Effects of SOA and flash pattern manipulations on ERPs, performance, and preference: Implications for a BCI system

Brendan Z. Allison a,*, Jaime A. Pineda b,c

a Department of Computer Information Systems, POB 4015, Georgia State University, Atlanta, GA 30303, United States
b Cognitive Science Department, University of California, San Diego, La Jolla, CA 92093, United States
c Neuroscience Department, University of California, San Diego, La Jolla, CA 92093, United States

Received 3 August 2004; received in revised form 2 February 2005; accepted 5 April 2005
Available online 27 July 2005

Abstract

P3 brain–computer interfaces (BCIs) are synchronous communication systems that allow users to communicate interest in a target event by choosing to attend to it while ignoring other events. In such a system, a cogneme refers to the user’s response to: ‘attend to the event’ or ‘ignore the event’. The present study examined subjects’ ability to generate more cognemes per minute (by varying stimulus onset asynchrony or SOA), or requiring fewer cognemes to convey a message (by varying the pattern of stimulus presentation). Both of these have implications for improved information throughput in a P3 BCI. SOAs of 125, 250, and 500 ms were used. Additionally, the conventional ‘single flash’ approach was compared to a new ‘multiple flash’ condition in which half of the stimuli in an 8/C2 grid were flashed simultaneously. In both conditions, P3-like component amplitudes decreased with faster SOAs at low target probabilities, but the trend did not hold for higher probabilities. The multiple flash condition produced more robust ERPs at the faster speeds. The results also indicate that attend/ignore differences were more apparent following multiple flashes for low target probabilities, but less apparent for high target probabilities. Although information throughput alone does not support the superiority of one approach over the other, only six cognemes are needed in the multiple flash conditions to identify a character, compared to sixteen cognemes in the single flash condition. This suggests that the former approach could operate more rapidly. Thus, the present results suggest that the multiple flash approach may be a more efficient and faster basis for a P3 BCI system.

© 2005 Elsevier B.V. All rights reserved.

Theme: Neural Basis of Behavior
Topic: Cognition

Keywords: P3; Synchronous BCI; Cognemes; EEG; Pattern recognition

1. Introduction

A brain–computer interface (BCI) is a real time communication system designed to allow users to voluntarily send messages or commands to an external device without using the brain’s normal output pathways (Wolpaw et al., 2002). While BCIs currently offer slower information throughput than conventional interfaces, like keyboards or mice, they may be the only means of communication for severely disabled individuals unable to use interfaces that require motor activity (e. g., Wolpaw et al., 1991; Birbaumer et al., 1999; Kennedy et al., 2000; Parker, 2003; Neuper et al., 2003, Wickelgren, 2003; see Kubler et al., 2001, Wolpaw et al., 2002, or Allison, 2003 for reviews and commentary). Two general classes of BCIs have been described in the literature. Asynchronous or non-cue-based BCIs allow a user to send information independent of any external event (e. g., Kennedy et al., 2000; Yom-Tov and Inbar, 2003; Gräf and Gräf, 2004; Mason et al., 2004; Townsend et al., 2004). Synchronous or cue-based BCIs allow the user to send messages or commands by producing one of two or more mental responses to external events (e. g. Farwell and Donchin, 1988; Sutter, 1992; Middendorf et al., 2002).
P3 BCIs are a category of synchronous BCIs in which the user conveys interest in a target event by choosing to attend to it while ignoring other stimuli. Since attended events yield larger P3s than ignored ones (Squires et al., 1975; Polich, 1998, 2004), P3 BCIs can determine which event produced the largest response and hence allow communication by sending that event to an output device.

In both synchronous and asynchronous BCIs, a user communicates by repeatedly engaging in specific mental activities that create distinct electroencephalogram (EEG) signatures. These discrete mental activities form the building blocks of a BCI language. Analogous to the words “grapheme” and “phoneme” which describe the smallest meaningful elements of a language, the term “cogneme” refers to the smallest quantifiable mental activity capable of producing a difference in meaning in a BCI. In a synchronous P3 BCI, for example, a cogneme is the user’s response to each event: “/attend to the event/” or “/ignore the event/” (Allison, 2003).

Different P3 BCI systems have been described in the published literature. In the first such article (Farwell and Donchin, 1988), users viewed a 6 × 6 grid containing English letters as well as other characters. Single rows or columns were sequentially flashed, and users were asked to count flashes containing the target character while ignoring other row or column flashes. Donchin and colleagues subsequently used a similar display to explore the use of discrete wavelet transform as a preprocessing approach (Donchin et al., 2000). In a different instantiation of a P3 BCI (Polikoff et al., 1995), users saw four letters (N, E, S, W), each of which represented a compass direction. Each letter was flashed in sequence, and users counted flashes of a target letter. Another approach (e.g., Bayliss and Ballard, 2000; Bayliss, 2003; Bayliss et al., 2004) described systems in which the user’s vocabulary contained icons instead of letters. Some of Bayliss’ studies utilized improved preprocessing and pattern recognition approaches, resulting in a substantial performance improvement over earlier approaches. Recent work has further suggested that improved pattern recognition approaches may substantially increase information throughput in P300 BCIs (e.g., Beverina et al., 2003; Kaper et al., 2004, Xu et al., 2003).

All BCIs depend not only on the pattern recognition approach used but also on the quantity, quality, and informativeness of the EEG data they receive. Because BCIs are very slow (at best about 1 bit/sec (Gao et al., 2003)), most research to date has been aimed at improving information throughput. In that regard, there are three general avenues toward improving information throughput in a synchronous BCI:

1) present stimuli more quickly, hence enabling the user to generate more cognemes per minute;
2) require fewer cognemes to convey a message, as could be done by obtaining more information from each cogneme or from a combination of cognemes, and/or by using a BCI with a larger vocabulary (one that recognizes a wider variety of cognemes); and
3) categorize cognemes more quickly and accurately, as might be done by using an improved classifier and/or by creating more robust differences between the EEGs associated with each cogneme.

