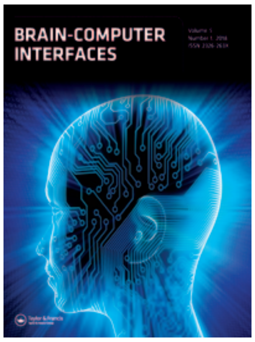
P300 brain-computer interface: comparing faces to size matched non-face stimuli

Article · January 2018
DOI: 10.1080/2326263X.2018.1433776

2 authors:

Marissa R Kellicut-Jones
East Tennessee State University

Eric W Sellers
East Tennessee State University



Taylor & Francis
Taylor & Francis Group

Brain-Computer Interfaces

ISSN: 2326-263X (Print) 2326-2621 (Online) Journal homepage: http://www.tandfonline.com/loi/tbci20

P300 brain-computer interface: comparing faces to size matched non-face stimuli

M. R. Kellicut-Jones & E. W. Sellers

BRAIN-COMPUTER INTERFACES, 2018
VOL. 5, NO. 1, 30–39
https://doi.org/10.1080/2326263X.2018.1433776

Taylor & Francis
Taylor & Francis Group

P300 brain-computer interface: comparing faces to size matched non-face stimuli

M. R. Kellicut-Jones and E. W. Sellers

Department of Psychology, East Tennessee State University, Johnson City, TN, USA

ABSTRACT
Non-invasive brain–computer interface (BCI) technology can restore communication for those unable to communicate due to loss of muscle control. Nonetheless, compared to augmentative and alternative communication (AAC) devices requiring muscular control, BCIs provide relatively slow communication. Therefore, implementing techniques improving BCI speed and accuracy is important. Previous studies indicate that facial stimuli elicit N170 and N400 components, in addition to the P300 component associated with P300 BCI. These additional components can increase speed and accuracy. Our study investigated the influence of image size and content using four conditions: large face, small face, large non-face, and small non-face. We predicted faces would provide higher accuracy than non-face stimuli and larger stimuli would provide higher accuracy than small stimuli. We found no significant difference in performance between conditions; however, significant waveform differences were found in each condition.

ARTICLE HISTORY
Received 14 September 2017
Accepted 25 January 2018

KEYWORDS
P300 speller; EEG; event-related potentials; communication; facial stimuli

1. Brain-computer interfaces

The uses of EEG recorded signals to operate brain-computer interfaces (BCIs) range from controlling external devices such as a robotic arm [1] to creating art [2]. BCIs have been found to be especially effective systems of communication for individuals who retain cognitive function but have lost the ability to communicate due to loss of muscle control. This loss of communication has a very negative impact on individuals with amyotrophic lateral sclerosis (ALS), brainstem stroke, or traumatic brain injury who now must rely on caretakers for activities of daily living. BCIs can be useful systems of communication for these individuals as they require no muscular movement [3]. The current state of BCI research is focused on making the systems used in laboratories more accessible and functional for in-home use [4]. While there has been great improvement in the operation of the P300 BCI, there is still more research required to increase functionality for those who may benefit from BCI use. Thus, speed and accuracy of word selection and increasing user friendliness remain primary goals of BCI research.

The P300 speller has been one of the most promising non-invasive BCIs; it is a modified oddball task and is fundamentally central to the P300 speller. Participants are presented a matrix on a computer monitor made up of characters such as letters, numbers, and various symbols, like that of a computer keyboard. Each of the characters in the matrix flashes randomly at different intervals. In order to select a character in the matrix, the participant pays attention to the character he or she wishes to select, and every time the character flashes the participant keeps a mental count. Each time the participant attends to the flashing of the desired character, a P300 component is elicited. By examining P300 responses, the P300 speller is able to discriminate target versus non-target items (i.e. letters the participant is trying to select versus letters the participant is not trying to select).

Different positive and negative ERPs are produced in response to different stimuli and events. The N170 and N400 components are two negative components that are considered to contribute to recognition and processing of faces. The N170 occurs around approximately 170 ms at the lateral temporal electrode sites following the presentation of facial stimuli [5–7]. The N400 occurs approximately between 200 ms and 600 ms [8] over the right hemisphere electrode sites. These components have been observed using unaltered facial images, inverted facial images, and even dummy faces [9,10].

Recently, face stimuli and the N170 and N400 components have been incorporated in the P300 speller. In 2011, Kaufmann et al. [11] investigated famous familiar faces as potential stimuli for P300 speller character matrices. This was done in the hope that in addition to the P300 component elicited by character flashes, using facial stimuli would produce an N170 and an N400 component

CONTACT M. R. Kellicut-Jones kellicut@etsu.edu

[11]. Using the familiar, famous image of Albert Einstein, and a familiar, famous image of Che Guevera, Kaufman et al. [11] superimposed the facial images so that they would flash over the characters in the matrix. Instead of the characters themselves in the matrix flashing, the Einstein image would flash. To create a control image, Kaufmann [11] took the facial images and pixelated the facial images to retain the physical characteristics of a face while still creating an indistinguishable and obviously different image from the original Albert Einstein face. A unique control condition was used based on the findings of [12], who found that physical features of stimuli such as chromatic differences, brightness, or contrast can impact BCI performance. Using the row-column flash pattern to investigate the influence of face stimuli on performance measures, they compared the face image paradigms to the pixelated image paradigm and the classic flashing of letter in the matrix typically used in the P300 speller [11]. The offline analyses conducted by Kaufmann et al. [11] indicated that flashing faces instead of flashing letters significantly decreased the number of flashes required to reach 100% accuracy. The reason behind this significant increase in performance is thought to be due to face-sensitive ERPs [13]. By eliciting face-sensitive components such as N170 and N400, as well as the P300 component, it is thought that discriminating between target and non-target items is made easier for the BCI system.

