



Long Term Use Effects of a P300-Based Spelling Application

Cristian-Cezar Postelnicu^{1,2}(✉), Florin Gîrbacia^{1,2},
Octavian Machidon^{1,2}, and Gheorghe-Daniel Voinea^{1,2}

¹ Department of Automotive and Transport Engineering,
Transilvania University of Braşov, Braşov, Romania
cristian-cezar.postelnicu@unitbv.ro

² Department of Electronics and Computers,
Transilvania University of Braşov, Braşov, Romania

Abstract. The P300-based Brain Computer Interfaces (BCIs) are widely used for communication purposes and proven to be fast and reliable. The present study aimed to identify the effects of a long term use of a P300-based spelling application for clinical healthy subjects under two display cases represented by a computer monitor and a head-mounted display (HMD). The typical 6×6 row-column spelling application was used by 7 participants under 12 recording sessions for both display methods. The results of the experiment showed that over time there is a statistical significant difference for the average accuracy rates for each display method. The average classification accuracies improved overall by a maximum of 21.43% for monitor display and 25.72% for HMD. However, the results of the experiment demonstrated that no statistical significant interaction existed between the two display methods.

Keywords: P300 speller · Brain computer interface · Electroencephalography

1 Introduction

Brain-computer interfaces (BCIs) allow people to use alternative methods of human-computer interfaces. These types of systems are proven to successfully replace the traditional computer control interfaces represented by keyboard and mouse. They have been initially developed in order to allow patients suffering of severe diseases like locked-in syndrome or amyotrophic lateral sclerosis (ALS) to communicate with others. Nowadays such interfaces can be also used by clinically healthy people as a supplementary communication and control channel.

BCIs are mainly based on the use of electroencephalography (EEG) recordings from the user's scalp, this solution being non-invasive and cost-effective compared with other solutions [1, 2]. Event related potentials (ERPs) extracted from the EEG signals are usually used as key features with various applications for spelling [3–6], drawing [7] or even to control smart houses [8, 9]. The most common ERP is the P300 potential, its usefulness being proven especially for spelling application but not only. The applications based on P300 have proven to work with accuracy levels of even 100% and with high information transfer rates [10].

There are relatively few long-term BCI usage studies, most of them being related to either independent home use of BCI-controlled applications for disabled patients [11–13] or clinical tests involving follow-ups after a certain period of time (usually months) [14, 15]. Also, there are a few studies regarding multiple-day usage of BCI for patients with various disabilities [16].

Apart from the clinical effects on the users, these long-term studies also provide a perspective on the accuracy and amplitude evolution compared to single-time experiments. Hence, while it is possible to obtain high accuracies in single-time experiments (during calibration and training), the results may not be conclusive. Instead, the long-term studies have shown that only after a certain period of time (several sessions, usually after a couple of weeks and months [11]) the accuracy levels tend to reach a certain value and remain stable [17]. It is also true that in the case of users suffering from serious disabilities like ALS, the accuracy levels may decline during a longer period of time (years) due to a decline in the users physical health caused by ALS. However, healthy user groups have shown to maintain a stable accuracy level over the entire long-term period.

Regarding the amplitude, studies have shown it to vary during longer period of times, with a short increase followed by a prolonged decrease in the study presented in [17].

Overall, the reviewed relevant research supports the fact that the long-term use of BCIs leads to high levels of accuracy and successful training for all healthy users and the vast majority of patients also (depending on their level of disablement due to disease). Also, studies have underlined the importance of ease of use, reliability and easy adjustment of the BCI system in the case of long-term usage [11, 12].

An important issue revealed by [16] is related to the need for spatially and frequency-range stable neural signs used for control, since it is not feasible to relearn the control parameters before every BCI session. Hence, it is important in the case of long-term applications to have as control features signals that are robust and stable over long periods of time [16].

Regarding potential side-effects of long-term usage, no significant reactions were found, just minor headaches or discomfort caused by fatigue in few cases [14].

One important BCI application is in the field of embodiment, i.e. using BCIs to control a robot (a humanoid robot or a robotic body part). Work in this direction shows that robotic humanoid whole or part body control can be achieved by using an (EEG) – based BCI [18]. However, one of the challenges encountered in this type of applications is related to reduced robot steering accuracy leading to the impossibility of positioning the robot in a precise position in space [19]. The authors of [19] propose a solution to improve BCI embodiment applications by using a Head-Mounted Display (HMD) to provide assisted steering, navigation and interaction through augmented reality.

