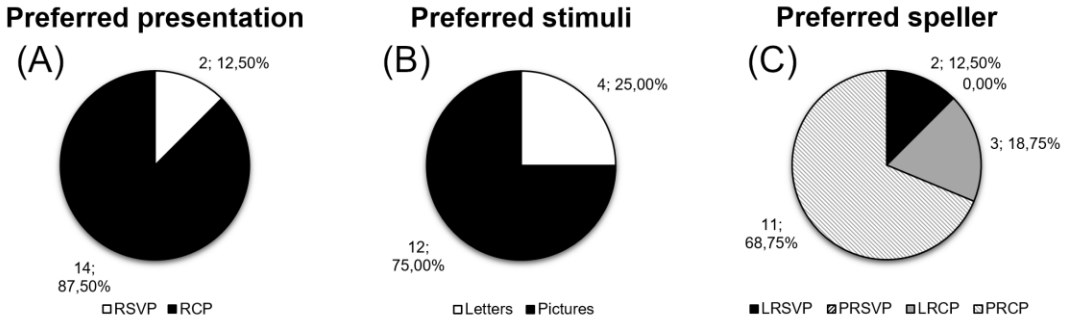

Fig. 8. A) Number and percentage of participants that preferred each presentation paradigm: rapid serial visual presentation (RSVP) or row–column paradigm (RCP). B) Number and percentage of participants that preferred each set of stimuli: letters or pictures. C) Number and percentage of participants that preferred each condition: letters under rapid serial visual paradigm (L-RSVP), pictures under rapid serial visual paradigm (P-RSVP), letters under row–column paradigm (L-RCP), and pictures under row–column paradigm (P-RCP).

4. Discussion

4.1. Accuracy and ITR in the calibration task

The main finding from analysing the calibration task data is a modulation effect of pictures to improve performance according to the presentation factor. Therefore, in general, the results found agree with the preliminary study previously cited [25].

In terms of accuracy, the results obtained can be considered good since all conditions obtained more than 90% average accuracy using five sequences. The most outstanding condition in the present work is P-RCP, since in the first sequence of the calibration it presented an average accuracy of 90.75%. In addition to the higher performance of the P-RCP condition (97.5 ± 6.47%) versus L-RCP (90.22 ± 9.63%), this condition under RCP (93.98 ± 10.09%) offered better performance than RSVP (83.72 ± 16.79%). On the other hand, the performance of the RSVP conditions was similar, showing no advantage due to the application of pictures under this presentation mode (L-RSVP, 83.34 ± 12.51%; P-RSVP, 84.13 ± 15.74%).

It should be noted that the accuracy does not take into account the time invested to make a selection, which is different for RSVP and RCP; however, this is considered by the ITR. Indeed, the ITR showed remarkable differences between RCP and RSVP paradigms that could be due to two factors: i) the shorter time needed to complete a sequence under RCP, and ii) the better accuracy provided by RCP (as shown above). These results are consistent with those obtained previously by [31]. As was reported for accuracy, the effect of the pictures was remarkable for RCP (L-RCP, 37.33 ± 17.93 bits/min; P-RCP, 47.39 ± 25.82 bits/min) but null for RSVP (L-RSVP, 18.07 ± 8.56 bits/min; P-RSVP, 17.8 ± 7.06 bits/min). As expected, the condition with the best ITR was P-RCP, ranging from 91.43 ± 15.6 bits/min for the first sequence to 23.14 ± 0.62 bits/min for the fifth sequence.

4.2. Event-related potential in the calibration task

The results found for the ERP signal agree with what was obtained during the calibration and online tasks. On the one hand, for the RSVP conditions (L-RSVP and P-RSVP), both stimuli presented quite similar ERP waveforms. Otherwise, under RCP, great differences between letters and pictures were found, especially for those components evoked around 400 ms after stimulus onset (such as N2, P300 and N400). Thus, the presentation paradigm modulates the amplitude difference in a different way for letters and pictures, which might explain the results obtained in terms of performance (i.e., accuracy and ITR).

Under RSVP, the P300 component was observed in all channels for both conditions (L-RSVP and P-RSVP). However, the RCP conditions (L-RCP and P-RCP) presented more complex results. In addition to a predominance of the P300 component, some negative peaks in amplitude difference were observed (temporally related to N2 and N400). Moreover, it was observed that the amplitude of these components depended on the stimuli employed (letters or pictures). In reference to P-RCP, the presence of N2 and N400 components in the frontal zone, along with positivity observed in the occipital areas, has been previously reported (e.g., [34, 35]).

