Set Size Effects on the P3b in a BCI Speller

Alessandro D'Amico (adamico@ucsd.edu)
Department of Cognitive Science, University of California, San Diego

Virginia R. de Sa (desa@ucsd.edu)
Department of Cognitive Science and Halıcıoğlu Data Science Institute, University of California, San Diego

Abstract

Data were collected from a brain-computer interface speller that utilized the P3b as a control signal. Stimuli consisted of letters and their "segments". Importantly, different letters were made up of different numbers of segments from a 10 segment library. Subjects were instructed to mentally note whenever segments from their letter (targets) were flashed. We found that P3b amplitudes of target segments decreased as the number of segments in a letter (target letter complexity) increased. In contrast, the P3b attenuation was not affected by the total number of letters a segment belonged to (segment frequency). These results may reflect higher task difficulty caused by increased working memory load with increased target letter complexity. Alternatively, it's possible that despite the target rate being fixed at 30% within each block, subjects erroneously believed the target rate increased with target letter complexity. Further work to disentangle these possibilities may enrich our understanding of the P3b.

Keywords: EEG; ERP; P300; P3b; BCI; brain-computer interfaces, speller

Introduction

The P3b has been used to create brain-computer interface (BCI) spellers, devices which allow people to communicate using their brains. The first speller, an electroencephalogram (EEG) system created by Farwell and Donchin (1988) known as the matrix or grid speller, exploited the large body of research surrounding the P300 component of the human event-related brain potential (ERP). By illuminating individual rows and columns containing letters and digits and by instructing subjects to keep a mental count every time their desired letter – the letter they're trying to spell – was illuminated, Farwell and Donchin gave subjects the power to "talk" off the tops of their heads. Various improvements were suggested and researched in regards to P3b spellers. By clustering stimuli more efficiently (Blankertz et al., 2006) or by using more human-salient stimuli (Kaufmann, Schulz, Grünzinger, & Kübler, 2011), researchers were able to create faster and more reliable systems alongside improvements in amplification and computation.

We studied data from a P3b speller with a novel paradigm originally designed by Stivers and de Sa (2017) to probe for letters in parallel without requiring eye movements (which are lacking in people who are completely locked in). Multiple letters are probed centrally by flashing letter segments that are part of multiple letters (Stivers & de Sa, 2017). The P3b is a useful control signal as it is only elicited by task-relevant stimuli (Polich, 2007). It is one of various P300 or P3 sub-components, so named because it originally appeared as a positive-going waveform that peaked approximately 300 ms after stimulus onset. Visual stimuli elicit a P3b that commonly peaks between 250 to 500 ms at parietal electrodes. Although it is not clear what the P3b component reflects, it is commonly associated with context-updating (Donchin, 1981) and working memory (for review, see Polich, 2007). Classically, researchers utilized variations of the oddball paradigm wherein a target or "oddball" stimulus is presented less frequently than the non-target or standard. Subjects are instructed to keep a mental count, make a mental response or make a physical response when they see the task-relevant oddball which will in turn elicit a P3b.

P3b amplitude is partially dependent on the probability of the target stimulus relative to the non-target (Duncan-Johnson & Donchin, 1977; Vogel, Luck, & Shapiro, 1998). P3b amplitudes are also affected by the sequence of stimuli (Donchin, 1981), and time between target stimuli (Gonsalvez & Polich, 2002; for review see Polich, 2007; Luck, 2014).

Furthermore, when multiple targets are present, the relative probability of the target class affects the amplitude of the P3b, not just the rarity of a specific target. Kutas, McCarthy, and Donchin (1977) originally discovered this phenomenon in two experiments where subjects were presented with female names in 20% of trials and male names the remaining 80%. In the condition where only a single female and male name were used, a strong P3b was elicited as expected. Interestingly, when random, one-off male and female names were shown, similarly potent P3bs were elicited, indicating the relative probability of the target class determines P3b amplitude.