Virtual reality (VR) is an emerging domain with multiple applications in almost all possible research fields. Also, it is currently completed by the mass production of commercial head-mounted displays (HMDs) like Oculus Rift [20], HTC VIVE [21] and Samsung Gear VR [22]. Virtual reality has already been used in BCIs and accurate results have been recorded. It was successfully used for games [23], virtual worlds navigation [24] and even for spelling applications [10].

It has been shown that applications fostering the synergy between HMD technology and BCIs have the potential to improve user experience and enable new types of immersive applications [25]. This is because such a symbiosis promises to surpass the common disadvantages of both technologies: HMDs can be optimized using the EEG signals from the BCI (e.g. for achieving 100% accuracy in detecting head rotation, like in [25]) and BCI systems can use the HMD as an adaptive screen to improve user experience during training on the fly.

BCI and HMDs are already used in applications shown to improve the control of 3D objects in VR environments. In [26] the authors proposed a system integrating an Oculus Rift HMD, an eye tracking system and a BCI interface that allows users to point the objects they are interested in by eye gazing and control them by thinking.

Moreover, BCIs and HMDs can be used to determine the quality of a VR application or scenario by monitoring the brain signal activity and comparing the results when the user observes real world and VR objects [27].

The current study aims to identify the effects of long term use of P300-based spelling applications for clinical healthy subjects in two display cases represented by a computer monitor and a HMD. The study is part of a research project which aims to offer feasible BCI-based communication and control channels for clinical healthy subjects for long-term use. Most studies are usually focused on single short use of spelling applications, while very few investigated the effects for a long-term purpose.

2 Materials and Method

2.1 Participants

Seven healthy, subjects (6 male and 1 female, mean age = 27.2 years, range = 22–40 years) with no prior experience in EEG or BCI were recruited as participants, while a single participant previously experimented HMDs. None of the subjects reported a history of psychiatric or neurologic symptoms. All the participants gave their informed consent and none of them was remunerated for the participation in the study.

2.2 Experimental Design and Procedure

The typical spelling matrix of 6×6 rows and columns filled with the 26 English letters, digits from 1 to 9 and the space character was used for the experiment [3]. The chosen flashing method was the row-column paradigm (RC), high accuracy classification rates being achieved in this case [4]. The RC paradigm was set to highlight items for 100 ms (flash time) with a short time between flashes of 60 ms (dark time) while all characters are grey. The number of repetitions was set to 7. This yields an interval of 13.44 s ($6 \text{ rows} \times 160 \text{ ms} \times 7 \text{ flashes} + 6 \text{ columns} \times 160 \text{ ms} \times 7 \text{ flashes}$) to select a character. After each character selection the matrix stopped flashing for 4 s (spelling pause - SP) which is enough time for the user to visualize the last selected item and to focus on the next character to be spelled.

The subjects were instructed to look at the character prompted to be spelled and to silently count each time it was highlighted. EEG signal patterns are known to be

influenced by factors like fatigue, motivation, mental state, level of attention or motivation [1]. Thus, before each experiment the users were requested to perform a calibration session to ensure that the classification was performed based on the users' current EEG recorded data.

Each subject was asked to spell a 4-character text, one letter at a time, without feedback from the P300 spelling application. The calibration data were then processed by linear discriminant analysis (LDA) to determine coefficients for online classification. Next, the users were asked to spell a 10-character. After each character selection the LDA was applied on the EEG data for each column and row. The application provided feedback to the user by indicating the character identified by the classifier (copy-spelling presentation method). The 10-character sequence used for copy-spelling was always different for each participant and balanced to cover as much as possible all the symbols in the 6×6 matrix.

The study was spanned over 12 weeks of recordings aiming to identify the effects of long-term use of a P300 spelling application. Two display cases have been considered for this study, the former represented by a 24 inches computer monitor (60 Hz), while the later by an Oculus Rift HMD system (see Fig. 1 for system architecture). The subjects were asked to participate for the long-term analysis for both display cases during each week, allowing them between 5 to 8 days between two consecutive recording sessions. A recording session consisted of a calibration session and two separate tests, one for the monitor display ("monitor") case and one for the VR system ("Oculus"), performed in a counter balanced manner to ensure there are few correlations between them.

