Flashing characters with famous faces improves ERP-based brain–computer interface performance

2011 J. Neural Eng. 8 056016

(http://iopscience.iop.org/1741-2552/8/5/056016)

IOP Publishing — Journal of Neural Engineering
J. Neural Eng. **8** (2011) 056016 (10pp) doi:10.1088/1741-2560/8/5/056016

Flashing characters with famous faces improves ERP-based brain–computer interface performance

T Kaufmann, S M Schulz[1], C Grünzinger and A Kübler

Department of Psychology I, University of Würzburg, Marcusstr. 9-11, 97070 Würzburg, Germany

E-mail: schulz@psychologie.uni-wuerzburg.de

Received 20 April 2011
Accepted for publication 4 August 2011
Published 20 September 2011
Online at stacks.iop.org/JNE/8/056016

Abstract
Currently, the event-related potential (ERP)-based spelling device, often referred to as P300-Speller, is the most commonly used brain–computer interface (BCI) for enhancing communication of patients with impaired speech or motor function. Among numerous improvements, a most central feature has received little attention, namely optimizing the stimulus used for eliciting ERPs. Therefore we compared P300-Speller performance with the standard stimulus (flashing characters) against performance with stimuli known for eliciting particularly strong ERPs due to their psychological salience, i.e. flashing familiar faces transparently superimposed on characters. Our results not only indicate remarkably increased ERPs in response to familiar faces but also improved P300-Speller performance due to a significant reduction of stimulus sequences needed for correct character classification. These findings demonstrate a promising new approach for improving the speed and thus fluency of BCI-enhanced communication with the widely used P300-Speller.

1. Introduction

Brain–computer interfaces (BCIs) provide an alternative communication channel that does not depend on muscular control and thus, have been found to be particularly useful for patients diagnosed with amyotrophic lateral sclerosis (ALS) who experience a gradual loss of control over voluntary muscular movement due to motor neuron degeneration (Birbaumer *et al* 1999, Kübler *et al* 2001a, Nijboer *et al* 2008).

The ERP-based spelling device, often referred to as P300-Speller, is one of the most commonly used BCIs with healthy participants as well as patients (Farwell and Donchin (1988), Nijboer *et al* (2008); for a review see Kleih *et al* (2011)). About 250 to 400 ms after stimulus onset a marked positive potential deflection can be observed in the event-related potential (ERP, i.e. the P300, Sutton *et al* (1965), Polich *et al* (1997); for a review see Picton (1992) and Polich (2007)). In BCI paradigms its latency can be even as early as 200 ms (Sellers *et al* 2010, Kleih *et al* 2010). P300 peak amplitudes can be recorded best at parietal and central electrodes (Picton 1992). Strong P300s are reliably elicited in the oddball paradigm, in which participants are required to attend to rare (odd) target stimuli in a series of many irrelevant stimuli. In the most common version of the P300-Speller, the participant attends a target character in a matrix consisting of the letters of the alphabet and digits. Then the rows and columns of the matrix are flashed consecutively in random order (row column paradigm (RCP), see also Farwell and Donchin (1988)). An increased P300 will follow flashing of the particular column and row where the attended character is located.

Although healthy participants produce highly accurate results with the P300-Speller (Guger *et al* 2009, Kleih *et al* 2010), accuracy is still an issue in patients, because P300 amplitudes are typically lower than in healthy controls (Hanagasi *et al* 2002). For efficient communication maximizing the speed of character selection without increasing error rates is the primary goal. Many researchers have addressed this issue by improving (1) the quality of signal processing and the development of novel classification techniques (e.g. Krusienski *et al* (2006), (2008), Blankertz *et al* (2011); for a review see Lotte *et al* (2007)) and (2)

[1] Author to whom any correspondence should be addressed.

procedural aspects of the oddball paradigm. Examples for the second approach are studies comparing how stimulus rate (McFarland *et al* 2011), inter-stimulus interval (ISI) and matrix size (Sellers *et al* 2006a), character size and inter-character distance (Salvaris and Sepulveda 2009), background color (Salvaris and Sepulveda 2009), chromatic differences in the flashing pattern (Takano *et al* 2009) and the flashing pattern itself (e.g. Townsend *et al* (2010), Frye *et al* 2011, Guger *et al* (2009), Jin *et al* (2010)) affect the accuracy level.

Altogether these studies suggest several modifications to the paradigm that may help to increase classification rates, bit rates or user acceptance, or both of the P300-Speller. Optimizing the core feature of the P300-Speller, that is reliably eliciting an ERP by means of flashing items in a matrix, has almost never been scrutinized. To our knowledge, only one study has tried to improve spelling performance by changing the immediate properties of this stimulus (Martens *et al* 2009). Instead of flashing characters, gray rectangles were flipped horizontally to elicit ERPs. This so called FLIP stimulus was less susceptible to refractory effects in ERPs. Thus, the bit rate may be slightly increased by decreasing the target-to-target intervals.

A more direct approach for improving classification accuracy is to optimize the immediate properties of the stimulus used for eliciting ERPs to increase the signal-to-noise ratio in the electro-encephalogram (EEG). This would allow for improved classification based on fewer ERPs, that is fewer stimulus repetitions (sequences). The latter would increase the bit rate and thus, enhance the fluency of communication. This goal may be accomplished by introducing a new stimulus type for flashing the characters which is known to elicit particularly strong ERPs. Ideal candidates for such a stimulus may be familiar faces. Numerous studies in the field of human face processing have revealed that the visual perception of familiar faces strongly involves several ERPs, specifically the N170, P300 and N400f. The N170 occurs between 130 and 200 ms post-stimulus and appears to be a face-specific ERP (Bentin *et al* 1996, Eimer 2000). It is assumed that the N170 is involved in rapid perception of faces but not in face recognition (Eimer 2000). Familiar faces further elicit a face-specific N400, the N400f, between 300 and 500 ms post-stimulus, which is located on parietal and central electrode sites (Eimer 2000). The recognition of familiar faces further involves a late positive component (P300/LPC, Henson *et al* (2003)). Classification in an ERP-BCI is not necessarily limited to the P300 component and other ERPs may also contribute (e.g. Bianchi *et al* (2010), Treder and Blankertz (2010), Allison and Pineda (2003, 2006)). Thus, presentation of familiar faces, which strongly involves different ERP components, may be exploited for improving classification. In particular using faces well known to everybody in a given culture should lead to strong and relatively stable effects across individuals.

