



Evaluating Brain-Computer Interface Performance Using Color in the P300 Checkerboard Speller

D.B. Ryan¹, G. Townsend², N. A. Gates, K. Colwell³, and E.W. Sellers¹

¹Department of Psychology, East Tennessee State University, Johnson City, TN, USA

²Department of Computer Science, Algoma University, Sault Ste. Marie, Ontario, Canada

³Department of Electrical and Computer Engineering, Duke University, Durham, NC, USA

Abstract

Objective—Current Brain-Computer Interface (BCI) systems typically flash an array of items from grey to white (GW). The objective of this study was to evaluate BCI performance using uniquely colored stimuli.

Methods—In addition to the GW stimuli, the current study tested two types of color stimuli (grey to color [GC] and color intensification [CI]). The main hypotheses were that in a checkerboard paradigm, unique color stimuli will: (1) increase BCI performance over the standard GW paradigm; (2) elicit larger event-related potentials (ERPs); and, (3) improve offline performance with an electrode selection algorithm (i.e., Jumpwise).

Results—Online results (n=36) showed that GC provides higher accuracy and information transfer rate than the CI and GW conditions. Waveform analysis showed that GC produced higher amplitude ERPs than CI and GW. Information transfer rate was improved by the Jumpwise-selected channel locations in all conditions.

Conclusions—Unique color stimuli (GC) improved BCI performance and enhanced ERPs. Jumpwise-selected electrode locations improved offline performance.

Significance—These results show that in a checkerboard paradigm, unique color stimuli increase BCI performance, are preferred by participants, and are important to the design of end-user applications; thus, could lead to an increase in end-user performance and acceptance of BCI technology.

Correspondence to: David B. Ryan, Mountain Home VA Health Care System, Audiology 126, Phone: +1-423-926-1171 Ext. 7575, Ryand1@goldmail.etsu.edu.

Conflict of Interest Statement

None of the authors have potential conflicts of interest to be disclosed.

Preliminary results have been presented at the 40th Annual Meeting of Society for Neuroscience, San Diego, CA, USA (November, 2010).

This research was conducted while working at the Department of Psychology, East Tennessee State University, Johnson City, TN, USA.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Keywords

Assistive Devices; Brain-Computer Interface; EEG; P300 Event-Related Potential; Color Processing

1 Introduction

A Brain Computer Interface (BCI) is a specific type of human-machine interaction. BCI is a direct link between the human brain and a computer. BCIs can be defined as invasive utilizing techniques requiring implantation such as electrocorticography (ECoG), or non-invasive techniques utilizing techniques requiring surface electrodes such as electroencephalography (EEG [Wolpaw et al. 2003]). A BCI can provide an important communication outlet for those who are “locked-in” by amyotrophic lateral sclerosis (ALS), brain stem stroke, or head trauma. Locked-In Syndrome (LIS) refers to a condition where all voluntary muscles, except those that control eye movement, are completely paralyzed. ALS is a progressive motorneuron disease that causes irreversible loss of motor function. As ALS symptoms progress, most individuals lose the ability to speak clearly or at all resulting in a locked-in state. Currently, people with ALS, particularly those in advanced stages, are the most likely to benefit from BCI technology because of their severe motor disability (Vaughan et al. 2006; Birbaumer et al., 2000).

One of the more common non-invasive forms of input for the BCI is the P300, an event-related potential (ERP) component (Sutton et al., 1965). Farwell and Donchin (1988) modified the oddball paradigm to use the P300 as a virtual typing or spelling device. Alternative classification algorithms (e.g., support vector machines and principal component analysis) and alternative stimulus presentation paradigms have improved BCI performance in terms of both speed and accuracy (Kaper, et al., 2004; Kaufmann et al., 2011; Kaufmann et al., 2013; Townsend et al., 2010).

1.1 Human Color Processing and P300 BCI

Anllo-Vento et al. (1998) examined the ERPs of selective attention to a color stimulus. ERP components associated with color processing were observed in the time window of 100–240ms. The majority of activation associated with color processing and can be seen in both hemispheres but consistently seen in the left hemisphere (McKeeffry and Zeki 1997; Zeki et al., 1991). Thus, the electrode locations and temporal windows used for P300 BCI classification should benefit from additional color processing components, as reported in previous research (Takano et al., 2009).

Takano et al. (2009) sought to improve BCI accuracy by changing the way the characters flashed with three different paradigms; a white/grey pattern (luminance condition), a green/blue isoluminance pattern (chromatic condition), and a green/blue luminance pattern (luminance chromatic condition). Of the ten participants, four had a significantly higher accuracy in the chromatic condition and five had a significantly higher accuracy in the luminance chromatic condition. The authors took advantage of the activity associated with large groups of neurons in the parietal, occipital, and temporal brain areas involved in color

processing, and the occipital and parietal areas involved in luminance processing. Presumably, this additional activity enhanced the EEG signal and provided a stronger response to the target items. Similar results were found when a grey/white stimulus and green/blue stimulus were tested in a sample of people with spinal cord injuries (Ikegami et al., 2011).