These avenues toward improvement often exhibit trade-offs with each other. For example, presenting stimuli more quickly may result in less robust differences in the EEG. The current study explores this question, as well as the effect of a new “multiple flash” approach on EEG measures. In previous P3 BCIs utilizing a matrix display (Farwell and Donchin, 1988; Donchin et al., 2000), only one row or column was flashed at a time (see Fig. 1). This sequential “single flash” method of stimulus presentation results in a low probability for the flashing target character. Because the target-to-target interval (TTI) is long, P3 amplitudes are typically large (Gonsalvez and Polich, 2002). However, it also means that a large number of flashes are needed to identify the target character. If the target could be identified with perfect accuracy, the number of flashes required equals the number of rows plus the number of columns. In an 8 × 8 grid containing 64 stimuli, sixteen flashes would be required to identify the target (eight flashes to identify the row, and eight to identify the target column).

Since a P3 BCI allows one of two cognemes, the maximum amount of information available from each flash is one bit. Hence, the number of flashes required may be as low as the log2 of the number of rows and columns. The present study compared the sequential “single flash” approach to a “multiple flash” approach based on binary decomposition of the choices (other articles have described spelling with non-P3 BCIs by decomposing the alphabet into a series of binary or other choices, e.g. Perelmouter and Birbaumer, 2000; Vaughan et al., 2001; Tregoubov and Birbaumer, 2005). In this new method, half of the stimuli on the screen (either four rows or four columns) are flashed at any one time (see Fig. 2). The top three images in Fig. 2 show the three row flashes used in this study, and the bottom three images show the three column flashes. In the case of an 8 × 8 matrix with 64 elements, identifying which of eight rows is the target would be possible using only three events, assuming perfect accuracy. For example, if the subject’s target is in the seventh row, only one of the three row flashes (shown in the top right of Fig. 2) will result in a P3. If the target were in any other row, a different subset of the three row flashes would evoke a P3. If none of the row flashes evokes a P3, the target must be in the bottom row, the only row that is not illuminated by any row flash. Similarly, it could also be possible to recognize the target column using only three flashes. Thus, using the multiple flash approach, as few as six events would be required to identify the target character. However, since targets would be more frequent
and probable, P3 amplitude may be reduced, which may impair the system’s accuracy.

The present study examines the effects of different display parameters that might be used in a P3 BCI. Using a methodology similar to that used in the “Donchin BCI” (Farwell and Donchin, 1988; Allison and Pineda, 2003), this study explored the relationship between two independent variables (stimulus onset asynchrony (SOA) and flash patterns) and three dependent variables (EEG measures, performance, and subjective reports) to assess implications for a P3 BCI. Using the “avenues toward improving information throughput” presented above, the SOA manipulation addresses the first avenue, creating more cognemes per minute, and the “flash pattern” manipulation addresses the second avenue, requiring fewer cognemes to convey a message. It was expected that both manipulations would affect EEG measures, thus affecting the third avenue of signal characteristics.

Fig. 1. “Single Flash” Condition: The 8 × 8 grid consisted of upper and lower case letters and special characters. The top row of matrices shows three of the eight row flashes, while the bottom row of matrices shows three of the eight column flashes used.

Fig. 2. “Multiple Flash” Condition: The top row of matrices shows the three row flashes used in this study, while the bottom row of matrices shows the three column flashes used.
2. Materials and methods

2.1. Subjects

Subjects were 13 undergraduate students at the University of California, San Diego (6 females, age range 18–20 years, mean = 18.9, S.D. = ±0.7). All subjects were free from neurological or psychiatric disorders and were rested and alert. Subjects signed an informed consent approved by the University’s Institutional Review Board and were awarded course credit for participation in the study. Subjects completed a brief questionnaire before EEG preparation and at the conclusion of the study.

2.2. EEG recording

EEG activity was recorded using Ag/AgCl electrodes prepositioned in a standard recording cap from F3, Fz, F4, C3, Cz, C4, P3, Pz, P4, T3, T4, T5, T6, O1, and O2 sites of the International 10–20 System. Active electrode sites were referenced to linked mastoids with a forehead ground. The filter bandpass was 0.01–100 Hz. Eye (EOG) activity was monitored by one electrode placed over the right orbit and filtered at 0.03–100 Hz — also referenced to linked mastoids. All impedances were kept below 5 kΩ, except the eye and forehead sites, which were below 10 kΩ. EEG data were sampled at 256 Hz and analyzed offline.

2.3. Experimental paradigm

Following EEG preparation, subjects were seated in a comfortable chair inside an acoustically- and electrically-shielded isolation chamber. They viewed a monitor containing an 8 x 8 matrix of green characters against a black background (see Figs. 1 and 2). The matrix occupied the central 9 cm of a monitor placed approximately 95 cm from the subject (subtending the central 2.7° of user-centered space). Elements in the matrix consisted of uppercase and lowercase English letters and twelve common symbols. Subjects were visually cued, prior to each run, as to which element in the array was the target for that run. Pseudorandomly selected rows or columns were then flashed sequentially by replacing the matrix of green characters with one of the bitmaps from Figs. 1 and 2 for 100 ms. Thus, from the subject’s perspective, a “flash” consisted of a brief color change from green to yellow lasting 100 ms. The same bitmap was never presented twice in sequence. Subjects counted the number of times the target flashed while ignoring other flashes. Subjects were told that, if they lost count, they should continue counting target flashes and estimate the target count.

