

The Changing Face of P300 BCIs: A Comparison of Stimulus Changes in a P300 BCI Involving Faces, Emotion, and Movement

Jing Jin^{1*}, Brendan Z. Allison^{2,5}, Tobias Kaufmann³, Andrea Kübler³, Yu Zhang^{1,4}, Xingyu Wang¹, Andrzej Cichocki⁴

1 Key Laboratory of Advanced Control and Optimization for Chemical Processes, Ministry of Education, East China University of Science and Technology, Shanghai, P.R. China, **2** Laboratory of Brain-Computer Interfaces, Institute for Knowledge Discovery, Graz University of Technology, Graz, Austria, **3** Department of Psychology I, University of Würzburg, Würzburg, Germany, **4** Laboratory for Advanced Brain Signal Processing, Brain Science Institute, RIKEN, Wako-shi, Japan, **5** Cognitive Neuroscience Laboratory, Department of Cognitive Science, University of California San Diego, La Jolla, California, United States of America

Abstract

Background: One of the most common types of brain-computer interfaces (BCIs) is called a P300 BCI, since it relies on the P300 and other event-related potentials (ERPs). In the canonical P300 BCI approach, items on a monitor flash briefly to elicit the necessary ERPs. Very recent work has shown that this approach may yield lower performance than alternate paradigms in which the items do not flash but instead change in other ways, such as moving, changing colour or changing to characters overlaid with faces.

Methodology/Principal Findings: The present study sought to extend this research direction by parametrically comparing different ways to change items in a P300 BCI. Healthy subjects used a P300 BCI across six different conditions. Three conditions were similar to our prior work, providing the first direct comparison of characters flashing, moving, and changing to faces. Three new conditions also explored facial motion and emotional expression. The six conditions were compared across objective measures such as classification accuracy and bit rate as well as subjective measures such as perceived difficulty. In line with recent studies, our results indicated that the character flash condition resulted in the lowest accuracy and bit rate. All four face conditions (mean accuracy >91%) yielded significantly better performance than the flash condition (mean accuracy = 75%).

Conclusions/Significance: Objective results reaffirmed that the face paradigm is superior to the canonical flash approach that has dominated P300 BCIs for over 20 years. The subjective reports indicated that the conditions that yielded better performance were not considered especially burdensome. Therefore, although further work is needed to identify which face paradigm is best, it is clear that the canonical flash approach should be replaced with a face paradigm when aiming at increasing bit rate. However, the face paradigm has to be further explored with practical applications particularly with locked-in patients.

Citation: Jin J, Allison BZ, Kaufmann T, Kübler A, Zhang Y, et al. (2012) The Changing Face of P300 BCIs: A Comparison of Stimulus Changes in a P300 BCI Involving Faces, Emotion, and Movement. PLoS ONE 7(11): e49688. doi:10.1371/journal.pone.0049688

Editor: Laura Frishman, University of Houston, United States of America

Received: June 8, 2012; **Accepted:** October 11, 2012; **Published:** November 26, 2012

Copyright: © 2012 Jin et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported in part by the Grant National Natural Science Foundation of China, under Grant Numbers 61074113 and 61203127, and supported in part by Shanghai Leading Academic Discipline Project, Project Number B504. This work was also supported by the Information and Communication Technologies Collaborative Project action "BrainAble" within the Seventh Framework of the European Commission, Project number ICT-2010-247447, and Fundamental Research Funds for the Central Universities, WH1114038. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: Jinjingat@gmail.com

Introduction

Brain-computer interface (BCI) systems allow for communication without movement. Users perform specific mental tasks that each has distinct patterns of brain activity. An artificial system categorizes these patterns to identify the associated mental tasks and thereby infer the messages or commands that the user meant to convey. Most modern BCIs rely on the EEG, which is recorded noninvasively via electrodes on the surface of the head [1]. While a variety of mental tasks have been explored for BCI control, most

BCIs rely on motor imagery or visual attention ([2–3], [4], [5], [6], [7], [8]).

"P300" BCIs

One type of BCI that usually relies on visual attention is called a P300 BCI ([9], for review, see [10], [11], [12]). This BCI is so named because it relies heavily on the P300, which is a well-known component of the event-related potential (ERP) that is largest when elicited by events that the subject considers important ([13], [14]). Therefore, subjects can generate larger, more distinct P300s by choosing to pay attention to specific events while ignoring

others. For example, a user who wants to spell the letter “P” could silently count each time the “P” flashes while ignoring any flashes that do not include the “P”. If the BCI signal processing software correctly recognizes that the largest P300 is elicited whenever the “P” flashes, then the BCI system would spell the letter “P” on a monitor and move on to the next letter. Although spelling is the most common application, P300 BCIs have been used to control other applications such as a “BrainPainting” system, internet browser, robot arm, or environmental control ([15], [16], [17], [18]). Recently, different P300 BCI systems have even been implemented with standard assistive technology (AT) software for text entry, emailing, and internet surfing and evaluated by severely impaired end-users in terms of effectiveness (accuracy), efficiency (bit rate and subjective workload) satisfaction and other factors ([19], [20], [21]). The P300-BCI driven AT software proved effective and efficient was judged reliable and easy to learn, and users were satisfied with the BCI. However, the low information transfer rate limited the practical use of the system by end-users. Thus, improving speed while maintaining reliability is a major issue when bringing BCIs to the patients’ bedsides [22].

P300 BCIs could be improved by enhancing the difference between attended and ignored events - which could entail more recognizable differences in the P300 and/or other components. In the past few years, different groups have focused on changes to stimulus presentation paradigms that generally seek to increase other components of the ERP that occur before or after the P300. P300 BCIs typically rely on not only the P300, but also other visual ERPs such as the N100, N200, and N400 components ([23], [24], [25], [26], [27], [28]).

For example, one study proposed a P300 BCI with stimuli that move instead of flash, which could elicit motion visual evoked potentials (M-VEPs) [29]. Hong et al. (2009) developed an offline M-VEP based spelling system, and showed that it might offer superior performance to a conventional P300 BCI [30]. Liu et al. (2010) validated the first online M-VEP BCI [31]. Jin et al. (2012) developed a combined system using moving flashes to improve the P300 BCI [32].

Another example of a study that focused on non-P300 components is Kaufmann et al. (2011b), which introduced stimuli that are transparently overlaid with famous faces. This approach accentuated the N170 and in particular the N400f, an ERP component involved in face recognition [28]. The resulting ERPs had a higher signal to noise ratio, which significantly improved classification accuracy. Zhang et al. (2012) reported that N170 and vertex positive potentials (VPP) also improve classification accuracy in a P300 BCI with stimuli that change to faces [33]. In addition to changing characters to faces, other types of changes may also offer advantages over the conventional “flash” approach in P300 BCIs ([25], [34], [35], [36], [37]).