In 2013, Kaufmann et al. [14] compared familiar personally known faces, familiar famous faces, and a classic character flash condition. In healthy participants, performance was significantly better for the face conditions than the classic character flash condition, but there was no significant difference between the two face conditions. In the patient sample, there were two patients who were unable to use the BCI using the character flash stimuli. However, when using the face stimuli, one patient was able to spell with an average accuracy greater than 80%. The other patient was able to spell with no errors [14]. These findings are important because it indicates that those who seem to be inefficient in using the BCI may be able to use the system when face images are used as stimuli.

To increase the speed of target selection, Kaufmann and Kubler [15] introduced a paradigm that implemented a simultaneous presentation of two very different stimuli in the four quadrants of the matrix. The image of Einstein was presented in the top left and bottom right quadrants, and a yin-yang symbol was presented in the top right and bottom left quadrants. The two-stimulus presentation was compared to the standard row-column. The results showed that the two-stimulus paradigm was able to make selections more quickly than the one-stimulus paradigm, despite a decrease in accuracy [15].

Since the initial use of faces as stimuli in P300 speller paradigms, the positive effects of faces on BCI performance has continued to be supported. Geronimo and Simmons [16] examined whether the ERPs produced by the use of facial stimuli could maintain the high-performance effects shown in previous studies even with cognitive decline. Although the P300 speller using facial stimuli did result in high performance despite cognitive decline, they suggest the result may not be due to face-specific components such as the N170, which appears to decline in amplitude.

1.1. Present study

In the past few years, the face paradigm has shown to be a viable alternative to the more typical letter flash paradigm. The most common explanation for these results, as discussed above, is that the additional N170 and N400 components are responsible for improved performance. Nonetheless, there are possible alternative explanations for the improved performance. In addition, not all of these studies have provided definitive evidence that the face paradigm produces significantly different N170 and N400 responses from other stimuli. For example, Kaufmann et al. in 2011 [11] found that facial stimuli produced significantly different P300 ERPs compared to control stimuli, but observed no differences in the N170 and N400 responses produced by the facial images and a pixelated control. Kaufmann et al. in 2013 [14] observed similar results, with differences observed in the P300 response between facial stimuli and classic character flashing, but no significant differences in the N170 or N400.

We propose that other physical properties of the face stimuli may be responsible for the observed increases in performance. Oftentimes the size of the Albert Einstein face image is much larger than the letters in the matrix (e.g. Figure 1). Therefore, we propose to examine the relationship between image size and stimulus type. If the face stimulus is producing the effects, then performance should remain the same despite the size of the image. Similar to Kaufmann et al. [11], we use a control condition that takes the face image we are interested in looking at and changes the image so that it retains the brightness, contrast, and chromatic information of the original image, but is unrecognizable as a face. In the present study, we compare four conditions: large face stimuli, small face stimuli, large non-face stimuli, and small non-face stimuli. Based on the results of previous studies, we predicted that faces would provide higher accuracy than non-face stimuli and larger stimuli would provide higher accuracy than small stimuli.