Additionally, Anllo-Vento et al., (1998) discovered that changes in color can be selectively attended to within multi-feature stimuli, suggesting a facilitation of attention to a uniquely colored item in a P300 speller matrix. These studies show additional patterns of activation and facilitated attention in response to color over monochromatic stimuli. In a color P300 BCI, this additional activation is expected to be present as an increase in amplitude on the electrodes over the occipital and parietal lobes (Oz, O1, O2, POz, PO3, PO7), with a higher increase in the left hemisphere.

1.2 Current Study

The current study combined two successful previous paradigms, a checkerboard paradigm and a color paradigm (Ikegami et al., 2011; Takano et al. 2009; Townsend et. al., 2011). The checkerboard paradigm already controlled for double flashes and four of the eight adjacent non-target flashes that surround the target (i.e. above, below, left, and right of target item; Townsend et al., 2010). The current color paradigm has expanded on the previous color paradigms with the addition of unique colors to adjacent items. By defining each item in the matrix with one of nine colors (e.g., a 3×3 of uniquely colored items), the target stimulus is a different color than all adjacent stimuli. The adjacent non-target flashes provided limited distraction through the unique color assignment. We hypothesized that the checkerboard color matrices will produce higher accuracy, faster communication, and higher personal preference over the traditional gray-to-white (GW) checkerboard paradigm. We hypothesized that the color stimuli in the GC and CI conditions would enhance ERPs in the parietal/occipital electrodes, which capture brain activity of color processing. The additional color processing features increase signal-to-noise ratio and improve classification accuracy. Additionally, we hypothesized that personal preference of a certain paradigm will facilitate attention and result in higher performance in the preferred paradigm over a paradigm of lesser preference.

2 Methods

2.1 Participants

Thirty-six participants, 11 male and 25 female (mean age = 24 years, age range = 19–47, 6 had previous BCI experience), were recruited from the participant pool at East Tennessee State University. All participants reported normal vision (with habitual correction) and had no color deficiencies based on the Ishihara colorblindness (Ishihara, 1917). All participants signed informed consent. The study was approved by the East Tennessee State University Institutional Review Board.

2.2 Stimuli and Materials

The traditional BCI visual presentation is a character matrix that has grey characters; each stimulus presentation changes a subset of the matrix characters from grey to white (Grey to White, GW; Figure 1a). The present study examined the traditional stimulus presentation method and two novel stimulus presentation methods that manipulate the color of the matrix items and the stimulus presentations. Here we will refer to a change in the color of a stimulus as a “flash.” All the paradigms use the checkerboard paradigm (Townsend et al., 2010).

The second condition is referred to as the color intensified (CI) condition (see Figure 1b). The matrix items are one of nine different colors so that no adjacent characters have the same color. Each character intensifies its assigned color (e.g., Figure 1b, red ‘C’ [stimulus off] to bright red ‘F’ [stimulus on]), instead of flashing from grey to white.

The third condition was Grey to Color (GC, Figure 1c). GC used grey stimulus off items but they change to an assigned color for stimulus on that is distinct from all adjacent flashes, similar to the CI paradigm.

2.3 Procedure

Each session consisted of two phases. Participants were seated in a chair approximately 1m from a computer monitor that displayed the 8×9 matrix (see Fig. 1). Stimuli in the matrix were presented in groups of six (randomized by the checkerboard paradigm), the standard onset asynchrony (SOA) was 125.0 ms and the inter-stimulus interval (ISI) was 62.5 ms. Data from five six letter words were collected. Paradigms were presented in a pseudorandom order; each participant was randomly assigned to one order of the three visual presentations: GW, CI, and GC.

Phase 1 consisted of a copy-spelling task to collect calibration data for the classifier, which was used for online classification of each character selection in Phase 2. During calibration the participant was given a word to spell and prompted to attend to each letter of the word in sequential order.

During Phase 2, the participant engaged in the same copy-spelling task as Phase 1 with the addition of feedback of the item selected. That is, the participant attended to the letter “A” and then the system presented the item selected by ERP classification (e.g, “A” correctly classified or any other selection as incorrectly classified). Each condition consisted of Phase 1 and 2. Each condition was administered on a separate day. No more than one week passed between the administration of each condition.

2.4 EEG Acquisition and Processing

EEG was recorded with a 32-channel electrode cap using tin electrodes (Electro-Cap International, Inc.). All channels were grounded to the left mastoid and referenced to the right mastoid. Two g.tec (Guger Technologies) 16-channel USB biosignal amplifiers were used to amplify the minute electrical activity recorded from the scalp. The electrical signal from the 32 channels was amplified ($\pm 2V$ before ADC), digitized at 256 Hz, high-pass filtered at 0.5 Hz, and low-pass filtered at 30 Hz. Only eight electrodes were used for

classification during Phase 2 of the study (Fz, Cz, P3, Pz, P4, PO7, PO8, and Oz; (Krusienski et al., 2006). All channel impedances were reduced below 10.0 k Ω before recording. BCI2000 (Schalk et al., 2004) controlled stimulus presentation, data collection, and online processing.