Subjects participated in six blocks, each of which contained six runs. Each run consisted of about 240 total flashes, with a different target for each run, for a total of 1440 flashes per block. The total number of flashes in a run was varied within 10% of 240 to ensure that the correct target count varied across runs. The order in which blocks were presented was determined randomly by rolling a six-sided die. Two independent variables were manipulated across the blocks: mean SOA (125, 250, and 500 ms) and flash pattern (single and multiple). The actual SOA between each flash was varied randomly within 10% of the mean. Each of the six blocks utilized a different combination of SOA and flash pattern. Subjects were given a brief break after the third run of each block and after each block.

In the single flash condition, target probability was always the same: one in eight, or 12.5% (single condition). The multiple flash conditions allowed for seven different target probabilities: 0 (none of the six flashes illuminated the target; this applied only to the “=” character on the bottom right); 17% (M17%); 33% (M33%); 50% (M50%); 67% (M67%); 83% (M87%); and 100% (all of the six flashes illuminated the target; this applied only to the “A” character on the top left). The six target letters per block were chosen a priori to ensure a good distribution of different target probabilities. The 0% probable and 100% probable stimuli were designated as targets less often than the other stimuli, as there was only one of each such stimulus.

2.4. Data analysis

Data were sorted, averaged, and event related potentials (ERPs) computed for attended and ignored flashes during each session, and for different flash patterns and SOAs. ERPs consisted of the period from 100 ms before to 900 ms after the onset of each flash, baselined to the period 100 ms before flash onset. Any trial in which the voltage on any channel exceeded ±50 μV was automatically rejected from further analysis. Approximately 15% of trials were thus rejected. Because the P3 and other ERP measures are affected by target probability, the ERPs evoked by different probability targets in the multiple flash conditions were grouped in separate bins for later analysis. That is, ERPs evoked by 17% probable targets were analyzed separately from ERPs evoked by 33% probable targets, and so on. As a result, the single flash bin contained six times more trials than the bins for each multiple flash probability. To maintain an equal number of trials across flash patterns, only one-sixth of the single trial ERPs in the single flash condition were selected at random for statistical analysis and for display purposes.

N1 was scored as the most negative peak between 140 and 200 ms following stimulus onset; P2 was the most positive peak in the 170–260 ms interval; N2 was the most negative peak from 230 to 320 ms; and P3 was the most positive peak from 300 to 500 ms poststimulus. The scored data were then analyzed with SPSS 9.0 using repeated-measures analysis of variance (ANOVA) with Bonferroni’s correction for degrees of freedom. The single flash condition was compared with the multiple 17%, multiple
33%, and multiple 50% conditions, respectively. Each individual analysis consisted of a 4-way ANOVA with factors of flash pattern (single and multiple), SOA (125, 250, and 500 ms), attention (attend or ignore), and electrode (Fz, Cz, and Pz).

Counting accuracy was examined using a three-way ANOVA with Bonferroni’s correction. Factors consisted of run position (whether it was the first, second, third, fourth, fifth, or sixth run in the block), SOA (125, 250, or 500 ms) and flash type (single and multiple).

3. Results

3.1. Counting accuracy

As shown in Fig. 3, counting accuracy, in terms of the mean error rate1, showed a statistically significant main effect of run block position (first = 5.6 ± 2.3%; second = 5.4 ± 1.1%; third = 8.2 ± 1.4%; fourth = 6.1 ± 2.0%; fifth = 7.9 ± 2.3%; sixth = 9.8 ± 2.6%, p = 0.018). Errors generally increased as a function of block. Counting accuracy also showed a main effect of SOA, declining with faster presentation speeds such that the smallest error rate occurred for the slowest speed at 500 SOA (3.7 ± 1.4%); intermediate at 250 SOA (6.2 ± 2.2%); and largest error rate at 125 SOA (11.5 ± 2.8%), (p = 0.037).

Counting accuracy was substantially better in the single flash condition (mean error rate = 4.3 ± 0.02%) compared to the multiple flash condition (10 ± 0.2%), (p = 0.001). In both conditions, errors increased with increasing SOA, with maximal errors for the fastest SOA. In the single condition at 500, 250, and 125 ms SOAs the error rates were 0.9% (±0.5%), 3.1% (±2.2%), and 4.0% (±2.2%), respectively while for the multiple (M17%) condition the errors at 500, 250, and 125 ms were 1.8% (±1.0%), 9.4% (±2.7%), and 18.9% (±3.6%), respectively. As illustrated in Fig. 4, the flash type × SOA interaction also reached statistical significance (p = 0.020), indicating that in the multiple flash condition there is a greater cost (i.e., increased errors) for faster speeds.

All subjects were asked how many hours per day they played electronic games with rapidly changing displays. Eleven subjects reported playing less than 1 h/day, and two subjects (both male) reported playing for more than 3 h/day. The two subjects with more than 3 h/day of gaming experience showed above the mean counting accuracy, especially in the fast SOA. This effect was not statistically significant.

3.2. Subjective reports

The feedback given via the entrance and exit questionnaires is consistent with counting accuracy. All subjects reported more difficulty with the faster speeds, which seemed to require greater attentional effort. Ninety-two percent of subjects (12 out of 13) reported the multiple–fast SOA as more difficult and absorbing than the single–fast SOA, and three subjects stated their dislike for the multiple–fast condition. Eighty-five percent of the subjects (11 out of 13) reported that targets on the edges and corners of the matrix were easier to detect compared to central targets, while the remaining two subjects voiced no preference. Forty-six percent of subjects (6 out 13) reported that punctuation marks were easiest to detect, while one subject felt letters were easier to detect and the remaining six voiced no preference.