The primary goal of this study was to parametrically compare a canonical “flash” condition with two other conditions in which the characters move or change to faces. Extending prior work, we explored both objective factors such as bit rate and classification accuracy as well as subjective factors such as whether subjects considered each condition difficult or tiring. This addressed the need for a direct, within-subject comparison of the two new types of stimulus changes that have gained the most attention in the literature (flashing, moving, and changing to faces), with consideration of personal preferences as well as performance. A secondary goal was to explore three new conditions in which the characters change to different types of faces. We hypothesized that faces that move and/or have an emotional expression could result in performance differences, with face stimuli leading to significantly better performance.

Methods

Subjects

Ten healthy subjects (6 male, aged 21–26 years, mean 24) participated in this study. All subjects signed a written consent form prior to this experiment and were paid for their participation. The local ethics committee approved the consent form and experimental procedure before any subjects participated. All subjects’ native language was Mandarin Chinese, and all subjects were familiar with the Western characters used in the display. Four subjects had used a BCI before this study.

The face images were obtained by photographing the first subject. All subjects were familiar with this face, since subject 1 was a fellow student of subjects 2–9. This face was chosen partly because pilot testing using the same face as in Kaufmann et al., (2011) (Albert Einstein presenting his tongue) indicated that some subjects preferred the face of the first subject [28]. Subject 1 provided written consent to use his images within the study and in Figure 1 of this article, and understood that his face images would be published.

Stimuli and procedure

After being prepared for EEG recording, subjects were seated about 105 cm in front of a monitor that was 30 cm high (visual angle: 16.3 degrees) and 48 cm wide (visual angle: 25.7 degrees). During data acquisition, subjects were asked to relax and avoid unnecessary movement. Figure 1 shows the display presented to all subjects, which was a 6×6 matrix with gray English letters and other characters against a black background. There were six conditions in the study, which differed only in the way that these characters changed. In different conditions, some of the characters would change by changing color, moving, and/or changing to different types of faces. We use the term “sub-trial” throughout this paper to refer to each individual event, such as a flash in the flash condition. The six conditions changed the stimuli as follows:

- (1) In the flash paradigm, the characters became white for 200 ms and then reverted to gray. (See Figure 1. Flash).
- (2) In the moving paradigm, each character became white and immediately moved 1 cm (visual angle: 0.5 degrees) to the right for 200 ms at a constant speed of 2.5 degrees/s. Next, the characters immediately reverted back to their original positions and reverted to gray. (See Figure 1. Flash, Move-Early and Move-Final).
- (3) In neutral face paradigm, each character changed to the “neutral face image” for 200 ms. (See Figure 1. Face-Neutral), and then reverted to gray characters.
- (4) In the smiling face paradigm, the only difference from the neutral face paradigm was that the “smiling face image” was used instead of the “neutral face image”. (See Figure 1. Face-Smiling).
- (5) In the shaking neutral face paradigm, the face rotated from the left to its front (see Figure 1. Face-Neutral-R1, Face-Neutral-R2, Face-Neutral-R3, Face-Neutral-R4 and Face-Neutral). Specifically, five images were shown, each for 25 ms, and each successive image was rotated 18 degrees toward the left side of the monitor. Hence, the rotation lasted 125 ms, after which the front face (see Figure 1. Face-Neutral) was shown for 75 ms.
- (6) In the shaking smiling face paradigm, the only difference from the shaking neutral face condition was the “smiling face image” was used instead of the “neutral face image”. (See Figure 1, Face-Smiling).

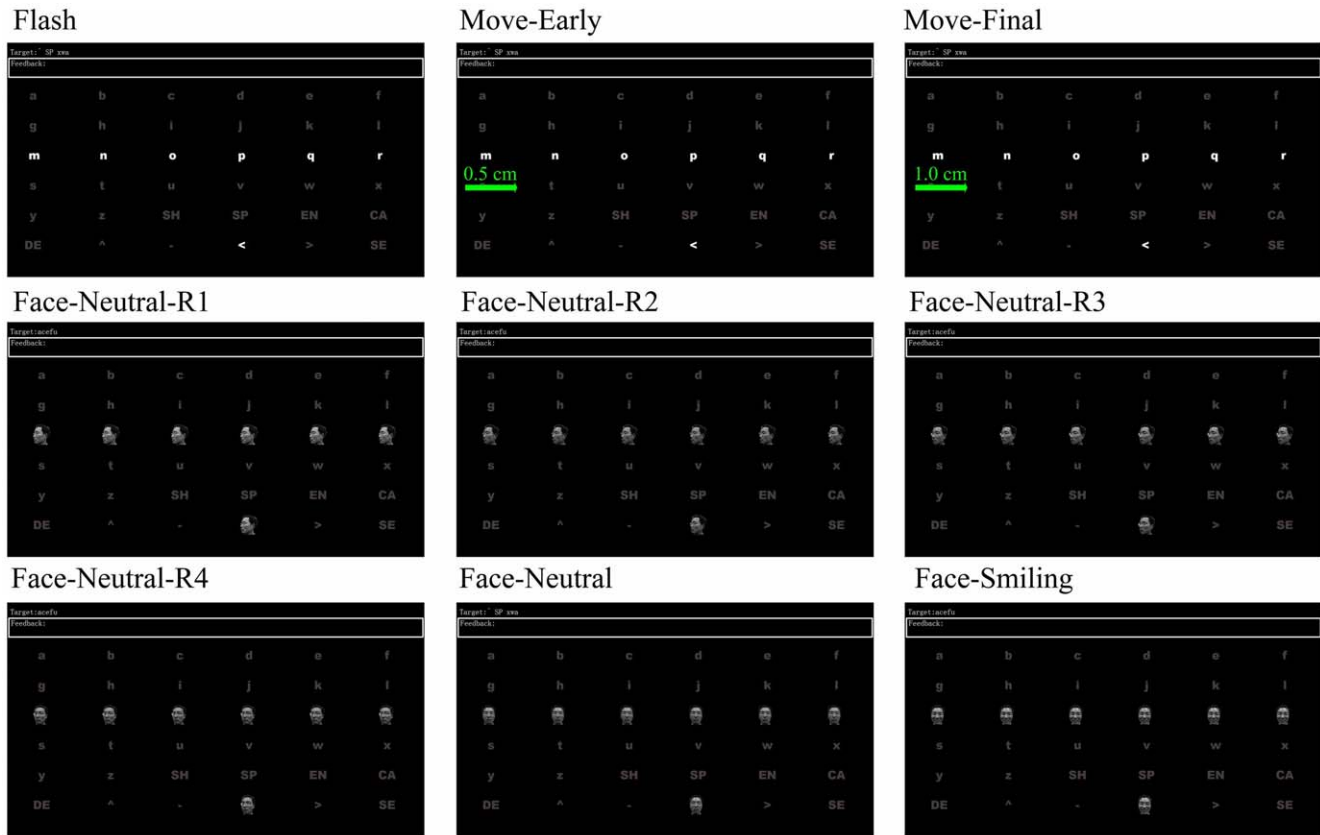


Figure 1. The display during the online runs. The five-letter target sequence is presented at the top of the screen, and the feedback is presented below it. Please see the text for a description of the different panels in this figure.
doi:10.1371/journal.pone.0049688.g001

In all six conditions, after the 200 ms sub-trial, all characters reverted to their usual background state for 50 ms before the next flash began. Hence, in all conditions, there were four flashes per second.