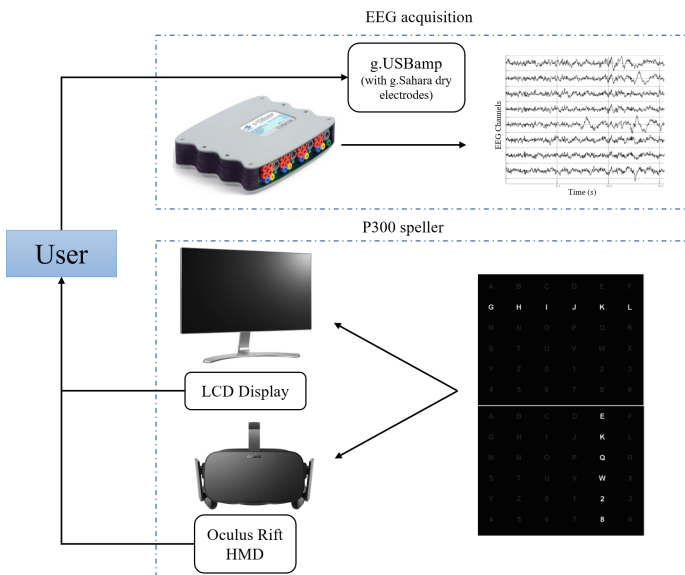


Fig. 1. System architecture

2.3 Data Acquisition, Calibration and Classification

The EEG signals were measured by eight g.Sahara active dry electrodes positioned at Fz, Cz, P3, P4, POz, PO7, PO8 and Oz according to the 10/20 International System. All channels were referenced to the right earlobe and the ground electrode was positioned at FPz. The data were recorded by a g.USBamp (g.tec Medical Engineering GmbH, Austria) biopotentials amplifier and sampled at 256 Hz. Also, the data were filtered using a high pass of 0.5 Hz and a low pass of 30 Hz. A notch filter around 50 Hz was used to reject the power line noise.

For each session the subjects were requested to perform the calibration procedure lasting maximum 3 min during which they were spelling a 4-character word. The data acquired during the calibration session is down-sampled to 64 Hz and the EEG signals are segmented in 800 ms epochs starting at each highlighted stimulus. The extracted trial epochs for all stimuli are entered in the LDA classifier to derive the weighting coefficients, separating them into two classes: target and non-target. During the online copy-spelling task the LDA will associate each visual stimulus with one of the two classes. Finally, the LDA will choose the character with the highest sum of the weighted parameters.

2.4 Effectiveness and Efficiency Metrics

One of the most used and relevant efficiency metric is the online spelling accuracy. It is calculated by dividing the number of correctly spelled characters by all characters that needed to be spelled. For the present study online classification accuracy rates have been recorded for both display methods, “monitor” and “Oculus”, for each recording session and each participant. A statistical analysis was performed to identify whether significant correlations between the display methods and time were present.

Another important measure that is used to grade the efficiency of the system is the Information Transfer Rate (ITR). Firstly, Shannon’s [28] formula was applied to calculate the bitrate:

$$B = \log_2 N + P \log_2 P + (1 - P) \log_2 \left(\frac{1 - P}{N - 1} \right) \quad (1)$$

where: N is the number of possible selections (36 in case of the 6×6 matrix) and P represents the probability that the desired item will actually be selected (classification accuracy). The bitrate is then multiplied by the number of possible decisions per minute (M) to obtain the ITR in bits/min.

Parameter M is calculated from:

$$M = 60 / (N_{rep} \times N_{groups} \times ISI + N_{groups} \times SP) \quad (2)$$

where: N_{rep} represents the number of repetitions for each character (7 repetitions), N_{groups} is the number of groups (6 rows and 6 columns), ISI stands for the time interval between two consecutive flashes (100 ms flash time + 60 ms dark time) and SP represents the spelling pause between two consecutive characters (36 s = 9 pauses \times 4 s).