Subjects were asked to rate their level of tiredness before and after the study. Fifty-four percent (7 out of 13) reported feeling more tired after the study than before it, while four voiced no change in fatigue. However, all subjects thought they would be able to perform additional runs if desired.

To further explore differences between gamers and healthy subjects, an additional assessment was performed in which five new subjects were presented with the multiple–125 SOA condition and asked whether they had difficulty identifying individual flashes. Three of the five

---

1 Mean error rate reflects the percent deviation between the subject’s count and the actual target count.
subjects (all of whom reported playing action-intensive computer games for more than 3 h/day) reported that this was not difficult, while the other two subjects (both unfamiliar with computer games) reported this as being more difficult.

3.3. Electrophysiology

3.3.1. P3 amplitude

As shown in the grand averages of Figs. 5 and 6, P3 amplitudes to attended events in the single and M17%
condition elicited very large positivities in the 300–500 ms post stimulus window. These responses showed a statistically significant main effect of attention in all three comparisons (single vs. M17%: $F(1,12)=61.3$, $p=0.000$; single vs. M33%: $F(1,12)=95.2$, $p=0.000$; single vs. M50%: $F(1,12)=67.2$, $p=0.000$). The flash pattern × attention interaction was also significant in all three comparisons (single vs. M17%: $F(2,24)=7.8$, $p=0.016$; single vs. M33%: $F(2,24)=10.4$, $p=0.012$; single vs. M50%: $F(2,24)=15.3$, $p=0.004$; means and standard errors in Table 1). As can be observed in Table 1, attend vs. ignore differences were most pronounced in the M17% condition, although the single flash condition had a lower target probability. As shown in Fig. 7, attend vs. ignore differences were less pronounced in the multiple flash condition at higher probabilities (M33% and M50%).

P3 amplitude showed a statistically significant flash pattern × electrode interaction in two comparisons (single

<table>
<thead>
<tr>
<th>Flash pattern</th>
<th>Attention</th>
<th>P2 amplitude</th>
<th>S.E. (±)</th>
<th>N2 amplitude</th>
<th>S.E. (±)</th>
<th>P3 amplitude</th>
<th>S.E. (±)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S12.5%</td>
<td>Attend</td>
<td>4.2</td>
<td>0.49</td>
<td>0.37</td>
<td>0.76</td>
<td>5.4</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td>Ignore</td>
<td>1.9</td>
<td>0.42</td>
<td>-0.85</td>
<td>0.27</td>
<td>1.1</td>
<td>0.31</td>
</tr>
<tr>
<td>M17%</td>
<td>Attend</td>
<td>6.0</td>
<td>1.6</td>
<td>0.63</td>
<td>1.14</td>
<td>8.1</td>
<td>1.8</td>
</tr>
<tr>
<td></td>
<td>Ignore</td>
<td>1.3</td>
<td>0.29</td>
<td>-1.3</td>
<td>0.28</td>
<td>0.80</td>
<td>0.31</td>
</tr>
<tr>
<td>M33%</td>
<td>Attend</td>
<td>4.7</td>
<td>0.48</td>
<td>0.72</td>
<td>0.54</td>
<td>3.7</td>
<td>0.27</td>
</tr>
<tr>
<td></td>
<td>Ignore</td>
<td>2.1</td>
<td>0.30</td>
<td>-0.89</td>
<td>0.20</td>
<td>1.6</td>
<td>0.20</td>
</tr>
<tr>
<td>M50%</td>
<td>Attend</td>
<td>3.7</td>
<td>0.60</td>
<td>0.60</td>
<td>0.36</td>
<td>3.6</td>
<td>0.61</td>
</tr>
<tr>
<td></td>
<td>Ignore</td>
<td>2.3</td>
<td>0.40</td>
<td>-0.75</td>
<td>0.35</td>
<td>2.4</td>
<td>0.46</td>
</tr>
</tbody>
</table>

S.E. = standard error.
Fig. 7. Grand Average ERPs evoked in the multiple flashes (M50%) — 500 ms SOA condition. The bottom lines in each figure represent ERPs evoked by attended flashes, and the top lines in each figure show ERPs evoked by ignored flashes. The y-axis reflects amplitude in mV, and the x-axis reflects time from stimulus onset in ms.

vs. M33%: $F(2, 24) = 7.6, \ p = 0.018$; and single vs. M50%: $F(2, 24) = 10.6, \ p = 0.008$; means and standard errors in Table 2). While P3 amplitude increased from front to back for the single and M17% conditions, the reverse was true for higher probabilities. P3 amplitude also varied significantly as a function of SOA in two of three comparisons, with the third comparison approaching marginal significance (single vs. M17%: $F(2, 24) = 6.0, \ p = 0.031$; single vs. M33%: $F(2, 24) = 3.8, \ p = 0.076$; single vs. M50%: $F(2, 24) = 5.3, \ p = 0.040$; means and standard errors in Table 3). As shown

