Changing the P300 Brain Computer Interface

JESSICA D. BAYLISS, Ph.D., SAMUEL A. INVERSO, M.S., and ALEKSEY TENTLER, M.S.

ABSTRACT

Brain–computer interfaces (BCIs) are now feasible for use as an alternative control option for those with severe motor impairments. The P300 component of the evoked potential has proven useful as a control signal. Individuals do not need to be trained to produce the signal, and it is fairly stable and has a large evoked potential. Even with recent signal classification advances, on-line experiments with P300-based BCIs remain far from perfect. We present two potential methods for improving control accuracy. Experimental results in an evoked potential BCI, used to control items in a virtual apartment, show a reduced response exists when items are accidentally controlled. The presence of a P300-like signal in response to goal items means that it can be used for automatic error correction. Preliminary results from an interface experiment using three different button configurations for a yes/no BCI task show that the configuration of buttons may affect on-line signal classification. These results will be discussed in light of the special considerations needed when working with an amyotrophic lateral sclerosis (ALS) patient.

INTRODUCTION

RECENT ADVANCES in signal processing have made brain–computer interfaces (BCIs) feasible for use as an alternative control option for those with severe motor impairments. Two main classes of BCIs exist: implanted devices and those that use electrodes on the scalp. Philip Kennedy’s group uses an implanted device for basic patient communication, although the eventual hope is that implants will allow patients that are completely locked-in to move again with prosthetic devices driven from the implant.

BCIs that use electrodes on the scalp currently have bit rates similar to those of implanted devices and have the added benefit of noninvasiveness. They also use well-studied electroencephalography (EEG) signals for classification. BCIs are most commonly used for communication—being a patient’s number one priority when completely locked-in, without even remaining eye movement abilities. Two types of externally based BCIs exist: (1) those based on operant conditioning and (2) evoked potential devices. The first type uses operant conditioning in order to teach subjects control of their ongoing EEG signals. These devices have the benefit of being subject-controlled so that over time subjects may control items faster as their control over specific EEG signals increases. For example, the Cyberlink system may use alpha and/or beta waves and is in use with several patients. Unfortunately, experiments commonly state one out of five subjects have difficulty learning control, and it may take several months for an ALS patient to learn control.

Evoked potential devices are simpler for subject’s to learn to control, as they are based on inherent evoked responses to known stimuli. Our current results indicate that a locked-in subject can learn to control a simple yes/no application in a single session using the P300 component of the evoked potential for control. Around 1964, Chapman and Bragdon independently discovered a positive wave peaking at around 300 msec after task-relevant stimuli. This component is known as the...
P300, and its general form is shown in Figure 1. While the P300 is evoked by many types of paradigms, the most common factors that influence it are the frequency of stimulus occurrence (less frequent stimuli produce a larger response) and task relevance. The P300 has been shown to be fairly stable in locked-in patients, re-appearing even after brain stem injuries. This makes it possible as a control choice for locked-in individuals. Farwell and Donchin first showed that this signal may be successfully used in a BCI. The P300 is a non-specific response, meaning that it occurs in response to a wide variety of stimuli types. Using a broad cognitive signal, like the P300, has the benefit of enabling control through a variety of modalities, because the P300 may enable discrete control in response to both auditory and visual stimuli.

BCIs remain slow and prone to error with system bit rates ranging from 6 bits/min in the Thought Translation Device as used by actual patients, to much higher speeds of around 85 bits/min for an off-line experiment with an evoked potential-based BCI. Given that the bit rate of a typist who can type around 60 words/min is approximately 1200 bits/min, there is much room for improvement.

There are two main ways that the bit rate of a BCI may be improved: increase the speed at which a BCI operates, or increase the accuracy of the BCI system. The speed of the BCI may be changed by making clever interfaces, by increasing the speed of making a control decision, or even by predicting what decision the user wants to make. Increasing the accuracy of the BCI system is another possibility and has been looked at the most heavily. Many algorithms have been looked at to increase the classification rate of both evoked potential-based BCIs and spontaneous EEG BCIs.