Three of these conditions corresponded to conditions in our two previous papers ([28], [32]). The flash condition was similar to the flash condition in these papers, as well as the canonical flash condition in nearly all P300 BCI papers (e.g., [9], [23], [38]). The moving condition was similar to the “flash and move” condition in [32], and the neutral face condition was similar to the face conditions in [28]. The remaining three conditions modified this neutral face condition by modifying two independent variables: emotion (joy vs. neutral) and motion (motionless vs. shaking).

Instead of grouping the flashed characters into rows and columns, we developed an alternate flash pattern approach based on binomial coefficients ([35], [39]). In this paper, we used the set of k combinations ($k=2$) from set $n=12$. Hence, there were twelve flash patterns, and between four and seven characters changed during each flash. Please see our prior work for additional details [25].

Experiment set up, offline and online protocols

EEG signals were recorded with a g.USBamp and a g.EEGcap (Guger Technologies, Graz, Austria) with a sensitivity of 100 μ V, band pass filtered between 0.1 Hz and 30 Hz, and sampled at 256 Hz. We recorded from EEG electrode positions Fz, Cz, Pz, Oz, F3, F4, P3, P4, P7, P8, O1, and O2 from the extended International 10–20 system.

In contrast to other P300 BCI approaches, the F3 and F4 channels were additionally selected since it was reported that F3 and F4 contain large N400 waves called fN400 [40].

The right mastoid electrode was used as the reference, and the front electrode (FPz) was used as a ground. Data were recorded and analyzed using the ECUST BCI platform software package developed through East China University of Science and Technology.

As noted, each **sub-trial** reflected each time a stimulus changed from a background stimulus, such as each flash in the flash condition. One **trial** contained all sub-trials with each of the twelve flash patterns. Since all conditions had 200 ms sub-trials followed by a 50 ms delay, each trial lasted three seconds. A **trial block** referred to a group of trials with the same target character. During offline testing, there were 16 trials per trial block. During online testing, the number of trials per trial block was variable, because the system adjusted this number to optimize performance as described in part 2.7. Each **run** consisted of five trial blocks, each of which involved a different target character. Subjects had a five minute break after each offline run, and a two minute break between each online run, since each online run was about more than four minutes shorter than each offline run.

There were six conditions, which were presented to each subject in pseudorandom order. For each condition, each subject first participated in three offline runs. Subjects had five minutes rest between each offline run. After the three offline runs, there were four online runs for each condition, which were presented in the same order as the offline runs. Each subject participated in the online and offline runs within each condition on the same day, and

each subject participated in three conditions in each day to decrease fatigue. Hence, each subject participated in two **sessions** that each consisted of nine offline runs followed by twelve online runs.

At the beginning of each run, subjects viewed a 6×6 matrix containing gray English letters and other characters against a black background (see Figure 1). The top of the screen presented a five-character sequence containing the five targets for that run. The subject's task was to silently count each time the target character flashed. In the offline runs, subjects never received feedback. In the online runs, whenever the classifier identified the target character, the system displayed it in the "Feedback" line on the top left of the screen, then the run ended. After a run ended, the matrix remained on the screen without any changes for one second.

Feature extraction procedure

A third order Butterworth band pass filter was used to filter the EEG between 0.1 Hz and 12 Hz. Although the P300 is primarily in the band 0.1–4 Hz [41], it can also be found in higher bands [42]. The EEG was downsampled from 256 Hz to 36.6 Hz by selecting every seventh sample from the filtered EEG. Single sub-trials lasting 800 ms were extracted from the data. The size of the feature vector was 12×29 (12 channels by 29 time points).

Classification scheme

Bayesian linear discriminant analysis (BLDA) is an extension of Fisher's linear discriminant analysis (FLDA) that avoids over fitting. The details of the algorithm can be found in [43]. BLDA was selected because of its demonstrated classification performance in P300 BCI applications [43]. Data acquired offline were used to train the classifier using BLDA and obtain the classifier model. This model was then used in the online system.

Practical bit rate

In this paper, we used two bit rate calculation methods called practical bit rate (PBR) and raw bit rate (RBR). The PBR is used to estimate the speed of the system in a real-world setting. Unless otherwise stated, all analyses in this paper are based on PBR; we only present RBR to facilitate comparisons with other studies. These two bit rate measures differ from each other in two ways. First, the PBR incorporates the fact that every error requires two additional selections to correct the error (a backspace followed by the correct character). The practical bit rate is calculated as $RBR \cdot (1 - 2 \cdot P)$, where RBR is the raw bit rate and P is the online error rate of the system [39]. Second, the RBR and PBR also incorporate the time between selections (1 second). Raw bit rate calculated with selection time was used to show the online information transfer rate of P300 BCIs which use other error correction methods [44] except "Backspace".

Adaptive system settings

The number of trials per average was automatically selected based on the classifier output. After each trial, the classifier would determine the target character based on data from all trials in that run block. If the classifier decided on the same character after two successive trials, then no new flashes were presented, and that character was presented as feedback to the subject.

For example, assume that the classifier decided that the letter "J" was the target based on the data from the first trial. The system would then present a second trial. The data from the first and second trials would be averaged, and the classifier would again decide which character was the target. If the classifier again selected "J", then the BCI assumed that "J" was the correct target,

and presented that feedback to the user. If the classifier did not select the "J", then another trial would begin, and so on until the classifier chose the same letter two consecutive times or until 16 trial blocks elapsed. That is, after 16 trials with the same intended target letter, the classifier would automatically select the last output of classifier as the target character [45].

Subjective report

After completing the last run, each subject was asked three questions about each of the six conditions. These questions could be answered on a 1–5 scale indicating strong disagreement, moderate disagreement, neutrality, moderate agreement, or strong agreement. The three questions were:

- 1) Was this paradigm annoying?
- 2) Did this paradigm make you tired?
- 3) Was this paradigm hard?

Statistical analysis

Before statistically comparing classification accuracy and practical bit rate, data were statistically tested for normal distribution (One-Sample Kolmogorov Smirnov test) and sphericity (Mauchly's test). Consecutively, repeated measures ANOVAs with stimulus type as factor were conducted. Post-hoc comparison was performed with Tuckey-Kramer tests. The alpha level was adjusted according to Bonferoni-Holm with $\alpha = .0083$ (significant). Non-parametric Kruskal-Wallis tests were computed to statistically compare the questionnaire replies.

Results

Online performance

Table 1 shows the results from the online runs. Classification accuracy ($F(5,45) = 7.78$, $p < .0001$, $\eta^2 = 0.46$) and practical bit rate ($F(5,45) = 8.54$, $p < .0001$, $\eta^2 = 0.49$) were significantly different across the six conditions.