<table>
<thead>
<tr>
<th>Flash pattern</th>
<th>Electrode</th>
<th>P3 amplitude</th>
<th>S.E. (±)</th>
<th>P3 latency</th>
<th>S.E. (±)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S12.5%</td>
<td>Fz</td>
<td>2.4</td>
<td>0.29</td>
<td>357</td>
<td>9.2</td>
</tr>
<tr>
<td></td>
<td>Cz</td>
<td>3.3</td>
<td>0.35</td>
<td>384</td>
<td>9.0</td>
</tr>
<tr>
<td></td>
<td>Pz</td>
<td>4.0</td>
<td>0.41</td>
<td>377</td>
<td>8.8</td>
</tr>
<tr>
<td>M17%</td>
<td>Fz</td>
<td>3.7</td>
<td>0.75</td>
<td>366</td>
<td>10.9</td>
</tr>
<tr>
<td></td>
<td>Cz</td>
<td>4.7</td>
<td>0.94</td>
<td>379</td>
<td>11.5</td>
</tr>
<tr>
<td></td>
<td>Pz</td>
<td>5.0</td>
<td>0.90</td>
<td>388</td>
<td>11.4</td>
</tr>
<tr>
<td>M33%</td>
<td>Fz</td>
<td>2.9</td>
<td>0.37</td>
<td>351</td>
<td>5.4</td>
</tr>
<tr>
<td></td>
<td>Cz</td>
<td>2.6</td>
<td>0.17</td>
<td>364</td>
<td>8.3</td>
</tr>
<tr>
<td></td>
<td>Pz</td>
<td>2.4</td>
<td>0.27</td>
<td>378</td>
<td>7.3</td>
</tr>
<tr>
<td>M50%</td>
<td>Fz</td>
<td>3.3</td>
<td>0.61</td>
<td>360</td>
<td>6.0</td>
</tr>
<tr>
<td></td>
<td>Cz</td>
<td>3.0</td>
<td>0.50</td>
<td>371</td>
<td>6.9</td>
</tr>
<tr>
<td></td>
<td>Pz</td>
<td>2.7</td>
<td>0.42</td>
<td>384</td>
<td>7.9</td>
</tr>
</tbody>
</table>

S.E. = standard error.

<table>
<thead>
<tr>
<th>Flash pattern</th>
<th>SOA</th>
<th>P3 amplitude</th>
<th>S.E. (±)</th>
<th>P3 latency</th>
<th>S.E. (±)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S12.5%</td>
<td>500</td>
<td>3.8</td>
<td>0.36</td>
<td>356</td>
<td>9.1</td>
</tr>
<tr>
<td></td>
<td>250</td>
<td>3.5</td>
<td>0.37</td>
<td>359</td>
<td>6.9</td>
</tr>
<tr>
<td></td>
<td>125</td>
<td>2.4</td>
<td>0.48</td>
<td>404</td>
<td>13.2</td>
</tr>
<tr>
<td>M17%</td>
<td>500</td>
<td>5.9</td>
<td>1.6</td>
<td>368</td>
<td>10.8</td>
</tr>
<tr>
<td></td>
<td>250</td>
<td>4.7</td>
<td>1.3</td>
<td>371</td>
<td>18.2</td>
</tr>
<tr>
<td></td>
<td>125</td>
<td>2.8</td>
<td>0.40</td>
<td>394</td>
<td>8.1</td>
</tr>
<tr>
<td>M33%</td>
<td>500</td>
<td>3.4</td>
<td>0.51</td>
<td>357</td>
<td>11.7</td>
</tr>
<tr>
<td></td>
<td>250</td>
<td>1.7</td>
<td>0.38</td>
<td>338</td>
<td>11.7</td>
</tr>
<tr>
<td></td>
<td>125</td>
<td>2.8</td>
<td>0.45</td>
<td>399</td>
<td>14.6</td>
</tr>
<tr>
<td>M50%</td>
<td>500</td>
<td>4.6</td>
<td>0.94</td>
<td>361</td>
<td>10.4</td>
</tr>
<tr>
<td></td>
<td>250</td>
<td>1.3</td>
<td>0.53</td>
<td>350</td>
<td>6.1</td>
</tr>
<tr>
<td></td>
<td>125</td>
<td>3.1</td>
<td>0.54</td>
<td>404</td>
<td>11.3</td>
</tr>
</tbody>
</table>

S.E. = standard error.
in Fig. 5, in the single and M17% conditions, P3 amplitude increased with slower SOAs. However, that trend did not hold for the higher probabilities, where the 250 ms SOA produced the smallest P3 potentials.

3.3.2. P3 latency

P3 latency varied significantly with electrode in all three comparisons (single vs. M17%: \( F(2,24)=7.4, p=0.015 \); single vs. M33%: \( F(2,24)=13.9, p=0.004 \); single vs. M50%: \( F(2,24)=10.7, p=0.007 \); see Table 3). P3 latency was typically longest at parietal sites and shortest at frontal sites. While this distinct scalp distribution is widely reported in the literature, it has not yet been used to discriminate different P3s, as could be useful in a BCI. P3 latency also varied significantly with SOA in all three comparisons (single vs. M17%: \( F(2,24)=-6.3, p=0.026 \); single vs. M33%: \( F(2,24)=8.0, p=0.016 \); single vs. M50%: \( F(2,24)=11.5, p=0.006 \); see Table 3). P3 latency was typically longest for the fastest SOA.

3.3.3. N1 amplitude

N1 amplitude was marginally significant as a function of SOA in two of the three comparisons (single vs. M33%: \( F(2,24)=4.1, p=0.065 \); single vs. M50%: \( F(2,24)=4.3, p=0.062 \)). There was also a statistically significant interaction between SOA and flash pattern in all three comparisons (single vs. M17%: \( F(2,24)=4.5 p=0.048 \); single vs. M33%: \( F(2,24)=5.3, p=0.039 \); single vs. M50%: \( F(2,24)=4.5, p=0.055 \); see Table 3). N1 amplitude was always larger in the 250 SOA condition, particularly in the multiple flash conditions.