We propose two ways of increasing the accuracy of an evoked potential P300 BCI: automatic error correction and altering the user interface. The first method uses existing information embedded in the EEG signal in order to correct errors. Off-line analysis of BCI data collected from controlling items in a virtual apartment indicates that a P300 signal exists after a goal item has been controlled and a reduced response occurs after a control mistake. With a theoretical mean improvement in recognition from 78% to 85%, we show a statistically significant improvement ($p < 0.008$, Wilcoxon signed rank test) in accuracy of 3% using the correlation algorithm. This improvement is shown using two simple and well-understood algorithms: peak picking and correlation. The simplicity of the algorithms and resulting improvements show the robustness of the effect.

Since automatic error correction is based on the response to control an item and not on the algorithm used in classification for BCI control, the possibility exists for using this technique in multiple types of BCIs. We discuss future work in this context.

Another method for increasing the accuracy of a P300 BCI is through the interface used to display the stimuli for control. An effect was shown with preliminary experiments in the VR experiment, but not enough data was collected to verify the significance of the effect. Another experiment has since been run with a simple yes/no control interface. The interface contains a condition with a single button that flashes yes/no, a two-button condition with separate buttons for yes/no that stimulate the P300 response by a red color flash, and a two-button condition where the color of the flashes changes to green midway through the condition. Results again are inconclusive due to a timing glitch with display in a Windows environment; however, results indicate that subjects subjectively prefer the two-button to the one-button application. It is of note that the single ALS patient who participated in the yes/no button experiment had no preferences other than that the device worked for her. Future work will be discussed within this context.

**MATERIALS AND METHODS**

**Automatic error correction**

We can make use of extra information embedded in the signals recorded by a BCI in order to make decisions. This possibility has not been heavily studied, since increasing accuracy through using more

**FIG. 1.** (Solid line) The general form of the P3 or P300 component of the evoked potential (EP). The P3 is a cognitive EP that appears approximately 300 msec after a task-relevant stimulus. (Dotted line) The general form of a non-task-related response.
information may actually decrease the speed of a system and the end bit rate.

Wolpaw et al. suggested the use of response verification (RV) for applications that need high accuracy more than they do high bit rate.\textsuperscript{19} Response verification occurs when the classification of a choice for an individual has been made and the individual is then asked to verify that choice by making a further choice. Two types of errors are possible in a BCI: missing the correct choice (a false-negative error) or making a choice accidentally (a false-positive error). RV may improve accuracy by reducing the occurrence of false positive mistakes in the data. Experimentally, response verification has led to higher accuracies at the cost of a lower overall bit rate.

We propose to use automatic error correction to increase BCI accuracy without decreasing the bit rate. We are not the first to suggest automatically correcting errors. Parra et al. perform automatic response error correction for users accidentally hitting a mouse button in a visual discrimination task.\textsuperscript{18} They use the error-related negativity (ERN) event-related potential for automatic correction. This potential is known to occur when subjects mistakenly respond to a stimulus and realize that they’ve made a mistake. Unfortunately, the ERN signal is difficult to use in a BCI since subjects do not normally “accidentally” make a mistake—instead a combination of faulty signal classification and faulty subject control over EEG signals combine to cause mistakes.

Since the ERN is difficult to use in a BCI, we propose an error correction method that depends on the use of the evoked potential P300 component. In order to use this component, it must be shown that the component exists in an error situation and that it can be reliably classified to increase classification.

The VR versus non-VR experiment

The experimental set-ups for the VR and non-VR environments were almost identical. The main difference was that all graphics for the non-VR environment were rendered on a 21-inch Silicon Graphics monitor, whereas the VR environment conditions were displayed in a head-mounted display (HMD).