P3b amplitudes and latencies are also affected by the number of items a person is required to memorize. In 1969, Dr. Saul Sternbeg devised an item-recognition task to probe human memory-scanning capabilities (Sternberg, 1969). The item-recognition task consists of a set of stimuli divided into a positive and negative set. By having this set contain relatively simple items such as digits, a subject can be given a "positive" set for memorization before being serially presented with items of the positive and negative sets. An experimenter can then ask subjects to respond to only items in the positive or negative set and record reaction times. Sternberg showed that as the size of the positive set increased, so did subjects' reaction times. Not long after Sternberg's work, EEG researchers began utilizing the item-recognition

task to investigate the effects of increased memory load on ERP components such as the P300. Roth et al. (1975) appear to be the first to have extended Sternberg's work with EEG recording by having subjects complete blocks where the target set size was one, two, three, or four digits sampled from 0 to 9. Roth et al. did not find that P300 latencies were affected by set size, but did find that P300 amplitudes decreased as the positive set size increased. Gomer, Spicuzza, and O'Donnell (1976) conducted a similar experiment with Latin alphabet characters and with set sizes of one, two, four, or six and found, in contrast to Roth et al., that P300 amplitudes to positive sets did not vary significantly as a function of set size, while P300 latencies increased with positive set size. Whether the set size influences amplitude or latency was explored by (Pelosi, Hayward, & Blumhardt, 1995), who suggested that decrease in amplitude as a function of increased target set size may cause (secondary) apparent latency variations. All three of these studies had balanced positive and negative item frequencies, e.g. the target stimulus rate was 50% and each required subjects to memorize a novel list of items shortly before probing. To the best of our knowledge, no subsequent work has explored the impacts of lower target rates, sets which have been previously learned, or target stimuli with different global frequencies throughout an experiment.

We discovered a similar trend of P3b amplitudes decreasing with increasing target set size in the segment-based P3b BCI speller. The intent of the speller was to optimize spelling performance by probing for multiple letters in parallel centrally as to be usable by patients unable to voluntarily control ocular muscles. The speller is able to probe for letters in parallel by showing segments that are part of multiple letters. While the speller operated successfully, the performance was not as high as expected. In the process of analyzing the speller's performance, we became interested in exploring the effects of global segment frequency and letter complexity on the P3b. We discovered that as the number of segments in the target letter, i.e. complexity, increased, P3b amplitudes decreased. In contrast, segment frequency —the number of letters a segment is in —did not significantly affect P3b amplitudes. This paper explores these findings and discusses why these differences may have occurred and how they can be utilized or avoided in future studies.

Methods

All data were collected as part of a larger brain-computer interface study aimed at designing a system usable by people suffering from completely locked-in syndrome. Supplementary data, figures, and code can be found at our GitHub repository (`https://github.com/desa-lab/cogsci22-damico-desa`).

Participants

7 subjects (5 naïve) were recruited to participate in the experiment. While the EEG cap was being prepared, participants were given a familiarization task where they had to identify whether a stimulus was a target or non-target (explained below). Prior to data collection it was decided that any participant with accuracy lower than 90% on this task would be excluded. One subject was excluded following this criterion by having an accuracy of 81% (mean accuracy was 98% with a standard deviation of 1.5%), resulting in a final pool of 6 subjects.

Experimental Design

Participants were enrolled in a visual speller brain-computer interface study in which they would see letters and components (segments) of letters serially in one central location (Stivers & de Sa, 2017; D'Amico, Mousavi, & de Sa, 2021). The goal of the speller was to spell words one letter at a time. The paradigm was designed to probe for multiple letters simultaneously by showing segments, which should be theoretically more efficient than showing only letters. Each of the 26 letters in the Latin alphabet in the study was represented by a unique combination of 10 "segments" (with the exception of the letters "O" and "D"; see Figure 2). Segments consisted of a unique color and spatial location in a five-by-seven dot matrix (see Figure 2). One letter or "macroblock" at a time, participants completed three blocks wherein they were given instructions and assigned a target letter (see Figure 1). Targets were defined as any segment belonging to the target letter (in blocks 2 and 3) or the target letter itself (in blocks 1 and 3), but not letters contained within letters (e.g. the letter "I" was not a valid target if the target letter was "T"). Non-targets consisted of all other stimuli allowed in the block. Participants were instructed to keep a mental note of when they saw a target.

