Proceedings of the
8th Graz Brain-Computer Interface Conference 2019 DOI: 10.3217/978-3-85125-682-6-46

INTRODUCING A MOTIVATING TRAINING STUDY DESIGN TO COMPARE AUDITORY AND TACTILE STREAMING-BASED P300 BCIS

P. Ziebell[1], J. Stümpfig[1], S.C. Kleih[1], M.E. Latoschik[2], A. Kübler[1], S. Halder[3]

[1] Institute of Psychology, University of Würzburg, Würzburg, Germany
[2] Institute of Computer Science, University of Würzburg, Würzburg, Germany
[3] School of Computer Science and Electronic Engineering (CSEE), University of Essex, Colchester, United Kingdom

E-mail: philipp.ziebell@uni-wuerzburg.de

ABSTRACT: Since neurodegenerative diseases or brain injuries causing locked-in-syndrome (LIS) might lead to loss of vision, different approaches of vision-independent BCIs were developed, for example P300 BCIs using auditory or tactile stimuli. However, BCI-inefficiency has been reported in these approaches as well, with high workload proposed as an underlying problem.

In this contribution, a motivating training study design is introduced to compare auditory with tactile P300 BCIs using a streaming-based approach of stimulus presentation that was proposed as a relatively low-workload alternative to classic approaches of sequential stimulus presentation.

First performance results of N = 6 healthy participants were examined case-wise – all participants were able to use at least one BCI version successfully. The preliminary results indicate high motivation and show that there is no superior modality per se but individual preferences in stimulus modalities. This should be considered for future research with healthy as well as LIS users.

INTRODUCTION

The P300 BCI enables users to control various applications, for example text spelling devices, via the non-muscular pathway of electroencephalography (EEG), primarily relying on event-related potentials (ERP), in particular the P300 component [1].

To elicit the P300, an external stimulation is necessary – in P300 BCIs, visual stimulation has until recently been the major focus of research. However, since the eye-gaze dependence of these visual approaches has been pointed out – which is especially problematic for a main target population of potential BCI users, namely severely paralyzed patients suffering from locked-in-syndrome (LIS), a condition often accompanied by a drastic loss of gaze control – eye-gaze independent BCIs exploring alternative P300 stimulation methods have become a new focus of research [2].

Promising results with LIS patients have been reported in a training study using a multi-class auditory P300 BCI; however, cases of BCI-inefficiency occurred as well, with high workload of the auditory multi-class stimulation approach being identified as a problematic factor [3]. A multi-class tactile P300 BCI also showed encouraging results in a training study with elderly healthy participants aged between 50 and 73, but BCI-inefficiency occurred in this approach, too [4]. Eventually, it also yet has to be tested with LIS users and might be problematic with regard to a potentially high workload for this user group in a similar way as the auditory sequential multi-class approach.

To decrease workload problems, an alternative approach, the streaming-based auditory P300 BCI, has been developed. This approach alters the common sequential multi-class presentation of stimuli in one stimulus stream by arranging them in a two-class stimulus presentation with two stimulus streams. First positive results with healthy and LIS users could be shown as well, indicating its potential as an intuitively usable communication device [5].

The current study aimed to replicate the positive results of this intuitive two-class auditory streaming-based P300 BCI with alternative stimuli and furthermore to create a two-class streaming-based tactile P300 BCI in a detailed training case-study approach. Since potential pitfalls of BCI training studies have been pointed out [6], attention was focused on developing a user-friendly study design. Therefore, outcome measures were selected based on the user-centered design approach to evaluate exhaustively the usability of BCI-controlled applications [7]. To ensure high motivation during the whole study, the guidelines by [8] were followed and ideas from BCI gaming literature were taken into account (e.g. [9]), resulting in a “Star Wars”-themed study design.

We hypothesized that the streaming-based P300 BCI version with auditory stimulation as well as the version with tactile stimulation are both intuitively usable and that user performance would further increase via training. Subsequently, preferences for one of the BCI versions and a potential transfer of learning were explored in an additional transfer session after three training sessions, where users switched from the auditory version to the tactile version or vice versa.