A direct comparison of the three established paradigms within subjects (flash-only condition [9], moving-flash condition [32], and neutral face condition [27]), revealed significant differences for classification accuracy ($F(2,18) = 9.61$, $p < .0015$, $\eta^2 = 0.52$) and practical bit rate ($F(2,18) = 10.74$, $p < .0010$, $\eta^2 = 0.54$). Post-hoc comparison of the neutral face and flash-only condition showed significant differences in terms of classification accuracy ($p < .0017$) and significant differences in terms of PBR ($p < .0083$). No significant difference was found between the moving flash and flash-only condition as well as between moving flash and neutral face condition (all $p > .0083$; classification accuracy and PBR). However, average classification accuracy and both bit rate measures were larger for moving flash than flash-only (see Table 1).

In a separate comparison, we compared the four different face conditions to each other to explore the secondary goal of investigating the effect of movement and emotion in face stimuli. Neither comparison was significant ($F(3,27) = 0.12$, $p = 0.9500$, $\eta^2 = 0.01$ for classification accuracy and $F(3,27) = 0.80$, $p = 0.5000$, $\eta^2 = 0.08$ for practical bit rate).

Non-P300 components

Figure 2 shows that grand averaged ERPs across subjects 1–10 over sites F3, Fz, F4, Cz, P7, P8, O1, Oz and O2. Figure 3 and Figure 4 clearly indicate that non-P300 components play an important role in classification, as expected. Because this work entails numerous issues with different ERP components, we have submitted these results separately, as a companion paper. This

Table 1. Performance from online feedback runs.

		S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	Average
Flash	Acc (%)	60	55	50	80	85	80	90	80	95	75	75±15.1
Move		65	90	75	95	90	95	80	85	90	75	84±9.9
N_face		90	80	85	100	100	100	90	100	85	95	92.5±7.5
S_Face		90	80	90	90	95	100	80	90	100	100	91.5±7.5
SN_Face		95	90	90	100	85	100	100	90	95	85	93±5.9
SS_Face		100	90	90	90	90	100	90	85	90	95	92±4.8
Flash	RBR (bit/min)	8.0	9.7	7.8	23.2	26.1	22.6	27.9	21.1	33.7	14.6	19.5±9.0
Move		9.8	29.9	15.4	26.4	25.0	26.1	22.8	23.3	26.2	14.5	21.9±6.5
N_face		25.4	24.0	28.1	46.0	29.1	38.3	34.9	25.2	27.6	30.8	30.9±6.9
S_Face		27.0	25.8	27.5	34.2	28.9	40.5	18.8	21.2	37.6	33.4	29.5±6.9
SN_Face		30.8	27.4	26.2	44.0	26.6	46.0	46.0	27.5	40.2	28.6	34.3±8.6
SS_Face		47.0	29.4	32.2	34.9	29.4	46.0	34.2	25.7	37.2	32.5	34.9±6.9
Flash	PBR (bit/min)	1.5	0.9	0	12.5	16.4	14.4	20.1	11.5	27.0	6.8	11.1±8.9
Move		2.7	21.4	7.1	21.7	18.2	21.4	12.3	14.8	19.0	6.7	14.5±7.0
N_face		18.4	12.9	17.5	40.0	26.6	34.1	24.5	23.1	17.2	25.0	23.9±8.2
S_Face		19.5	13.8	19.8	24.1	23.6	35.9	10.3	15.6	33.5	30.1	22.6±8.5
SN_Face		25.0	19.8	19.0	39	16.7	39.3	40.0	19.8	31.6	17.8	26.8±9.7
SS_Face		40.8	21.0	22.8	24.5	21.0	40.0	24.1	16.2	25.9	26.2	26.3±8.0
Flash	Avt	5.35	3.9	4.1	3.0	2.9	3.6	3.0	3.25	2.75	4.2	3.6±0.8
Move		5.0	2.8	4.0	3.5	3.35	3.6	3.05	3.25	3.2	4.25	3.6±0.7
N_face		3.3	2.8	2.7	2.3	3.6	2.7	2.4	4.1	2.75	3.0	3.0±0.6
S_Face		3.1	2.7	3.05	2.5	3.2	2.6	3.65	3.95	2.75	3.1	3.1±0.5
SN_Face		3.0	3.1	3.2	2.4	2.9	2.30	2.25	3.05	2.3	2.65	2.7±0.4
SS_Face		2.2	2.9	2.6	2.4	2.9	2.25	2.45	2.95	2.25	2.85	2.6±0.3

“Flash” is the flash-only paradigm, “Move” is moving-flash paradigm, “N_Face” is neutral face paradigm, “S_Face” is smiling face paradigm, “SN_Face” is neutral shaking face paradigm and “Face_SS” is smiling shaking face paradigm. “Acc” is classification accuracy, “RBR” is raw bit rate and “PBR” is practical bit rate (bits/min). “Avt” is average trials used for average.

doi:10.1371/journal.pone.0049688.t001

paper includes a review of relevant literature, peak amplitude and latency measurements for different ERP components, statistical analyses, and discussion.

However, we did conduct another type of analysis within this study to explore the importance of different time periods within each ERP. We performed 15-fold cross validation with three offline data sets. Figure 4 shows the results obtained from data which was extracted from 0–299 ms, 300–450 ms, 451–700 ms, 700–1000 ms, 0–800 ms, and 0–1000 ms after flash onset. These results indicate that data before and after the P300 contribute to correct classification. This outcome supports our assertion that non-P300 components are underappreciated in so-called “P300” BCIs.

Target to target interval

Target to target interval (TTI) is a critical determinant of P300 amplitude [46]. The TTI typically varies within P300 BCIs, and these variations can affect performance ([47], [48]). Furthermore, with very short TTIs, late components might actually reflect activity from subsequent trials.

In this study, very few ERPs were influenced by a second target stimulus that occurred during the 1000 ms after each target flash. With the flash pattern approach used here, the minimum TTI was 500 ms. Table 2 shows the distribution of TTIs. 99% of TTIs are longer than 750 ms, and 86% of TTIs are longer than 1000ms (see

Table 2). This reflects that ERPs evoked by current stimulus are rarely influenced by the ERPs evoked by next stimulus.

Subjective report

Table 3 presents the subjects’ responses to the three questions. None of the participants reported any of the stimulus conditions as tiring, difficult or annoying, since they reported neutrality, moderate disagreement or strong disagreement with these questions. No significant difference was found between conditions ($H(5) = 7.37$, $p = .19$ for annoying; $H(5) = 3.93$, $p = .56$ for difficult; $H(5) = 13.66$, $p = .02$ for tired; adjusted alpha level: $\alpha = .0083$).