Five objects or commands could be controlled by the user in the virtual apartment as shown in Figure 2: a lamp, a stereo system, a television set, a “Hi” command, and a “Bye” command. The lamp, stereo, and television all worked as toggle switches to turn the items on/off. The Hi and Bye commands made a three-dimensional graphics figure appear (for “Hi”) or disappear (for “Bye”). All responses to commands were visual—for instance, musical notes appeared over the stereo when the stereo was on.

A sphere associated with each controllable object blinked in the environment; when visible, it had a semi-transparent red coloring. Semi-transparency was used so that blinking spheres would be less distracting to subjects concentrating on one specific sphere for a task.

Approximately once per second, a stimulus was provided when the sphere on a randomly chosen item appeared. The stimulus would last for approximately 250 msec. The stimulus presentation rate varied by up to 16 msec in a random manner. The P300 response occurs for task-relevant stimuli. To make the red sphere flashes on the controllable object task-relevant, subjects had to count the flashes on a goal item.

Seven electrode sites were arranged on the heads of nine subjects with a linked mastoid reference. Sites Fz, Cz, Pz, P3, P4, as well as an upper and lower vertical electro-oculographic (EOG) channel were used from the International 10–20 system of placement.\textsuperscript{9} For on-line recognition and analysis, EOG artifacts were regressed out of the signals of interest using the algorithm by Semlitsch.\textsuperscript{10}

The EEG signals were amplified using Grass amplifiers with an analog bandwidth of 0.1–100 Hz. Electrode impedances were 2–10 kOhms for all subjects. An epoch size from $-100$ msec (prior to stimulus onset) to 1500 msec was specified for a total epoch size of 1600 msec. The data were recorded continuously and saved to a file.

The experiment consisted of four tasks:

1. Calibration: The subject counted the number of sphere flashes located on the virtual lamp on a monitor.
2. VR condition: The subject was fully immersed in the virtual apartment while wearing a HMD.
3. Monitor condition: The subject looked at the virtual apartment on a monitor.
4. Fixed display condition: The subject looked at the virtual apartment on a fixed screen inside of the HMD.

The calibration task was used to train a signal processing algorithm on a particular subject’s P300 signal response. A total of 300 stimulus presentations were presented to each subject. In this task, subjects were told to count only the lamp sphere flashes; thus, in this task only the lamp sphere flashes were task-relevant, and these flashes should have caused a P300 response. Since the spheres flashed randomly over the five controllable items, approximately 60 lamp flashes occurred over the course of 5 min.

Tasks 2–4 were accomplished in a randomized block order and lasted for approximately 5 min each (250 stimulus presentations with the sphere flashed randomly on items). These tasks involved on-line single trial classification of the P300 in order to control the different items in the apartment. The time taken for these trials depended on how many items were controlled, as subjects received feedback for each item for which the signal classification algorithm classified the trial as a P300 component trial.

Due to the difficulty of signal classification, false-positive mistakes (accidental control) were possible as well as true goal control. In tasks 2–4, the subjects received instructions in English at the bottom of the screen indicating what goal to achieve, and each subject attempted to achieve that goal by counting the number of flashes on the sphere located on that particular goal item. During each task, the goal was chosen randomly, and the subject tried to achieve the goal for up to 50 presentations of the goal stimulus. When the goal was achieved, an action involving visual feedback occurred in the virtual apartment (e.g., the room was lightened when the light was turned on). During the waiting period for this visual feedback (1.5 sec), no new stimuli were presented. Then, the next goal was randomly chosen.

P300 component existence

While the experiment involved on-line classification and feedback, an off-line analysis was done to compare the obtained P300 components between different conditions. Only epochs with a maximum vertical EOG signal of less than 50 microvolts were used. This reduced the possibility of EOG contamination of the averages. From this analysis, it was discovered that the P300 signals obtained in the different environments were not significantly different from each other.23

Since all items have a discrete control, evoked responses where items are successfully controlled may be examined. The P300 component occurs when subjects choose a goal item, and a reduced signal occurs when a mistake in classification is made. Grand averages showing this result appear in Figure 3.