2.5 Classification

The EEG data from the 36 item selections served as training data for a stepwise linear discriminate analysis (SWLDA) solution, which was then used for online response classification and feedback for an additional 36 selections. To estimate the optimal number of stimulus flashes for Phase 2 online testing, a cross validation procedure was applied to the data. Specifically, the optimal written symbol rate (WSR) was determined for each participant utilizing Phase 1 data and the resulting number of flashes was implemented in Phase 2. WSR is an offline estimate of how many correct selections a participant will make in one minute, including correcting erroneous selections. This measure only takes into account correctly selected items and does not count the deletion of a character as this does not contribute to the message (see Townsend et al., 2010 for WSR explanation in detail). The WSR provides a conservative estimate of the fewest number of character flashes to maximize accuracy. Thus, in Phase 2, the most accurate participants were presented with fewer flashes than their less accurate counterparts.

2.6 Performance Measures

The three conditions' online (i.e., Phase 2) performance was measured by three metrics: 1) accuracy, number of correct selections divided by the total selections, 2) information transfer rate (ITR), an objective measure that incorporates accuracy, number of items in a matrix, and selections per minute, and 3) theoretical information transfer rate (tITR), ITR with time between selections removed to allow comparison between studies.

2.7 Jumpwise Channel Selection

An offline analysis was conducted using the Jumpwise channel selection algorithm (Colwell et al., 2011). The algorithm uses a SWLDA technique to analyze a channel's worth of EEG features in a regression model. Each channel is added in turn to the model and the channels that are most relevant, as determined by a partial *f*-test, are selected. After Jumpwise selects the eight optimal channel locations, the data from these eight channels are used as input to a SWLDA that determines the optimal set of temporal feature weights.

2.8 Self-report measures

Following the final session, the participant was asked to report 1) the condition they preferred, and 2) the reasons for their preference. Condition preference was recorded by ranking the conditions in order of preference (e.g., 1-GW, 2-GC, 3-CI) and was used as a covariate in performance measure analysis. Participants' reasons for a preference were not analyzed; they were only used to help understand what informed their preference. These reasons for preference will be used to help develop surveys for use in future studies. To measure eye strain, a self-report measure was administered at the beginning and at the end of each session on a scale of 1 to 5 (i.e., with a score of 1 indicating no eye strain, and a

score of 5 indicating very strained eyes), the difference between the before and after scores provided a level of eye strain caused by each type of presentation.

2.9 Statistical Methods

Phase 2 data were used to compare performance across the three paradigms. Accuracy, ITR, and tITR were all computed and compared. Statistical comparisons were conducted with repeated a measures analysis of variance (ANOVA) and planned contrasts were conducted with least significant difference (LSD) tests. Greenhouse-Geisser corrections were used if there was a violation of sphericity. We hypothesized that the personal preference of an individual may affect performance. Thus, condition preference (GC, CI, or GW) was entered into a repeated measures ANOVA as a covariate for performance measures.

Waveform analyses were only performed on Phase 1 data (i.e., calibration); in Phase 2 data sets of the participant contained unequal numbers of ERPs due to the WSR optimization. That is, a condition that required fewer flashes would have fewer waveforms than a condition required more flashes with more waveforms. Unequal data points between conditions could alter the statistical test results and visualization of ERP waveforms. Therefore, calibration data of Phase 1 was used for analysis due to stimulus presentations being held constant across conditions. Statistical comparisons were conducted with repeated measures ANOVAs and planned contrasts were conducted with least significant difference (LSD) tests.

3 Results

3.1 Performance Measures

Prior to personal preference being entered as a co-variate into the repeated measures ANOVA for all performance measures, personal preference was entered into a repeated measures ANOVA as a variable to determine if the results vary with preference. The ANOVA was significant, $F(2,70) = 3.28$, $p = 0.044$, $\eta^2 = 0.086$.

3.1.1 Accuracy—Significant differences were observed, with Greenhouse-Geisser correction, in accuracy between the three conditions (see Table 1). Mean accuracy was 85.97% (SD = 9.67) in the GC condition, 81.17% (SD = 15.57) in the CI condition, and 81.81% (SD = 12.53) in the GW condition. Planned contrasts revealed that accuracy in the GC condition was significantly higher than accuracy in the CI condition and the GW condition.

3.1.2 Information Transfer Rate (ITR)—Similar to the accuracy results, ITR was significantly different in the three conditions. Mean ITR was 21.82 (SD = 7.52) in the GC condition, 19.04 (SD = 7.38) in the CI condition, and 19.73 (SD = 7.43) in the GW condition. Planned contrasts revealed that ITR in the GC condition was significantly higher than ITR in the CI condition and the GW condition.

3.1.3 Theoretical ITR—Theoretical ITR also was significantly different among the three groups. Mean tITR was 28.54 (SD = 11.79) in the GC condition, 24.44 (SD = 10.96) in the CI condition, and 25.63 (SD = 11.38) in the GW condition. Planned contrasts revealed that

tITR in the GC condition was significantly higher than ITR in the CI condition and the GW condition.

3.2 Jumpwise Channel Selection

A subsequent offline performance analysis was carried out to compare the Jumpwise selected channels to the standard set of channels. Data from two participants could not be entered into the offline analysis due to an online recording error, which was required to predict offline Jumpwise performance; however, the recording error did not affect online performance. Thus, data from 34 of the 36 participants were used in the offline analysis.

Each participants' data was separately entered into the Jumpwise algorithm for each of the three conditions. The algorithm analyzed all 32 recorded channels and selects the eight channels that account for the most variance for frequency of channel selection (see Figure 2).