Discussion

The primary goal of this study was to compare two emerging stimulus approaches with each other and with a canonical “flash only” stimulus approach. In line with previous results, the moving-flash [32] and face stimuli [28] improved classification accuracy and bit rate relative to flash-only, and the face stimulus was the most effective approach. Our moving flash condition did not significantly differ from the flash only condition. However, we previously reported that the moving-flash condition was significantly better than the flash-only condition [32]. Unlike that study, we did not use a blue/green paradigm in the current study, and we did not improve the moving and flash times through pilot testing. Instead, to facilitate comparison to other conditions, we used a

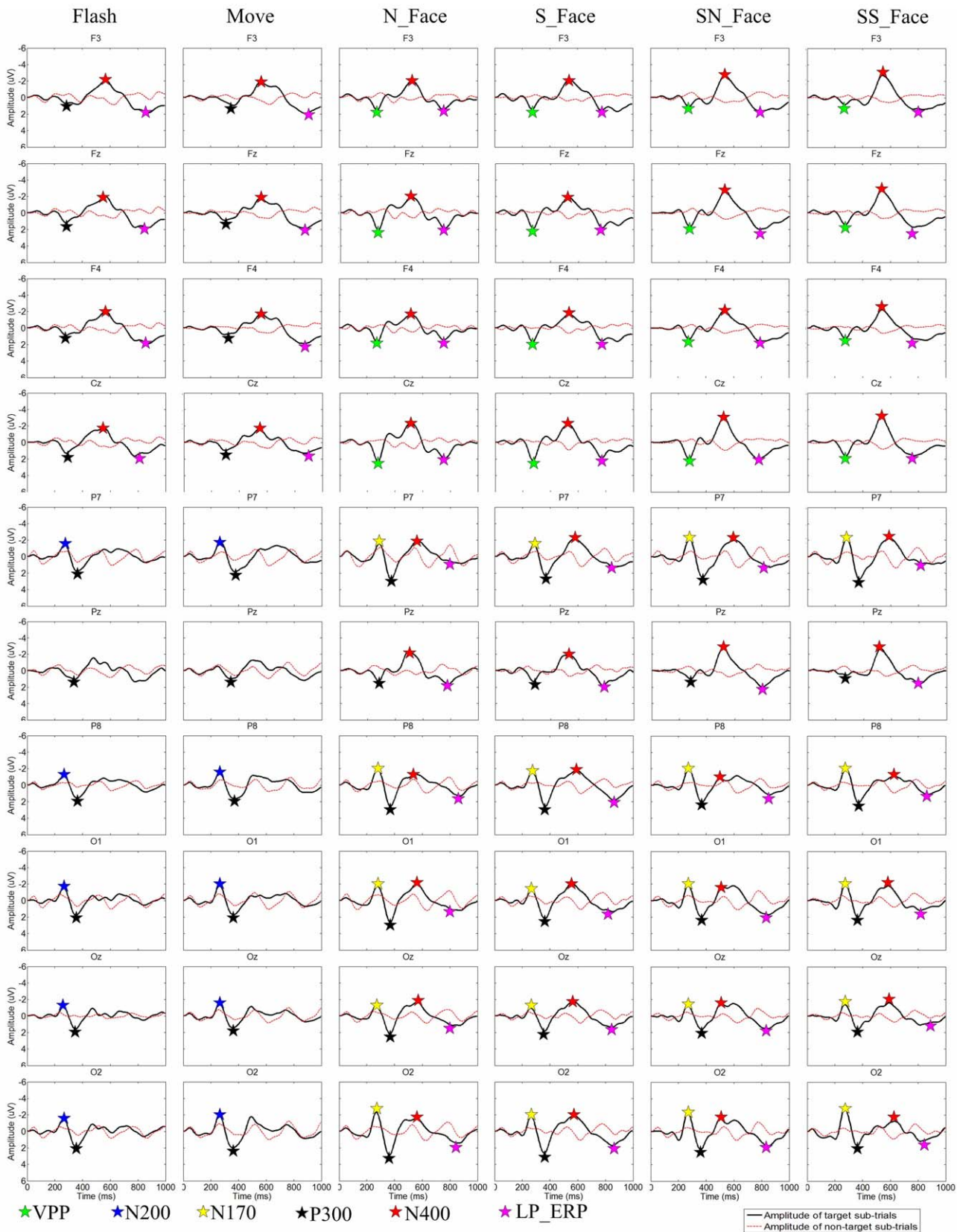


Figure 2. Grand averaged ERPs. Grand averaged ERPs across subjects 1–10 over sites F3, Fz, F4, Cz, P7, P8, O1, Oz and O2, LP_ERP is late positive ERP. doi:10.1371/journal.pone.0049688.g002

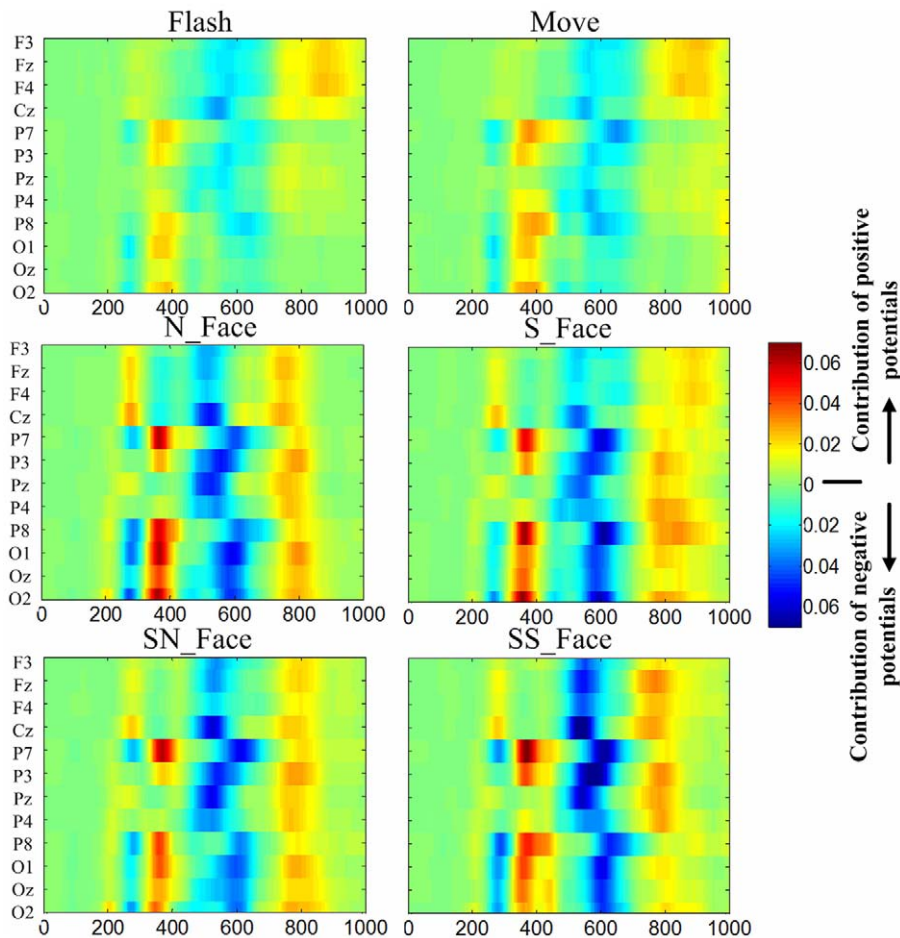


Figure 3. R-squared values of ERPs. R-squared values of ERPs from all paradigms at 0-1000ms averaged from subject 1–10 on site F3, Fz, F4, Cz, P7, P3, Pz, P4, P8, O1, Oz, O2.
doi:10.1371/journal.pone.0049688.g003

white/grey pattern, and the movement duration was as long as the flash duration. The present results suggest that these modifications might be crucial for further enhancing performance in the moving-flash condition.