FIG. 3. (A) The grand averages for control of goals at site Pz (solid lines) shown with the grand averages for non-goals (dashed lines) in each experimental condition. (B) The grand averages over nine subjects for responses to goal item control (the solid lines) and mistakes in item control (the dashed lines).
Since subjects could have blinked or moved directly after an item was controlled, all trials used in the grand averages had a maximum recorded vertical eye movement of less than 50 microvolts. The grand average for vertical eye movement is shown, and while it is not flat, there are no peaks around 400 msec, when the maximal P300 component appears.

The maximum signal for Fz and Cz are slightly larger than the signal for Pz. It is possible that subjects found controlling items in a virtual apartment to be “novel,” and that would lead to a more frontal response. Controlled goal items are task relevant because the subject achieves control and may go on to the next task. It is hypothesized that false-positive mistakes do not generally cause this response, since subjects do not always have to correct errors, so they are ignored.

Using an evoked potential for automatic error correction

If a BCI informs the user of his selection with a response, the response should cause an evoked potential in the user. For example, in the virtual apartment an evoked potential should occur when an item is controlled, such as the lamp lighting. An evoked potential’s presence can then indicate if the user intended to select the stimuli and be used to correct erroneous non-goal selections.

The method to use an evoked potential for response verification (RV) is similar to the response verification procedure described in Wolpaw et al.\textsuperscript{19} However, instead of prompting the user to verify the selection with a Yes/No trial, the system automatically accepts or discards the selected stimuli if the evoked potential is present or absent.

The predicted RV accuracy, \( C \), is the probability that a decision is correct.

\[
C = \frac{pq}{pq + (1 - p)(1 - q)}
\]

The P300 is a natural choice for evoked-potential response verification. Responses are task-relevant stimuli that should evoke the P300. In this way, selected goals and non-goals can be distinguished based upon the presence or absence of the P300 component.

Classification algorithms

For analysis, the epochs corresponding to the nine subjects’ responses from the virtual apartment experiment monitor condition were divided into goal and non-goal responses. For example, when lamp lighting was the goal, it was a goal response, while turning on the TV when the lamp was a goal was a non-goal response. These were further divided into responses with P300s present and P300s absent, yielding four categories of responses: goal P300 present, goal P300 absent, non-goal P300 present, and non-goal P300 absent.

Two algorithms (peak picking and correlation\textsuperscript{20}) were considered separately to recognize a P300 component in the Fz, Cz, and Pz channels. Response epochs containing signal amplitude greater than 90 mV were ignored. This accounted for artifacts not corrected through regression. Ignoring a response resulted in the stimuli remaining triggered. For example, if the lamp was lit, it would stay lit.

Two methods were experimented with for accepting and discarding responses. The first method compared accepting the response if the P300 component was present on Fz, Cz, Pz, or any of the three channels, while the second method accepted the response if the P300 component was absent on Fz, Cz, Pz, or all of the channels. Both methods were tested with the two different P300 classification algorithms.

Peak picking

Peak picking is a simple algorithm to classify a P300 component using the difference between the minimum and maximum amplitude in an epoch. An epoch with a prototypical P300 signal contains a large peak around 300 msec, peak picking recognizes a P300 when the amplitude difference is greater than or equal to a specified voltage difference between the minimum and maximum voltage points within a specified time window.

\[
\max(x) - \min(x) \geq \tau \quad \{ \begin{array}{ll} 1: & \text{P3 Present} \\ 0: & \text{P3 Absent} \end{array} \}
\]

where \( x \) is a vector that represents the data for a single EEG response and \( \tau \) represents the threshold voltage difference required to accept a P300. For recognition, the time window with the best result was 300–600 msec. The voltage difference threshold was varied in experiments to yield the best result.
Correlation

A slightly more complex algorithm, but still easily calculated, is correlation. Correlation may be looked at as template matching when the correlation is performed between single trials and a template of what each kind of trial should look like. EEG responses were correlated with P3 and non-P3 averages using the following formula:

\[ \rho_{x,y} = \frac{\text{cov}(x, y)}{\sigma_x \sigma_y} \]

where \( x \) is a vector which represents the data for an EEG response, \( y \) is a vector which represents the P3 or non-P3 average to compare against, is the covariance of \( x \) and \( y \), and \( \sigma \) is the standard deviation of the appropriate signal.