A 2×3 repeated measures analysis was performed using the three stimulus conditions (GC, GW, CI) and the two channel sets, Jumpwise and standard (i.e., channels selected by the algorithm and original classification channels, respectively) as independent variables. The preferred condition was entered as a covariate using a LSD post hoc analysis. The main effect of condition and channel set were not significant in the accuracy measure or in the ITR measure. There was not a significant interaction of channel set and condition. The main effect for channel set was significant in the tITR measure $F(1,32) = 4.63, p = 0.039, \eta^2 = 0.126$. The mean tITR for the Jumpwise channel set was 28.1 (SD=2.03) and for the standard channel set was 25.5 (SD=2.11).

3.3 Preference and Eye Fatigue

Nineteen of the 36 participants preferred the GC condition (52.8%). Eight preferred the CI condition (22.2%), and 9 preferred the GW condition (25.0%). A chi-square test of goodness-of-fit was performed to determine if the three conditions were equally preferred. Preference was not equally distributed across conditions, $X^2(2, N=36) = 6.167, p < .05$. Participants who preferred the GC condition said that the colors facilitated the maintaining of their attention to the target and found the GW non-target flashes distracting. The opposite was reported by the participants that preferred GW, that the GC non-target flashes were distracting. Those who preferred the CI condition liked that the target letter never changed color, making it easier to keep their attention on the target item even when it was not flashing.

The self-reported eye fatigue measure revealed no differences between conditions, or between pre and post task reports. This could be due to the low sensitivity of the scale (i.e., 1–5) or the short duration of the task. Further studies should consider a higher resolution of a 1–10 scale or longer task duration.

3.4 Waveform Analysis

The waveform analyses were limited to seven of the electrode locations used for online classification. The seven electrodes were placed into two averaged waveform groups, which

were (1) a Midline group comprised of Fz, Pz, and Oz; and, (2) a parietal/occipital group (P/O) comprised of P3, P4, PO7, and PO8. This grouping limited the number of statistical tests to two electrode groups rather than seven individual tests per measure and allowed for a representation from the classification electrodes and limited the chance of a Type I Error. Amplitude and latency analyses were conducted on time windows determined by visual inspection of the group-averaged waveforms (see Figure 3). Specifically, four amplitude and latency measures were compared via repeated measures ANOVAs (positive peak amplitude, positive peak latency, negative peak amplitude, and negative peak latency) for the two electrode groups (Midline and P/O) for the three conditions (GW, GC, CI).

3.4.1 Positive Peak Amplitude—Mean positive peak amplitude of the Midline group (all conditions) was 3.33 μV (SD=1.70) and for the P/O group 2.75 μV (SD=0.95). Mean positive peak amplitude in the Midline group for the GC condition was 3.55 μV (SD=1.35), for the GW condition was 3.23 μV (SD=1.42), and for the CI condition was 3.21 μV (SD=1.12). Mean positive peak amplitude in the P/O group for the GC condition was 2.99 μV (SD=1.03), for the GW condition was 2.65 μV (SD=1.03), and for the CI condition was 2.61 μV (SD=0.75). The Group x Condition interaction was not significant $F(2,70)=0.07$, $p=0.938$. The ANOVA revealed a significant main effect of electrode group $F(1,35)=70.97$, $p<0.001$, $\eta^2=0.670$ and condition (see Table 2). Planned contrasts revealed that amplitude in the GC condition was significantly higher than amplitude in the GW condition and the CI condition. The GW and CI conditions were statistically equivalent.

3.4.2 Positive Peak Latency—Mean positive peak latency of the Midline group (all conditions) was 276.97 ms (SD=24.32) and for the P/O group 281.13 ms (SD=35.06). Mean positive peak latency in the Midline group for the GC condition was 269.60 ms (SD=35.07), for the GW condition was 282.90 ms (SD=32.95), and for the CI condition was 278.4 ms (SD=21.50). Mean positive peak latency in the P/O group for the GC condition was 273.40 ms (SD=38.22), for the GW condition was 287.00 ms (SD=37.74), and for the CI condition was 283.00 ms (SD=32.48). The Group x Condition interaction was not significant $F(2,70)=0.019$, $p=0.981$. There was no main effect of electrode group $F(1,35)=3.81$, $p<0.059$; however, there was a significant effect of condition. Planned contrasts revealed that latency in the GC condition was significantly shorter than GW.

3.4.3 Negative Peak Amplitude—Mean negative peak amplitude of the Midline group (all conditions) was $-3.06 \mu\text{V}$ (SD=1.66) and for the P/O group $-2.25 \mu\text{V}$ (SD=0.98). Mean negative peak amplitude in the Midline group for the GC condition was $-3.28 \mu\text{V}$ (SD=1.48), for the GW condition was $-2.99 \mu\text{V}$ (SD=1.18), and for the CI condition was $-2.86 \mu\text{V}$ (SD=1.17). Mean negative peak amplitude in the P/O group for the GC condition was $-2.41 \mu\text{V}$ (SD=1.15), for the GW condition was $-2.26 \mu\text{V}$ (SD=0.93), and for the CI condition was $-2.08 \mu\text{V}$ (SD=0.85). The Group x Condition interaction was not significant $F(2,70)=1.71$, $p=0.189$. The main effect for electrode group was significant, $F(1,35)=139.23$, $p<0.001$, $\eta^2=0.80$; however the main effect of condition was not significant $F(2,70)=2.43$, $p=0.095$.