Certain letters such as "I", "V", and "X" contained a single segment, while the letter "B" contained six. In the case of single-segment letters, 100% of the target stimuli in block 2 (segments only; Figure 1) will be the single segment. For all letters with two or more segments, the random selection procedure is unconstrained, meaning a letter could have five segments, yet only one segment is selected for each of the three target trials in block 2. Consequently, any letter containing four or more segments will never have each segment selected as a target in a single block since there are fewer allowed targets than possible candidate targets. Furthermore, not all segments are present in an equal number of letters, thus the global frequency of each segment is inconsistent . For example, the blue, vertical segment 3 is present in 18 letters, while the purple, diagonal segment 0 is present in only two. While analyzing the data off-line, we noticed larger P3b amplitudes elicited by the dark-green, vertical segment 9 present in "I", "T", and "Y", and thus decided to do further analyses into the effects of target set size and global frequency on the P3b (see Figure 2).

Stimuli were presented on a Dell P992 with a resolution of 1280×1024 pixels and a refresh rate of 75 Hz using PsychoPy3 (improved version of PsychoPy2; Peirce et al., 2019) PsychoPy3 was selected for its frame-perfect control of stimulus presentation. In combination with a photoresistor, it was

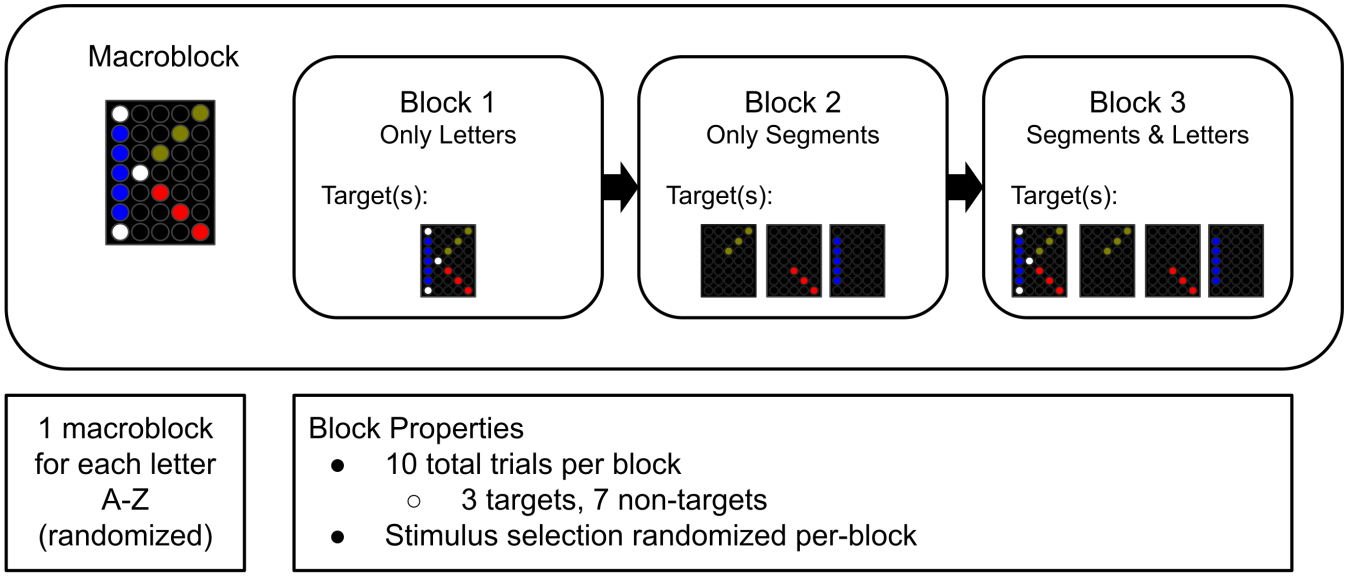

Figure 1: Overview of task design. Each participant completed 26 unique macroblocks, one for each letter in the alphabet. The order of blocks 1, 2, and 3 is fixed across all macroblocks. In each block, 10 trials are shown in total, of which 3 are always targets and 7 non-targets. In total, each participant viewed 780 stimuli; 234 targets and 546 non-targets. All target and non-target stimuli were randomly selected during each block.

possible to precisely synchronize the visual stimuli with EEG recordings. Target and non-target stimuli were presented for 30 frames (approximately 400 ms) followed by an interstimulus interval of 14 frames (approximately 187 ms). While the EEG cap was being prepared, participants completed an online task designed to familiarize them with the letters and segments. During this task, participants were presented with a letter stimulus for approximately 500 ms, after which they were presented with a random segment and asked to identify whether or not the segment belonged to the previously shown letter. Target rate was set at 50% during this task and the primary goal was to assess whether or not participants could accurately identify target and non-target segments. Each participant completed 100 trials with each letter shown approximately four times (mean 3.85, median 4, mode 4).