To determine if an EEG response contains a P3, the correlation between the EEG response epoch and the P3 average was compared with the EEG response epoch and non-P3 average according to the algorithm:

\[
\text{if } \rho_{P3} > \tau \text{ and } \rho_{\text{nonP3}} \geq \rho_{\text{nonP3}} \text{ then P3 Present else P3 Absent}
\]

where \( \rho_{P3} \) is the correlation between the EEG response and P3 average, \( \rho_{\text{nonP3}} \) is the correlation between the EEG response and non-P3 average, and \( \tau \) is the threshold set to determine the extent of the desired correlation. The threshold was varied in experiments to yield the best results. If the EEG response did not correlate with a P3 or non-P3 average the algorithm is indeterminate. During experimentation, all indeterminate responses were treated as not having P3s.

Classification

Table 1 compares the best RV accuracies achieved by the two algorithms against the original accuracy. Peak picking and correlation achieved the same overall RV accuracy, 81%; however, correlation more significantly impacted the subjects accuracies, \( p < 0.008 \), than peak picking, \( p < 0.04 \) (\( p \) values were derived using the Wilcoxon signed rank test).

In addition, Table 1 presents the theoretically best accuracy that can be achieved in this experiment when error correcting mistakes are present. Overall, it was possible to change the accuracy from 78% to 85%. This indicates that 15% of the errors were caused by missed goals, while 7% were caused by mistakes, and of those mistakes, 3% were corrected using correlation response verification.

Table 2 shows the bits per minute that the subjects achieved when keeping the response if a P3 was present. The bits per trial equation used was derived by Pierce\(^2\) (and originally from Shanon and Weaver\(^2\)):

\[
B_t = \log_2 N + P \log_2 P + (1 - P) \log_2 \frac{1 - P}{N - 1}
\]

where there are \( N \) possible selections, each of which are equally probable, and \( P \) accuracy probability. Bits per minute \( B_m \) is the bits per trial \( B_t \) multiplied by the average number of trials per minute \( T_m \):
In this the virtual apartment experiment, $N$ was 5 stimuli and $T_m$ was 12 trials/min.
The overall increase in bit rate was 1.26 bits/min, and the greatest increase was 4.01 bits/min by subject 5. Subject 5’s change reflects the benefit in using automatic response error correction. From Table 1, we see subject 5’s original accuracy was 67%, with a large amount of errors from mistakes 13%. Through response error correction, the subject achieved 77% accuracy.

Improving the interface

When classification accuracy is improved in a BCI, often off-line experimental results do not match on-line results. This may be due to the on-line feedback that occurs when subjects use a BCI. As in most interfaces, speed and accuracy is related to the design of the interface, although developing usable BCIs is a studied challenge for system designers. The complexity of the hardware, real-time processing demands of the software, and small number of test users can leave little time for usability testing and tuning.

This is changing. Moore and Kennedy overview human–computer interface and training issues for an implanted BCI.12 The user moved a cursor by imagining movements in his left hand. This resulted in a communication rate of about three letters per minute.

When classification accuracy is improved in a BCI, often off-line experimental results do not match on-line results. This may be due to the on-line feedback that occurs when subjects use a BCI. As in most interfaces, speed and accuracy is related to the design of the interface, although developing usable BCIs is a studied challenge for system designers. The complexity of the hardware, real-time processing demands of the software, and small number of test users can leave little time for usability testing and tuning.