3.3.4. N1 latency

N1 latency differed significantly with attention in all the three comparisons (single vs. M17%: \( F(1,12)=20.4, p=0.017 \); single vs. M33%: \( F(1,12)=22.0, p=0.002 \); single vs. M50%: \( F(1,12)=9.4, p=0.016 \). N1 latency was roughly 10 ms longer in the ignore condition in all flash patterns studied. N1 latency differed significantly with electrode in two of three comparisons with the third bordering on marginally significant (single vs. M17%: \( F(2,24)=8.0, p=0.016 \); single vs. M33%: \( F(2,24)=15.2, p=0.003 \); single vs. M50%: \( F(2,24)=3.9, p=0.072 \). N1 latencies were generally longer at posterior sites. This effect was slight in the single flash condition, barely visible in the multiple 17% condition, yet became more pronounced with higher probabilities.

3.3.5. P2 amplitude

P2 amplitude showed main effects of attention, electrode, and SOA in all three comparisons. The flash pattern \( \times \) attention interaction was marginally significant in one of the three comparisons (single vs. M50%: \( F(2,24)=5.269, p=0.051 \). P2 amplitude was larger to attended flashes than ignored flashes. In the M17% and M33% conditions, the P2 amplitude difference was greater than in the single flash condition (single vs. M17%: \( F(1,12)=46.2, p=0.000 \); single vs. M33%: \( F(1,12)=61.9, p=0.000 \); single vs. M50%: \( F(1,12)=39.6, p=0.000 \); see means and standard errors in Table 2). For electrode comparisons, P2 was typically largest at central sites and decreased both anteriorly and posteriorly (single vs. M17%: \( F(2,24)=6.0, p=0.033 \); single vs. M33%: \( F(2,24)=11.0, p=0.007 \); single vs. M50%: \( F(2,24)=4.8, p=0.049 \). P2 amplitude generally decreased with faster SOAs and was typically larger over frontal and central than parietal sites (single vs. M17%: \( F(2,24)=7.1, p=0.019 \); single vs. M33%: \( F(2,24)=6.3, p=0.026 \); single vs. M50%: \( F(2,24)=6.0, p=0.031 \).

3.3.6. P2 latency

No statistically significant effects of P2 latency were noted. Though P2 latency was longer in ignored than attended trials in all three bins, this effect was only significant in the single vs. M33% comparison (\( F(1,12)=6.784, p=0.031 \).

3.3.7. N2 amplitude

N2 amplitude varied significantly with attention in two of three comparisons (single vs. M33%: \( F(1,12)=5.5, p=0.047 \); single vs. M50%: \( F(1,12)=6.0, p=0.040 \). Ignored events produced larger N2s than attended events. N2 amplitude varied significantly with SOA in two of the three comparisons (single vs. M33%: \( F(2,24)=7.5, p=0.018 \); single vs. M50%: \( F(2,24)=5.9, p=0.032 \). N2 amplitude becomes more positive with slower SOAs. This is most likely a result of the overlap with positive components, including P2, P3 or both.

3.3.8. N2 latency

The flash pattern \( \times \) electrode interaction was significant in two of three comparisons (single vs. M33%: \( F(2,24)=9.3, p=0.011 \); single vs. M50%: \( F(2,24)=6.5, p=0.025 \). N2 latency was generally longer in posterior sites in the multiple flash conditions, while the opposite trend appeared in the single flash condition.

4. Discussion

The purpose of this study was to explore two different methods of structuring stimulus presentation in a standard P3 BCI paradigm. One method enabled the user to generate more cognemes per minute by varying stimulus presentation speed, while the second resulted in requiring fewer cognemes to convey a message by varying the pattern of stimulus presentation. These manipulations affected signal robustness, subject performance, and subjective preferences. Although the behavioral performance and subjective preferences argue that the standard single flash methodology is easier to use and preferred by untrained subjects, especially at faster SOAs, the multiple flash methodology at lower target probabilities yielded a more robust difference between
target and nontarget flashes that may lead to easier classification of target ERPs and thus improved information throughput.

One of the most important factors in a P3 BCI is a robust difference between ERPs evoked by attended versus ignored cognemes. The more apparent this difference is to an artificial pattern recognition system, the greater the system’s accuracy and speed in identifying target cognemes. Alternatively, information throughput can be improved by requiring fewer trials to attain a specified accuracy threshold. Results from this study suggest that attend vs. ignore differences are more apparent after multiple flashes than single flashes at low target probabilities (M17% versus S12.5%). This difference is less apparent at high target probabilities. Only six cognemes are needed in the multiple flash conditions to identify a character, compared to sixteen cognemes in the single flash condition, which means that the former approach would operate much more quickly if the resulting ERPs could be categorized as accurately. Furthermore, the magnitude of the signal at faster SOAs was more robust in the multiple flash conditions. Finally, the multiple flash conditions showed differential spatial distributions of P3 amplitude as a function of stimulus probability. That is, like the single flash condition, P3s to M17% showed increasing magnitude at more centroposterior sites. However, this trend was reversed for higher probability conditions (M33% and M50%). Such differences in spatial distribution could facilitate pattern recognition. Thus, the present results provide some support for multiple flash presentations as a better basis for a P3 BCI system. While the ultimate goal is to assess these benefits in an on-line BCI, studies such as the present one help identify avenues toward improvement that may prove useful.

P3 amplitude is proportional to flash speed and target probability (see Gonsalvez and Polich, 2002) both of which influence target-to-target intervals (TTI). Stimuli presented more quickly, whether due to a faster SOA or a higher target probability, may generate more cognemes per minute. However, they also produce a reduction in P3 amplitude, which results in less robust EEG differences between flash patterns. Hence, the SOA manipulation reflects a tradeoff between stimulus speed and signal robustness. Similarly, P3 amplitude is inversely proportional to target probability; and the single flash approach ensures that target probability will always be low. Hence, different flash patterns also present a tradeoff between the number of cognemes and signal robustness avenues for improvement. Since the multiple flash approach may require fewer events to identify which of 64 grid elements is the target, fewer cognemes are necessary to send a message. However, the multiple flash approach also increases the mean target probability, thus reducing TTI and hence P3 amplitude (Fig. 8).