The secondary goal was to compare an established face paradigm with three novel face paradigms. We hypothesized that manipulating facial emotion (smiling faces) or motion (shaking faces) might lead to more distinct ERPs and thereby improve performance. Although these manipulations did appear to influence the expected ERP components with some subjects, any ERP differences were not robust enough to lead to statistically significant performance differences within the face conditions.

However, the results do not necessarily discourage future research involving facial movement and/or emotion. For examples, angry faces elicit larger ERPs than happy faces [49], and inverted faces could evoke a larger N170 and thereby improve the performance of a BCI [33]. Other paradigmatic changes could improve performance by eliciting more ERP differences between attended and ignored stimuli.

For example, presenting new faces or other stimuli with each flash might enhance the P3a or other components ([14], [50], [51]). Presenting the same face each time could yield a small reduction in negative potentials, but could enhance positive potentials. The merits of novel face presentation should be explored in an online BCI study, since (at least) P300 habituation

from offline non-BCI studies may differ from habituation in online studies with feedback ([24], [52]). Different instructions to subjects could also encourage more attention to faces and thus enhance face related components. For example, instead of counting stimuli, subjects could be asked to count only certain types of faces, such as happy, male, or familiar faces. These possibilities underscore that further research is needed to translate these speculative options into actual research results.

As none of the subjects considered any of the conditions especially tiring, annoying, or difficult, we conclude that the new stimulus material is feasible for BCI use. However, non-significant trends indicated a slight preference for simpler conditions, and that the smiling face condition was more tiring. These trends should be explored in further studies. Some types of subjects, such as elderly users or persons with attentional disorders, might have more trouble with some conditions.

This work is significant for two general reasons. First, we provide further support for an increasingly obvious conclusion: the canonical “flash-only” paradigm, which has dominated P300 BCIs for almost 25 years, does not outperform other approaches. Alternate paradigms are superior in terms of objectively measurable dependent variables like accuracy and bit rate, and do not differ significantly in terms of subjective report. Since these approaches can be easily implemented, there seems to be no reason to continue using “flash-only” P300 BCIs, specifically with

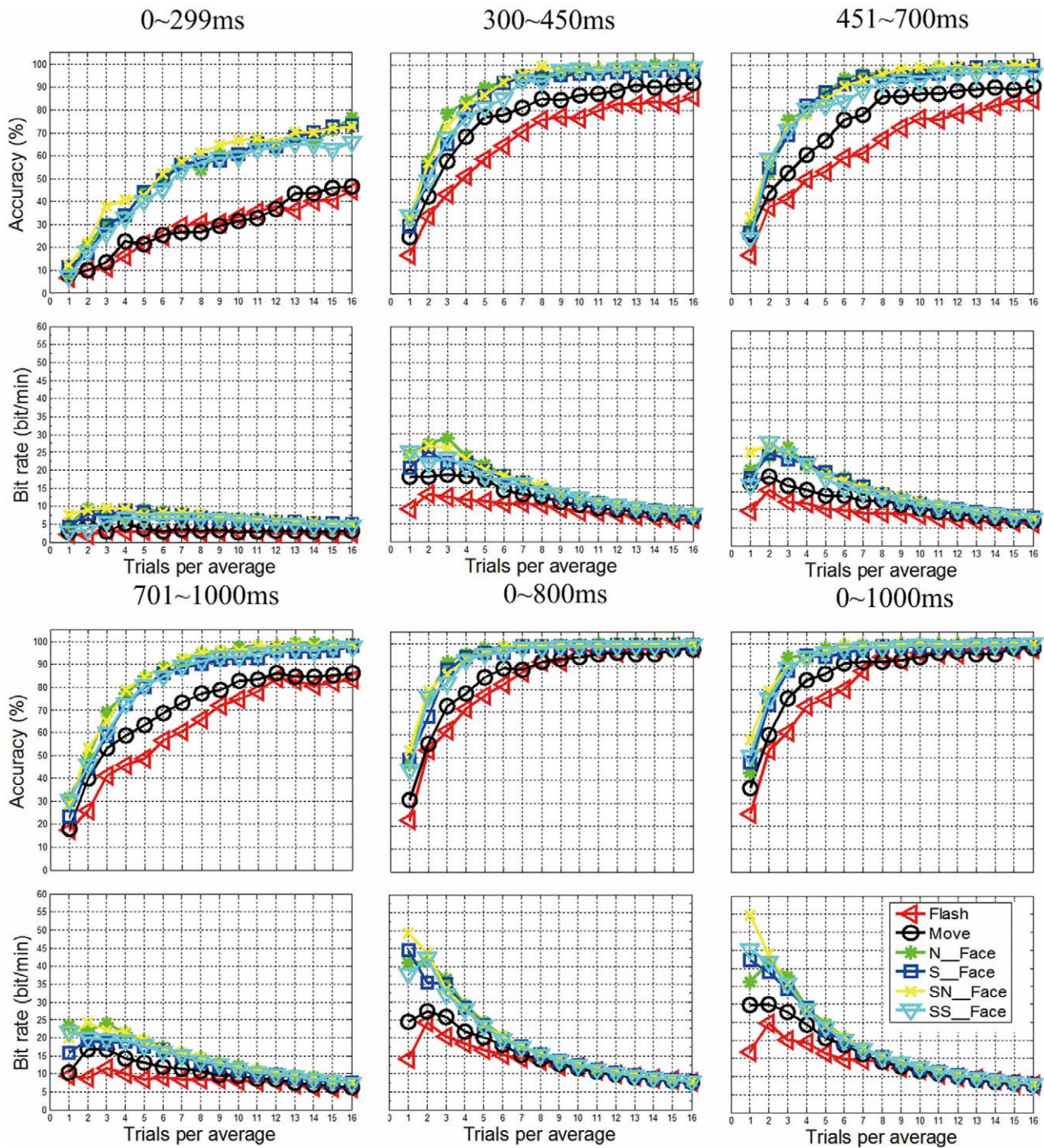


Figure 4. Classification accuracy and bit rate based on offline data. The two columns from left to right present the classification accuracy and bit rate based on offline data. The four rows from top to bottom present these measures based on data from 0–299 ms, 300–450 ms, 451–700 ms, 701–1000 ms, 0–800 ms and 0–1000 ms after flash onset, respectively. doi:10.1371/journal.pone.0049688.g004

Table 2. The percentages of TTIs between sub-trials.