Thus, we investigated whether familiar faces of famous individuals may serve as applicable stimuli for flashing items in the character matrix of the classic P300-Speller. We hypothesized that eliciting ERPs involved in face processing would lead to better classification rates as compared to simple letter highlighting, due to better signal determination within the response window.

2. Methods

2.1. Participants

Participants were $N = 21$ university students with no known neurological disorders and normal or corrected-to-normal vision. Due to noisy EEG, data from 1 participant were discarded. The remaining sample comprised $N = 20$ participants (10 men; mean age 24.6 years, $SD = 4.55$, range 19–41). All were BCI novices and signed informed consent prior to the study. The experiment was conducted in accordance with standard ethical guidelines as defined by the Declaration of Helsinki (World Medical Association) and the European Council's Convention for the Protection of Human Rights and Dignity of the Human Being with regard to the Application of Biology and Medicine (Convention on Human Rights and Biomedicine).

2.2. Equipment and data acquisition

EEG was recorded from 12 passive Ag/AgCl electrodes, as shown in figure 1, with mastoid ground and reference. Electro-oculogram (EOG) was acquired from 2 vertical (vEOG) and 2 horizontal (hEOG) passive Ag/AgCl sintered electrodes. All signals were amplified using a 16-channel g.USBamp amplifier (*g.tec Medical Engineering GmbH*, *Austria*) and recorded at a sampling rate of 512 Hz. Data were collected with the software tool BCI2000 (Schalk *et al* 2004).

2.3. The spelling paradigm

The spelling paradigm was implemented as a multi-thread application in Python 2.5 and connected to BCI2000 via user datagram protocol. Spelling characters were presented in a 6 × 6 matrix and rows and columns of the matrix were flashed consecutively in random order (RCP). We compared the classic character flashing (CF) to a novel face flashing (FF) condition. To ensure equal timing of stimulus presentation, both conditions were implemented and controlled with the same hardware and software implementation. The CF condition defaulted to the traditional approach (i.e. as originally described by Farwell and Donchin (1988)). In the FF condition, we overlaid rows and columns of the characters and digit matrix with well-known pictures of famous faces (Albert Einstein or Ernesto 'Che' Guevara). Pictures were semi-transparent to allow for uninterrupted focusing on the target letter while flashing the face stimuli. In addition to the goal of increasing the psychological salience of the stimulus, facial stimuli may differ from flashing characters also with regard to physical parameters such as brightness, contrast or chromatic differences. Chromatic differences, for example, were recently found to influence performance in the P300-BCI (Takano *et al* 2009). Therefore, we also implemented a control condition using a random reorganization of all pixels of the FF pictures, resulting in meaningless images with the exact same brightness, contrast and chromatic properties as compared to the FF condition. Consequently, the remaining difference between this pixelated face flashing (PFF) condition and FF is the configuration and semantic properties of the picture that

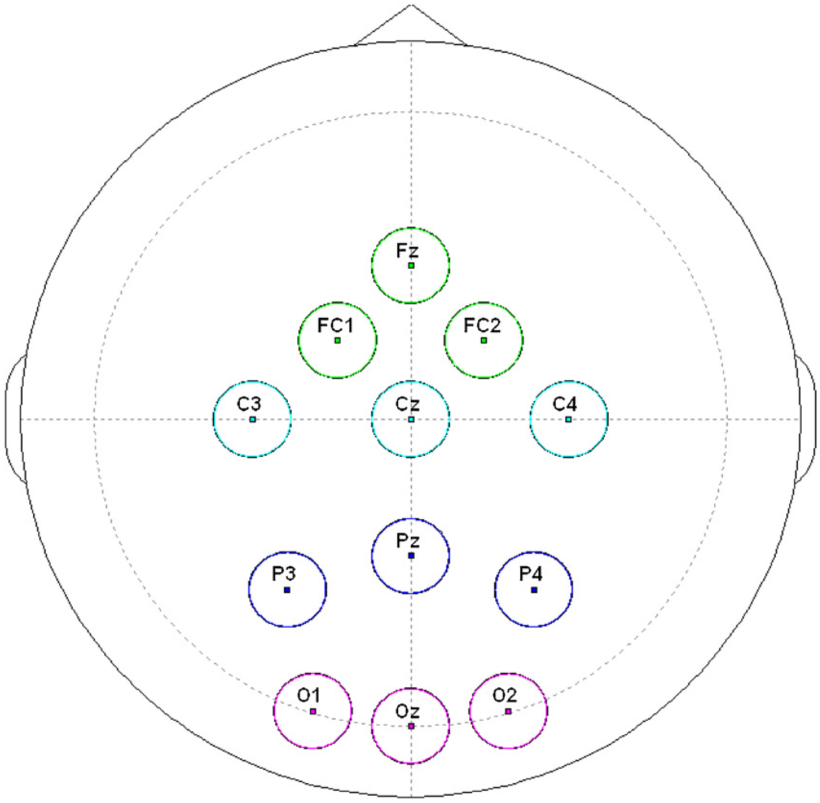

Figure 1. Electrode setup consisting of 12 passive Ag/AgCl electrodes. Locations were Fz, FC1, FC2, C3, Cz, C4, P3, Pz, P4, O1, Oz and O2.

is a prominent face. In total this resulted in five conditions: one CF condition, two FF conditions (Einstein versus 'Che' Guevara) and two PFF conditions (pixelated Einstein versus pixelated 'Che' Guevara). Figure 2 shows examples of the CF, FF and PFF conditions.

2.4. Procedure

Participants were seated comfortably in a chair approximately 70 cm distant from a computer screen (size: 19″, resolution: 1280 × 1024 pixel, refresh rate: 60 Hz, size of presented matrix: 30 × 20 cm^2). After informed consent was signed, they were instructed to focus on the target character, avoid blinking during stimulus presentation and to count the number of stimulus flashes. A trial necessary to select a target (i.e. letter to spell) comprised 15 flashes of each row and each column (sequences). The speller software was set to default values for the ISI of 125 ms and a flash duration of 31.25 ms, resulting in 33.33 ms real flash duration and 133.33 ms ISI at 60 Hz screen refresh rate. All participants spelled the word 'BRAIN' in each of the five conditions. The order of condition was permuted across participants to control for potential habituation effects. This procedure was repeated once, resulting in 2 × 5 runs in total. No feedback on the correctness of spelling was provided to the participants during any of the runs.

To avoid novelty effects, prior to each run participants were familiarized for 15 s with the stimulus used for flashing characters, i.e. the face picture in an FF condition. After the run, they rated on a 10-point Likert-like scale from 0 (not familiar at all) to 10 (extremely familiar) how familiar the stimulus appeared to them. Scales were presented with E-Prime® (*Psychology Software Tools*, *USA*).