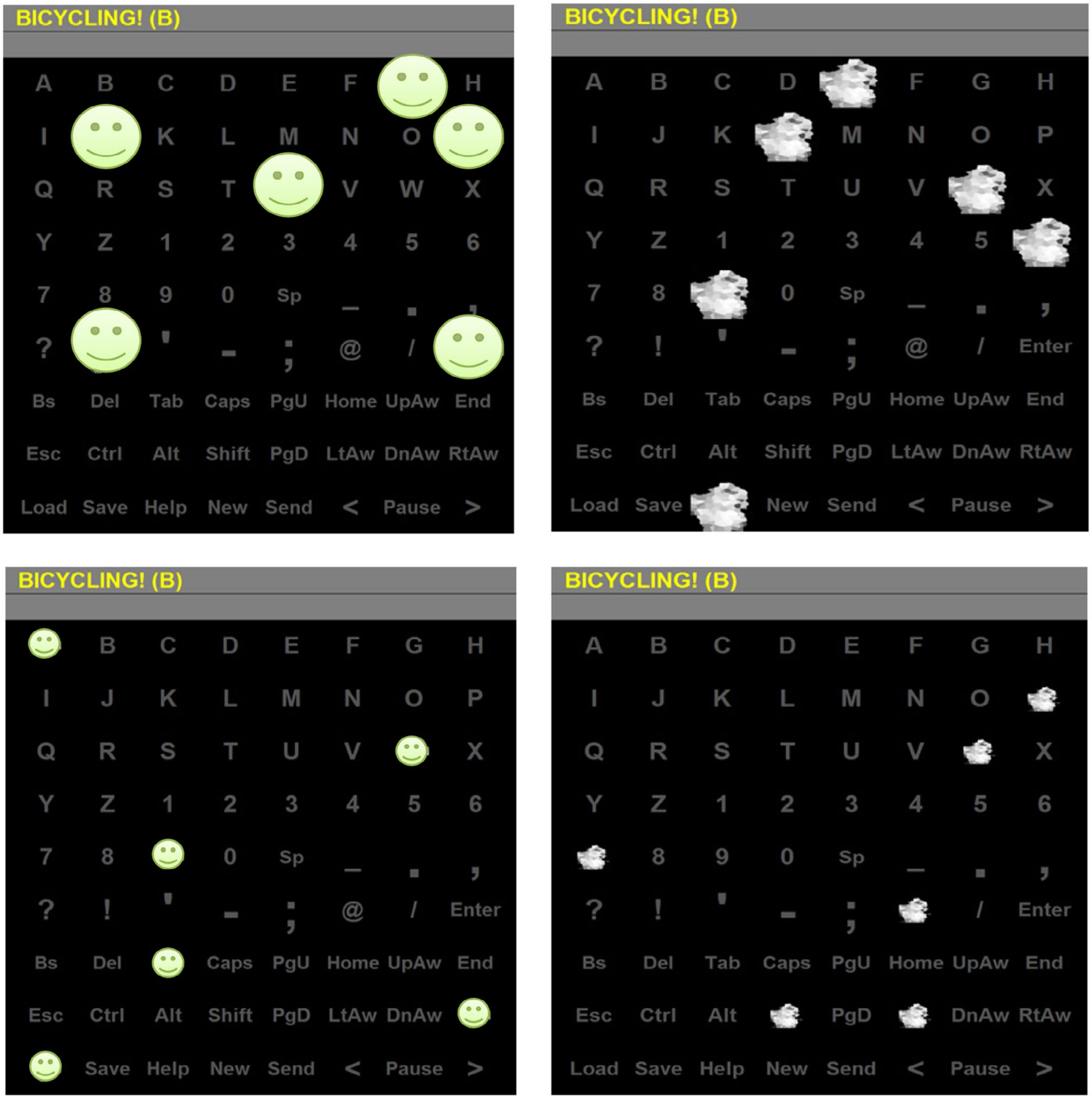


Figure 1. Example of the four stimulus presentation matrices: Large Face (top left), Small Face (bottom left), Large Crystallized (top right), and Small Crystallized (bottom right). Using a checkerboard flash pattern, the face images or crystallized images are intensified or 'flash' over the characters in the matrix. Participants are instructed to keep a mental count of the number of times the image flashes over the desired character in the matrix. (The picture of Einstein was used as the stimulus in the face conditions. The picture is not used in the present figure due to copyright restrictions.)

The overarching goal of the current study is not to introduce a new paradigm that will increase overall speed and accuracy of BCI performance. The goal is to examine hypotheses related to why the Face Speller is superior to the traditional flashing paradigm. By examining the role of the ERPs produced by facial images, as well as the effect that size may have on performance and ERPs, we may have a further understanding of the relationship between the use of faces as stimuli and increased BCI performance.

2. Methods

2.1. Participants

Twenty able-bodied participants (7 men, 13 women; age range 18–49) completed the study. Three of the participants had prior BCI experience, the rest were naïve to BCI use. The study was approved by the East Tennessee State Institutional review board and each participant gave informed consent prior to the beginning of the study.

2.2. Data acquisition and processing

Electroencephalograph (EEG) signals were recorded using a cap (Electro-Cap International, Inc.) embedded with 32 tin electrodes. The right mastoid was used as the reference electrode and the left mastoid was used as the ground electrode. Only eight electrodes were used for online classification. The eight electrodes were subject-specific and determined by the jumpwise algorithm (described in Section 2.3; Colwell et al. [17]). The signal was digitized at 256 Hz and bandpass-filtered to [0.5 Hz, 30 Hz] by two16-channel g.tec g.USBamp amplifiers. BCI2000 was used for stimulus presentation and data collection [18].

2.3. Classification and determining the optimal number of sequences

The classification technique known as stepwise linear discriminate analysis (SWLDA) as described by [19] was used to determine classification coefficients. SWLDA has previously been shown to be an efficient method of classification for BCI research [17,20,21]. By using forward regression and backward regression, the SWLDA algorithm selects the spatiotemporal features that account for the most unique variance. Features are weighted by ordinary least-squares regression. Beginning with no features in the model, the feature that is the most statistically significant ($p < .1$) is added. A backward stepwise analysis then removes features that are found to be the least significant ($p > .15$). There is a predetermined number of features included in the model, so that the process is repeated until the maximum number of features are included in the model (which is 60 features), or no more features meet the criterion for entry or removal. Once the classification coefficients for each feature are produced, they are applied to the ERPs corresponding to each of the characters in the matrix. The character that has the highest sum of the coefficients is selected as the target item.

To improve upon classification performance, multiple electrodes at various locations distributed over the scalp are used. A limitation to using a system with a larger number of electrodes is that the systems are more expensive and require more time to set up for at-home users [17]. A filter method known as jumpwise selection is used to improve upon classification through optimal channel selection. Jumpwise selection uses a variant of SWLDA that selects electrodes instead of electrode-specific features. Once the most statistically significant channels are selected, a SWLDA, as described above, is conducted on the specific features contained within the jumpwise-defined electrode set. The advantage of jumpwise selection is that it reduces the feature space to a unique set of features that are chosen from electrode locations that are optimized for each individual participant.

During calibration the number of target flashes was fixed at 14. For online testing, the number of target flashes was determined by each participant's maximum written symbol rate (WRS), which was calculated from the calibration data [22,23]. The WSR is a conservative measure of the optimal number of target flashes because it assumes the BCI user would choose to correct all errors.

2.4. Experiment stimuli and matrix

In previous studies examining the use of facial stimuli, the famous image of Albert Einstein sticking his tongue out has been used. To remain consistent with previous study, we also used the familiar image of Albert Einstein. Based on the results of previous studies, we predicted that faces would provide higher accuracy than non-face stimuli. To test this hypothesis, we designed four conditions: large face stimuli, small face stimuli, large non-face stimuli, and small non-face stimuli (see Figure 1). For the large face condition, we used the same image size found in previous studies. For the small conditions the images used were 25% the size of the large condition images. The two large images presented were 77×96 pixels in size, and the two small images were 39×48 pixels in size. The non-face image was constructed using a crystallize filter (Photoshop CS5.5) and rotating the image by 180 degrees, thereby preserving the image content while making the stimulus unrecognizable as a face. The large non-face image was the size equivalent of the large face image and the small non-face image was the size equivalent of the small face image.

The row-column flash pattern used by Kaufmann [11,14] has been shown to be associated with higher error rates due to the flashing of adjacent non-target rows or columns when a user is attempting to make a target selection. To reduce the errors associated with adjacent flashing, the checkerboard paradigm developed by [22] was used.

2.5. Experiment design

Participants completed two experimental sessions. Two of the four conditions were tested in each session. To control for order effect a Latin square design was used to determine the order in which the conditions would be presented. Every session contained a calibration phase for each of the two conditions being tested in the session. Each calibration phase was followed by an online copy-spelling phase that corresponded to the calibration condition. For each of the four phases, participants spelled three 6-letter words from the English Lexicon Project [24]. A Matlab script was used to randomly select the words.