The expected effect of faster flashes producing smaller P3s was only apparent in the multiple flash low probability condition (M17%) but not in the M33% and M50% conditions, in which P3 amplitude was actually larger in response to faster flash. The multiple flashes thus have an advantage over the single flash condition, in that a BCI using this methodology may produce EEG signals that are more robust at faster speeds. The fact that attend versus ignore differences were most pronounced in the multiple 17% compared to the single 12.5% condition suggests that some other facet of the multiple flash condition, besides global stimulus probability, contributes to the larger P3 responses. One possibility may be increased stimulus and task complexity. The relationship between complexity and P300 amplitude is not a simple one (Ullsperger and Mecklinger, 1996; Ullsperger et al., 1987, 1990; Katayama and Polich, 1998; Wintink et al., 2001) and requires further research.

The P3 was typically about 40 ms slower in the fast condition than in the other conditions. P3 latency is often correlated with task difficulty (Ullsperger et al., 1990; Leuthold and Sommer, 1998; Smulders et al., 1999). Again,
subjects indicated in their questionnaires that they perceived the fast SOAs, especially in the multiple flash condition, as the most difficult. This difference in latency would probably have only a minor effect on a pattern recognition system.

The results do not provide strong support for a particular SOA. As expected, there is a tradeoff between stimulus presentation speed and ERP measures. Faster flashes allow more cognemes per minute, but in some cases produce less distinct ERPs. Some subjects had trouble with the multiple flash approach at the fastest speed. Therefore, the question of which SOA is ideal depends on a number of factors, e.g., the user’s preference and/or facility for processing fast flashes, and the pattern recognition approach used. One pattern recognition approach might perform best when given the large but noisy attend vs. ignore difference apparent after one or two slow flashes, while another might be better able to identify the smaller but more reliable difference seen after four or eight fast flashes. Although some classic studies have addressed this important question of SOA in the context of a P3 BCI (e.g., Farwell and Donchin, 1988; Donchin et al., 2000, Meinicke et al., 2003), it seems unlikely that any SOA will prove best for all users and all conditions.

The discussion has so far focused only on the P3. However, other components of the event-related potential may also vary with selective attention and may provide additional information to a pattern recognition system. In the present study, the N1, P2, and N2 amplitudes and latencies showed changes related to the task. N1 amplitude changed as a function of SOAs, with largest responses evoked by the 250 ms SOA in the multiple flash conditions. N1 latencies to higher probability targets in the multiple flash conditions exhibited a more distinct spatiotemporal distribution across midline sites, being slower at more centroposterior than frontal sites. Attend/ignore differences reflected in P2 are more pronounced in the multiple flash condition than the single flash condition. Despite the relatively low probability of single flashes, this suggests that P2 amplitudes do not vary with probability alone. The increased P2 amplitude difference in multiple flash conditions is likely due to the increased task and stimulus complexity seen in that situation. This difference becomes less pronounced at higher probabilities, as attended P2 amplitude decreased and ignored P2 amplitude increased, a result similar to those for P3 amplitude. Finally, N2s to ignored events tended to be larger than to attended events. This may have occurred because the N2 overlaps two components known to be more positive in attended trials: the P2 and P3.

P3 BCIs have focused primarily on the P3 component as the primary correlate of the cognemes: “/attend to the event/” or “/ignore the event/”. However, all P3 BCIs published to date except (Polikoff et al., 1995) used pattern classification parameters that could be influenced by ERPs preceding the P3. Hence the term “P3 BCI” may be problematic. It is also possible that target events produce other EEG changes in addition to effects on ERP measures described above. For example, attended vs. ignored events can produce changes in rhythmic EEG activity (mainly alpha) in addition to ERP changes (Spencer and Polich, 1999; Yordanova et al., 2001; Jung et al., 2001). This underscores the important of improved signal processing and pattern recognition of data acquired in a so-called “P3” BCI. (See Farwell and Donchin, 1988; Allison, 2003; Kaper et al., 2003; or Xu et al., 2003 for comments and examples regarding non-P3 components in a P3 BCI.)

Furthermore, the variable probability seen in the multiple flashes condition could be of value to an advanced pattern recognition system. Consider the following situation: a user wishes to send the “/%” icon. This icon is only illuminated in one of the six flashes (the top right image in Fig. 2). The flash containing the “/%” icon produces a large P300 complex, but another flash (for example, the flash shown in the bottom right image in Fig. 2) produces a small and indeterminate result. The BCI thus must decide whether the user attended to one of the six flashes only, meaning the “/%” is the target, or whether the user attended to two of the flashes, meaning the “$” is the target. Because the ERPs evoked by 17% probable and 33% probable targets look distinctly different, the BCI could use this knowledge to help identify the target. In this case, the “$” icon is probably not the target because had the subject been counting two of the six flashes, target probability would have been 33% and the large response evoked by the “/%” icon would have been smaller. This “probability matching” approach is not possible with the single flash approach, as the probability never changes and thus the ERPs evoked by different flashes all look very similar. Thus, the multiple flash approach allows a new avenue for resolving uncertainty. “Probability matching” would require a sophisticated pattern recognition system.