2.5. Artifact rejection and data processing

EEG data were filtered with 0.1 Hz high and 30 Hz low pass. EEG was corrected for ocular artifacts with both hEOG and vEOG recordings using an algorithm introduced by Gratton *et al* (1983). These preprocessing steps were performed in Brain Vision Analyzer® (*Brain Products GmbH*, *Germany*). Data were then exported to Matlab® (*The MathWorks*, *USA*) for all further data analysis including signal classification. ERP amplitudes were determined as the maximum (positive potentials) or minimum (negative potentials) voltage in a defined time window.

This study used short ISIs of 125 ms, which bears the risk of overlapping ERPs. Regarding BCI performance, however, it could be demonstrated that short (175 ms) ISIs yield better performance results as compared to longer (350 ms) ISIs (Sellers *et al* 2006a). Furthermore, immediate consecutive flashing of the same target is rare due to the properties of the oddball paradigm. The reason for and advantage of short ISIs is rapid character selection which fosters the applicability of the P300-Speller. However, as we were not only interested in the effect of stimulus presentation on classification accuracy, but also on ERP amplitudes and latencies in different conditions, only target stimuli separated

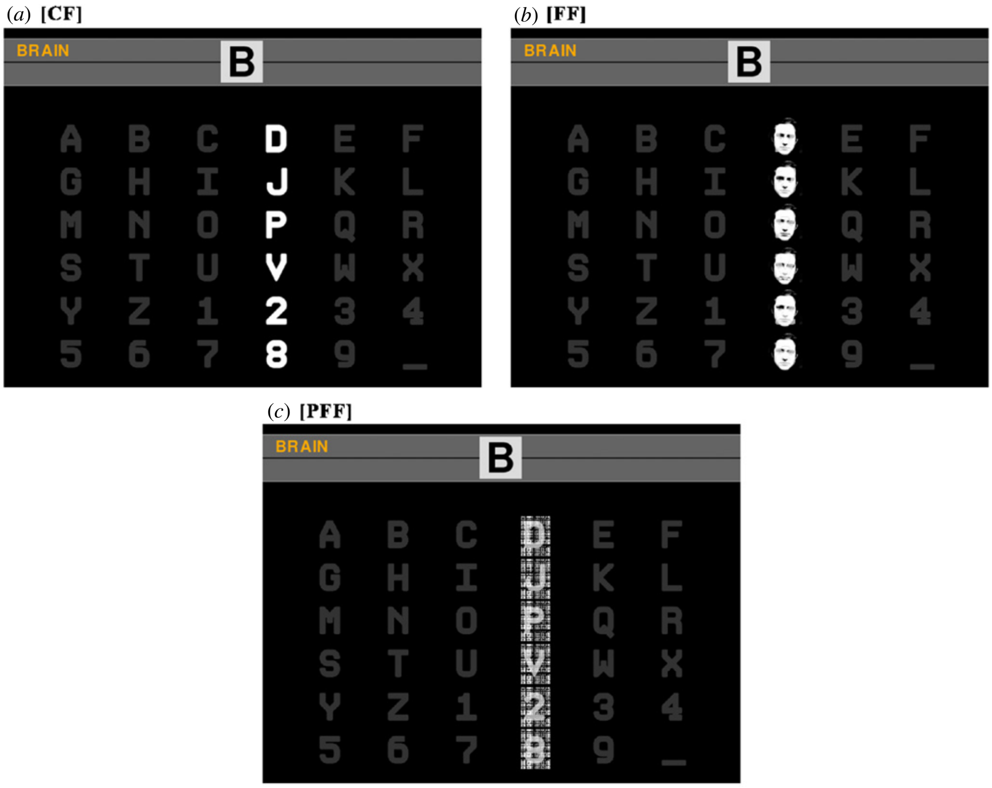

Figure 2. The different conditions of the paradigm. (*a*) The classic P300-Speller (CF), (*b*) the characters with flashing familiar faces (FF) and (*c*) the control condition with all pixels of the face picture after spatial randomization (PFF). Both FF and PFF stimuli were presented semi-transparent such that the characters were still visible. In the experiment we used famous faces which are not shown in this figure due to print license.

by more than two non-target stimuli were included for offline analysis of ERPs, resulting in a minimum time distance of 466 ms between two target stimuli. This correction was performed for ERP analysis only. The classification results comprise the full data set without this correction, thus reflecting the real P300-Speller condition. Classification of target versus non-target stimuli was performed using stepwise linear discriminant analysis (Krusienski *et al* 2006).

2.6. Statistical analysis

Differences in the number of sequences needed for offline classification between the CF and FF conditions and differences in ERP amplitudes and determination coefficients between CF, FF and PFF conditions were statistically validated using repeated measures analysis of variance (ANOVA) (with 'condition' as within subject factor). *Post-hoc* comparisons were performed with Tukey–Kramer tests.

3. Results

All participants completed all scheduled runs. The familiarity of the face stimuli as rated on a 10-point Likert-like scale was high ($M = 8.5$, $SD = 0.7$ for Einstein (FF); $M = 8.2$, $SD = 0.9$ for Che Guevara (FF)). In contrast, the familiarity of the pixelated control was rated low ($M = 2.4$, $SD = 2.3$ for pixelated Einstein (PFF); $M = 2.6$, $SD = 2.0$ for pixelated Che Guevara (PFF)).

3.1. Classification results

In the offline analysis a classifier was trained for the CF and FF conditions. As expected classifier accuracy was increased when faces were used as flashing stimuli (FF) as compared to the classic approach (CF), see figure 3. In the FF conditions, on average fewer sequences ($M = 2.9$, $SD = 1.8$ (Einstein) and $M = 3.0$, $SD = 1.7$ (Che Guevara)) were necessary for achieving the goal of stable 100% classifier accuracy as compared to the classic flashing paradigm ($M = 5.3$, $SD = 3.0$). A repeated measures ANOVA (within factor condition: CF, FF Einstein, FF Che Guevara) confirmed high significance ($F(2, 38) = 7.69$, $p < 0.002$, $\eta^2_{\text{partial}} = 0.29$). Besides the 100% accuracy level (maximum accuracy) we assessed two further levels of accuracy, as comprehensibility of communication is not only possible with the maximum accuracy level. A level of $\geqslant$70% may be regarded as a minimum level of communication (Kübler *et al* 2001b, Kübler and Birbaumer 2008), although communication accuracy not only depends on the amount of falsely spelled characters but also on the influence of the error on the word meaning. We therefore use the $\geqslant$70% as a reference for comparison to demonstrate the potential for improvement of communication with our novel approach. This $\geqslant$70% level was achieved with fewer stimulus sequences in the FF condition ($F(2, 38) = 4.95$, $p < 0.012$, $\eta^2_{\text{partial}} = 0.20$), for the majority of participants already after the first sequence. Secondly, the same but even more pronounced effect was found for a