This is changing. Moore and Kennedy overview human–computer interface and training issues for an implanted BCI.12 The user moved a cursor by imagining movements in his left hand. This resulted in a communication rate of about three letters per minute.

### Table 2. Change in the Bit Rate from the Original Bit Rate when Keeping Responses with P3 Signals Using the Peak Picking and Correlation Algorithms to Recognize P3s

<table>
<thead>
<tr>
<th>Subject</th>
<th>Original, bits/min</th>
<th>Peak pick</th>
<th>Correlation</th>
<th>Theoretical, bits/min</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>14.45</td>
<td>13.44</td>
<td>14.98</td>
<td>17.8</td>
</tr>
<tr>
<td>2</td>
<td>14.23</td>
<td>15.4</td>
<td>15.4</td>
<td>20.48</td>
</tr>
<tr>
<td>3</td>
<td>17.66</td>
<td>19.53</td>
<td>19.81</td>
<td>20.38</td>
</tr>
<tr>
<td>4</td>
<td>14.45</td>
<td>14.72</td>
<td>15.27</td>
<td>18.86</td>
</tr>
<tr>
<td>5</td>
<td>9.09</td>
<td>12.87</td>
<td>13.1</td>
<td>14.3</td>
</tr>
<tr>
<td>6</td>
<td>14.62</td>
<td>16.13</td>
<td>15.74</td>
<td>16.92</td>
</tr>
<tr>
<td>7</td>
<td>12.37</td>
<td>13.13</td>
<td>13.91</td>
<td>16.46</td>
</tr>
<tr>
<td>8</td>
<td>12.92</td>
<td>13.15</td>
<td>12.69</td>
<td>15.62</td>
</tr>
<tr>
<td>9</td>
<td>11.63</td>
<td>11.88</td>
<td>11.88</td>
<td>12.63</td>
</tr>
<tr>
<td>Mean</td>
<td>13.49</td>
<td>14.47</td>
<td>14.75</td>
<td>17.05</td>
</tr>
<tr>
<td>SD</td>
<td>2.38</td>
<td>2.32</td>
<td>2.22</td>
<td>2.65</td>
</tr>
</tbody>
</table>

The best theoretical accuracy occurs if all selected goals are kept while all selected non-goals are rejected.

\[
B_m = B_i T_m
\]

In this the virtual apartment experiment, $N$ was 5 stimuli and $T_m$ was 12 trials/min.

The overall increase in bit rate was 1.26 bits/min, and the greatest increase was 4.01 bits/min by subject 5. Subject 5’s change reflects the benefit in using automatic response error correction. From Table 1, we see subject 5’s original accuracy was 67%, with a large amount of errors from mistakes 13%. Through response error correction, the subject achieved 77% accuracy.

In work with brain injury patients, Cole et al. note that system developers’ design processes must change.23 They note that “our work with brain injury patients has shown that patient-system performance is extremely sensitive to . . . minor design parameters: furthermore, that the brain injury survivor needs to be viewed as a relatively sensitive component, while the computer system design needs to be the most flexible.” In fact, about two thirds of user interface changes and three quarters of the functionality were requested by patients or clinicians. They conclude “it is clear that at least some . . . changes would not have been suggested by those with systems expertise [developers] because those changes were either counter-intuitive or violated accepted guidelines.”

The accepted P300 component BCI interface violates the simple guideline that items should not consistently and continuously flash at users. This flashing is necessary since it is required to indicate a stimulus onset for users. Even in this situation, some interfaces are more difficult to use than others. This has to do with the cognitive nature of the P300a and P300b components. The P300a is more frontal and is larger due to novelty effects, while the P300b is the more task-relevant component of the P300. This information may be used in a P300-based BCI.