3.4.4 Negative Peak Latency—Mean negative peak latency of the Midline group (all conditions) was 498.90 ms (SD=50.58) and for the P/O group 509.27 ms (SD=49.81). Mean negative peak latency in the Midline group for the GC condition was 506.30 ms (SD=49.49), for the GW condition was 495.2 ms (SD=51.03), and for the CI condition was 495.2 ms (SD=44.52). Mean negative peak latency in the P/O group for the GC condition was 507.90 ms (SD=52.26), for the GW condition was 505.6 ms (SD=54.48), and for the CI condition was 514.3 ms (SD=46.72). There was a significant interaction between electrode group and condition $F(2,70) = 5.45, p=0.006, \eta^2=0.135$ (Figure 4). The main effect of electrode group was significant $F(1,35) = 12.60, p=0.001, \eta^2=0.265$; however there was no significant effect of condition $F(2,70) = 0.055, p=0.580$.

In summary, the GC condition had higher positive peak amplitude than CI and GW and shorter latency than GW. The P/O group had higher negative peak amplitude and longer latency than the Midline group. There was a significant interaction of electrode group and condition.

4 Discussion

We hypothesized that by defining each item in the matrix with one of nine colors; the adjacent non-target flashes provided limited distraction thus, improving performance and preference over the standard GW condition. This discussion will cover the attention and color processing enhancement properties of the GC stimulus, the role of personal preference, and Jumpwise channel selection. The GC condition resulted in higher performance measures than the CI and GW conditions. The key to the success of the GC condition was the facilitation of attention through stimuli manipulation, resulting in an enhanced ERP response. This technique has shown to improve ERP based BCI performance (Takanio et al., 2009). Enhancing the ERP components provides the classifier more salient features to categorize targets and non-targets. Enhanced components result in the classifier requiring less data (i.e., fewer flashes) to make a correct selection, as indicated by higher ITR and tITR compared to the other conditions. In this study, the more salient stimuli of the GC condition elicited ERP components that the classifier could categorize with higher accuracy and fewer stimulus presentations. If selections require fewer ERPs, then the end-user will use fewer attentional resources to carry out each item selection, thereby providing a higher rate of communication (per each unit of time), BCI use of a longer duration, and higher user satisfaction. It could be argued that a stronger focus of attention was utilized to obtain better performance in fewer flashes. Nonetheless, the GC condition utilized the enhanced ERPs of passive color processing, as the task did not require participants to identify the color of the item's flash.

The color stimulation of GC condition resulted in two positive peaks within the time window for positive peak analysis. The initial positive peak in the GC condition was higher in amplitude and shorter in latency than the positive peak in the GW condition. The initial positive peak in the GC condition was also higher in amplitude than the CI condition. The additional peak in the GC condition was most likely due to color processing and not a P3a, as the topography of the initial peak was most prominent in the parietal/occipital electrodes and attenuated over the frontal/central electrodes. This topography is opposite that expected

of a P3a and P3b (Polich, 2009). The waveform differences could be due to underlying occipitotemporal activity. It is possible that the waveform differences are a result of lateral excitation. In the GC condition, the color processing neurons have been inhibited by single color stimulus (e.g., stimulus off, grey); thus, enhancing their response to a color stimulus (e.g., stimulus on, yellow). In the GW and CI conditions, the color processing neurons are only processing a change in intensity of the same color, suggesting that habituation of color processing neurons in the GW and CI conditions and excitation in the GC condition resulted in a change in waveforms. The GC stimulus was similar to the stimulus used in previously mentioned color processing studies, a black to color checkerboard was used in both Allison et al. (1993) and Anllo-Vento et al. (1998). The time window of the positive response in the current study (113–310 ms) overlaps the color processing activity described by Allison et al. (1993) and Anllo-Vento et al. (1998). This finding suggests that in the GC condition, the activity of these color processing areas was responsible for the initial positive peak and did not elicit an initial positive peak for the GW condition. Therefore, the waveform analysis resulted in higher amplitude and shorter latency of the positive peak in the GC condition compared to the GW condition. Moreover, the activity associated with the color processing of GC stimuli could enhance the classification process resulting in higher classification accuracy.

Although there is color processing associated with CI, the reduced activity in the CI condition, compared to GC, could be a result of habituation of the color processing neurons. Tailby et al. (2008) describe the habituation of V1 color processing neurons in single unit recordings of macaque monkeys after 30s of single color stimulus duration. For each selection, participants in CI would have to focus on a single color (e.g., target) for 15s. For this 15s duration the foveated color processing neurons would be receiving only one color, the stimulus off (e.g., red) and stimulus on (e.g., bright red). The extended exposure to one color could have habituated the corresponding color neurons enough to lower amplitude recorded at the scalp thus, provided fewer salient features and affected classification. Nevertheless, the CI condition did not have lower accuracy than GW condition. Interestingly, the non-target waveforms in the CI condition did not have the prominent 8Hz oscillation found in GW or GC conditions. The lack of activity in response to non-targets could have compensated for the lower amplitude features of the targets. Future analysis including a signal-to-noise ratio examination would aid in clarifying this issue.