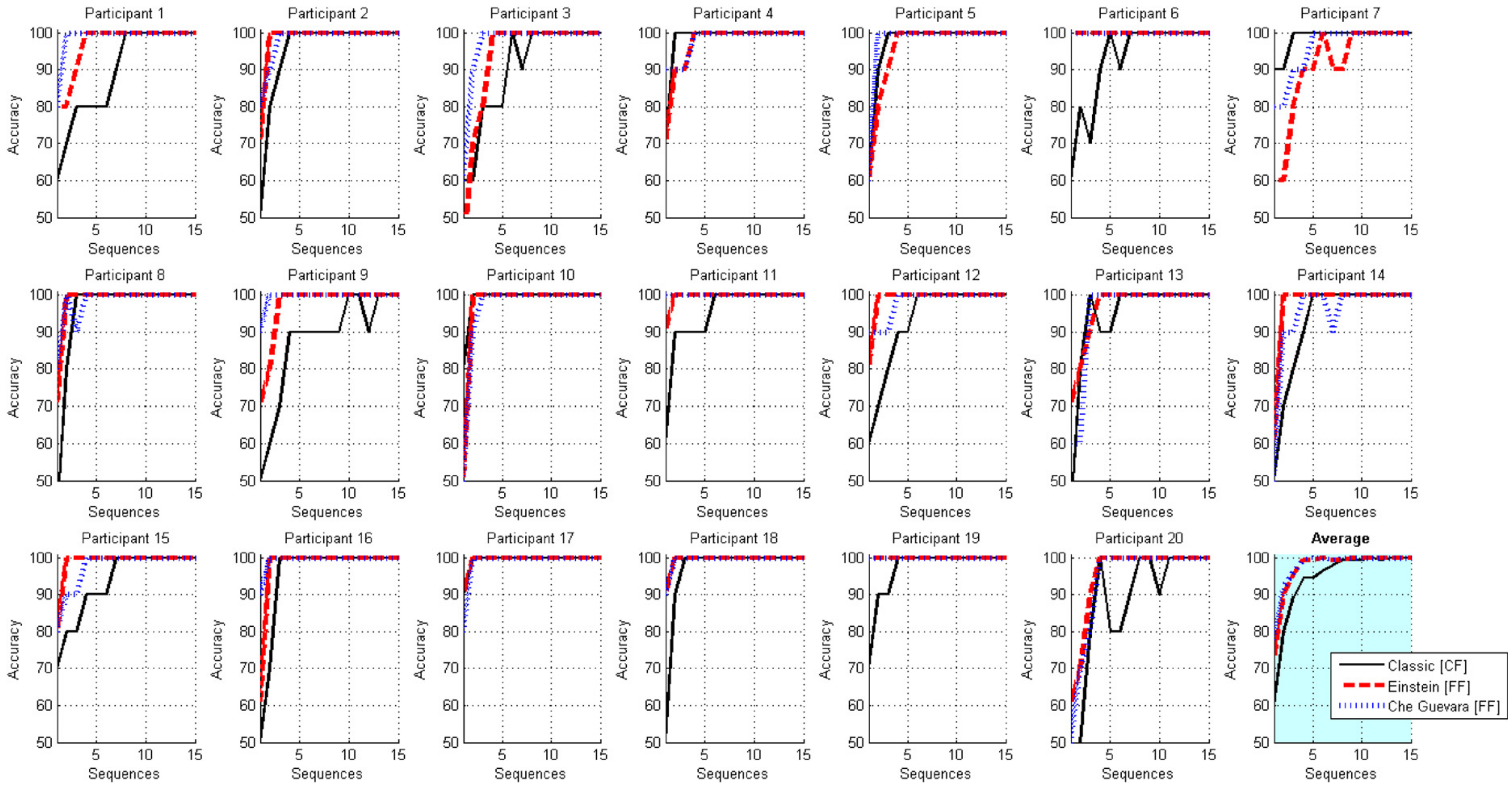

Figure 3. Classifier performance curves for all 15 stimulus intensification sequences. The three conditions ('classic "speller" (CF)', 'Einstein (FF)' and 'Che Guevara (FF)') are plotted for each participant. On the bottom right a grand average over participants is presented. As can be seen, for the majority of participants fewer sequences were necessary to achieve stable 70%, 90% and 100% accuracy levels in the FF conditions compared to CF.