EEG Collection and Processing

EEG data were collected from 64 electrodes placed on an actiCAP (BrainVision, n.d.-a) following the 10-5 a.k.a. extended 10-20 system (Oostenveld & Praamstra, 2001). All analyses were conducted using electrode site Pz. Mastoid electrodes were placed directly on subjects and were not attached to the cap. Impedance values of each electrode were taken below 25 kΩ before the experiment began. Cz was used as the online reference.

EEG and auxiliary data were recorded using BrainProducts' actiCHamp (BrainVision, n.d.-b) digitized at a rate of 50 kHz and were acquired using PyCorder. Within PyCorder, the data were downsampled and recorded at a rate of 500 Hz. Using the BrainVision RDA server, a single labstreaming layer (LSL; Kothe, 2014) stream was created for the combination of EEG and auxiliary data on the local machine. Data were recorded and saved in PyCorder using the BrainVision header format and in LabRecorder using the extensible data format (XDF). The LSL marker stream created by stimuli presentation was synchronized with the EEG and auxiliary LSL stream and recorded in XDF format.

Prior to filtering, data were re-referenced to the average of the mastoids. EEG data were filtered using a non-causal second-order (functionally fourth-order) Butterworth filter with cutoff frequencies of 0.1 Hz and 15 Hz. Filter coefficients were generated as second-order sections using the function `butter()` and were applied using the function `sosfiltfilt()` from the SciPy toolbox (Virtanen et al., 2020). For computing the amplitude of the P3b, we processed each trial of data independently. Data were baseline corrected by subtracting the mean activity 100 ms prior to stimulus onset from all time points in the epoch. The epoch was defined as all activity from 100 ms prior to 1000 ms following stimuli onset. We then performed artifact rejection using simple voltage thresholding with a threshold of 75 μV. We averaged voltage over a pre-defined window spanning 200 to 600 ms after stimulus onset to represent P3b amplitude. A combination of a photoresistor and an LSL marker stream were utilized for epoching wherein the LSL marker provided information about the stimulus shown such as the stimulus identity and whether or not it was a target, while the photoresistor was used for precise synchronization at the individual trial level. Finally, we computed grand average ERPs of target, non-target and difference waves for each block (1, 2, or 3) and stimulus type (character or segment). Within each of these categories, six independent grand averages were computed,