In addition, some BCI users may not easily be able to focus on the multiple elements used for BCI control on a computer screen. These subjects are completely locked-in and don’t have reliable control over their eye movements. We wanted to test the
efficacy of flashing all choices in the same area of the screen versus having more than one choice flashing spatially on the computer screen. This was done with a simple yes/no interface that consisted of a one-button yes/no flash condition and a two-button condition where one button contained the yes command and the other contained the no command.

Novelty and the P300-based BCI

In the VR experiment previously described, some subjects commented that the semi-transparent red bubbles that flashed were too strong and needed to be toned down. A pilot experiment was thus begun with three of the nine subjects. This experiment added a fifth condition: the different color condition. In this condition, buttons flashed a variety of pastel colors and were more transparent than the original red bubble flashes. The fifth condition was done at a different time in the regular experiment for each subject, although it was always done after the training task. The performance results of these subjects are shown in Table 3. The task with different colored bubbles is labeled “monitor with different colors” as it is basically the monitor condition without red colored flashing bubbles. It may be seen that all subjects who tried this task performed better than they did on the monitor condition, but the results are not statistically significant due to the small number of subjects. These preliminary results were surprising from such a small interface change and are most probably due to the novelty of the different button flashes.

Yes/no button experiment

Experimental data was collected to ascertain whether introducing novelty into the stimuli or presenting all the stimuli in the same area, as opposed to different areas, had an effect on the P3 component for a BCI control device. The basic application design was to flash a specified amount of randomly interspersed “yes” and “no” stimuli, recording and saving the data during the process.

Recordings were acquired using Ag/AgCl electrodes from Fz, Pz, and Cz sites, along with P3, P4, C3, C4, and F3 and F4 in some of the experiments, according to the International 10–20 system of placement. In addition, two linked mastoid reference electrodes were attached to the mastoid processes behind both ears, a ground electrode was attached to the forehead, and an electrode was attached above and below each eye to filter eye-blink artifacts.

There were nine total subjects (eight male high school/college-age, and one female ALS patient). When a subject came into the lab, (s)he was asked to sit down in front of the computer screen and the background and goals of the experiment were explained. The subject was then fitted with an electrode cap and face electrodes, the impedance values checked, and the electrodes adjusted until impedance values below $5\text{kOhms}$ were achieved.

In the experiment, all subjects were asked to use the three application GUIs described below. The order of each GUI was determined randomly. For each GUI, seven trials were conducted consecutively, with each trial composed of 34 total flashes of the stimuli. The target response for each trial alternated between “yes” and “no.” For each trial, the subject was told they could do one of two tasks to make the target stimulus task relevant. They could either count the number of times that the target response flashed, or they could ask themselves a question that had the target response as the answer and visualize achieving the answer each time the target response flashed. Counting the number of times the target flashes is the more traditional method for obtaining a P3 component, but in pilot experiments several subjects had described trying to boost the meaning of the yes/no answer by coming up with their own meaningful questions that had the goal answer. As an example, one subject visualized a Porsche car to make the “yes” answer meaningful.

To analyze the ease in getting P3 potentials with different presentation methods, three applications were created:

1. The base application (two-button) consisted of two large “yes” and “no” buttons that flash randomly, by turning the color of the flashed button’s text from black to blue. This application is shown in Figure 4.

2. A second application (novelty two-button) was made to look exactly like the two-button application, with the exception that at a random point during the flashing of the stimuli, the flashing

<table>
<thead>
<tr>
<th>Conditions</th>
<th>Subject</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>6  8  9</td>
</tr>
<tr>
<td>Monitor</td>
<td>1.2  1.3  2.9</td>
</tr>
<tr>
<td>Monitor with different colors</td>
<td>1.8  1.7  5.0</td>
</tr>
</tbody>
</table>
color of both buttons changed from blue to dark green.

3. A third application (one-button) was made with just one stimulus display area, on which either “yes” or “no” could be displayed.