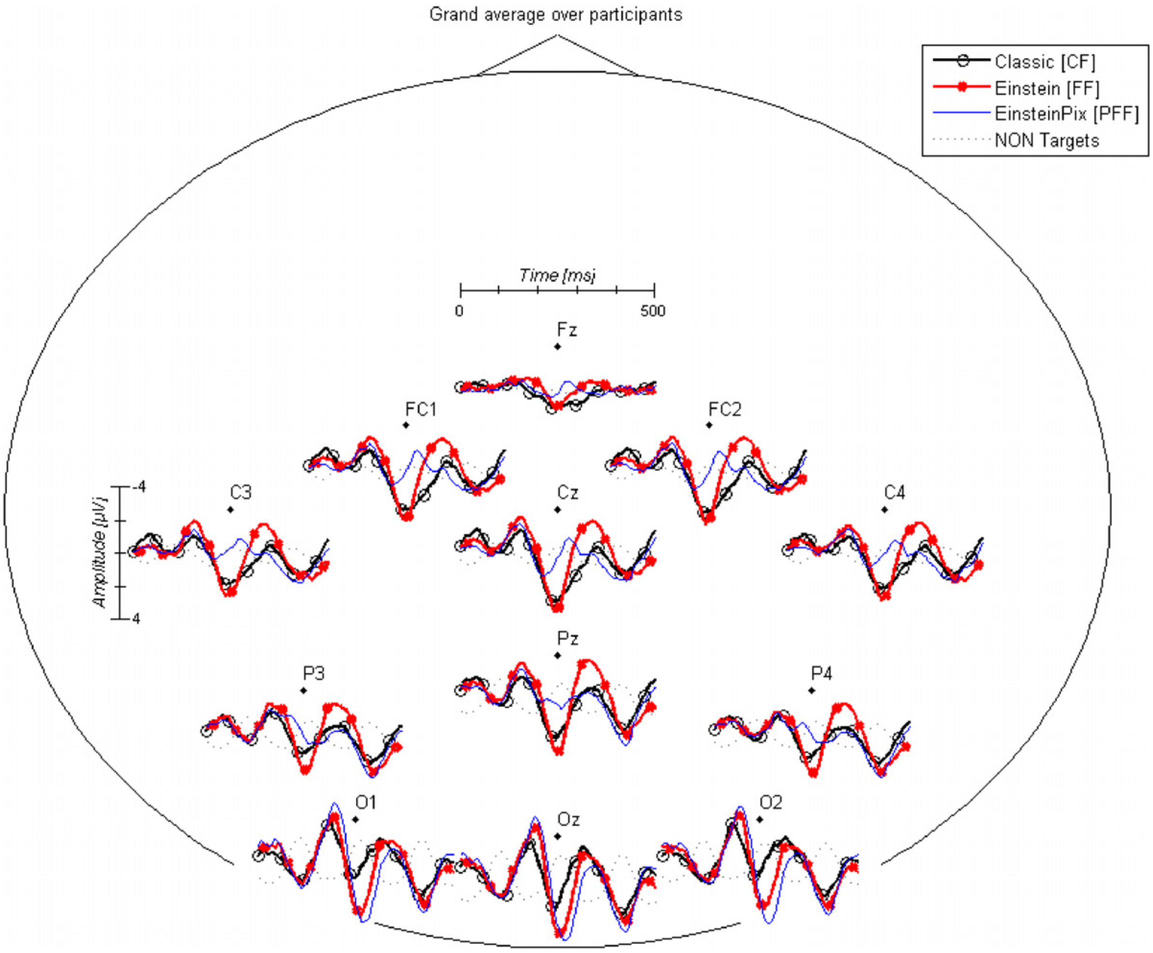

Figure 4. Topographic plot of all 12 electrodes for the three conditions CF, FF and PFF as a grand average over participants. In this case the FF and PFF conditions for the face of Einstein are presented. In the FF condition an N400f was elicited between 300 and 400 ms on centro-parietal electrode sites that was not present in any other condition.

high (>90%) accuracy level ($F(2, 38) = 11.7$, $p < 0.0001$, $\eta^2_{\text{partial}} = 0.38$). *Post-hoc* Tukey–Kramer comparisons revealed no significant difference between the two FF conditions for any of the accuracy levels. The difference between the CF condition and any of the FF conditions was highly significant (Tukey–Kramer: $p < 0.01$) for the 90% and 100% accuracy levels and significant (Tukey–Kramer: $p < 0.05$) for the 70% accuracy level. In sum, the presentation of flashing faces transparently superimposed on characters in the spelling matrix significantly decreased the number of sequences required as compared to simple highlighting of the characters.

3.2. Effects on ERPs

We further examined which ERPs contribute to this improvement in classification accuracy. Figure 4 shows a topographic plot of the event-related brain potential at different electrode sites averaged across all participants. The CF condition elicited a clear P300 with the largest amplitude at Cz. The same held true for the FF conditions (Einstein face in the example presented in figure 4). P300 amplitudes were not significantly different between CF and any of the FF conditions, but significantly decreased for the pixelated control of Einstein ($F(4, 76) = 9.34$, $p < 0.0001$, $\eta^2_{\text{partial}} = 0.32$). Presentation of familiar faces elicited two negative deflections, one between 130 and 200 ms (further referred to as N170 time window) and one between 300 and 400 ms (further referred to as N400f time window) with the maximum peak amplitudes at Pz. For further analysis the maximum peak amplitudes within these time windows were computed at Pz for all five conditions.

In the N170 time window mean peak amplitudes for both FF and PFF conditions were increased as compared to the CF condition ($F(4, 76) = 6.51$, $p < 0.0001$, $\eta^2_{\text{partial}} = 0.25$). In line with our expectations, this difference was highly significant for the Einstein FF condition (Tukey–Kramer: $p < 0.01$) and not significant for the corresponding PFF condition. However, the Che Guevara FF condition failed to reach significance and furthermore the corresponding PFF condition was highly significantly increased (Tukey–Kramer: $p < 0.01$).

For the N400f time window both FF conditions yielded highly significantly increased potentials as compared to CF and to both PFF conditions ($F(4, 76) = 24.17$, $p < 0.0001$, $\eta^2_{\text{partial}} = 0.56$), see figure 4. *Post-hoc* comparison confirmed high significance between CF and FF conditions