4.3. Accuracy and ITR in the online task

Despite expectations, the accuracy in the online task did not show any significant difference. However, it should be noted that the P-RCP condition offered higher accuracy than the others, by a difference of more than 10% (L-RSVP, 84.61 ±

21.88%; P-RSVP, 81.49 ± 20.45%; L-RCP, 82.74 ± 13.84%; P-RCP, 96.21 ± 5.7%). In addition, according to the review by [11], the present work presents similar levels of accuracy, except for those for the P-RCP condition, which are higher than what is normally offered by other proposals. However, due to the small set size of our proposal – which was selected to establish a comparison with the same number of stimuli in RCP and RSVP without fatiguing the user – and the selection criteria for sequences in the online phase, an adequate comparison between studies is not an easy task. Alternatively, unlike the accuracy, the ITR allows better control of these variables – matrix size, accuracy and writing speed – to offer a fairer comparison between interfaces.

In the case of the ITR, the expected differences have been obtained. On the one hand, as has been reported in previous works, RCP has been shown to be significantly better – in terms of ITR – than RSVP [31, 36]. On the other hand, the pictures provoked an improvement in RCP, but not in RSVP. These results in the online phase regarding the ITR are explained by the better performance offered in the calibration phase, since users reached their maximum accuracy with fewer sequences needed (3.86 ± 0.92 and 2.38 ± 0.96 sequences for P-RSVP and P-RCP, respectively). Also, under the RCP conditions, it should be noted that the ITR obtained for the P-RCP condition was more than twice that for the L-RCP condition (49.05 ± 13.8 bits/min versus 21.11 ± 9.51 bits/min, respectively). For the P-RCP condition, the ITR even surpassed the online ITR reported by [23] in their proposal using green faces, which obtained an average of 39 ± 5 bits/min. Therefore, it can be affirmed that the pictures could be as good stimuli as the green faces in a P300-based speller. However, there are other works that – through the use of more complex paradigms than the classic RCP – have managed to show better ITR results (e.g., [37, 38]). Therefore, it could be interesting to adapt the use of these images to said presentation paradigms.

4.4. Subjective questionnaire

Both in terms of effort required and fatigue produced, the participants seem to agree that the RCP-based paradigms presented the best results. In addition, for effort, an interaction effect was found, which indicates that the presentation of pictures had a different effect depending on which mode of presentation was applied. This interaction effect indicates a clear improvement due to the use of pictures in RCP, while this effect may be even the opposite under RSVP. These results would fit with P-RSVP not being chosen as the preferred speller by any of the users.

The best score for RCP for both variables, effort and fatigue, may be observed because the user does not have to discriminate every 288 ms (the duration of the SOA) if the stimulus presented is the target or not. This effort could produce mental fatigue due to the need for concentration, as well as physical fatigue due to the eye strain of focusing on a stimulus. In RCP, the user only has to concentrate on the target stimulus area, and the onset of the target stimuli could be more discriminable.

Despite the lack of significant differences, the pictures in RSVP seem to have required more effort and caused more fatigue than expected, even versus letters in RSVP. This may have happened because the user had to match the desired picture with the corresponding letter (although the letter was still in the figure, and the user could use a document to check the picture–letter arrangement). These results lead us to hypothesize that, maybe in P-RSVP, although the user knows the picture–letter association, he/she has to make the effort to categorize each picture.

These previous results have been corroborated by the users' preference for RCP. On the other hand, this effect was not found for pictures since, despite the good performance in RCP, the pictures did not present a high performance in RSVP versus the other conditions. However, it should be noted that 75% of participants preferred pictures. While no participant selected the P-RSVP interface as a favourite, 75% of the participants chose P-RCP as their preferred one; this shows the important role of presentation paradigm in the use of pictures.

5. Conclusions

The present work has shown a robust interaction effect between the presentation paradigm and the stimuli used in a visual P300-based speller. This finding is especially relevant if we consider that most researches have been focused on RCP, which requires eye control, a skill that the main population of these devices – that is, patients with severely affected motor skills– do not have. Therefore, it could be convenient for future works to test previous proposals under RCP in other paradigms that are more convenient for patients, such as RSVP.

Likewise, the presented results may be tested in larger sets of stimuli, with more than nine items. However, the possibility of using reduced matrices is highlighted, since they tend to offer better performance in accuracy [19] and could be a more realistic objective for future patients. It should be remembered that, to date, no visual P300-based BCI has been controlled by a patient in CLIS [39]. In addition, the benefit of larger matrices to calculate the ITR might not be enough to compensate the loss in accuracy, as can be seen in the Discussion section.