Participants were seated approximately 90 cm away from a computer monitor that displayed an 8×9 matrix of letters and numbers. The speller matrix was adapted from a BCI2000 system used in a patient's home. After the participant was fitted for the electrode cap, the 8 × 9 matrix of letters and characters was presented on the computer monitor. For the calibration phase, participants were asked to focus their attention on a specific character in the matrix and to count how many times the letter flashed, while ignoring the images flashing over the other characters in the matrix. For example, as shown in Figure 1, the top left side of the display would show a word (e.g. BICYCLING!) and the letter they should attend to was shown in parentheses at the end of the word. After a predetermined amount of flashes of each character (in this case 14) the matrix would stop flashing and after a 4-s pause the letter in parentheses would change to the next letter in the word (e.g. (I)).

During the calibration phase, participants did not receive feedback on character selection. Once the calibration phase was completed, the participants completed the online copy-spelling phase. During the online copy-spelling phase, feedback was presented directly under the word the subject was spelling. Participants were instructed not to correct errors made by the BCI. Following the calibration phase of each condition in the two sessions, each participant completed a survey to indicate preference of condition. Each survey consisted of three questions evaluated on a Likert scale ranging from 1 to 7, with 1 indicating strongly disagree and 7 indicating strongly agree. The items on the survey were: (1) I was able to focus on the target item, (2) Non-target items were distracting, (3) It was difficult to see all the target flashes. These surveys were given following the calibration phase to eliminate potential bias for a condition based on the participants' accuracy of character selection.

2.6. Statistical analyses

Comparisons between the four conditions were made based on the online accuracy, bit rate, selections per minute, and ERP amplitudes and latencies. Spelling accuracy of online testing was calculated as the percentage of correctly selected letters for each word. To determine the number of target selections required to make a selection, sets per sequence for online testing were optimized for each participant using calibration data. Selections per minute were calculated based on total selections made correctly or incorrectly in a given minute. Bit rate was calculated using the formula described by Pierce [25]:

$$BR = \log_2 N + P\log_2 P + (1-P)\log_2\left[(1-P)/(N-1)\right]$$

Bit rate takes into account the number of possible targets (N) and the probability that the target is classified accurately (P). Dividing B by the trial duration in minutes results in the bit rate. The comparisons were made using repeated-measures analysis of variance (ANOVA) within participants with Greenhouse-Geisser corrections.

3. Results

Mean scores on the surveys were calculated for each of the four conditions, and the mean scores were statistically similar across all four of the conditions. Overall, the participants rated each condition a 5 on the 7-point Likert scale.

A two-way repeated-measures analysis of the variance (ANOVA) was conducted to compare the effect of stimulus condition on accuracy, bitrate, and selections per minute. Additional offline analyses were conducted to determine if the performance measures could be improved upon; however, these analyses were statistically equivalent to the online analyses. Thus, the offline data are not reported. A repeated-measures ANOVA was also used to investigate the effects of stimulus condition on the waveforms by examining the positive amplitudes and latencies as well as the negative amplitudes and latencies. Online means and standard deviations for performance measures are shown in Table 1.

3.1. Accuracy

In the analysis of effect of condition on accuracy, Mauchly's test of sphericity was found to be significant (p = .031), indicating a violation of the assumption of sphericity. Using Greenhouse-Geisser corrections, the analysis indicated that condition had no significant effect on accuracy $F(2.172,41.259) = 1.372$, p = .266. Post-hoc tests using

Table 1. Online means and standard deviations (in parentheses) for performance measures for the Large Face (LF), Small Face (SF), Large Non-Face (CL), and Small Non-face (CS) conditions in the online test phase of the experiment.

	Accuracy	Target flashes	Selections per minute	Bit rate	Survey
LF	98.30 (3.6)	3.75 (1.2)	4.00 (0.77)	23.97 (5.01)	4.90 (1.22)
SF	95.45 (7.4)	3.65 (1.2)	4.14 (1.03)	23.73 (7.12)	4.67(1.16)
CL	97.45 (4.2)	3.80 (1.2)	4.00 (0.98)	23.63 (6.07)	5.18 (0.98)
CS	94.65 (8.6)	3.65 (1.2)	4.06 (0.74)	23.01 (5.07)	5.40 (0.96)

Bonferroni corrections revealed no significant differences in accuracy between any of the four conditions.

3.2. Target flashes

The analysis of effect of condition on target flashes required to make a character selection, Mauchly's test of Sphericity was not significant (p = .954). With sphericity assumed, the analyses showed no significant effect of condition on target flashes required to make a selection, $F(3,57) = .111$, $p = .954$. Post-hoc tests using Bonferroni corrections also revealed no significant differences in target flashes between any of the four conditions.

3.3. Selections per minute

For the analysis of effect of condition on selections per minute, Mauchly's test of sphericity was not significant ($p = .341$). With sphericity assumed, the analysis showed that there was also no significant effect of condition on selections per minute, $F(3,57) = .158$, $p = .924$. Post-hoc tests using Bonferroni corrections also revealed no significant differences in accuracy between any of the four conditions.

3.4. Bit rate

For the analysis of effect of condition on bit rate, Mauchly's test of sphericity was not significant (p = .436), indicating no violation of the assumption of sphericity. With sphericity assumed, the ANOVA showed that there was no significant difference in bit rate between conditions $F(3,57) = .135$, $p = .939$. Post-hoc tests using Bonferroni corrections revealed no significant differences in accuracy between any of the four conditions.

