2024 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) | 979-8-3503-7149-9/24/$31.00 ©2024 IEEE | DOI: 10.1109/EMBC53108.2024.10782158

Assessing the Impact of Environment and Electrode Configuration on P300 Speller Performance and EEG Signal Quality

1st Sandra-Carina Noble
Department of Electronic Engineering
Maynooth University
Maynooth, Ireland
sandracarina.noble.2017@mumail.ie

2nd Tomás Ward
Insight SFI Research Centre for Data Analytics
Dublin City University
Dublin, Ireland
tomas.ward@dcu.ie

3rd John V. Ringwood
Department of Electronic Engineering
Maynooth University
Maynooth, Ireland
john.ringwood@mu.ie

***Abstract*—Recent years have seen extensive use of brain-computer interfaces (BCIs) using electroencephalography (EEG). A critical element in BCI research is electrode selection, which influences performance, experiment duration, resource utilization, and consequently, cost. Electrode choice is partly dictated by the study location, as environmental electrical noise can impact EEG signal quality. This study evaluates the performance of a P300 speller and EEG signal quality using 4-, 6-, 8-, and 16-electrode configurations in two different office environments. Ten healthy adults participated in a single session, using a P300 speller to spell three words with each electrode set. Participants were split between two locations, with five individuals in each. Significant performance disparities were observed between the locations. Notably, within each location, the performance differences among 4-, 6-, and 8-electrode sets were minimal; only the 16-electrode set outperformed the others in both settings. The location associated with poorer performances also exhibited lower P300 amplitudes and higher levels of mains electricity noise.**

***Index Terms*—P300 speller, brain-computer interface, BCI, electroencephalography, EEG**

I. Introduction

The P300 speller, introduced by Farwell and Donchin in the 1980s [1], is a significant application in the field of brain-computer interfaces (BCIs). Originally designed as a communication tool, its use has extended to controlling computers and devices [2], [3].

Electroencephalographic (EEG) electrode selection is a crucial factor in the effectiveness of P300 spellers and BCIs in general. The number and placement of electrodes directly affect BCI performance, as well as the setup time, resource needs, and overall cost of the study.

Several studies have investigated how electrode choice influences P300 speller performance, such as [4]–[6]. These studies suggest that a lower number of electrodes can perform similarly to larger electrode sets, though these were all conducted in single locations. To the best of our knowledge, only one study [7] has compared P300 speller performance across different locations. However, that study utilized the same 8-electrode set.

The effect of different electrode setups on P300 speller performance in various locations remains under-explored. This is vital to investigate as an effective setup in one location may not be as efficient in another.

Our study aims to address this by exploring how different electrode configurations and locations affect P300 speller performance. This information is essential for designing BCI experiments. The study design is outlined in Section II, our findings are presented in Section III, followed by a discussion of these results in Section IV. The study conclusions are summarized in Section V.

This research is supported by the Irish Research Council under project ID GOIPG/2020/692 and Science Foundation Ireland under grant number SFI/12/RC/2289_P2.

II. Study Design

This study was carried out at Maynooth University and received approval from the Maynooth University Ethics Committee (BSRESC-2023-36713).

A total of 10 healthy adults, all over 18 years of age, participated in a single experimental session. During the session, participants used a P300 speller with varying electrode sets to classify target and non-target rows. Details about these electrode sets and the task will be elaborated in subsequent subsections.

The experiments took place in two different office locations. The selection of these locations was intentional, aimed at collecting real-world data outside a laboratory setting and contrasting different environmental conditions. Location 1 was minimally equipped with just the necessary laptop and screen, using natural light from a window. Location 2, in contrast, had multiple electrical devices, relied on bright artificial lighting due to the absence of windows, and had an electric ventilation system. The first five participants completed the experiment in Location 1, while the latter five were in Location 2.