In short, in addition to validating the use of pictures for a visual P300-based BCI speller, this study has shown that not all

stimuli work in the same way under RCP and RSVP presentation modes. Thus, the present work has shown the importance of testing the findings achieved regarding the stimuli under RCP in RSVP as well.

Acknowledgements

This work was partially supported by the project SICCAU: RTI2018-100912-B-I00 (MCIU/AEI/FEDER, UE). Moreover, the authors would like to thank all participants for their cooperation.

References

1. Patterson JR, Grabois M (1986) Locked-in syndrome: A review of 139 cases. Stroke 17:758–764. https://doi.org/10.1161/01.STR.17.4.758
2. Bauer G, Gerstenbrand F, Rumpl E (1979) Varieties of the locked-in syndrome. J Neurol 221:77–91. https://doi.org/10.1007/BF00313105
3. Murguialday AR, Hill J, Bensch M, et al (2011) Transition from the locked in to the completely locked-in state: A physiological analysis. Clin Neurophysiol 122:925–933. https://doi.org/10.1016/j.clinph.2010.08.019
4. Fuchino Y, Nagao M, Katura T, et al (2008) High cognitive function of an ALS patient in the totally locked-in state. Neurosci Lett 435:85–89. https://doi.org/10.1016/j.neulet.2008.01.046
5. Lemoignan J, Ells C (2010) Amyotrophic lateral sclerosis and assisted ventilation: How patients decide. Palliat Support Care 8:207–213. https://doi.org/10.1017/S1478951510000027
6. Wolpaw JR, Birbaumer N, McFarland DJ, et al (2002) Brain-computer interfaces for communication and control. Clin Neurophysiol 113:767–791. https://doi.org/10.1016/S1388-2457(02)00057-3
7. Birbaumer N (2006) Breaking the silence: Brain-computer interfaces (BCI) for communication and motor control. Psychophysiology 43:517–532. https://doi.org/10.1111/j.1469-8986.2006.00456.x
8. Farwell LA, Donchin E (1988) Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. Electroencephalogr Clin Neurophysiol 70:510–523. https://doi.org/10.1016/0013-4694(88)90149-6
9. Corralejo R, Hornero R, Álvarez D (2011) A domotic control system using Brain-Computer Interface (BCI)
10. Zhang R, Wang Q, Li K, et al (2017) A BCI-based environmental control system for patients with severe spinal cord injuries. IEEE Trans. Biomed. Eng. 64:1959–1971
11. Rezeika A, Benda M, Stawicki P, et al (2018) Brain–computer interface spellers: A review. Brain Sci 8:. https://doi.org/10.3390/brainsci8040057
12. Sutton S, Braren M, Zubin J, John ER (1965) Evoked-potential correlates of stimulus uncertainty. Science (80-). https://doi.org/10.1126/science.150.3700.1187
13. Treder MS, Blankertz B (2010) (C)overt attention and visual speller design in an ERP-based brain-computer interface. Behav Brain Funct 6:. https://doi.org/10.1186/1744-9081-6-28
14. Brunner P, Joshi S, Briskin S, et al (2010) Does the "P300" speller depend on eye gaze? J Neural Eng 7:56013. https://doi.org/10.1088/1741-2560/7/5/056013
15. Acqualagna L, Treder MS, Schreuder M, Blankertz B (2010) A novel brain-computer interface based on the rapid serial visual presentation paradigm. In: 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC'10. pp 2686–2689
16. Acqualagna L, Treder MS, Blankertz B (2013) Chroma Speller: Isotropic visual stimuli for truly gaze-independent spelling. Int IEEE/EMBS Conf Neural Eng NER 1041–1044. https://doi.org/10.1109/NER.2013.6696115
17. Sato H, Washizawa Y (2016) An N100-P300 Spelling Brain-Computer Interface with Detection of Intentional Control. Computers 5:31. https://doi.org/10.3390/computers5040031
18. Aricò P, Aloise F, Schettini F, et al (2011) GeoSpell : an alternative P300-based speller interface towards no eye gaze required. Int J Bioelectromagn 13:152–153
19. Sellers EW, Krusienski DJ, McFarland DJ, et al (2006) A P300 event-related potential brain-computer interface (BCI): The effects of matrix size and inter stimulus interval on performance. Biol Psychol 73:242–252. https://doi.org/10.1016/j.biopsycho.2006.04.007
20. Kaufmann T, Schulz SM, Grünzinger C, et al (2011) Flashing characters with famous faces improves ERP-based brain-computer interface performance. J Neural Eng 8:056016. https://doi.org/10.1088/1741-2560/8/5/056016
21. Townsend G, LaPallo BK, Boulay CB, et al (2010) A novel P300-based brain-computer interface stimulus presentation

paradigm: Moving beyond rows and columns. Clin Neurophysiol 121:1109–1120. https://doi.org/10.1016/j.clinph.2010.01.030