The two-button condition is the base condition. The novelty two-button condition was done to study the effects of novelty on the P300 component in a BCI while the one-button interface was to study the possibility of stimulus location. It is possible that if one stimulus is more in the center of the subject’s field of view, that answer could be unintentionally given bias over the other answer which might be more in the periphery of the field of vision. Previous experiments\textsuperscript{20,23} show that, if an object can only be seen in peripheral vision, the accuracy has a slight drop.

Before the official trials started, the subject was given four sample runs to get used to the system and to get an idea of what concentration technique they felt more comfortable with (envisioning the answer or counting the targets). The sample runs were done with whichever application GUI was chosen to be used first. At the end of each trial, the system’s response was shown to the subject. At the end of the experiment, the subject would be disconnected from the equipment, and asked for his/her opinion on which of the three GUIs was the easiest or hardest to use and for what reason(s).

RESULTS

Results from the experiment loosely mirrored the VR experiment, although the data could not be shown to be significant due to a variable length glitch in stimuli time stamping due to using the Windows Operating System. The subjective experiences of the subjects were tabulated. Of the five individuals with an application preference, four out of them preferred the two-button application, citing the one-button application as “hard to use.” According to subjects, this difficulty was due to the predictability of the two-button application, since if the subject saw the left or the right button flash, (s)he instantly knew whether it was a “yes” or a “no” that had flashed, while with the one-button application, more time was needed to actually read the sign before knowing which stimulus it had been. The ALS

FIG. 4. The two-button task application. The boxes on the right of the screen normally contain information about subject performance for automatic subject feedback.
patient had no preferences in yes/no interface, although it is the first application she has felt she has some measure of control over.

The two-button GUI with the color change was almost the same as the regular two-button GUI according to subjects. Participants were split on the efficacy of the application with some saying that the color change seemed to boost their attention and make it easier for them to concentrate and others saying that the found the color change distracting.

**DISCUSSION**

We have shown two methods for improving P300-based BCI accuracy that do not rely on improving the signal classification of single trial evoked potential data. These methods may be used in combination with alternate classification schemes for an overall improved BCI system. In order to use the automatic error correction method presented, it is necessary to know when device control occurs in order to measure whether or not a P300 occurs in response to the control signal. This may not be relevant for all BCI devices. Future work needs to show if this response occurs when the main BCI is not P300-based and more sophisticated algorithms for classification of the response should be used.

Our first work with an ALS subject took place for the yes/no button experiment. There are many differences between working with healthy individuals and an ALS subject. The first one discovered was system interference between the ALS subject’s respirator and the EEG signal acquisition. A 30-Hz low-pass filter was added to the system to help reduce respirator-related noise. Usually, EOG artifacts contaminate some portion of the data and we often don’t count this data for off-line analyses. The ALS subject had no such artifacts, and we have since stopped recording EOG-related activities due to their infrequency.

Work is currently in progress to reproduce the one- and two-button experiment without the timing glitch discovered in previous recordings. While the timing glitch interfered with on-line BCI classification and responses for the yes/no interface, this interface still represents the first interface that the ALS subject felt control over. All previous interfaces used by this subject had been experimental operant control-based BCIs in other labs.

For future work, the ALS individual and her family have suggested a three-button interface might be the most useful. This type of application would have a yes, no, and accept button combination as well as a morse code spelling interface with a dot, dash, and space combination. In addition, the subject’s family likes to see her direct evoked potential responses along with her answer. This gives the family a chance to verify correct operation of the classification algorithm in the BCI and makes them feel “in the loop.” We are planning on implementing such an application for the future.

**ACKNOWLEDGMENTS**

This research was supported by NIH/PHS grant 1-P41-RR09283. It was also facilitated in part by a National Physical Science Consortium Fellowship and by stipend support from NASA Goddard Space Flight Center.

**REFERENCES**


Address reprint requests to:
Dr. Jessica Bayliss
102 Lomb Memorial Dr.
Rochester, NY 14623-5608
E-mail: jdb@cs.rit.edu