A. EEG Electrode Sets

In this study, we employed four different electrode sets. These sets were informed by a previous data analysis study we

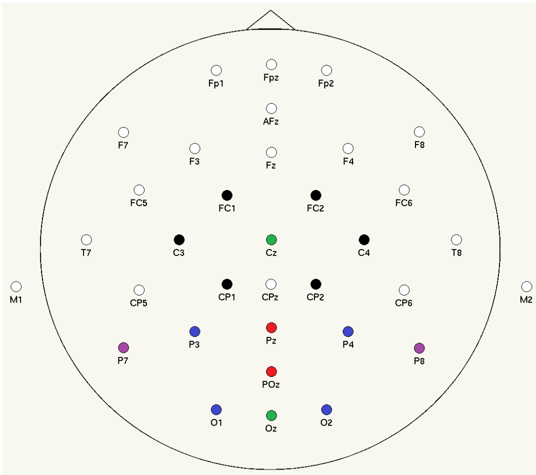

Fig. 1: Electrode configurations for each set. 4-electrode set: red and purple electrodes, 6-electrode set: red and blue electrodes, 8-electrode set: red, blue and green electrodes, 16-electrode set: electrodes of any colour.

conducted [8], which itself utilized data from a comprehensive study involving the P300 speller [9], [10]. The electrode sets comprised 4, 6, 8, and 16 electrodes. Fig. 1 illustrates the specific electrodes included in each set. Based on our previous study findings, these electrodes were identified as the most influential for the P300 speller performance.

B. P300 Speller Task

We used a standard P300 speller, implemented using Openvibe [11]. Participants were presented with a 6x6 grid displaying all letters of the alphabet and numbers 1-9. Each row and column in the grid was highlighted, or ‘flashed’, by enlarging the font size and brightening the color of the relevant characters. Participants counted the number of times the row and column containing their target symbol were flashed. This process elicited a P300 response to the flashes, enabling the computer to identify the targeted letter.

Initially, to calibrate the speller, all participants spelled the words ‘THE’ and ‘QUICK’, with each row and column being flashed 12 times.

To compare the four different electrode sets, participants were then asked to copy-spell the word ‘DANCE’ three times for each electrode configuration. The order of electrode set use varied for each participant to minimize learning effects and fatigue. In the first attempt, with a specific electrode set, the system used 10 flashes per row and column. This was reduced to 5 flashes in the second attempt, and in the third and final attempt, only 3 flashes were used. This strategy was designed to collect data under various conditions. Therefore, each participant spelled ‘DANCE’ a total of 12 times during the experiment.

C. EEG Acquisition and Processing

The EEG signals were acquired using the ANT Neuro eego rt amplifier [12], paired with a 32-channel waveguard cap [13]. In this setup, AFz served as the ground electrode, while CPz was designated as the reference electrode (see Fig. 1). Electrode impedance was maintained below 10 kOhms.

For the P300 speller task, the EEG signals underwent band-pass filtering within the range of 1 to 20 Hz, followed by downsampling by a factor of 4. An xDAWN spatial filter was applied for the 16- and 8-electrode sets, reducing the number of channels to 3 and 2 components, respectively. Subsequently, a Linear Discriminant Analysis (LDA) classifier was employed in all electrode sets to discriminate between target and non-target trials.

For offline EEG processing, the EEG signals were band-pass filtered to 1-20 Hz. Subsequently, all trials were epoched, capturing data from 150ms to 550ms post-stimulus onset. Baseline correction was applied, using the 150ms period preceding each stimulus as the baseline.

D. Data Analysis

In this study, we conducted an analysis of two key metrics. The first metric, spelling accuracy, serves as a measure of performance for the P300 speller. Spelling accuracy represents the percentage of correctly identified letters during a run. For instance, if a participant spells ‘DFBCE’ instead of ‘DANCE’, the spelling accuracy for that run is 60%. The ability to effectively control the speller is crucial for its usability and user acceptance, making performance a central focus of this study. We calculated the mean spelling accuracy across the three words spelled for each participant and each electrode set.