TTI (ms)	500	750	1000	1250	1500	1750	2000	2250
(%)	0.42	12.9	13.1	16.9	12.9	17.1	13.3	13.3

doi:10.1371/journal.pone.0049688.t002

Table 3. Subjects' responses to three questions for each of the six conditions.

		S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	average
Flash	Tired	1	1	1	1	3	1	3	3	1	1	1.6±1.0
	Diff	1	2	1	1	1	1	1	1	1	1	1.1±0.3
	Annoy	1	1	1	1	1	1	2	1	1	1	1.1±0.3
Move	Tired	1	1	1	1	2	2	3	3	3	1	1.8±0.9
	Diff	1	1	1	1	2	2	2	1	1	1	1.3±0.5
	Annoy	1	1	1	1	1	1	3	1	1	1	1.2±0.6
N_Face	Tired	2	2	1	1	3	1	3	2	1	2	1.8±0.8
	Diff	1	2	1	1	1	1	2	1	1	1	1.2±0.4
	Annoy	1	1	1	2	1	1	2	1	1	1	1.2±0.4
S_Face	Tired	2	2	2	1	3	1	3	3	3	2	2.2±0.8
	Diff	2	1	1	1	1	1	2	1	1	2	1.3±0.5
	Annoy	2	1	1	1	1	1	2	1	1	2	1.3±0.5
SN_Face	Tired	2	2	3	1	2	1	1	3	2	2	1.9±0.7
	Diff	2	2	2	1	2	2	1	1	2	2	1.7±0.5
	Annoy	2	1	2	2	1	1	2	1	1	2	1.5±0.5
SS_Face	Tired	2	2	3	1	2	2	1	3	3	2	2.1±0.7
	Diff	2	2	2	1	2	1	2	1	2	2	1.7±0.5
	Annoy	2	1	2	2	1	1	2	1	1	2	1.5±0.5

Subjects were asked whether each condition seemed tiring, difficult, or annoying. Questions and responses are translated from Mandarin Chinese. A "1" means strong disagreement, a "2" means moderate disagreement, a "3" reflects a neutral response, a "4" means moderate agreement or a "5" means strong agreement.

doi:10.1371/journal.pone.0049688.t003

patients. While this work helps elucidate the objective and subjective effects of different approaches, further research is needed to better identify which alternate approaches are most effective.

This work is also significant in drawing attention to the importance of non-P300 components. Consistent with earlier work, this article also shows that the "P300" BCI could be improved when other ERP components are included. Indeed, numerous distinct ERPs, both before and after the P300, can improve classification accuracy.

Conclusion

This paper further investigated how P300 BCIs could be improved when combined with other ERPs. Face emotion and motion were not found to increase performance as compared to neutral face stimulation; however, faces significantly outperformed

other stimulation approaches. Future work should further explore ways to enhance any differences between target and non-target flashes. This major challenge could entail not only ERP components but other differences such as alpha power ([23], [53]). Future work could also assess whether adding alternate tasks (such as imagined movement) can further improve performance in a hybrid BCI ([54], [55]). The next practical step should be implementing faces in ERP-based applications such as text entry, emailing and internet surfing with standard assistive technology [56].

Author Contributions

Conceived and designed the experiments: JJ BA TK AK. Performed the experiments: JJ YZ. Analyzed the data: JJ BA TK YZ. Contributed reagents/materials/analysis tools: JJ TK XW AC. Wrote the paper: JJ BA TK AK AC.

References

- Mason MF, Norton MI, Van Horn JD, Wegner DM, Graffton S, et al. (2007) Wandering minds: the default network and stimulus-independent thought. *Science magazine* 315: 393–5.
- Kübler A, Kotchoubey B, Kaiser J, Wolpaw J, Birbaumer N (2001) Brain-computer communication: unlocking the locked in. *Psychol Bull* 127:358–75.
- Wolpaw JR, Birbaumer N, McFarland DJ, Pfurtscheller G, Vaughan TM (2002) Brain-computer interfaces for communication and control. *Clin Neurophysiol* 113: 767–91.
- Allison BZ, Wolpaw EW, Wolpaw JR (2007) Brain-computer interface systems: progress and prospects. *Expert Rev Med Devices* 4: 463–74.
- Allison BZ, McFarland DJ, Schalk G, Zheng SD, Jackson MM, et al. (2008) Towards an independent brain-computer interface using steady state visual evoked potentials. *Clin. Neurophysiol* 119: 399–408.
- Birbaumer N, Cohen LG (2007) Brain-computer interfaces: communication and restoration of movement in paralysis. *J Physiol* 15: 621–36.
- Mak JN, Wolpaw JR (2009) Clinical application of brain-computer interface: Current state and future prospects. *IEEE Rev Biomed Eng* 2: 187–99.
- Ortner R, Allison BZ, Korisek G, Gaqq H, Pfurtscheller G (2011) An SSVEP BCI to control a hand orthosis for persons with tetraplegia. *IEEE Trans Neural Syst Rehabil Eng.* 19: 1–5.
- 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: 510–23.
- Mak JN, Arbel Y, Minett JW, McCane LM, Yuksel B, et al. (2011) Optimizing the P300-based brain-computer interface: current status, limitations and future directions. *J. Neural Eng* 8: 025003.
- Kleih SC, Riccio A, Mattia D, Schreuder M, Tangermann M, et al. (2011) Motivation affects performance in a P300 Brain Computer Interface. *International Journal of Bioelectromagnetism* 13: 46–7.
- Fazel-Razai R, Allison BZ, Guger C, Sellers EW, Kleih SC, et al. (2012) P300 brain computer interface: current challenges and emerging trends. *Front Neuroeng* 5: 14.
- Johnson R Jr, Donchin E (1978) On how P300 amplitude varies with the utility of the eliciting stimuli. *Electroencephalogr Clin Neurophysiol* 44: 424–37.
- Polich J (2007) Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology* 118: 2128–48.