 DOI: 10.3217/978-3-85125-682-6-46

MATERIALS AND METHODS

Participants: Herein reported are data from *N* = 6 healthy participants between 20 and 41 years (2 female). Participants were recruited via internet advertisement and compensated financially for participation. Exclusion criteria were: no auditory or neurological impairments, no use of psychotropic substances, no left-right disorientation and no previous BCI-use. All participants spoke German at native-speaker level. The study was conducted in accordance with the ethical guidelines of the Declaration of Helsinki [10].

EEG data collection: EEG was recorded using a g.USBamp-amplifier and 16 Ag/AgCl active electrodes (g.Ladybird electrodes on a g.Gamma cap; products of g.tec Medical Engineering GmbH, Austria, http://www.gtec.at/). Electrodes were placed on the following 10-20-system positions (modified international standard [11]): AF7, Fpz, AF8, F3, Fz, F4, FC3, FCz, FC4, C3, Cz, C4, CP3, CPz, CP4 and Pz. The reference electrode was attached to the right earlobe, the ground electrode to position AFz. The signal was band-pass filtered between 0.1 to 30 Hz, notch-filtered at 50 Hz and recorded with a sampling rate of 256 Hz. A Hewlett–Packard ProBook 6460b with Dual-Core-CPU, 4 GB RAM and 64-Bit Windows 7 was used for data collection, BCI2000 was used for P300 stimulus presentation as well as recording and processing of the signal [12]. These EEG data formed the basis for calculating objective measures for successful BCI-use, covering the BCI-usability criteria effectiveness and efficiency.

Design of the streaming-based P300 BCIs: As a first step, the intuitive streaming-based P300 BCI version with auditory stimulation was designed as a slight modification of the one introduced by [5]. The streaming-based tactile P300 BCI was inspired by the auditory version and furthermore incorporated ideas from earlier tactile P300 BCI research [4]. A major focus while designing the two BCI versions was to maximize their comparability. In both versions, an object on a computer screen had to be moved toward a target, via a predetermined pathway of single steps leading left, right, up or down. The participants' task was to choose the correct direction using the BCI. In the current study only left and right selections had to be actively chosen. Figure 1 shows a schematic of the task accompanied by a more detailed description.

For the auditory streaming-based P300 BCI version, one stimulus-channel was presented to the right ear and the other stimulus-channel to the left ear, via stereo-headphones (Sennheiser HD 280 PRO, http://de-de.sennheiser.com/hd-280-pro). On the right ear, the German word "rechts" (meaning "right") served as a target-stimulus and the English word "right" served as a non-target-stimulus, while on the left ear, the German word "links" (meaning "left") served as target-stimulus and the English word "left" served as non-target stimulus. The words "links" and "left" were spoken by a female voice, the words "rechts" and "right" by a male voice, duration of each word was standardized to 500 ms, volume and tone of voice were kept neutral. At the end of the setup, volume was individually adjusted for each participant, so that all stimuli were pleasantly audible, easy to discriminate and none more salient in comparison to the other stimuli.