The total task completion time was calculated as the total number of flashes used for the copy-spelling task, also including the pauses between the selections of two consecutive characters ($170.4\text{ s} = 13.44\text{ s} \times 10\text{ characters} + 36\text{ s}$).

3 Results

Average accuracies obtained for the spelling tasks with the two display modalities over the entire period are grouped per session and are presented in Fig. 2 and Table 1. The average accuracy rate for the entire period was of 86.07% for “monitor” and 83.21% for “Oculus” display methods. The ITR had an average value of 16.54 bits/min (SD = 1.47) for “monitor” and 16.23 bits/min (SD = 1.54) for “Oculus”, with a minimum of 13.19 bits/min and a maximum of 18.2 bits/min for both of them.

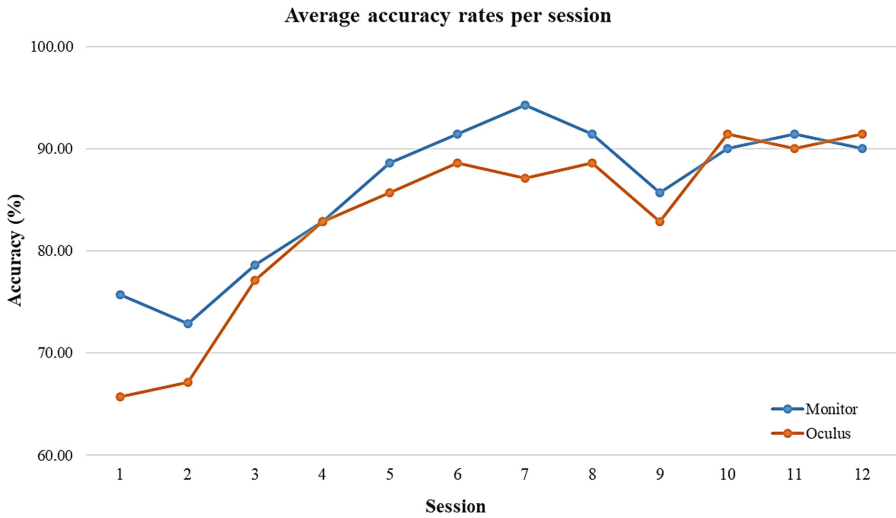


Fig. 2. Average accuracies for both display methods and all recording sessions

The repeated measures ANOVA revealed a main effect over “time” $F(4.821, 57.857) = 7.138$, Greenhouse-Geisser corrected $p < 0.05$, but no main effect of “display” on spelling performance $F(1, 12) = 0.509$, $p = 0.489$ and no significant interaction between “time” and “display” $F(4.821, 57.857) = 0.338$, $p = 0.882$.

For the “monitor” display case statistically significant differences have been found during “time” especially between sessions $S1 \div 2$ compared with $S5 \div 8$, $S10 \div 12$ ($p < 0.05$), $S3$ compared with $S5 \div 8$, $10 \div 11$ ($p < 0.05$), and $S4$ and $S6$ ($p < 0.05$) (see Table 1 for accuracy values). For the “Oculus” case statistically significant differences have been found between sessions $S1$ and $S5 \div 8$, $10 \div 12$ ($p < 0.05$), $S2$ and $S4 \div 8$, $10 \div 12$ ($p < 0.05$), and $S3$ and $S6$, $10 \div 12$ ($p < 0.05$).

Figure 3 displays the ERP waveforms for one participant for all electrodes under “monitor” condition for targets and non-targets to facilitate comparison. By analyzing

Table 1. Average accuracies for the two display methods and all recording sessions

Session	Display method			
	Monitor		Oculus	
	Average accuracy (%)	Standard deviation	Average accuracy (%)	Standard deviation
S1	75.71	15.91	65.71	14.00
S2	72.86	17.50	67.14	8.81
S3	78.57	14.57	77.14	14.85
S4	82.86	13.85	82.86	13.85
S5	88.57	13.55	85.71	10.50
S6	91.43	9.90	88.57	12.45
S7	94.29	4.95	87.14	10.30
S8	91.43	6.39	88.57	9.90
S9	85.71	14.00	82.86	16.66
S10	90.00	9.26	91.43	11.25
S11	91.43	6.39	90.00	5.35
S12	90.00	9.26	91.43	9.90