15. Allison BZ, Moore MM (2004) Field validation of a P300 BCI under adverse conditions. Society for Neuroscience Conference. Program No. 263.9. San Diego, CA.
16. Mugler E, Bensch M, Halder S, Rosenstiel W, Bogdan M, et al. (2008) Control of an internet browser using the P300 event-related potential. *Int. J. Bioelectromagn* 10: 56–63.
17. Münbinger JI, Halder S, Kleih SC, Furdea A, Raco V, et al. (2010) Brain Painting: first evaluation of a new brain-computer interface application with ALS-patients and healthy volunteers. *Front Neurosci* 4: 18.
18. Aloise F, Schettini, Arico P, Leotta F, Salinari S, et al. (2011) P300-based brain-computer interface for environmental control: an asynchronous approach. *J neural Eng* 8: 025025.
19. Huggins JE, Wren PA, Gruis KL (2011) What would brain-computer interface users want? Opinions and priorities of potential users with amyotrophic lateral sclerosis. *Amyotroph Lateral Scler* 12: 318–24.
20. Zickler C, Riccio A, Leotta F, Hillian-Tress S, Halder S, et al. (2011) A brain-computer interface as input channel for a standard assistive technology software. *Clinical EEG and Neuroscience* 42: 222–30.
21. Blain-Moraes S, Schaff R, Gruis KL, Huggins JE, Wren PA (2012) Barriers to and mediators of brain-computer interface user acceptance: focus group findings. *Ergonomics* 55: 516–25.
22. Holz EM, Kaufmann T, Desideri L, Malavasi M, Hoogerwerf EJ, et al. (2012) User centred design in BCI development. In B.Z. Allison, S. Dunne, R. Leeb, J.D.R. Millán, & A. Nijholt (Eds.), *Towards Practical Brain-Computer Interfaces*. Berlin: Springer 155–72.
23. Allison BZ, Pineda JA (2003) ERPs evoked by different matrix sizes: Implications for a brain computer interface (BCI) system. *IEEE Trans Neural Syst Rehabil Eng* 11: 110–3.
24. Sellers EW, Donchin E (2006) A P300-based brain-computer interface: initial tests by ALS patients. *Clin Neurophysiol* 117: 538–48.
25. Jin J, Horki P, Brunner C, Wang X, Neuper C, et al. (2010) A new P300 stimulus presentation pattern for EEG-based spelling systems. *Biomed Tech* 55: 203–10.
26. Treder MS, Blankertz B (2010) (C)overt attention and visual speller design in an ERP-based brain-computer interface. *Behav Brain Funct* 28: 28.
27. Kaufmann T, Hammer EM, Kübler A (2011) ERPs Contributing to Classification in the P300 BCI. 5th International Brain-Computer Interface Conference. Graz, Austria, 136–9.
28. Kaufmann T, Schulz SM, Grünzinger C, Kübler A (2011) Flahing characters with famous faces improves ERP-Based brain-computer interface performance. *J Neural Eng* 8: 056016.
29. Guo F, Hong B, Gao X, Gao S (2008) A brain-computer interface using motion-onset visual evoked potential. *J Neural Eng* 5: 011.
30. Hong B, Guo F, Liu T, Gao S (2009) N200-speller using motion-onset visual response. *Clin Neurophysiol* 120: 1658–66.
31. Liu T, Goldberg L, Gao S, Hong B (2010) An online brain-computer interface using non-flashing visual evoked potentials. *J Neural Eng* 7: 036003.
32. Jin J, Allison BZ, Wang X, Neuper C (2012) A combined brain computer interface based on P300 potentials and motion-onset visual evoked potentials. *J Neurosci Methods* 205: 265–76.
33. Zhang Y, Zhao Q, Jin J, Wang X, Cichocki A (2012) A novel BCI based on ERP components sensitive to configural processing of human faces. *J Neural Eng* 9: 026018
34. Martens SMM, Hill NJ, Farquhar J, 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.
35. Jin J, Allison BZ, Sellers EW, Brunner C, Horki P, et al. (2011) Optimized stimulus presentation patterns for an event-related potential EEG-based brain computer interface. *Med Biol Eng Comput* 49: 181–91.
36. Frye GE, Hauser CK, Townsend G, Sellers EW (2011) Suppressing flashes of items surrounding targets during calibration of a P300-based brain-computer interface improves performance. *J neural Eng* 8: 025024.
37. Shishkin SL, Ganin IP, Kaplan AY (2011) Event-related potentials in a moving matrix modification of the P300 brain-computer interface paradigm. *Neurosci Lett* 496: 95–9.
38. Kleih SC, Nijboer F, Halder S, Kübler A (2010) Motivation modulates the P300 amplitude during brain-computer interface use. *Clin Neurophysiol* 121:1023–31.
39. Townsend G, LaPallo BK, Boulay CB, Krusienski DJ, Frye GE, et al. (2010) A novel P300-based brain-computer interface stimulus presentation paradigm: moving beyond rows and columns. *Clin Neurophysiol* 121: 1109–20.
40. Curran T, Hancock J (2007) The FN400 indexed familiarity-based recognition of faces. *NeuroImage* 36: 464–71.
41. Jansen BH, Allam A, Kota P, Lachance K, Osho A, et al. (2004) An exploratory study of factors affecting single trial P300 detection. *IEEE Trans Biomed Eng* 51: 975–78.
42. Kolev VH, Demiralp T, Yordanova J, Ademoglu A, Isoglu-Alka U (1997) Time-frequency analysis reveals multiple functional components during oddball P300. *NeuroReport* 8: 2061–5.
43. Hoffmann U, Vesin JM, Ebrahimi T, Diserens K (2008) An empirical bayesian framework for brain computer interface. *J Neurosci Meth* 167: 115–25.
44. Dal Seno B, Matteucci M, Mainardi L (2010) Online detection of P300 and error potentials in a BCI speller. *Comput Intell Neurosci* 2010: 307254.
45. Jin J, Allison BZ, Sellers EW, Brunner C, Horki P, et al. (2011) An adaptive P300-based control system. *J Neural Eng* 8: 036006.
46. Gonsalvez CL, Polich J (2002) P300 amplitude is determined by target-to-target interval. *Psychophysiology* 39: 388–96.
47. Fazel-Rezai R (2007) Human error in P300 speller paradigm for brain-computer interface. 29th Annual Int. Conf. of the IEEE EMBS 2516–9.
48. Jin J, Sellers EW, Wang X (2012) Targeting an Efficient Target-to-Target Interval for P300 Speller Brain-Computer Interfaces. *Med Biol Eng Comput* 50: 289–96.
49. Martens U, Leuthold H, Schweinberger SR (2010) On the temporal organisation of facial identity and expression analysis: Inferences from event-related brain potentials. *Cogn Affect Behav Neurosci* 10: 505–22.
50. Squires NK, Squires KC, Hillyard SA (1975) Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalogr Clin Neurophysiol* 38: 387–401.
51. Itier RJ, Taylor MJ (2002) Inversion and contrast polarity reversal affect both encoding and recognition processes of unfamiliar faces: a repetition study using ERPs. *NeuroImage* 15: 353–72.
52. Ravden D, Polich J (1998) Habituation of P300 from visual stimuli. *Int J Psychophysiol* 30: 359–65.
53. Kelly P, Lalor EC, Reilly RB, Foxe JJ (2005) Visual spatial attention tracking using-density SSVEP data for independent brain-computer communication. *IEEE Trans Neural Syst Rehabil Eng* 13: 172–8.
54. Brunner C, Billinger M, Vidaurre C, Neuper C (2011) A comparison of univariate, vector, bilinear autoregressive, and band power features for brain-computer interfaces. *Med Biol Eng Comput* 49: 1337–46.
55. Li Y, Long J, Yu T, Yu Z, Wang C, et al. (2010) An EEG-based system for 2-D cursor control by combining Mu/Beta rhythm and P300 potential. *IEEE Trans Biomed Eng* 57: 2495–2505.
56. Zickler C, Riccio A, Leotta F, Hillian-Tress S, Halder S, et al. (2011) A brain-computer interface as input channel for a standard assistive technology software. *Clin EEG Neurosci* 42: 236–44.