We also assessed performance in terms of the P300 amplitude at POz, a common electrode in all sets. Here, the amplitude is defined as the difference between the positive and negative peaks in the target epochs defined in II-C.

Our analysis focused on comparing spelling accuracy and P300 amplitude across different electrode sets within each subject and between locations within each electrode set. For between-set comparisons, we employed repeated measures ANOVA for normally distributed data and Friedman tests for data not meeting the normality assumption [14]. Where necessary, post-hoc pairwise comparisons were conducted using paired t-tests (for normally distributed data) or Wilcoxon signed rank tests (for non-normally distributed data), both with Bonferroni adjustment [14]. For within-set location comparisons, we used Student’s t-tests (for normally distributed data) or Wilcoxon rank sum tests (for non-normally distributed data), again with Bonferroni adjustment [14].

To investigate potential factors contributing to the differing performances across the two locations, we analyzed the power at 50Hz (the frequency of mains electricity) in the raw EEG signals relative to EEG activity. Initially, we calculated the power spectral density at POz from the raw EEG recordings for all runs, sets, and subjects to determine the power at 50Hz. We then computed the EEG activity’s power as the mean power between 1 and 20Hz in target epochs. The power at 50Hz was then divided by the mean power of EEG activity and subsequently converted to decibels (dB). Given

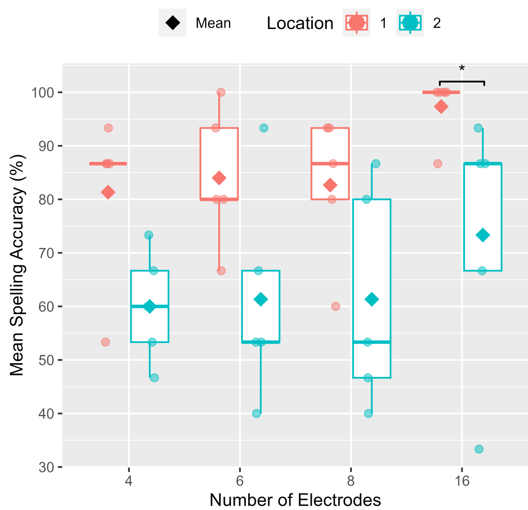

Fig. 2: Mean spelling accuracy across all runs, illustrated by location and number of electrodes. Statistical analyses include within-location comparisons between sets and within-set location comparisons, using (paired) t-tests and Wilcoxon rank tests, * $p < 0.05$.

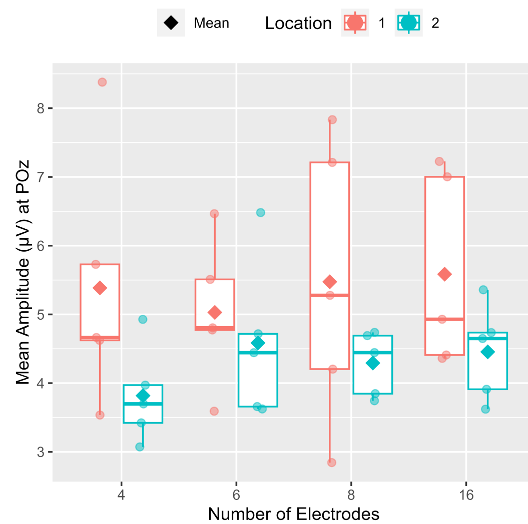

Fig. 3: Mean peak-to-peak P300 amplitude (μV) across all trials and runs, illustrated by location and number of electrodes. Statistical analyses include within-location comparisons between sets and within-set location comparisons, using (paired) t-tests and Wilcoxon rank tests.

that these ratios were normally distributed in both locations, we conducted a Student's t-test [14] to compare the locations.

Given the study small sample size, we calculated effect sizes for all pairwise comparisons using Cohen's D, d, [15] for normally distributed data and the correlation coefficient, r, [16] for non-normally distributed data.