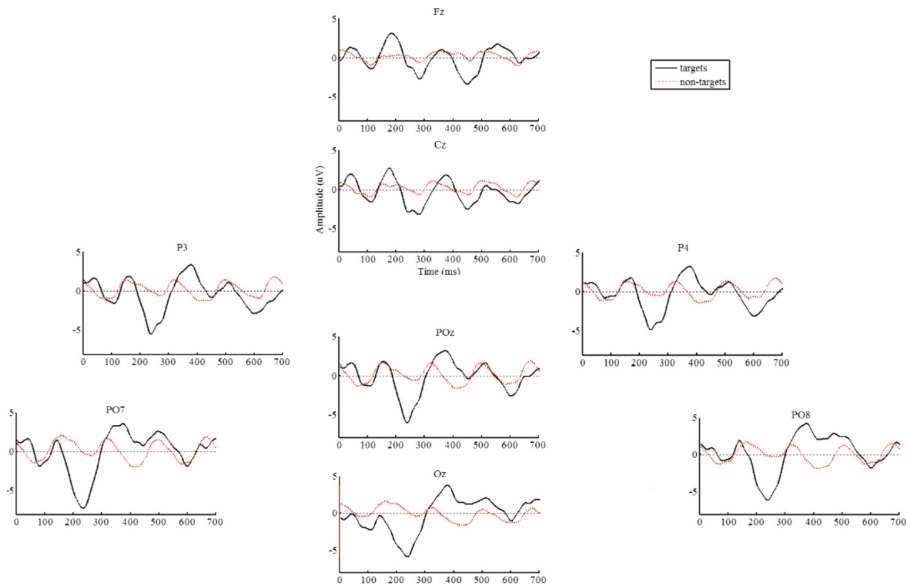


Fig. 3. ERP waveforms average for targets and non-targets for one user under “monitor” case

all the waveforms obtained for all the participants it was found that the P300 potential was most visible in the time frame between 250 and 500 ms after the stimulus onset. Also, the P300 had higher amplitude values for the P3, P4, POz and Oz channels in general, but for most of the participants the POz channel yielded the highest values being further analyzed for amplitude and latency parameters.

The ANOVA yielded no significant results for the P300 latency on the POz channel for “time” $F(11, 55.222) = 1.439, p = 228$, Greenhouse-Geiser corrected, nor for “display” $F(1, 12) = 0.038, p = 0.849$, nor a significant interaction between “time” and “display” $F(4.602, 55.222) = 1.321, p = 0.271$. The P300 latency had an average value of 343.65 ms ($SD = 37.85$) for “monitor” and 357.59 ms ($SD = 36.17$) for “Oculus” display methods.

The P300 amplitude on the POz channel had an average value of 3.53 μV ($SD = 1.33$) for “monitor” and 3.65 μV ($SD = 1.23$) for “Oculus” display methods. On average the “Oculus” display method was higher by 0.12 μV compared with the “monitor” method, but no significant effects have been found during “time” $F(11, 132) = 1.084, p = 0.379$ nor for the “display” method $F(1, 12) = 0.025, p = 0.878$.

4 Discussion

The study intended to discover whether significant effects can be found after a long term use of a P300 speller application under two display methods conditions. It was shown that the average accuracies achieved during online spelling did not statistically differ between the Oculus Rift VR and the 22 inches monitor even if the experiment was spanned over 12 recording sessions.

By analyzing the obtained results over the 12 sessions it can be seen that for both display cases the average accuracy rates increased between sessions achieving a maximum average of 94.29% for “monitor” case during S7 and 91.43% for “Oculus” case during S10 and S12 (see Table 1).

In general, most of the participants were comfortable using both Oculus Rift and EEG cap. One of the participants didn't feel too comfortable during S1 ÷ 3 with Oculus Rift and/or the electrodes, but in time he got used with them. He managed to achieve an average accuracy of 71.66% ($SD = 16.24$), with a minimum of 50% during S1 and a maximum of 90% during S7. Also, it should be stated that for a few sessions a couple of participants took a while to accommodate with the g.Sahara electrodes.

For S9 the average accuracy decreased by 5.72% for “monitor” and 5.71% for “Oculus” cases, due to the poor performance of two participants which declared to have problems to concentrate on the flashing items.