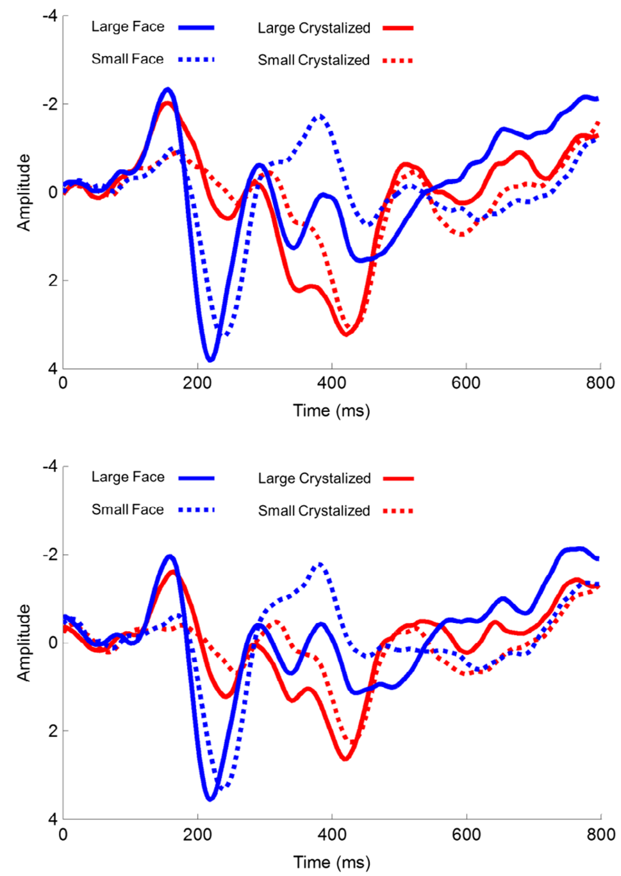


Figure 2. Participant waveforms: target waveforms for all 20 participants for electrode locations Cz, Pz, Po7, and Po8 (from left to right) for each Large Face (LF; solid blue line), Small Face (SF; dashed blue line), Large Crystallized (CL; solid red line), and Small Crystallized (CS; dashed red line) conditions.

3.5. Waveforms

Significant differences were not observed between any of the conditions on BCI performance; however, each of the four stimuli did produce distinctly different waveforms. Figures representing the waveforms averaged across all participants for each of the four conditions are shown in Figures 2 and 3. As with the analyses of performance measures, a two-way repeated measures ANOVA was used to examine the statistical differences between the waveforms elicited by each of the four conditions at electrode locations Pz and Cz. The specified time windows to examine the P300 amplitudes and latencies were set to 190–290 ms. The specified time windows to examine the N170 and N400 amplitudes and latencies were set to 128–195 ms for the N170 component, and 351–425 ms for the N400 component. The time windows were determined by examination of the grand mean waveforms. Means and standard deviations for amplitudes and latencies are shown in Table 2. All analyses used a 2 × 2 repeated-measures ANOVA with Greenhouse-Geisser correction, and post-hoc tests used Bonferroni corrections.

3.5.1. P300

The ANOVA examining P300 amplitudes at electrode Pz indicated significant differences between the four conditions at electrode Pz ($F(3,57) = 31.529$, $p < .001$). Post-hoc tests indicated large face amplitude to be significantly higher than both crystallized conditions (Large Crystallized(p = .021; Small Crystallized p < .001). In addition, Small Face amplitude was higher than Small Crystallized amplitude ($p < .001$), and the large Crystallized condition was significantly higher than the Small Crystallized condition, $p < .001$.

The ANOVA examining positive amplitudes at electrode Cz found significant differences between the four conditions, $F(3,57) = 25.569$, $p < .001$. The pattern of results was the same as that found on electrode Pz.

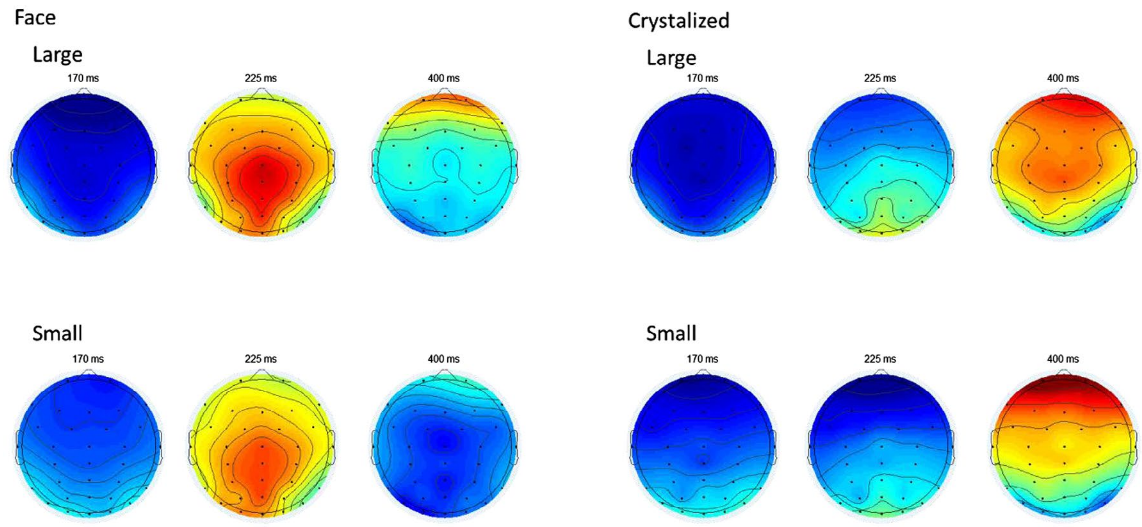


Figure 3. Grand average ERP response: topographical representation averaged across all 20 participants for Large Face (LF), Small Face (SF), Large Crystallized (CL), and Small Crystallized (CS) conditions.

Table 2. Waveform amplitudes and latencies produced by the Large Face (LF), Small Face (SF), Large Crystallized (CL), and Small Crystallized (CS) conditions in positive time window 190–290 ms (P300), and negative time windows 128–195 (N170) ms and 351–425 ms (N400).