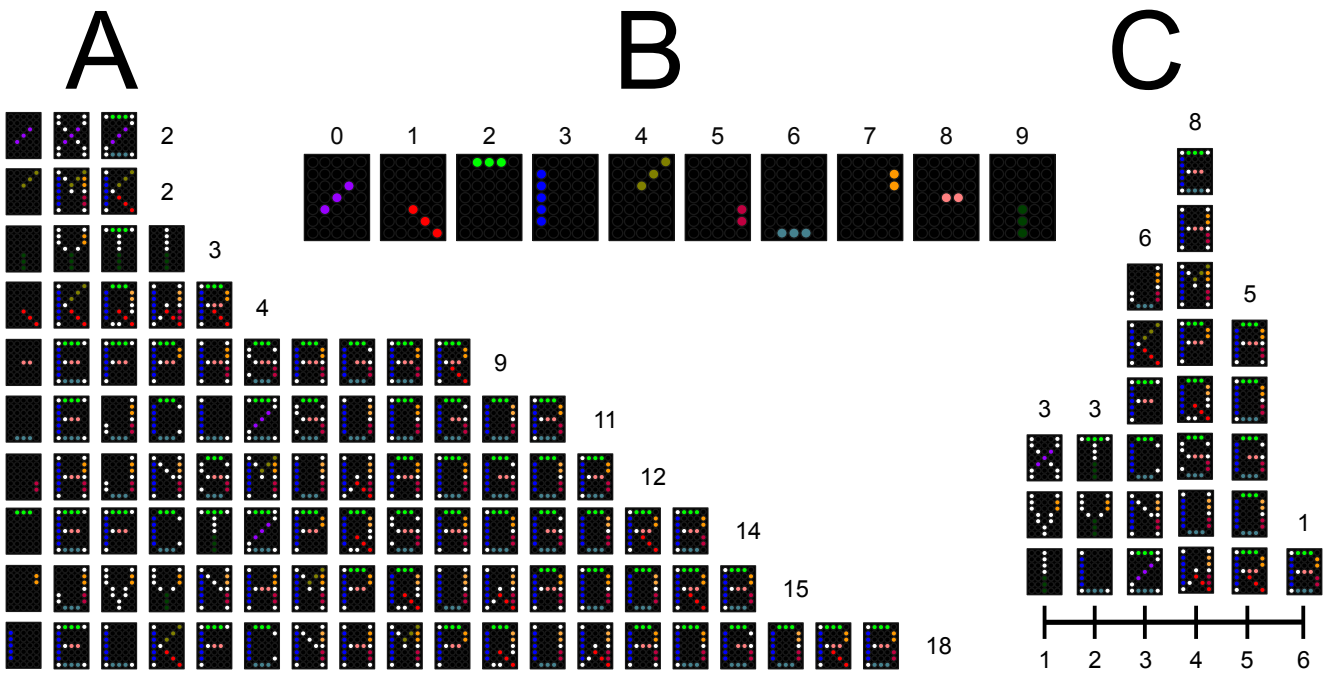

Figure 2: The segments and characters used in the BCI experiment. (**A**) Global segment frequency. Each row shows one segment and all the letters that contain that segment; right-most number indicates the count. (**B**) The ten segments with upper-text representing names. (**C**) Letter complexity. Letters are ordered left-to-right by the number of segments they contain with the top number indicating number of letters in each category.

one for each letter complexity level (i.e. letters containing between one and six segments; see Supplementary Figure 1). We also created a figure that illustrates the effects of target letter complexity on target segment-locked ERPs (see Figure 3).

We computed onset latencies of the P3b by first filtering the data using an identical filter as described above but with cutoff frequencies of 0.1 and 10 Hz. Target and non-target averages were computed within each subject and within each letter complexity (see Supplementary Figure 1, right-most column). In other words, epochs were categorized by the number of segments the target letter had before averaging. These averages were created for segment and character trials differentiated by block number. Difference waves were then computed by subtracting the non-target average from the target average. These difference waves were also computed separately for each of the twelve categories for all subjects. In the event that there were no target or non-target trials for a specific segment, as was often the case when the target letter was "B", no onset latency was recorded. We utilized a version of the fractional peak latency algorithm (see Luck, 2014; Kiesel et al., 2008) to determine P3b onsets. Peaks were searched within the ERP window of interest using `signal.find_peaks()` from SciPy. We then selected the peak closest to the average of the ERP window, which in this case was 400 ms. The amplitude of this peak was recorded and divided by two (50% fractional peak). An iterative algorithm then searched backwards from the selected peak until it found a point that was within 0.005 μV of the fractional peak. This point's latency was determined to be the P3b onset latency. All onset latency analyses were conducted on channel Pz.

Statistics

In order to understand how the global frequency of segments and the complexity of target letters impacted the P3b, we performed a series of linear mixed-effects regressions on target trials using the R packages lme4 (Bates, Mächler, Bolker, & Walker, 2015) and lmerTest (Kuznetsova, Brockhoff, & Christensen, 2017). Our base segment model consisted of three fixed effects: block number (2 or 3), global frequency (number of letters the segment was present in) and letter complexity (total number of segments in the target letter). We had two random effects; the subject and the segment shown, and one dependent variable: the mean amplitude of Pz from 200 ms to 600 ms after segment onset. Another model which included an interaction between segment frequency and letter complexity was created in order to compare to the base model. Two additional models were created by isolating target and non-target stimuli. For character stimuli, we ran similar models, with the only significant difference being the omission of global segment frequency (which is not applicable to characters).