The interaction of condition and group resulted in longer negative peak latency of the GC than the GW and CI conditions in the Midline group. In this study GW was used as control and by comparison the addition of color stimuli elicited an early positive color processing component in the P/O electrode group. The color processing component precedes the P300 thus, increasing the P300 latency and subsequent negative peak over the Midline. The CI condition negative peak latency over the Midline was the same as GW condition. The lack of difference could be a result of habituation of the color processing neurons. Additionally, the habituation in the CI condition resulted in longer negative peak latency in the P/O electrode group than GW and GC.

4.1 Personal Preference Survey

Participants preferred the GC condition compared to traditional (i.e., GW condition) flash paradigms. Three participants stated that the GW paradigm was “hard to use”, and the GC condition was “easy to use”, two other participants said GC was “fun”. This suggests that the GC condition was user friendly and participants considered GC a better BCI experience than the GW condition. Regrettably, this data was collected prior to Kübler et al. (2014) publication which presents measures of BCI usability. The self-report measures used for this study were developed to simply capture which paradigm was preferred by the participant. The lack of a developed measure to allow inter-study comparison and to measure workload is a limitation to the current study.

4.2 Jumpwise Channel Optimization

The Jumpwise channel set revealed a higher tITR than the standard channel set. This difference was not present in the measures of accuracy and ITR. The different type of stimulus (i.e., each condition was a different stimulus) activated different areas of the brain and optimizing the channels to capture this activity resulted in better performance. The grey to color stimuli elicited new information from the ERPs. These results suggest that when a new stimulus is examined in a BCI paradigm, a channel selection analysis should be conducted to properly capture the activation associated with the new stimulus. These results also suggest that an optimized channel set for each individual improves performance. Changes in individual responses, however, vary across the scalp. A channel set built to capture the more salient responses of the participant contributes to better performance.

5 Conclusion

The grey to color condition has shown promise as a paradigm that results in better BCI performance over traditional grey to white flash paradigms. The results of the current study demonstrate the importance of basic psychophysiological research in understanding the underlying processes of implementing BCI paradigms. GC is easy to implement, and can be readily combined with other techniques to improve BCI performance, (e.g. Jumpwise channel selection and predictive spelling [Ryan et al., 2011]). The enhanced ERPs provide more salient features for target classification and reduce potential distraction from non-target flashes. This enhanced response to color stimulus can be further improved upon when channels are optimized. Moreover, these performance results suggest that grey to color stimuli provide more effective BCI paradigms and have higher participant preference.

Acknowledgments

National Institute on Deafness and Other Communication Disorders, National Institute of Health (R21/R33 DC010470-01); National Institute of Biomedical Imaging and Bioengineering/National Institute of Neurological Disorders and Stroke, National Institute of Health (EB00856).

References

Allison T, Begleiter A, McCarthy G, Roessler E, Nobre AC, Spencer DD. Electrophysiological studies of color processing in human visual cortex. *Electroencephalogr Clin Neurophysiol.* 1993; 88(5): 343–355.