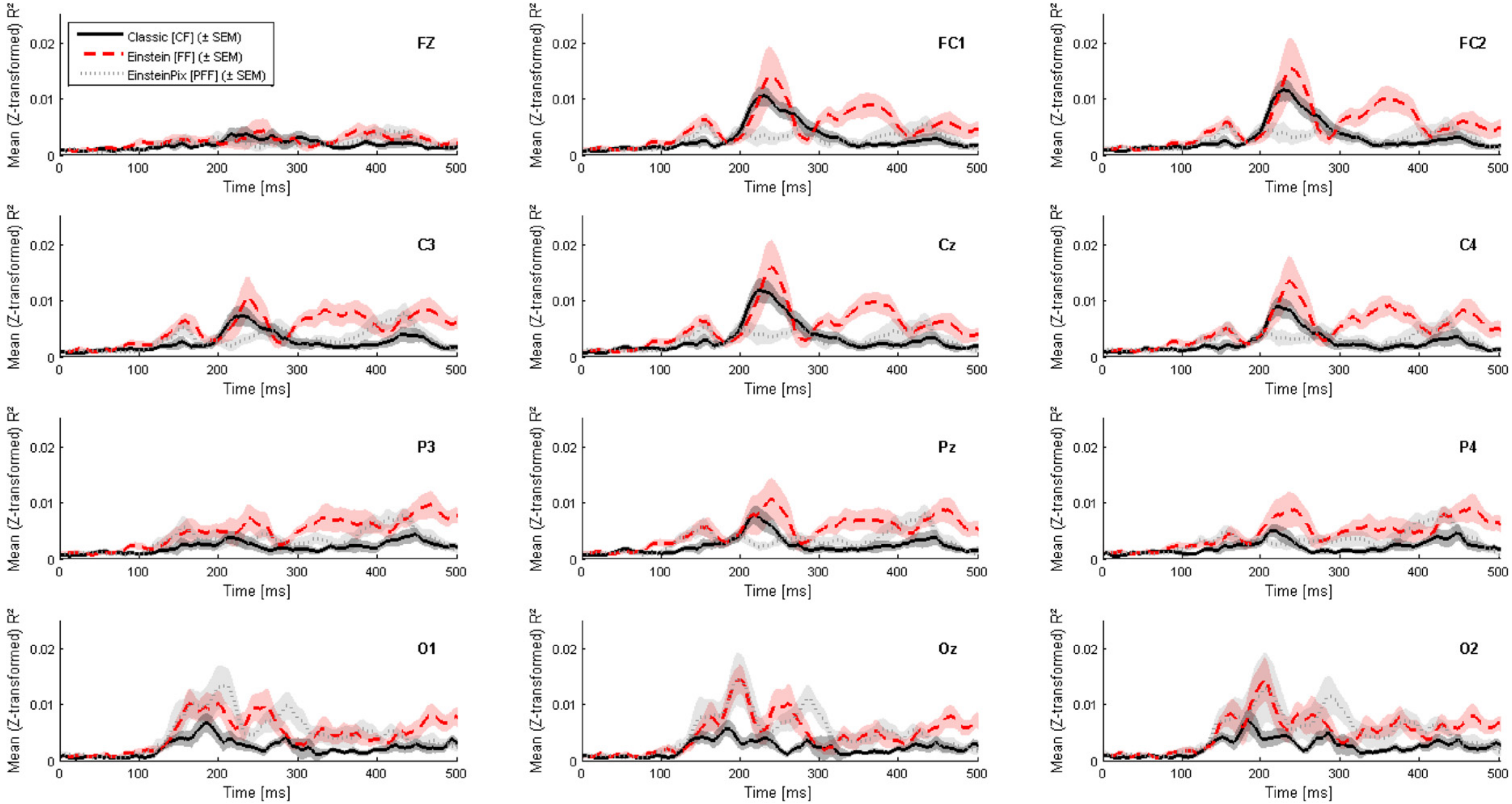

Figure 5. Mean determination coefficients over time for all electrode sites: values were Fisher-*Z* transformed and averaged. Here, the FF and PFF conditions for the face of Einstein are presented. As can be seen for FF and CF conditions the maximum determination was reached in the P300 time window. However, for the FF condition the N400f (i.e. Cz, 300–400 ms) further contributes significantly to determination.

(Tukey–Kramer: $p < 0.0001$). We further assessed this difference by comparing the maximum determination coefficients for the different conditions. Both FF conditions yielded significantly increased (highly significant in the case of Einstein, Tukey–Kramer: $p < 0.01$) maximum determination coefficients as compared to CF and PFF conditions ($F(4, 76) = 4.1$, $p < 0.01$, $\eta^2_{\mathrm{partial}} = 0.18$). As can be seen from figure 5 the maximum determination was found between 200 and 300 ms (the P300) for both FF and the CF condition as well. However the FF conditions (Einstein in figure 5) further yielded significantly increased determination in the N400f time window and a slightly increased determination in the N170 time window.

Taken together, ERP amplitudes and classification accuracy were significantly increased when famous faces were used for flashing as compared to the classic CF.

4. Discussion

In this study we compared famous faces (popular photographic representations of Albert Einstein and Che Guevara) superimposed transparently on the character matrix of the P300-Speller interface (FF) to the classic approach of simply flashing the characters (CF). In addition, we compared ERPs elicited in the FF condition to those elicited by stimuli with equal physical properties but no semantic or configuration information (PFF).

With such optimized stimulus presentation for eliciting ERPs, this study demonstrated increased speed of character selection in the widely used P300-Speller. In particular, the findings suggest that psychological salience of a stimulus can be exploited to elicit particularly strong ERPs not restricted to the P300. Consequently, 100% offline classification accuracy was achieved with significantly fewer sequences. Presentation of famous faces, rated as very familiar, produced (1) more ERPs that contributed to classification and (2) ERPs of larger amplitude which also contributed to better classification between attended versus unattended characters of the spelling matrix.

The significance of this approach becomes apparent when considering that the fluidity of communication by 'typing' clearly depends on the bit rate of character selection (number of correctly selected characters per time unit). Thus, achieving high bit rates is one of the main goals in BCI-based communication. The bit rate depends on both accuracy and speed of character selection. Speed is basically increased by decreasing ISIs (Sellers *et al* 2006a, 2006b) and decreasing the number of stimulus sequences used for averaging (Donchin *et al* 2000, Halder *et al* 2010, Kleih *et al* 2010). Sellers *et al* (2006a) achieved better classification accuracy and bit rate with small ISIs. Unfortunately, decreasing ISIs leads to decreased target-to-target intervals, which resulted in smaller P300 amplitudes and larger latencies (Gonsalvez and Polich 2002). As described in the introduction, Martens *et al* (2009) suggested an attempt to solve this problem. However, this approach is limited by the fact that to capture a certain ERP, a minimum number of data needs to be acquired and refractory periods cannot be completely eliminated. Thus, reducing the duration of single trials is inherently limited. Therefore, to further increase the speed of character selection one has to focus on reducing the number of stimulus sequences used for averaging. Unfortunately, this reduction inevitably decreases the signal-to-noise ratio and thus typically entails a drop in performance. As demonstrated in this study, using stimulus types that elicit particularly pronounced ERP responses serves well as a countermeasure to this issue, leading to an improved signal-to-noise ratio when reducing the number of sequences to average.

Our findings indicate that famous faces could reduce the overall time needed to spell a character on average by a factor of 1.8. Furthermore, at least twice as many of the participants (i.e. two-thirds of all participants) achieved a performance level of 70% already after the first stimulus sequence when using any of the FF paradigms as compared to the CF paradigm. This level has been described as the minimum level for communication (Kübler *et al* 2001b) and illustrates the potential of our new approach. Thus, in the majority of participants the FF paradigm would even allow for increasing the bit rate to maximum, that is flashing rows and columns only once.