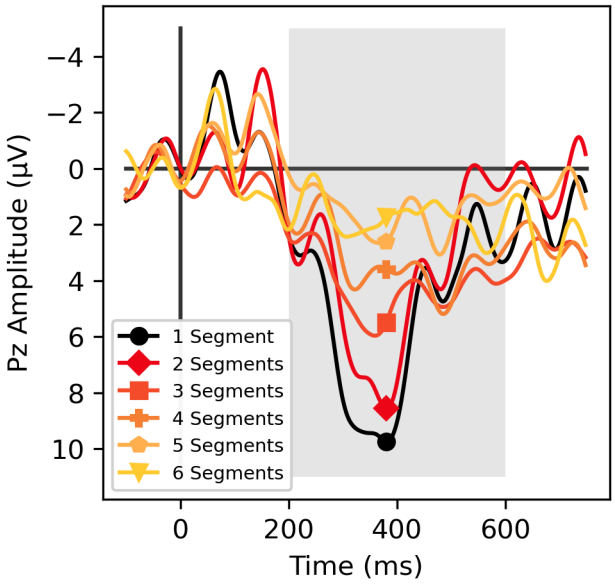

Figure 3: Grand average of target segments. Averages computed using segment-locked target ERPs from blocks 2 and 3 across all subjects. There is a monotonic, near-linear amplitude decrease in the P3b window (gray) with increasing target letter complexity (number of segments in the letter).

For both segment and character target-only trials, we also created "α models," which encoded the number of trials since the previous target stimulus (T) was shown. The number of trials since previous target was computed across all targets within a single block. The only stimuli between targets were non-targets (N). When P3b amplitudes are plotted by distances between targets, there is a non-linear trend. In order to better approximate a linear trend for our linear mixed effects model, we merged all distances of three and greater into a single category, leaving us with three categories: target immediately followed by a target (TT), one non-target spacer between two targets (TNT), and two or more non-target spacers between targets (TNN*T).

Results

P3b amplitudes decreased as the number of segments in the target letter increased (see Figure 3). This trend was visible with target segment stimuli, but not target characters (see differences between first two and last two rows, first column in Supplementary Figure 1). The base segment model revealed a significant effect of letter complexity (t = -2.53, $p < 0.05$), block (t = -2.70, $p < 0.01$), and whether the stimulus was a target or a non-target (t = 8.54, $p < 0.0001$), but not segment frequency. The interaction model returned insignificant differences on everything except whether the stimulus was a target ($p < 0.0001$). We compared both models using the `anova()` function and found that the models were similar; the base model fit the data better with an Akaike information criterion (AIC) of 14,717 while the interaction model had an AIC of 14,719. The Chi-squared test did not show significant results ($\chi^2 = 0.1744$, p = 0.6762).

The target-only model returned a significant effect of letter complexity (t = -2.665, $p < 0.01$), but nothing else, while the non-target-only model only returned a significant effect of block (t = 2.135, $p < 0.05$).

The model including the additional stimulus sequence information —"model α" —showed a significant effect of sequence (t = 4.903, $p < 0.0001$), letter complexity (t = -2.454, $p < 0.05$), and block (t = -2.299, $p < 0.05$), but not segment frequency. The model containing sequence information fits the target data significantly better than the base target-only model ($\chi^2 = 23.419$, $p < 0.0001$).

The base character model only showed a significant effect of whether or not the stimulus was a target (t = 7.822, $p < 0.0001$). Neither of the target-only or non-target-only models returned any significant results. The model accounting for sequence only found a significant effect of sequence (t = 2.880, $p < 0.01$).

Onset latencies were not affected by block, stimulus type, or number of segments in the target letter (all $p > 0.1$).

Discussion

P3b amplitudes to segments appear to significantly attenuate with an increase in letter complexity as shown by the base model and target-only models and visualized in Figure 3. We theoretically wouldn't expect the P3b amplitudes to change for non-target trials, which the models confirm. Since theory dictates that only target trials should elicit a P3b and that stimulus-order (Donchin, 1981) and time-to-target intervals affect P3b amplitudes (Gonsalvez & Polich, 2002), we believe our most accurate model is model α. The significance of sequence is supported by this model, which also shows significant effects of letter complexity and block on segment P3b amplitudes. Judging from the grand-averaged segment ERPs, both block 2 and block 3 exhibit the fairly linear attenuation of segment-locked P3b amplitudes (see Supplementary Figure 1).