- Anllo-Vento L, Luck S, Hillyard SA. Spatio-Temoral dynamics of attention to color: Evidence from human electrophysiology. *Hum Brain Mapp.* 1998; 6:216–238.
- Birbaumer N, Kubler A, Ghanayim N, Hinterberger T, Perelmouter J, Kaiser J, et al. The thought translation device (ttD) for completely paralyzed patients. *IEEE Trans Rehabil Eng.* 2000; 8(2):190–193.
- Colwell KA, Ryan DB, Throckmorton CS, Sellers EW, Collins LM. Channel selection methods for the P300 speller. *J Neurosci Methods.* 2014; 232:6–13.
- Ikegami S, Takano K, Saeki N, Kansaku K. Operation of a P300-based brain-computer interface by individuals with cervical spinal cord injury. *Clin Neurophysiol.* 2011; 122:991–6.
- Ishihara, S. Tests for Color-blindness. Handaya, Tokyo: Hongo Harukicho; 1917.
- Kaufmann T, Schulz SM, Grunzinger C, Kubler A. Flashing characters with famous faces improves ERP-based brain-computer interface performance. *J Neural Eng.* 2011; 8(5)
- Kaufmann T, Schulz SM, Köblitz A, Renner G, Wessig C, Kübler A. Face stimuli effectively prevent brain-computer interface inefficiency in patients with neurodegenerative disease. *Clin Neurophysiol.* 2013; 124(5):893–900.
- Kaper M, Meinicke P, Grossekhoefer U, Lingner T, Ritter H. BCI competition 2003 data set IIb: support vector machines for the P300 speller paradigm. *IEEE Trans Biomed Eng.* 2004; 51(6):107–6.
- Krusienski DJ, Sellers EW, Cabestaing F, Bayouhd S, McFarland DJ, Vaughan TM, et al. A comparison of classification techniques for the P300 Speller. *J Neural Eng.* 2006; 3(4):299–305.
- Krusienski DJ, Sellers EW, McFarland DJ, Vaughan TM, Wolpaw JR. Toward enhanced P300 speller performance. *J Neurosci Methods.* 2008; 167:15–21.
- Kübler A, Holz EM, Riccio A, Zickler C, Kaufmann T, Kleih SC, et al. The User-Centered Design as Novel Perspective for Evaluating the Usability of BCI-Controlled Applications. *PLoS One.* 2014; 9(12):e112392.doi: 10.1371/journal.pone.0112392
- Lee HW, Hong SB, Seo DW, Tae WS, Hong SC. Mapping of functional organization of human visual cortex: Electrical cortical stimulation. *Neurology.* 2000; 54(4):849–54.
- Polich J, Ellerson. Updating P300: An integrative theory of P3a and P3b. *Clin Neurophysiol.* 2007; 118(10):2128–2148.
- Polich J, Ellerson PC, Cohen J. P300, stimulus intensity, modality and probability. *Int J Psychophysiol.* 1996; 23:55–62.
- Ryan DB, Frye GE, Townsend G, Berry DR, Mesa GS, Sellers EW. Predictive spelling with a P300-based brain-computer interface: increasing the rate of communication. *Int J Hum-Comput Int.* 2011; 27(1):69–84.
- Schalk G, McFarland DJ, Hinterberger T, Birbaumer N, Wolpaw JR. BCI2000: a general-purpose brain-computer interface (BCI) system. *IEEE Trans Biomed Eng.* 2004; 51(6):1034–1043.
- Sutton S, Braren M, Zubin J, John ER. Evoked-potential correlates of stimulus uncertainty. *Science.* 1965; 150:1187–1188.
- Tailby C, Solomon SG, Dhruv NT, Lennie P. Habituation reveals fundamental chromatic mechanisms in striate cortex of macaque. *J Neurosci.* 2008; 28(5):1131–1139.
- Takano K, Komatsu T, Hata N, Nakajima Y, Kansaku K. Visual stimuli for the P300 brain-computer interface: A comparison of white/grey and green/blue flicker matrices. *Clin Neurophysiol.* 2009; 120(8):1562–1566.
- Townsend G, LaPallo BK, Boulay CB, Krusienski DJ, Frye GE, Hauser CK, et al. A novel P300-based brain-computer interface stimulus presentation paradigm: moving beyond rows and columns. *Clin Neurophysiol.* 2010; 121(7):1109–1120.
- Vaughan TM, McFarland DJ, Schalk G, Sarnacki WA, Krusienski DJ, Sellers EW, et al. The Wadsworth BCI research and development program: At home with BCI. *IEEE Trans Neural Syst Rehabil Eng.* 2006; 14(2):229–234.
- Wolpaw J, McFarland D, Vaughan T, Schalk G. The Wadsworth center brain-computer interface (BCI) research and development program. *IEEE Trans Neural Syst Rehabil Eng.* 2003; 11(2):204–207.

Zeki S, Watson JD, Lueck CJ, Friston KJ, Kennard C, Frackowiak RS. A direct demonstration of functional specialization in human visual cortex. *J Neurosci.* 1991; 11:641–9.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Highlights

1. The proposed methods improve Brain-Computer Interface performance.
2. Multiple color stimuli enhance event-related potentials (ERPs) and target classification.
3. Participants preferred the color stimuli over traditional gray-to-white stimuli.

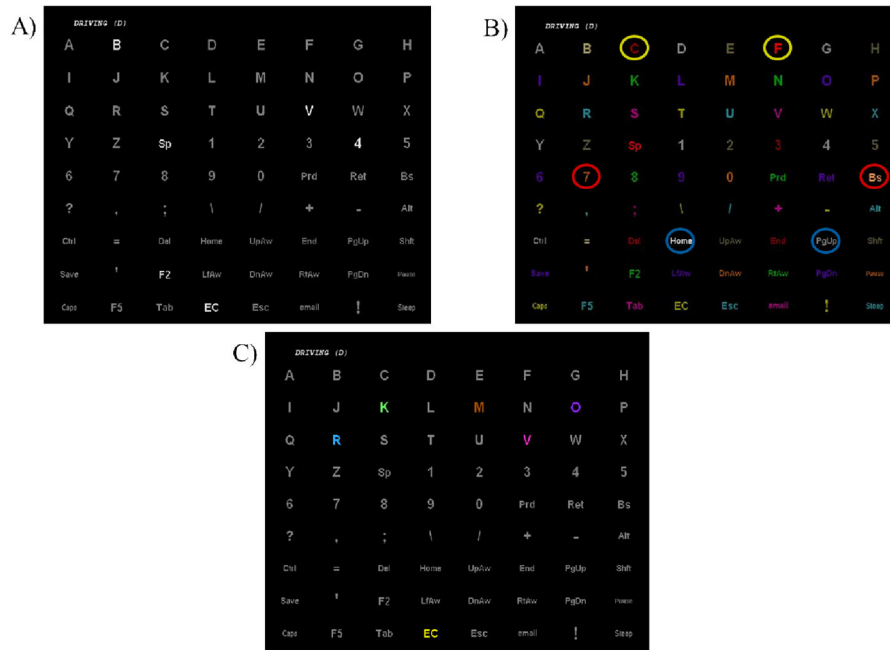


Figure 1. A) The 8×9 matrix with grey characters with white flash stimulus in the grey-to-white (GW) paradigm. B) The 8×9 matrix in color intensification (CI) condition. Each item color intensified for stimulus presentation (e.g. dark red “C” compared to bright red “F”, highlighted in yellow circles for description emphasis). The items are assigned a color (i.e., 1 of 9 colors) so that no adjacent items use the same color. Each colored circle pair displayed a stimulus on and stimulus off of that color. C) The 8×9 matrix in grey to color (GC) condition. Each item flashed the assigned color (1 of 9 colors) so that no item flash is the same color as adjacent flashes.