Two of the participants managed to achieve similar or even higher accuracies during the “Oculus” test for at least 8 sessions compared with the “monitor” case, even if only one of them had previous experience with HMDs.

All the participants managed to achieve during the entire experiment a maximum accuracy of at least 90%, while six of them reached 100% for 2 up to 5 sessions for both display methods.

The P300 amplitude and latency yielded on the POz channel the highest discriminative ERP values for all the participants, but no statistical significant effects were identified between the two display methods or sessions.

5 Conclusions

With the current study we were able to identify a few correlations that arise between two display methods for a long term experiment with a P300-based speller application. No statistical significant differences have been found between the “monitor” and “Oculus” display methods over time. Separately, each display method presented significant effects on the accuracy rates mainly by comparing initial ($S1 \div 3$) with middle ($S5 \div 8$) and last ($S10 \div 12$) recording sessions. This demonstrates that on average participants managed to increase the spelling accuracy during time. It was also demonstrated that most of the participants, 6 out of 7, can achieve the maximum spelling accuracy (100%) for both HMD and monitor display methods.

Acknowledgments. This work was supported by a grant of the Ministry of National Education and Scientific Research, RDI Programme for Space Technology and Advanced Research - STAR, project number 566.

References

1. Wolpaw, J.R., Birbaumer, N., McFarland, D.J., Pfurtscheller, G., Vaughan, T.M.: Brain-computer interfaces for communication and control. *Clin. Neurophysiol.* **113**(6), 767–791 (2002)
2. Allison, B.Z., Leeb, R., Brunner, C., Muller-Putz, G.R., Bauernfeind, G., Kelly, J.W., Neuper, C.: Toward smarter BCIs: extending BCIs through hybridization and intelligent control. *J. Neural Eng.* **9**(1), 7 (2012)
3. Donchin, E., Spencer, K.M., Wijesinghe, R.: The mental prosthesis: assessing the speed of a P300-based brain-computer interface. *IEEE Trans. Rehabil. Eng.* **8**(2), 174–179 (2000)
4. Guger, C., Daban, S., Sellers, E., Holzner, C., Krausz, G., Carabalona, R., Gramatica, F., Edlinger, G.: How many people are able to control a P300-based brain-computer interface (BCI)? *Neurosci. Lett.* **462**(1), 94–98 (2009)
5. Postelnicu, C.-C., Talaba, D.: P300-based brain-neuronal computer interaction for spelling applications. *IEEE Trans. Biomed. Eng.* **60**(2), 534–543 (2012)
6. Townsend, G.T., LaPallo, B.K., Boulay, C.B., Krusienski, D.J., Frye, G.E., Hauser, C.K., Schwartz, N.E., Vaughan, T.M., Wolpaw, J.R., Sellers, E.W.: A novel P300-based brain-computer interface stimulus presentation paradigm: moving beyond rows and columns. *Clin. Neurophysiol.* **121**(7), 1109–1120 (2010)
7. Botrel, L., Holz, E.M., Kubler, A.: Brain painting V2: evaluation of P300-based brain computer interface for creative expression by an end-user following the user-centered design. *Brain Comput. Interfaces* **2**, 135–149 (2015)
8. Edlinger, G., Holzner, C., Guger, C.: A hybrid brain-computer interface for smart home control. In: Jacko, J.A. (ed.) *HCI 2011. LNCS*, vol. 6762, pp. 417–426. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-21605-3_46
9. Postelnicu, C.-C., Covaci, A., Panfir, A.N., Talaba, D.: Evaluation of a P300-based interface for smart home control. In: Camarinha-Matos, L.M., Shahamatnia, E., Nunes, G. (eds.) *DoCEIS 2012. IAICT*, vol. 372, pp. 179–186. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-28255-3_20
10. Kathner, I., Kubler, A., Halder, S.: Rapid P300 brain-computer interface communication with a head-mounted display. *Front. Neurosci.* **9**, 207 (2015)