III. Results

A. Spelling Accuracy

Figure 2 presents the mean spelling accuracy across all electrode sets and both locations.

At Location 1, a Friedman test revealed significant differences in spelling accuracy among the electrode sets ($\chi^2_{(3)} = 9.133$, $p = 0.028$). Although pairwise comparisons showed no significant differences between any paired sets, large effect sizes were noted between the 16-electrode set and others (4-electrode: $r = 0.92$, 6-electrode: $r = 0.86$, 8-electrode: $r = 0.91$).

Conversely, at Location 2, the repeated measures ANOVA indicated no significant differences in spelling accuracy between sets ($F_{(3,12)} = 0.767$, $p = 0.534$), corroborated by non-significant pairwise comparisons. However, medium effect sizes were observed between the 16-electrode set and the others (4-electrode: $d = 0.69$, 8-electrode: $d = 0.65$).

A comparison between the two locations within each set highlighted a statistically significant difference only in the 16-electrode set ($Z = 23$, $p = 0.029$). Though not statistically significant, large effect sizes were observed in all sets (4-electrode: $r = 0.60$, 6-electrode: $d = 1.33$, 8-electrode: $d = 1.21$, 16-electrode: $r = 0.73$).

B. P300 Amplitude

Figure 3 shows the mean P300 amplitude for all sets and locations.

A repeated measures ANOVA test of mean P300 amplitude in Location 1 revealed no significant differences between electrode sets ($F_{(3,12)} = 0.53$, $p = 0.668$). While pairwise comparisons did not show any significant difference between pairs of sets, the effect size between the 16-electrode set and the 6-electrode set is moderate ($d = 0.72$).

In Location 2, a repeated measures ANOVA test also showed no significant differences between electrode sets ($F_{(3,12)} = 2.55$, $p = 0.105$). While there are no significant pairwise differences, the effect sizes between the 4-electrode set and all others are large (6-electrode: $d = 1.09$, 8-electrode: $d = 1.10$, 16-electrode: $d = 1.93$), and the effect size between the 8- and 16-electrode set is medium ($d = 0.53$).

Within each set, there are no significant differences between the locations according to t-tests. However, there are medium to large effect sizes for all sets apart from the 6-electrode set (4-electrode: $d = 1.12$, 8-electrode: $d = 0.79$, 16-electrode: $d = 1.01$).

C. Mains Electricity Noise

The mean raw mains electricity noise relative to EEG activity is 0.47dB in Location 1, compared to 3.95dB in Location 2. This difference, while not statistically significant ($t = -1.28$, $p = 0.239$), exhibited a large effect size ($d = 0.81$), suggesting a potential impact of mains electricity noise on performance.

IV. Discussion

Our study highlights two main findings. First, the performance advantage of the 16-electrode set over the 4-, 6-, and 8-electrode sets is evident in both locations, which is consistent with our expectations and previous study results. Interestingly, the differences in performance among the 4-, 6-, and 8-electrode sets are minor, suggesting limited benefit in choosing the 6- or 8-electrode sets over the 4-electrode set. This parallels the findings of our earlier study.

The second key observation is the impact of location on performance. In Location 1, performance was significantly better than in Location 2, with even the 4-electrode set achieving over 80% spelling accuracy. The EEG analysis confirmed lower P300 amplitude at Location 2. Furthermore, our analysis into mains electricity noise revealed higher levels in Location 2 compared to Location 1, which might have contributed to the poorer performance in Location 2. But, the performance differences between locations seem influenced by more than just electrical noise. Participants in Location 2 often reported eye strain and needed more breaks. One participant even requested dimmed lighting due to the bright glare affecting their concentration and causing eye fatigue.

Despite the small sample size, limiting the significance of statistical tests, the large effect sizes observed indicate meaningful differences between electrode sets and locations. However, as the study involved only healthy adults, results may vary with different participant groups, such as those with disabilities or children.