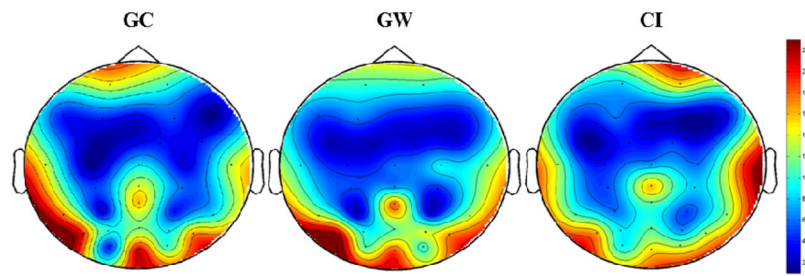


Figure 2.

Frequency of each channel selected by the Jumpwise algorithm (summed from 34 participants) for the GC, GW, and CI conditions in the Jumpwise channel set. The Jumpwise algorithm selects the eight channels that account for the most variance within each data set.

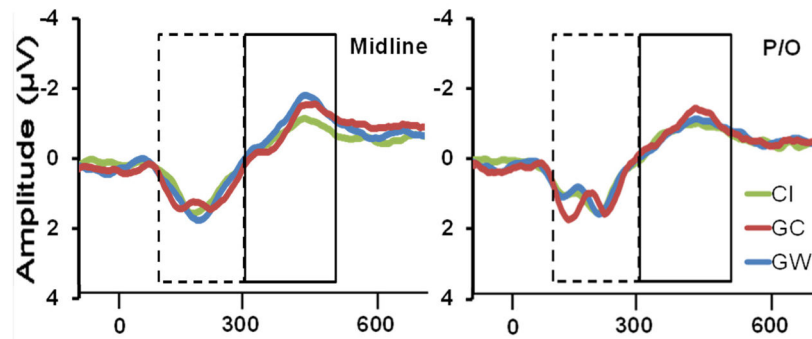


Figure 3.

Fz, Cz, and Pz were averaged together and analyzed as Midline waveform (Left panel). P3, P4, PO7, PO8, and Oz were averaged together and analyzed as parietal/occipital (P/O) waveform (right panel). Positive peak amplitude was analyzed in a time window of 113ms to 310ms (dotted boxes). Negative peak amplitude was analyzed in a time window of 310ms to 515ms (solid boxes). GC had higher amplitude in the Midline group than CI and GW. In the P/O group GC had higher amplitude than CI and GW. GC had a shorter latency than GW in the Midline group. GC also had shorter latency than GW in the P/O group.

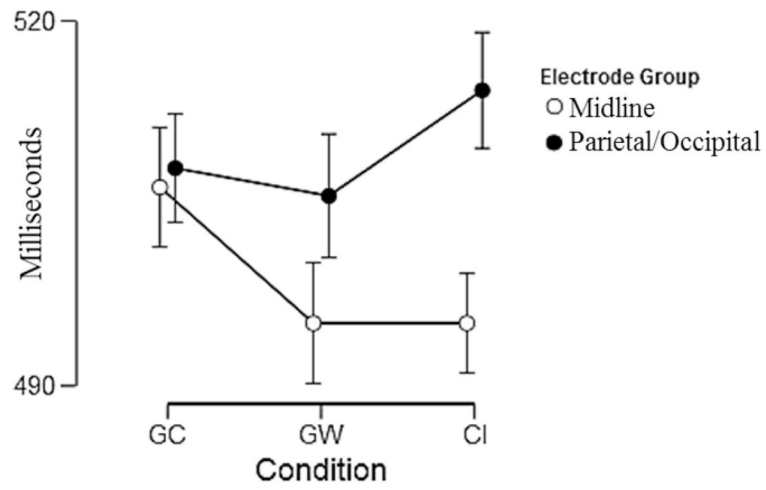


Figure 4. Line graph of the interaction of negative peak latency of the difference electrode groups and condition.

Table 1
Significant statistical results for Accuracy, ITR, and tITR with LSD pairwise comparisons.

Measure	F	p	η^2	Condition	LSD
Accuracy (%)	4.8	0.018	0.124	GC	CI 0.027
				GW	0.020
ITR	6.5	0.003	0.161	GC	CI 0.006
				GW	0.009
tITR	6.2	0.003	0.154	GC	CI 0.010
				GW	0.011

Table 2

Significant statistical results for the waveform analysis with LSD pairwise comparisons.

Measure	F	p	ηp^2	Condition	LSD
Positive Amplitude	4.66	0.013	0.117	GC	CI
				GC	CI
Positive Latency	3.68	0.030	0.095	GC	GW
				GC	GW