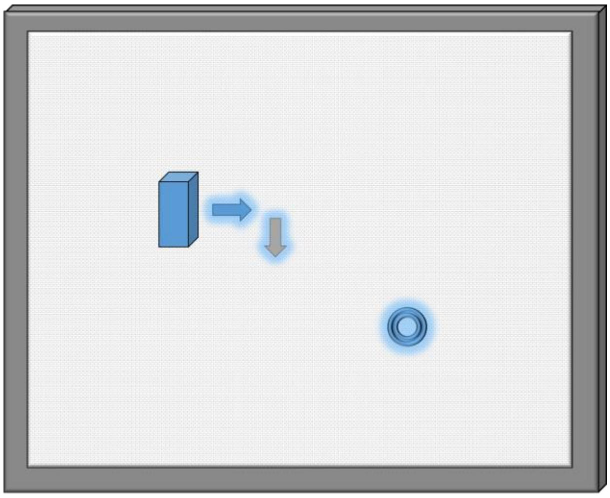

Figure 1: Schematic of the users' task. Since the current study involved only healthy participants without severe visual impairments, the direction selections that had to be chosen via BCI were announced visually on a computer screen. An object (cuboid) had to be moved along a predetermined pathway of glowing left/right arrows to a goal (glowing ring). In addition, up/down arrows were integrated, to create more diverse pathways and to potentially expand the task and the P300 BCI by including two more direction selection possibilities. In the current study the up/down arrows were automatically shown and selected with the previous left or right arrow. Only the currently relevant direction selection arrow was shown on the computer screen, to facilitate concentration on the currently relevant direction selection and to not forestall the whole pathway of arrows, which might have induced boredom. After every correct direction selection, users received auditory feedback in form of a positive sound. If users selected the wrong direction, the sound remained absent. Then the hitherto existing arrow vanished, the object was moved to the next position, the next and now currently relevant arrow appeared and the next direction selection began. The glowing goal ring was presented permanently on the screen. For the object, the screen wallpaper, the arrows and the goal, different pictures were inserted to increase task variability.

The tactile streaming-based P300 BCI version used the right forearm as one stimulus-channel and the left forearm as the other stimulus-channel. Therefore, four tactile stimulators which were able to induce tactile sensations via vibration were attached using elastic bands (C-2 Tactors, Engineering Acoustics, Inc., https://www.eaiinfo.com/). One stimulus-channel consisted of two tactile stimulators on the inner side of

 DOI: 10.3217/978-3-85125-682-6-46

the right forearm and the other of two tactile stimulators on the inner side of the left forearm. On each of those two stimulus channels one tactile stimulator was placed near the right/left wrist. These two tactile stimulators were defined as the emitters of the tactile target-stimuli to select a right direction selection (right wrist) or a left direction selection (left wrist). The remaining tactile stimulators were placed near the inner side of the elbow on the right/left forearm and defined as the emitters of the tactile non-target-stimuli. All four tactile stimulators vibrated with a stimulus duration of 125 ms and an inter-stimulus-interval of 250 ms from the end of one stimulus to the beginning of the next stimulus. A paper tissue was placed between the skin and each tactile stimulator to reduce EEG artifacts that could potentially originate from the tactile stimulators. As in the auditory version, participants were eventually asked, if the tactile stimuli were easy to discriminate from each other and if the caused vibrations were perceived as equally strong and pleasant.

In both versions, auditory as well as tactile, a stimulus-selection-sequence consisted of 10 stimuli in total with one target stimulus and four non-target-stimuli per channel. After the start of the first stimulus on one channel, the stimuli of the other channel were started in constant anti-phase to the stimulus-stream of the first channel – in both streaming paradigm versions, the stimuli of the second channel started 375 ms after the corresponding stimuli of the first channel, which lead to the total length of each stimulus-selection-sequence being 3.75 seconds. Within each stimulus-selection-sequence, the target-stimuli and non-target-stimuli were presented in a pseudo-randomized order to enlarge the P300 effect via heightened unpredictability. To give participants a chance to refocus, a break of four seconds was included after the last stimulus of each stimulus-selection-sequence, before the next stimulus-selection-sequence could get started.

Training protocol: Since recent training studies mainly showed increases until the third training session, but mostly plateau effects in the sessions thereafter (e.g. [3, 4]) and the streaming-based P300 BCI approach has been shown to be intuitively quick to learn [5], the current study included three training sessions, t1, t2 and t3, before the additional transfer session, t4, where users switched from the auditory version to the tactile version or vice versa, so all in all four sessions. The sessions were individually assigned on four different days with at least one day and at most six days without BCI-use between each session. Before the first session, the participants were pseudo-randomly assigned to one of two groups: Group AT, using the auditory BCI version at t1, t2 and t3, transferring to the tactile version at t4, and Group TA, using the tactile BCI version at t1, t2, and t3, transferring to the auditory version at t4.