Pz amplitudes	190–290 ms	128–195 ms	351–425 ms
LF	4.16 (2.08)	−2.56 (1.33)	−0.69 (1.99)
SF	4.22 (2.34)	−1.45 (0.09)	−2.3 (1.96)
CL	1.68 (1.70)	−2.31 (1.32)	0.40 (1.62)
CS	1.56 (1.69)	−1.1 (0.82)	−0.37 (1.30)
Pz			
Latencies	190–290 ms	128–195 ms	351–425 ms
LF			
230.6 (12.8)	166.6 (15.1)	383.9 (26.4)	
SF	248.8 (16.8)	172.0 (23.4)	386.7 (27.3)
CL	254.6 (18.8)	168.5 (19.8)	377.1 (25.2)
CS	260.7 (37.0)	166.2 (26.2)	378.1 (29.7)
Cz amplitudes	190–290 ms	128–195 ms	351–425 ms
LF	4.64 (2.70)	−2.88 (1.6)	−0.45 (2.52)
SF	4.37 (2.66)	−1.8 2(1.15)	−2.33 (2.51)
CL	1.65 (1.84)	−3.02 (1.43)	1.19 (1.86)
CS	1.59 (1.69)	−1.67 (1.00)	0.43 (1.98)
Cz			
Latencies	190–290 ms	128–195 ms	351–425 ms
LF	236.6 (22.6)	166.9 (16.5)	388.8 (29.0)
SF	255.8 (18.4)	172.6 (22.1)	385.9 (26.9)
CL	255.0 (28.1)	168.5 (21.8)	381.0 (27.3)
CS	264.4 (35.6)	168.7 (24.0)	369.7 (25.1)

Specifically, Large Face amplitude was significantly higher than both Crystallized conditions (Large Crystallized $p < .001$; Small Crystallized $p < .001$). Small Face amplitude was higher than Small Crystallized amplitude ($p < .001$), and the Large Crystallized amplitude was significantly higher than Small Crystallized amplitude, $p < .001$.

The ANOVA used to examine P300 latencies at electrode Pz indicated significant differences between the four conditions, $F(3,57) = 10.3328$, $p < .001$. Post-hoc tests indicated significantly earlier latencies for the Large Face condition than both Crystallized conditions (Small Face condition $p < .001$; Large Crystallized $p < .001$) and the Small Crystallized, $p = .002$). Significant differences were not found, however, between the remaining conditions.

The ANOVA, to examine P300 latencies at electrode Cz indicated significant differences between the four conditions, $F(3,57) = 25.569$, $p < .001$. Post-hoc tests indicated significantly earlier latencies for the Large Face condition than the Small Face condition ($p < .001$), and both Crystallized conditions (Large Crystallized $p < .001$; Small Crystallized $p = .002$). Significantly earlier latencies were also observed for the Small Face condition than the two crystallized conditions (Large Crystallized $p < .001$; Small Crystallized $p < .001$). No significant differences were found between the Large Crystallized and Small Crystallized conditions.

3.5.2. N170

The ANOVA used to examine N170 amplitudes at electrode Pz indicated significant differences between the four conditions, $F(3,57) = 19.903$, $p < .001$. Post-hoc tests indicated significantly larger amplitude for the Large Face condition than the two small image conditions (Small Face $p < .001$; Small Crystallized $p < .001$), and larger amplitude for the Large Crystallized condition than the two small image conditions (Small Face $p < .008$; Small Crystallized $p = .001$). No significant differences in amplitude were found between the Large Face and Large Crystallized conditions or between the Small Face and Small Crystallized conditions. Furthermore, no significant differences were observed in the N170 latencies for any of the four conditions.

The ANOVA used to examine N170 amplitudes at electrode Cz indicated significant differences between the four conditions, $F(3,57) = 15.819$, $p < .001$. Post-hoc tests indicated significantly larger amplitude for the Large Face condition than the two small image conditions (Small Face $p < .002$; Small Crystallized $p < .004$), and larger amplitude for the Large Crystallized condition than the two small image conditions (Small Face $p <. 008$; Small Crystallized $p <.001$). No significant differences in amplitude were found between the Large Face and Large Crystallized conditions or between the Small Face and Small Crystallized conditions. Furthermore, no significant differences were observed in the N170 latencies for any of the four conditions.

3.5.3. N400

The ANOVA used to examine N400 amplitudes at electrode Pz showed significant differences between the four conditions, $F(3,57) = 18.260$, $p < .001$. Post-hoc tests indicated a significantly larger amplitude for the Small Face condition than the other three conditions (Large Face condition $p < .001$; Large Crystallized condition $p < .001$; Small Crystallized $p = .001$). No significant differences were observed, however, in latencies for any of the four conditions.

The ANOVA used to examine N400 amplitudes at electrode Cz indicated significant differences between the four conditions, $F(3,57) = 31.397$, $p < .001$. Post-hoc tests indicated a significantly larger amplitude for the Small Face condition than the other three conditions (Large Face $p < .001$; Large Crystallized $p < .001$; Small Crystallized $p = .001$). Significantly larger amplitude was also observed for the Large Crystallized condition than the Large Face condition, ($p = .002$), and the Small Crystallized condition, $p = .008$.

The ANOVA was used to examine N400 latencies at electrode Cz between the four conditions, $F(3,57) = 4.579$, $p = .010$. Post-hoc tests, however, indicated no significant differences between the four conditions.

4. Discussion

Previous studies have proposed that the additional information that facial stimuli provide, the N170 and N400 components in addition to the P300 component, should increase the speed and accuracy of character selection [11,14,15]. To further investigate the factors contributing to these effects, we conducted a study manipulating and testing the size of facial stimuli used in previous studies, and compared these stimuli to a control image. We observed that the facial stimuli used in previous studies were larger than the characters in the matrix themselves, and hypothesized that perhaps the size of the image had an impact on performance. Furthermore, we felt that the control stimuli used in other studies may not be appropriate because the stimuli were not matched to the facial stimuli for brightness and contrast. By creating a new control image, we believe we were better able to understand how important the face itself is for increasing speed and accuracy. Having four conditions, Large Face, Small Face, Large Crystallized, and Small Crystallized, we could examine the possible effect of size of stimuli on BCI performance. Based on previous literature, we predicted that that the facial stimuli would lead to better performance than the non-facial stimuli. We also predicted that the larger stimuli would lead to better performance than the smaller stimuli.

Contrary to our hypotheses, we found no significant differences in performance between any of the four conditions. Measures of accuracy, target flashes required to make a character selection, selections per minute, and bit rate all indicated that the facial stimuli did not produce significantly better BCI performance than non-facial stimuli, and that the larger stimuli did not produce significantly better BCI performance than smaller stimuli. Although the performance results were not consistent with our hypotheses, our findings were also not consistent with previous literature that has found that facial stimuli lead to significantly better BCI performance than non-facial stimuli. Furthermore, we conducted offline analyses to determine whether a dynamic stopping algorithm would reveal differences between conditions. Offline analyses supported our findings and did not result in statistical differences in any of the four conditions.