11. Holz, E.M., Botrel, L., Kaufmann, T., Kübler, A.: Long-term independent brain-computer interface home use improves quality of life of a patient in the locked-in state: a case study. *Arch. Phys. Med. Rehabil.* **96**(3), S16–S26 (2015)
12. Sellers, E.W., Vaughan, T.M., Wolpaw, J.R.: A brain-computer interface for long-term independent home use. *Amyotrophic Lateral Scler.* **11**(5), 449–455 (2010)
13. Shahriari, Y., Vaughan, T.M., Corda, D.E., Zeitlin, D., Wolpaw, J.R., Krusienski, D.J.: EEG correlates of performance during long-term use of a P300 BCI by individuals with amyotrophic lateral sclerosis. In: *Proceedings of the Fifth International Brain-Computer Interface Meeting* (2013). <https://doi.org/10.3217/978-3-85125-260-6-29>
14. Onose, G., Grozea, C., Anghelescu, A., Daia, C., Sinescu, C.J., Ciurea, A.V., Popescu, C.: On the feasibility of using motor imagery EEG-based brain-computer interface in chronic tetraplegics for assistive robotic arm control: a clinical test and long-term post-trial follow-up. *Spinal Cord* **50**(8), 599 (2012)
15. Nandrajog, P., Idris, Z., Azlen, W.N., Liyana, A., Abdullah, J.M.: The use of event-related potential (P300) and neuropsychological testing to evaluate cognitive impairment in mild traumatic brain injury patients. *Asian J. Neurosurg.* **12**(3), 447 (2017)
16. Blakely, T., Miller, K.J., Zanos, S.P., Rao, R.P., Ojemann, J.G.: Robust, long-term control of an electrocorticographic brain-computer interface with fixed parameters. *Neurosurg. Focus* **27**(1), E13 (2009)
17. Botrel, L., Holz, E.M., Kübler, A.: Using brain painting at home for 5 Years: stability of the P300 during prolonged BCI usage by two end-users with ALS. In: Schmorrow, D.D., Fidopiastis, C.M. (eds.) *AC 2017. LNCS (LNAI)*, vol. 10285, pp. 282–292. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-58625-0_20
18. Gergondet, P., Kheddar, A., Hintermüller, C., Guger, C., Slater, M.: Multitask humanoid control with a brain-computer interface: user experiment with HRP-2. In: Desai, J., Dudek, G., Khatib, O., Kumar, V. (eds.) *Experimental Robotics. Springer Tracts in Advanced Robotics*, vol 88. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-319-00065-7_16
19. Petit, D., Gergondet, P., Cherubini, A., Meilland, M., Comport, A.I., Kheddar, A.: Navigation assistance for a BCI-controlled humanoid robot. In: *2014 IEEE 4th Annual International Conference on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER)*, pp. 246–251. IEEE (2014)
20. Oculus Rift. www.oculus.com/rift/. Accessed 09 Jan 2018
21. HTC VIVE. www.vive.com. Accessed 09 Jan 2018
22. Samsung Gear VR. www.samsung.com/global/galaxy/gear-vr/. Accessed 09 Jan 2018
23. Finke, A., Lenhardt, A., Ritter, H.: The MindGame: a P300-based brain-computer interface game. *Neural Netw.* **22**(9), 1329–1333 (2009)
24. Groenegrass, C., Holzner, C., Guger, C., Slater, M.: Effects of P300-based BCI use on reported presence in a virtual environment. *Presence* **19**(1), 1–11 (2010)
25. Brouwer, A.M., van der Waa, J.S., Hogervorst, M.A., Cacace, A., Stokking, H.: A feasible BCI in real life: using predicted head rotation to improve HMD imaging. In: *Proceedings of the 2017 ACM Workshop on an Application-Oriented Approach to BCI Out of the laboratory*, pp. 35–38. ACM (2017)
26. Chun, J., Bae, B., Jo, S.: BCI based hybrid interface for 3D object control in virtual reality. In: *2016 4th International Winter Conference on Brain-Computer Interface (BCI)*, pp. 1–4. IEEE (2016)
27. Chun, J., Kaongoen, N., Jo, S.: EEG signal analysis for measuring the quality of virtual reality. In: *15th International Conference on Control, Automation and Systems (ICCAS)*, pp. 1801–1804. IEEE (2015)
28. Shannon, C.E., Weaver, W.: *The Mathematical Theory of Communication*. University of Illinois Press, Champagne, IL (1964)