For systematic further improvement of the conceptual idea behind the FF condition (i.e. exploiting semantic or psychological salience of stimuli to improve the ERP signal-to-noise ratio), it is important to better understand the reasons for a superior classification performance with well-known faces. Figures 4 and 5 indicate the main difference between all conditions in a time window of 300 to 400 ms at frontal, central and parietal electrodes with significantly higher amplitudes in the FF condition. This represents a face-specific N400f (Eimer 2000). The fact that the amplitudes in the PFF condition remain low supports this notion. Furthermore, there is an increased negative component between 150 and 200 ms in the FF condition. This component may well represent the expected N170. However, the PFF condition also showed increased negativity, which impedes a clear conclusion about this ERP component. Nevertheless, as this component does not appear as strong in the CF condition it constitutes a further signal difference that might allow for improved classification. Finally, the FF condition elicited a P300 at least of the same amplitude as the typical P300 classified in the classic CF paradigm (higher amplitude for Einstein). This effect is further underlined by a trend of higher determination coefficients in the P300 time window in the FF as compared to the CF condition (see figure 5). The fact that PFF conditions show decreased P300 amplitudes may be interpreted as an effect of smaller stimulus complexity and missing stimulus value, which has been shown to also influence P300 amplitudes (Johnson 1984, 1993).

4.1. Limitations and future experimentation

To further establish the significance of our findings, online application and studies with the target patient groups (e.g. those with ALS or other diseases leading to severe motor impairment) are mandatory. Since it has been shown that in a face recognition test ALS patients did perform equally well as age matched healthy controls (Abrahams *et al* 1996) we speculate that our results may well transfer to these patients. Although faces are typically recognized even when heavily degraded, a failure to focus eye-gaze may reduce effects in patients (e.g. Brunner *et al* (2010), Treder and Blankertz (2010)). Furthermore, covert attention to faces should be considered particularly when using the BCI in patient communication. Treder and Blankertz (2010) have suggested resolving issues of covert attention with a specialized character arrangement that clearly outperformed the classic matrix presentation in a covert attention condition. This approach could easily be adapted to our FF paradigm.

The generalizability of our results may be limited because the spelled words were rather short and did not comprise a representative set of characters from the complete alphabet. However, the stability of improvements across individual participants is promising with regard to generalization of effects to other samples.

Furthermore, as ERP latencies vary between participants, overlaps between them might occur. Hence, the N400f may partially cancel out the P300. The early peaking of the P300 in our data as well as reported in other BCI studies in general renders strong interaction relatively unlikely. Moreover, our findings indicate that the N400f contributed in combination with the P300 to classification accuracy in the FF conditions. Therefore, eliciting a N400f proved useful even if it would have cancelled out some P300 activity.

While there is ample evidence that the general P300-BCI paradigm is applicable for a wide range of individuals including patients, it is less clear whether all individuals will show reliably increased ERPs, in particular the N400f, in response to well-known faces. For example, our results indicate minor differences between the two FF conditions tested in this experiment. This may be attributable to physical properties of the stimuli. Thus a broader variety of face stimuli should be validated in terms of (1) familiarity, (2) brightness, (3) contrast and (4) chromatic differences. However, stimulus content may also influence the ERPs that are the semantic or psychological properties of the stimuli. Thus, stimuli varying in terms of individual relevance with regard to psychological (e.g. valence, arousal) or semantic properties (e.g. individual attraction) should be examined. For example, a recent study demonstrated that celebrities elicit a stronger N400f than unknown faces, but the strongest ERPs were elicited by pictures of family members (Touryan *et al* 2011). This could not only serve the goal of further improving classification but might as well increase patient acceptance of the BCI because it can be performed with personally meaningful and liked pictures. Importantly, implementing personally meaningful stimuli in the P300-BCI is straightforward and feasible, and could thus be provided for all BCI users, allowing an individual to achieve high classification at the maximum bit rate.

5. Conclusion

This study demonstrated that superimposing characters of the P300-Speller with famous faces significantly increases the bit rate. The results from $N = 20$ participants indicated a highly significant improvement of classification rates, primarily based on ERP components that are typically elevated when familiar faces are processed, in particular the N400f. This may have a clinically significant impact by increasing communication speed and accuracy with the P300-Speller in patients with severe motor impairment.

Acknowledgments

This work is supported by the European ICT Programme Projects FP7-224631. This paper only reflects the authors' views and funding agencies are not liable for any use that may be made of the information contained herein. SMS had the idea of using flashing faces in the ERP-BCI; SMS, TK and AK designed the study; TK programmed the speller paradigm; CG and TK collected the data; TK and SMS analyzed the data; TK, SMS, CG and AK discussed the results; TK drafted the manuscript, AK, SMS and CG critically revised the manuscript. All gave their final approval to the version to be published.

References

Abrahams S, Goldstein L H, Kew J J, Brooks D J, Lloyd C M, Frith C D and Leigh P N 1996 Frontal lobe dysfunction in amyotrophic lateral sclerosis. A PET study *Brain* **119** (Pt 6) 2105–20

Allison B Z and Pineda J A 2003 ERPs evoked by different matrix sizes: implications for a brain computer interface (BCI) system *IEEE Trans. Neural Syst. Rehabil. Eng.* **11** 110–3

Allison B Z and Pineda J A 2006 Effects of SOA and flash pattern manipulations on ERPs, performance, and preference: implications for a BCI system *Int. J. Psychophysiol.* **59** 127–40

Bentin S, Allison T, Puce A, Perez E and McCarthy G 1996 Electrophysiological studies of face perception in humans *J. Cogn. Neurosci.* **8** 551–65

Bianchi L, Sami S, Hillebrand A, Fawcett I P, Quitadamo L R and Seri S 2010 Which physiological components are more suitable for visual ERP based brain–computer interface? A preliminary MEG/EEG study *Brain Topogr.* **23** 180–5

Birbaumer N, Ghanayim N, Hinterberger T, Iversen I, Kotchoubey B, Kübler A, Perelmouter J, Taub E and Flor H 1999 A spelling device for the paralysed *Nature* **398** 297–8

Blankertz B, Lemm S, Treder M, Haufe S and Müller K-R 2011 Single-trial analysis and classification of ERP components—a tutorial *NeuroImage* **56** 814–25

Brunner P, Joshi S, Briskin S, Wolpaw J R, Bischof H and Schalk G 2010 Does the 'P300' speller depend on eye gaze? *J. Neural Eng.* **7** 056013

Donchin E, Spencer K M and Wijesinghe R 2000 The mental prosthesis: assessing the speed of a P300-based brain-computer interface *IEEE Transactions on Rehabilitation Engineering: a Publication of the IEEE Engineering in Medicine and Biology Society* **8** 174–9

Eimer M 2000 Event-related brain potentials distinguish processing stages involved in face perception and recognition *Clin. Neurophysiol.* **111** 694–705

Farwell L A and Donchin E 1988 Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials *Electroencephalogr. Clin. Neurophysiol.* **70** 510–23

Frye G E, Hauser C K, Townsend G and Sellers E W 2011 Suppressing flashes of items surrounding targets during calibration of a P-300-based brain–computer interface improves performance *J. Neural Eng.* **8** 025024