Despite no significant differences in BCI performance across the four conditions, each of the four images produced distinctly different ERPs. Waveform analyses indicated that the face conditions produced more robust P300 ERPs than the two crystallized conditions, similar to the results observed by Kaufmann et al. [11,14]. In the examination of the N170, which is traditionally considered a facial-specific component, we observed a more robust N170 to the large stimuli than the N170 elicited by the small stimuli. There was not a significant difference between the face images and the crystallized images; this finding is similar to the findings of Kaufmann et al. [11], who compared familiar faces to classic character flashing and a pixelated version of the face image. Although we did not include a comparison of N170 responses produced by classic character flashing, face and crystallized images, this finding is also consistent with the findings of Kaufmann et al. [11]. This finding was unexpected, as we observed N170 responses to the crystallized images. The N170 elicited by the non-face images may be explained by the fact that early ERPs are often considered to reflect processing of the physical properties of a stimulus [26]. It is possible when participants viewed the crystallized images they were trying to determine whether the images resembled a face, which may explain the observed early N170 response to the crystallized images.

In addition to a 'face-specific' N170, we only observed a significantly larger N400 response to the small face stimulus, which is only partially consistent with our hypothesis. The N400 response has traditionally been considered to be associated with processing semantic information [27], and has also been observed when viewing 'familiar' and 'unfamiliar' faces [28].

Thus, based on the results of the current study, the relationship between BCI performance and facial stimuli remains unclear. Further investigation into the use of novel stimuli to elicit additional ERPs for classification may provide a better understanding of how facial stimuli affect BCI performance.

4.1. Limitations

It is possible that using an undergraduate sample to test our hypotheses may have had an influence on the results of our findings, such that testing with an ALS population may lead to different findings and conclusions. ALS patients have been shown to have lower P300 amplitudes than healthy individuals [8]. We may have created a ceiling effect, as most healthy undergraduate participants are already good at operating the BCI without the added facial stimuli. Finding no differences in performance between the four conditions may be due to participants being already efficient enough at the task that they would achieve high levels of performance regardless of the type of stimuli presented.

Furthermore, previous studies using facial stimuli utilized row-column flash patterns, whereas we utilized the checkerboard flash pattern developed by [29]. The checkerboard flash pattern eliminates the problem of adjacent character flashes that can impact BCI performance when using row-column. Due to the increase in performance when using checkerboard compared to row-column, it is possible that the high performance produced by the checkerboard flash pattern could overshadow the effects of using facial stimuli, thus possibly making it unnecessary to use facial stimuli in addition to a checkerboard flash pattern.

4.2. Future directions

Replication of this study using a patient sample could provide greater insight into the use of facial stimuli for BCI operation. The healthy undergraduate sample showed comparable BCI performance in all four conditions, whereas using a sample of patients, who have been shown to produce ERPs different from healthy participants, could show greater performance due to facial stimuli.

Facial images appear to produce different P300 ERPs compared to traditional character flashes, but not N170 and N400 ERPs. Perhaps there are alternative classes of stimuli that may elicit more differential ERPs than images of faces. Kaufmann et al. [23] introduced the idea of a two-stimulus presentation paradigm using images of faces and symbols. The use of facial stimuli and an alternative class of stimuli used to elicit additional ERPs to the P300 may lead to better target discrimination in a two-stimulus paradigm than images of faces and stimuli.

5. Conclusions

The goal of the current study was to test hypotheses as to why the Face Speller is superior to the traditional flashing P300 speller. Most P300 speller studies have a narrow focus – to improve the speed and accuracy of the system. Improvements are achieved through paradigm manipulation or signal-processing techniques. Often, the work is exploratory in nature. The Face Speller is different than most previous work in that face stimuli have been shown to elicit components related to processing of facial information, specifically the N170 and N400 components. The rationale driving the Face Speller paradigm is that these additional components will be elicited by facial stimuli and provide additional discriminatory information to be used by the classification algorithm. Thus, the additional features will result in superior performance of the Face Speller as compared to the traditional flashing of characters.

The current work is a departure from previous work focusing on paradigm manipulation and signal-processing improvements. In this study we examined hypotheses as to why the Face Speller is superior to the standard flashing paradigm. Face Speller studies typically use facial stimuli that are larger than the flashing character stimuli. Thus, we hypothesized that the size of the stimuli – not the content – may be responsible for improved performance. To test this hypothesis we predicted, regardless of content, larger stimuli would provide higher performance. In addition, based on prior evidence, we also predicted that facial stimuli would provide higher performance than non-facial stimuli. To test these hypotheses we included four conditions in the study: large facial stimuli, small facial stimuli, large contrast-matched non-face stimuli, and small contrast-matched non-face stimuli. The results of our study are significant in that the face-specific ERPs that are hypothesized to contribute to better target discrimination were also observed in response to non-facial images and we did not find differences in speed and accuracy between any of the four conditions. Further investigation is necessary to determine why previous research has found facial stimuli to be so effective.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

[1] Velliste M, Perel S, Spalding MC, et al. Cortical control of a prosthetic arm for self-feeding. Nature. 2008;453(7198):1098–1101. doi:10.1038/nature06996. PubMed PMID: 18509337.

[2] Munssinger JI, Halder S, Kleih SC, et al. Brain painting: first evaluation of a new brain-computer interface application with ALS-patients and healthy volunteers. Front Neurosci. 2010;4:1–11. doi:10.3389/fnins.2010.00182. PubMed PMID: 21151375; PubMed Central PMCID: PMCPMC2996245.

[3] Kübler A, Kotchoubey B, Kaiser J, et al. Brain–computer communication: unlocking the locked in. Psychol Bull. 2001 May;127(3):358–375. PubMed PMID: 11393301.