22. Ryan DB, Frye GE, Townsend G, et al (2011) Predictive spelling with a P300-based brain-computer interface: Increasing the rate of communication. Int J Hum Comput Interact 27:69–84. https://doi.org/10.1080/10447318.2011.535754

23. Li Q, Liu S, Li J, Bai O (2015) Use of a green familiar faces paradigm improves P300-speller brain-computer interface performance. PLoS One 10:1–15. https://doi.org/10.1371/journal.pone.0130325

24. Fernández-Rodríguez Á, Velasco-Álvarez F, Medina-Juliá MT, Ron-Angevin R (2019) Evaluation of emotional and neutral pictures as flashing stimuli using a P300 brain-computer interface speller. J Neural Eng. https://doi.org/10.1088/1741-2552/ab386d

25. Fernández-Rodríguez Á, Medina-Juliá MT, Velasco-Álvarez F, Ron-Angevin R (2019) Preliminary Results Using a P300 Brain-Computer Interface Speller: A Possible Interaction Effect Between Presentation Paradigm and Set of Stimuli. In: Rojas I, Joya G, Catala A (eds) Advances in Computational Intelligence. Springer International Publishing, p 10

26. Schalk G, McFarland DJ, Hinterberger T, et al (2004) BCI2000: A general-purpose brain-computer interface (BCI) system. IEEE Trans. Biomed. Eng. 51:1034–1043

27. Velasco-Álvarez F, Sancha-Ros S, García-Garaluz E, et al (2019) UMA-BCI Speller: an Easily Configurable P300 Speller Tool for End Users. Comput Methods Programs Biomed 172:127–138. https://doi.org/10.1016/J.CMPB.2019.02.015

28. Cheng J, Jin J, Wang X (2017) Comparison of the BCI performance between the semitransparent face pattern and the traditional face pattern. Comput Intell Neurosci 2017:. https://doi.org/10.1155/2017/1323985

29. Lang, P. J., Bradley, M. M., & Cuthbert BN (2008) International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Technical Report A-8. University of Florida, Gainesville, FL.

30. Wolpaw JR, Ramoser H, McFarland DJ, Pfurtscheller G (1998) EEG-based communication: Improved accuracy by response verification. IEEE Trans Rehabil Eng 6:326–333. https://doi.org/10.1109/86.712231

31. Chennu S, Alsufyani A, Filetti M, et al (2013) The cost of space independence in P300-BCI spellers. J Neuroeng Rehabil 10:1–13. https://doi.org/10.1186/1743-0003-10-82

32. IBM Corp. (2016) IBM SPSS Statistics for Windows, Version 24.0. 2016

33. Delorme A, Makeig S (2004) EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods 134:9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

34. Kutas M, Sereno MI, Ganis G (1996) The Search for "Common Sense": An Electrophysiological Study of the Comprehension of Words and Pictures in Reading. J Cogn Neurosci 8:89–106

35. Garrison KE, Crowell AL, Finley AJ, Schmeichel BJ (2017) Effects of prior mental effort on picture processing: An ERP investigation. Psychophysiology 54:1714–1725. https://doi.org/10.1111/psyp.12914

36. Moghadamfalahi M, Orhan U, Akcakaya M, et al (2015) Language-model assisted brain computer interface for typing: A comparison of matrix and rapid serial visual presentation. IEEE Trans Neural Syst Rehabil Eng 23:910–920. https://doi.org/10.1109/TNSRE.2015.2411574

37. Kaufmann T, Kübler A (2014) Beyond maximum speed - A novel two-stimulus paradigm for brain-computer interfaces based on event-related potentials (P300-BCI). J Neural Eng 11:. https://doi.org/10.1088/1741-2560/11/5/056004

38. Yeom SK, Fazli S, Ller KRM, Lee SW (2014) An efficient ERP-based brain-computer interface using random set presentation and face familiarity. PLoS One 9:1–13. https://doi.org/10.1371/journal.pone.0111157

39. Guger C, Spataro R, Allison BZ, et al (2017) Complete locked-in and locked-in patients: Command following assessment and communication with vibro-tactile P300 and motor imagery brain-computer interface tools. Front Neurosci 11:251. https://doi.org/10.3389/fnins.2017.00251