Responses to character stimuli seem to differ considerably, with no reliable patterns emerging. We hypothesize that this is due to character stimuli being processed uniquely and more automatically than segment stimuli. Character stimuli are easily discriminable as they resemble standard, everyday stimuli that are processed automatically. Segments, on the other hand, are foreign outside the scope of this experiment, and therefore may employ different or additional cognitive systems. Since we hypothesize that users are processing full character stimuli as single-length objects, we would not expect there to be any significant trend in the block 1 (character only) trials.

We did not find any differences in P3b onset latencies across any examined condition. This could be due to the fact that we were unable to compute single-trial onsets due to unbalanced target and non-target classes. However, the method we used that exploits the difference wave may be a more reliable method for detecting P3b onsets (Luck, 2014; Kiesel et

al., 2008). One limitation of the difference wave technique is that the number of trials per condition are not fixed, which further complicates the fact that target to non-target ratio per letter complexity is also not fixed. Finally, our version of the fractional peak latency algorithm may make assumptions not typically made by similar approaches, thus leading to unstable results. Our implementation is available in the supplementary materials.

Global segment frequency, i.e. the number of letters a segment is in, does not appear to impact the P3b in any systematic way. This suggests that although certain segments are more common and may therefore be easier to identify than others, these differences did not meaningfully impact the P3b.

One parsimonious explanation for these exploratory findings is that P3b amplitudes decrease because it becomes more difficult for subjects to determine if a segment is a target when there are more possible targets, i.e. it is possible that P3b attenuation is caused by increased working memory requirements. With this view, the amplitude differences we found could be similar to those found by (Pelosi et al., 1995), which could also explain the lack of latency variation. However, the lack of onset latency variation does not necessarily indicate that there are no processing differences across letter complexities. Rather, it's possible that the categorization of targets in all cases is simply not slow enough to significantly delay the onset of the P3b. One way to tease this apart would be to create a modified "font" with a larger number of segments. Alternatively, it is possible that the limitation exists in visual working memory. Visual working memory capacity is approximately three or four objects (Luck & Vogel, 2013) and most of the letters in our font contain three, four, and five segments assuming that participants are chunking individual "pixels" of matching color as a single segment, which was the intent of the design and instructions. It is therefore unlikely that most participants are able to accurately store the segment representation of letters for the duration of the macroblock. Future iterations of a visual segmented font should include letters with fewer segments in order to reduce visual working memory requirements.

Another explanation is that the P3b amplitude decreases are caused by changes in subjective probabilities of targets. Although targets are shown in 30% of all blocks, when asked at the end of the experiment, some subjects did not realize they saw three targets every block. It could therefore be the case that when there are more possible target segments in a letter, that the participant expects those segments to be presented more frequently, i.e. they may expect that the letter "B" has a segment target rate of 60% while the letter "I" has a segment target rate of 10%. One way to test this hypothesis is by telling participants a priori what the target rate is, although these instructions themselves may modify ERP components.

In summary, we found that P3b amplitudes to target segments were affected by the number of segments in a target letter but not the global segment frequency. For subsequent iterations of similar spellers, our work has two significant implications. The first is that if the goal is to maintain consistency, all characters should ideally be composed of equal numbers of segments, preferably between one and three. The second is that the these differences in P3b amplitudes may be potentially used as additional information. For example, if a classifier is sufficiently trained to discriminate responses to different letter complexities, it may be possible to achieve better accuracy than simply discriminating between target and non-target segments.

Finally, determining the underlying explanation for the decreased P3b amplitude with increasing target set size may provide more insight into the P3b in general.

Acknowledgments

This project was supported by NSF grants SMA 1041775 and IIS 1528214 and 1817226. Thank you to Dr. Marta Kutas and Dr. Seana Coulson for their general advice, Shaan McGhie and Cameron Jones for their advice regarding statistics, Josh Stivers for originally prototyping the study, and Tianyu Ma, Zhengyu Wu, and Zhijian Wu for assisting in EEG data collection.