[4] Vaughan TM, Mcfarland DJ, Schalk G, et al. The Wadsworth BCI research and development program: at home with BCI. IEEE Trans Neural Syst Rehabil Eng. 2006;14(2):229–233. doi:10.1109/TNSRE.2006.875577. PubMed PMID: 16792301.

[5] Bentin S, Allison T, Puce A, et al. Electrophysiological Studies of Face Perception in Humans. J Cogn Neurosci. 1996;8(6):551–565. doi:10.1162/jocn.1996.8.6.551. PubMed PMID: 20740065; PubMed Central PMCID: PMCPMC2927138.

[6] Eimer M. The face-specific N170 component reflects late stages in the structural encoding of faces. NeuroReport. 2000;11(10):2319–2324. PubMed PMID: 10923693.

[7] Eimer M. Event-related brain potentials distinguish processing stages involved in face perception and recognition. Clin Neurophysiol. 2000;111(4):694–705. PubMed PMID: 10727921.

[8] Kutas M, Federmeier KD. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). Annu Rev Psychol. 2011;62(1):621–647. doi:10.1146/annurev.psych.093008.131123. PubMed PMID: 20809790; PubMed Central PMCID: PMCPMC4052444.

[9] Jin J, Daly I, Zhang Y, et al. An optimized ERP brain-computer interface based on facial expression changes. J Neural Eng. 2014;11(3):036004. doi:10.1088/1741-2560/11/3/036004. PubMed PMID: 24743165.

[10] Chen L, Jin J, Zhang Y, et al. A survey of the dummy face and human face stimuli used in BCI paradigm. J Neurosci Methods. 2015;239(239):18–27. doi:10.1016/j.jneumeth.2014.10.002. PubMed PMID: 25314905.

[11] Kaufmann T, Schulz SM, Grünzinger C, et al. Flashing characters with famous faces improves ERP-based brain-computer interface performance. J Neural Eng. 2011;8(5):056016. doi:10.1088/1741-2560/8/5/056016. PubMed PMID: 21934188.

[12] Takano K, Ora H, Sekihara K, et al. Coherent activity in bilateral parieto-occipital cortices during P300-BCI Operation. Front Neurol. 2014;5:1–17. doi:10.3389/fneur.2014.00074. PubMed PMID: 24860546; PubMed Central PMCID: PMCPMC4030183.

[13] Zhang Y, Zhao Q, Jing J, et al. A novel BCI based on ERP components sensitive to configural processing of human faces. J Neural Eng. 2012;9(2):026018. doi:10.1088/1741-2560/9/2/026018. PubMed PMID: 22414683.

[14] Kaufmann T, Schulz SM, Köblitz A, et al. Face stimuli effectively prevent brain-computer interface inefficiency in patients with neurodegenerative disease. Clin Neurophysiol. 2013;124(5):893–900. doi:10.1016/j.clinph.2012.11.006. PubMed PMID: 23246415.

[15] Kaufmann T, Kübler A. Beyond maximum speed–a novel two-stimulus paradigm for brain-computer interfaces based on event-related potentials (P300-BCI). J Neural Eng. 2014;11(5):056004. doi:10.1088/1741-2560/11/5/056004. PubMed PMID: 25080406.

[16] Geronimo AM, Simmons Z. The P300 'face' speller is resistant to cognitive decline in ALS. Brain-Computer Interfaces. 2017;4(4). doi:10.1080/2326263X.2017.1338013.

[17] Colwell KA, Ryan DB, Throckmorton CS, et al. Channel selection methods for the P300 Speller. J Neurosci Methods. 2014;232(232):6–15. doi:10.1016/j.jneumeth.2014.04.009. PubMed PMID: 24797224; PubMed Central PMCID: PMCPMC4106671.

[18] Schalk G, McFarland DJ, Hinterberger T, et al. BCI2000: a general-purpose brain-computer interface (BCI) system. IEEE Trans Biomed Eng. 2004;51(6):1034–1043. doi:10.1109/TBME.2004.827072. PubMed PMID: 15188875.

[19] Draper NR, Smith H. Applied regression analysis. New York, NY: Wiley; 1981.

[20] Farwell LA, Donchin E. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. Electroencephalogr Clin Neurophysiol. 1988;70(6):510–523. PubMed PMID: 2461285.

[21] Sellers EW, Donchin E. A P300-based brain–computer interface: initial tests by ALS patients. Clin Neurophysiol. 2006;117(3):538–548. doi:10.1016/j.clinph.2005.06.027. PubMed PMID: 16461003.

[22] Townsend G, LaPallo BK, Boulay CB, et al. A novel P300-based brain–computer interface stimulus presentation paradigm: Moving beyond rows and columns. Clin Neurophysiol. 2010;121(7):1109–1120. doi:10.1016/j.clinph.2010.01.030. PubMed PMID: 20347387; PubMed Central PMCID: PMCPMC2879474.

[23] Furdea A, Halder S, Krusienski DJ, et al. An auditory oddball (P300) spelling system for brain-computer interfaces. Psychophysiology. 2009;46(3):617–625. doi:10.1111/j.1469-8986.2008.00783.x. PubMed PMID: 19170946.

[24] Balota DA, Yap MJ, Hutchison Keith AMJ, et al. The English Lexicon project. Behav Res Methods. 2007;39(3):445–459. PubMed PMID: 17958156.

[25] Pierce J. An introduction to information theory. New York, NY: Dover Press; 1980.

[26] Hillyard SA, Kutas M. Electrophysiology of cognitive processing. Annu Rev Psychol. 1983;34:33–61. doi:10.1146/annurev.ps.34.020183.000341. PubMed PMID: 6338812.

[27] Kutas M, Hillyard SA. Reading senseless sentences: brain potentials reflect semantic incongruity. Science. 1980;207(4427):203–205. PubMed PMID: 7350657.

[28] Bentin S, Deouell LY. Structural encoding and identification in face processing: ERP evidence for separate mechanisms. Cogn Neuropsychology. 2000;17(1-3):35–55. doi:10.1080/026432900380472. PubMed PMID: 20945170.