Gonsalvez C L and Polich J 2002 P300 amplitude is determined by target-to-target interval *Psychophysiology* **39** 388–96

Gratton G, Coles M G and Donchin E 1983 A new method for off-line removal of ocular artifact *Electroencephalogr. Clin. Neurophysiol.* **55** 468–84

Guger C, Daban S, Sellers E, Holzner C, Krausz G, Carabalona R, Gramatica F and Edlinger G 2009 How many people are able to control a P300-based brain–computer interface (BCI)? *Neurosci. Lett.* **462** 94–8

Halder S, Rea M, Andreoni R, Nijboer F, Hammer E M, Kleih S C, Birbaumer N and Kübler A 2010 An auditory oddball brain–computer interface for binary choices *Clin. Neurophysiol.* **121** 516–23

Hanagasi H A, Gurvit I H, Ermutlu N, Kaptanoglu G, Karamursel S, Idrisoglu H A, Emre M and Demiralp T 2002 Cognitive impairment in amyotrophic lateral sclerosis: evidence from neuropsychological investigation and event-related potentials *Brain Res. Cogn. Brain Res.* **14** 234–44

Henson R N, Goshen-Gottstein Y, Ganel T, Otten L J, Quayle A and Rugg M D 2003 Electrophysiological and haemodynamic correlates of face perception, recognition and priming *Cereb. Cortex* **13** 793–805

Jin J, Horki P, Brunner C, Wang X, Neuper C and Pfurtscheller G 2010 A new P300 stimulus presentation pattern for EEG-based spelling systems *Biomed. Tech. Biomed. Eng.* **55** 203–10

Johnson R Jr 1984 P300: a model of the variables controlling its amplitude *Ann. New York Acad. Sci.* **425** 223–9

Johnson R Jr 1993 On the neural generators of the P300 component of the event-related potential *Psychophysiology* **30** 90–7

Kleih S C *et al* 2011 Out of the frying pan into the fire—the P300 based BCI faces real world challenges *Prog. Brain Res.* **194** at press

Kleih S C, Nijboer F, Halder S and Kübler A 2010 Motivation modulates the P300 amplitude during brain–computer interface use *Clin. Neurophysiol.* **121** 1023–31

Krusienski D, Sellers E, Cabestaing F, Bayoudh S, McFarland D, Vaughan T and Wolpaw J 2006 A comparison of classification techniques for the P300 Speller *J. Neural Eng.* **3** 299–305

Krusienski D J, Sellers E W, McFarland D J, Vaughan T M and Wolpaw J R 2008 Toward enhanced P300 speller performance *J. Neurosci. Methods* **167** 15–21

Kübler A and Birbaumer N 2008 Brain–computer interfaces and communication in paralysis: extinction of goal directed thinking in completely paralysed patients? *Clin. Neurophysiol.* **119** 2658–66

Kübler A, Kotchoubey B, Kaiser J, Wolpaw J R and Birbaumer N 2001a Brain–computer communication: unlocking the locked in *Psychol. Bull.* **127** 358–75

Kübler A, Neumann N, Kaiser J, Kotchoubey B, Hinterberger T and Birbaumer N P 2001b Brain–computer communication: self-regulation of slow cortical potentials for verbal communication *Arch. Phys. Med. Rehabil.* **82** 1533–9

Lotte F, Congedo M, Lécuyer A, Lamarche F and Arnaldi B 2007 A review of classification algorithms for EEG-based brain–computer interfaces *J. Neural Eng.* **4** R1–13

Martens S M M, Hill N J, Farquhar J and Schölkopf B 2009 Overlap and refractory effects in a brain–computer interface speller based on the visual P300 event-related potential *J. Neural Eng.* **6** 026003

McFarland D J, Sarnacki W A, Townsend G, Vaughan T and Wolpaw J R 2011 The P300-based brain–computer interface (BCI): effects of stimulus rate *Clin. Neurophysiol.* **122** 731–7

Nijboer F *et al* 2008 A P300-based brain–computer interface for people with amyotrophic lateral sclerosis *Clin. Neurophysiol.* **119** 1909–16

Picton T W 1992 The P300 wave of the human event-related potential *J. Clin. Neurophysiol.* **9** 456–79

Polich J 2007 Updating P300: an integrative theory of P3a and P3b *Clin. Neurophysiol.* **118** 2128–48

Polich J, Alexander J E, Bauer L O, Kuperman S, Morzorati S, O'Connor S J, Porjesz B, Rohrbaugh J and Begleiter H 1997 P300 topography of amplitude/latency correlations *Brain Topogr.* **9** 275–82

Salvaris M and Sepulveda F 2009 Visual modifications on the P300 speller BCI paradigm *J. Neural Eng.* **6** 046011

Schalk G, McFarland D J, Hinterberger T, Birbaumer N and Wolpaw J R 2004 BCI2000: a general-purpose brain–computer interface (BCI) system *IEEE Trans. Biomed. Eng.* **51** 1034–43

Sellers E W, Krusienski D J, McFarland D J, Vaughan T M and Wolpaw J R 2006a A P300 event-related potential brain–computer interface (BCI): the effects of matrix size and inter stimulus interval on performance *Biol. Psychol.* **73** 242–52

Sellers E W, Kübler A and Donchin E 2006b Brain–computer interface research at the University of South Florida Cognitive Psychophysiology Laboratory: the P300 Speller *IEEE Trans. Neural Syst. Rehabil. Eng.* **14** 221–4

Sellers E W, Vaughan T M and Wolpaw J R 2010 A brain–computer interface for long-term independent home use *Amyotrophic Lateral Scler.* **11** 449–55

Sutton S, Braren M, Zubin J and John E R 1965 Evoked-potential correlates of stimulus uncertainty *Science* **150** 1187–8

Takano K, Komatsu T, Hata N, Nakajima Y and Kansaku K 2009 Visual stimuli for the P300 brain–computer interface: a comparison of white/gray and green/blue flicker matrices *Clin. Neurophysiol.* **120** 1562–6

Touryan J, Gibson L, Horne J H and Weber P 2011 Real-time measurement of face recognition in rapid serial visual presentation *Front. Psychol.* **2** 42

Townsend G *et al* 2010 A novel P300-based brain–computer interface stimulus presentation paradigm: moving beyond rows and columns *Clin. Neurophysiol.* **121** 1109–20

Treder M S and Blankertz B 2010 (C)overt attention and visual speller design in an ERP-based brain–computer interface *Behav. Brain Funct.* **6** 28