V. Conclusions

This study assessed the performance of a P300 speller across different electrode sets (4, 6, 8, and 16 electrodes) and in 2 different environments. We discovered that, in low-noise environments with dim and natural lighting, all electrode sets yielded relatively high performance. Conversely, in high-noise environments with bright artificial lighting, performance across all sets was much lower. Across both settings, the 16-electrode set outperformed the others, while the 4-, 6-, and 8-electrode sets demonstrated comparable results.

These findings suggest that a minimal 4-electrode set may be adequate in controlled, low-noise, and dimly lit settings. However, for situations requiring near-perfect performance, or in uncontrolled, high-noise environments, opting for a 16-electrode set is advisable.

References

[1] L. A. Farwell and E. Donchin, "Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials," *Electroencephalogr. Clin. Neurophysiol.*, vol. 70, 1988.

[2] S. Halder, A. Pinegger, I. Käthner, *et al.*, "Brain-controlled applications using dynamic P300 speller matrices," *Artif. Intell. Med.*, vol. 63, no. 1, 2015.

[3] P. K. Shukla, R. K. Chaurasiya, and S. Verma, "Performance improvement of P300-based home appliances control classification using convolution neural network," *Biomed. Signal Process. Con.*, vol. 63, no. 102220, 2021.

[4] S. A. Ludwig and J. Kong, "Investigation of different classifiers and channel configurations of a mobile P300-based brain-computer interface," *Med Biol Eng Comput*, vol. 55, 2017.

[5] W. Speier, A. Deshpande, and N. Pouratian, "A method for optimizing EEG electrode number and configuration for signal acquisition in P300 speller systems," *Clin. Neurophysiol.*, vol. 126, no. 6, 2015.

[6] M. T. McCann, D. E. Thompson, Z. H. Syed, and J. E. Huggins, "Electrode subset selection methods for an EEG-based P300 brain-computer interface," *Disabil. Rehabil. Assist. Technol.*, vol. 10, no. 3, 2015.

[7] R. Ortner, R. Prückl, V. Putz, *et al.*, "Accuracy of a P300 Speller for different conditions : A comparison," in *5th Int. BCI Conf.*, Graz, Austria, 2011.

[8] S.-C. Noble, T. Ward, and J. V. Ringwood, "Comparing the Effect of Different Electrode Subsets on P300 Speller Performance," in *2023 IEEE Int. Conf. SMC*, Honolulu, USA, 2023.

[9] S.-C. Noble, E. Woods, T. Ward, and J. V. Ringwood, "Adaptive P300-based Brain-Computer Interface for Attention Training: Protocol for a Randomized Controlled Trial," *JMIR Res. Protoc.*, vol. 12, e46135 2023.

[10] S.-C. Noble, E. Woods, T. Ward, and J. V. Ringwood, "Accelerating P300-based Neurofeedback Training for Attention Enhancement Using Iterative Learning Control: A Randomized Controlled Trial," *J. Neural Eng.*, vol. 21, no. 2, 026006 2024.

[11] Y. Renard, F. Lotte, G. Gibert, *et al.*, "OpenViBE: An Open-Source Software Platform to Design, Test and Use Brain-Computer Interfaces in Real and Virtual Environments," *Presence Teleoperators Virtual Environ.*, vol. 19, no. 1, 2010.

[12] ANT Neuro, *eego (TM) rt*. [Online]. Available: https://www.ant-neuro.com/products/eego_rt.

[13] ANT Neuro, *waveguard (TM) original*. [Online]. Available: https://www.ant-neuro.com/products/waveguard_original.

[14] D. J. Denis, *Univariate, Bivariate, and Multivariate Statistics Using R: Quantitative Tools for Data Analysis and Data Science*. John Wiley & Sons, Inc, 2020.

[15] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*. Routledge, 1988.

[16] M. Tomczak and E. Tomczak, "The need to report effect size estimates revisited. An overview of some recommended measures of effect size," *Trends Sport. Sci.*, vol. 1, no. 21, 2014.