While the primary objective of this study was to explore the relationship between SOA, flash patterns, and ERP measures, other important variables were also examined. Since subjects were asked to count targets, it was possible to compare performance across different flash patterns, providing an objective measure of task performance. Counting accuracy is relevant to ERP measures. That is, if subjects are not counting accurately, it may imply they did not see each individual target, and hence would not generate robust ERPs. Subjects were also given brief questionnaires prior to and following the study to explore the potential role of background and lifestyle factors, as well as their subjective preferences for different flash patterns. Subjective reports regarding BCI ease of use are an important issue, especially for systems for the severely disabled that may be used for several hours a day.

Counting accuracy was excellent at the slowest SOA in both flash conditions. However, faster flashes resulted in greater number of errors in the multiple flashes compared to the single flash condition. Faster SOAs appeared to make the task more difficult and hence produce a greater cost in accuracy, especially with multiple flashes. Subjects in the single flash condition could count well at all speeds, but subjects in the multiple flash conditions had increasingly...
more difficulty with faster speeds. Poor counting accuracy may occur because subjects did not see each individual flash. In this case, P3 amplitude would be reduced since the P3 occurs only in response to detected targets. However, poor accuracy may also occur if subjects detect target flashes but lose count. In this case, P3 amplitude should remain large. The data suggest that subjects were able to detect targets but easily lost count in the fast SOA conditions. At least four subjects reported losing count in the multiple–125 SOA condition, and one of these subjects reported the same problem in the multiple–250 SOA condition. It is possible that alternate instructions, such as simply noting the target rather than maintaining a target count, would have produced different results, although differences would probably be minor.

The increasing error rate across runs within a single recording session was likely due to mild fatigue, as performance declined only within a block of runs and not across blocks or between recording sessions. This suggests that subjects using this P3 BCI would benefit from frequent breaks. Indeed, performance improved following a break after the third set of runs. Subjects with more experience and motivation would likely perform better over extended periods.

Subjects also reported a preference for peripheral targets. That is, they found it easier to detect targets when they occurred in the sides or corners of the $8 \times 8$ matrix. This is perhaps due to flanking effects, in that central targets were surrounded by eight nontarget distractors, while fewer nontargets surrounded targets on a corner or side. This suggests that future P3 BCI designers might consider placing commonly chosen letters in the sides and corners to maximize ease of use.

The subjective report that special characters were easier to detect compared to letters may have occurred because punctuation marks are more distinct icons than letters; hence, P3 BCIs should strive toward using distinct icons if possible. It seems likely that even more visually distinct grid elements, such as pictures, words, or different colored elements, would be easier to discriminate.

Finally, although subjects reported feeling tired at the end of the study, they were willing to continue. Hence, while the use of a P3 BCI in a non-distracting environment may be boring and even tiring (as suggested by decreasing counting accuracy within blocks), it is not so exhausting as to necessitate rest. Very recent work (Allison and Moore, submitted for publication) showed that it is possible to effectively use an on-line P300 BCI in a noisy, distracting environment for several hours. Thus, P3 BCIs may be more practical for sustained use than some other types of BCIs, such as those used by the Kennedy or Birbaumer groups, in which subjects often report the need for frequent breaks during and after BCI use (Kennedy et al., 2000; Hinterberger et al., 2003).

The observation that gamers are better at, and more comfortable with, tasks involving faster displays is not surprising (e.g., Green and Bavelier, 2003). This suggests two potentially useful conclusions for future BCI designers. First, individuals with prior experience with rapidly changing displays may be better suited to BCIs using faster stimulus presentation. Second, users can be trained to recognize faster flashes. No study to date has explored the effects of long-term use of a P3 BCI system, but it is likely that most users could be trained to perform well with multiple flash–fast SOA condition. Thus, while some subjects in this study disliked and performed poorly in such a condition, their preference and performance may well have changed had they used the BCI for a longer period. Long-term use of a P300 BCI might have other effects on EEG measures (e.g., Sommer and Schweinberger, 1992; Romero and Polich, 1996), performance, and preference.

The effects of both sustained and long term use – including issues such as training for both the user and computer, optimal feedback parameters, side effects of BCI use, customized interfaces and applications, and addressing the unique needs and abilities of each user – should ideally be explored in real-world environments, such as a patient’s home. This also ensures that other potential concerns, such as distractions, noise from medical or other electronic devices, or effects of medication, do not seriously affect results. Although some types of BCIs, especially those based on slow cortical potentials or SCPs, have been extensively tested in patients’ homes in long term studies (e.g., Birbaumer et al., 1999; Neuper et al., 2003), P300 BCIs have only very recently been validated with ALS patients in their homes2 (Bayliss et al., 2004; Mellinger et al., 2004; Sellers et al., 2004). There are no published long-term studies with P300 BCIs.3

This study sought to determine whether the multiple flash approach might be a viable alternative to the conventional single flash approach. The advantage of the multiple flash approach – the need for fewer events or cognemes to convey a message – must be considered along with the disadvantages of a less discriminable signal in some circumstances and a task that some users found significantly more difficult at faster speeds. Results suggest that the multiple approach may be preferable, especially in subjects comfortable with more complex and rapidly changing displays. To further explore this question, future research should explore the multiple flash approach in an online BCI. This should ideally be done in a long-term study with patients, in their home environments, using an advanced pattern recognition system capable of accounting for the different ERP and other EEG activity associated with targets of different probabilities.

---

2 None of these ALS patients were fully locked in (email communication with the three lead authors, Feb. 2005). Hence, P3 BCIs still need to be validated with locked in patients.

3 The first author (BZA) and colleagues are currently exploring training with P3 BCIs.
Acknowledgements

We would like to thank Drs. Andrey Vankov and John Polich for help with the design of the study.

References