References

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. doi: 10.18637/jss.v067.i01

Blankertz, B., Dornhege, G., Krauledat, M., Schröder, M., Williamson, J., Murray-Smith, R., & Müller, K.-R. (2006). The Berlin Brain-Computer Interface presents the novel mental typewriter Hex-o-Spell. *Clinical Neurophysiology*.

BrainVision. (n.d.-a). (The experiment used an actiCap not an actiCap slim or snap, but official documentation seems to have been removed)

BrainVision. (n.d.-b). (The experiment used an actiCHamp not an actiCHamp plus, but official documentation seems to have been removed)

D'Amico, A., Mousavi, M., & de Sa, V. R. (2021). Parallel Spelling using P300 and Feedback Response. In *Proceedings of the 8th international brain-computer interface meeting, organized by the BCI society* (p. 121).

Donchin, E. (1981). Surprise!... surprise? *Psychophysiology*, *18*(5), 493–513.

Duncan-Johnson, C. C., & Donchin, E. (1977). On quantifying surprise: The variation of event-related potentials with subjective probability. *Psychophysiology*, *14*(5), 456–467.

Farwell, L. A., & Donchin, E. (1988). Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and clinical Neurophysiology*, *70*(6), 510–523.

Gomer, F. E., Spicuzza, R. J., & O'Donnell, R. D. (1976). Evoked potential correlates of visual item recognition during memory-scanning tasks. *Physiological Psychology*, *4*(1), 61–65.

Gonsalvez, C. J., & Polich, J. (2002). P300 amplitude is determined by target-to-target interval. *Psychophysiology*, *39*(3), 388–396.

Kaufmann, T., Schulz, S. M., Grünzinger, C., & Kübler, A. (2011). Flashing characters with famous faces improves ERP-based brain–computer interface performance. *Journal of neural engineering*, *8*(5), 056016.

Kiesel, A., Miller, J., Jolicœur, P., & Brisson, B. (2008). Measurement of ERP latency differences: A comparison of single-participant and jackknife-based scoring methods. *Psychophysiology*, *45*(2), 250–274.

Kothe, C. (2014). *Lab streaming layer (LSL).*

Kutas, M., McCarthy, G., & Donchin, E. (1977). Augmenting mental chronometry: the P300 as a measure of stimulus evaluation time. *Science*, *197*(4305), 792–795.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1–26. doi: 10.18637/jss.v082.i13

Luck, S. J. (2014). *An introduction to the event-related potential technique*. MIT press.

Luck, S. J., & Vogel, E. K. (2013). Visual working memory capacity: from psychophysics and neurobiology to individual differences. *Trends in Cognitive Sciences*, *17*(8), 391-400. Retrieved from `https://www.sciencedirect.com/science/article/pii/S1364661313001265` doi: https://doi.org/10.1016/j.tics.2013.06.006

Oostenveld, R., & Praamstra, P. (2001). The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical Neurophysiology*, *112*(4), 713–719. doi: https://doi.org/10.1016/S1388-2457(00)00527-7

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., ... Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior research methods*, *51*(1), 195–203.

Pelosi, L., Hayward, M., & Blumhardt, L. (1995). Is "memory-scanning" time in the Sternberg paradigm reflected in the latency of event-related potentials? *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, *96*(1), 44–55.

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical neurophysiology*, *118*(10), 2128–2148.

Roth, W., Kopell, B., Tinklenberg, J., Darley, C., Sikora, R., & Vesecky, T. (1975). The contingent negative variation during a memory retrieval task. *Electroencephalography and Clinical Neurophysiology*, *38*(2), 171–174.

Sternberg, S. (1969). Memory-scanning: Mental processes revealed by reaction-time experiments. *American scientist*, *57*(4), 421–457.

Stivers, J., & de Sa, V. R. (2017). Spelling in Parallel: towards a Rapid, spatially Independent BCI. In *Proceedings of the 7th Graz Brain-Computer Interface Conference.* doi: 10.3217/978-3-85125-533-1-86

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., ... SciPy 1.0 Contributors (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, *17*, 261–272. doi: 10.1038/s41592-019-0686-2

Vogel, E. K., Luck, S. J., & Shapiro, K. L. (1998). Electrophysiological evidence for a postperceptual locus of suppression during the attentional blink. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(6), 1656.