Each training session started with an individual calibration to estimate the optimal parameters for EEG signal classification (based for example on [3]). First, users sat down in front of the computer screen in a comfortable position, then EEG, headphones, and in case of the tactile version, tactile stimulators were set up as described above. To prevent EEG artifacts, participants were instructed to keep their eyes open and to avoid unnecessary movements during the stimulus-selection-sequences. Thereafter the experimenters explained the BCI, recommending two strategies, first, focusing on the relevant stimulus-channel and blending out the irrelevant stimulus-channel, and second, counting the target-stimuli of the relevant stimulus channel as a way to focus attention on them. Then the participants could ask questions. The calibration consisted of 24 direction-selections, each consisting of 10 stimulus-selection-sequences, divided in two runs with 12 direction-selection-sequences and a one-minute break in between. Based on that, the optimal classification weights and the number of optimal stimulus-selection-sequences for the subsequent online-classification-runs were estimated using the heuristic described by [3]. The online-classification-runs had 12 direction selections and per training session, four online-classification-runs had to be completed, resulting in 48 direction selections. For both the calibration-runs as well as the online-classification-runs, the total number of left direction selections was exactly equal to the total number of right direction selections. Sequence-effects were balanced and each direction selection did not occur more than three times in a row, to avoid monotony which might lead to a weaker P300 response.

To embed the calibration-runs into the motivational context of “Star Wars”, they were named “Introduction”, like the introduction to the magical “Star Wars” power called “the Force”, which some “Star Wars” characters experience in their adventure stories. The object that needed to be moved was a “Star Wars” light saber and it was moved before a “Star Wars” background, a “Jedi Knight Temple”. The online-classification-runs were called “Missions”, where different objects out of the “Star Wars” movies had to be moved through different “Star Wars” backgrounds, with each object and background forming a scene from a “Star Wars” movie. To give immediate feedback on successful direction selection, a one-second long feedback sound was given via headphones after each successful selection, in the auditory as well as the tactile version. A positive sound of the popular “Star Wars” droid “R2D2” was chosen (“R2D2a.wav”, http://www.galaxyfaraway.com/gfa/1998/12/star-wars-sounds-archive/), to fit into the “Star Wars” context. If the selection was incorrect, no feedback sound was given. To further support the atmosphere, short pretexts inspired by the introduction text from the “Star Wars” movies were created and presented, to announce the upcoming task (example: “Move the Millennium Falcon through the clouds of Bespin…”). All object and background pictures were collected via internet search engines, each object-background combination was checked for perceptibility and the backgrounds were checked to be free of potential distractions (as recommended by [9]).

Data analysis: To evaluate the effects of our

 DOI: 10.3217/978-3-85125-682-6-46

experimental design consisting of two independent variables, group (AT, TA) and session (t1, t2, t3, t4), the following dependent variables were analyzed: EEG ERPs following targets and non-targets, online-selection-accuracy (P), information transfer rate (ITR) and the number of possible selections per minute (SPM), which was selected based on the users' performance during the offline-calibration.

EEG data were analyzed offline with the MATLAB toolbox EEGLAB and additional MATLAB scripts ([13], https://sccn.ucsd.edu/eeglab/). Artifacts were removed using artifact subspace reconstruction on the continuous data [14]. The data were segmented in epochs ranging from -200 ms to 800 ms surrounding stimulus-presentation, the epoch of -200 ms to 0 ms before stimulus presentation served for baseline correction. Epochs with extreme amplitudes or distributions were rejected. Epochs related to target-stimuli were grouped and averaged; the same procedure was applied to the non-target-stimuli. Since the P300 was expected to be most pronounced at electrode position Cz (e.g. [3, 4]), analysis was focused on Cz.

To determine the weights and the number of necessary stimulus-selection-sequences for online-run signal classification, a stepwise linear discriminant analysis (SWLDA) was applied to the EEG data from the calibration-runs, for a detailed description of the used algorithm see [15]. SWLDA has been used successfully in earlier auditory and tactile paradigms, with the best results obtained if calculated anew for each individual session, allowing the weights and number of stimulus-selection-sequences to change from session to session via training (e.g. [3]).

To explore successful BCI-use on an objective level, P and ITR were calculated. P served as a measure for the BCI-usability criterion effectiveness and was defined as the percent value of correct selections out of all given selections of a session. As a measure of BCI-inefficiency it was analyzed how many participants could exceed the threshold of 70% selection-accuracy introduced by [16]. ITR served as a measure for the BCI-usability criterion efficiency, and was defined as the amount of correctly transferred information during the time interval of one minute, using the following formula (1) as recommended by [17]:

$$B = \log 2\, N + P * \log 2\, P + (1 - P) * \log 2\, (\tfrac{1-P}{N-1}) \qquad (1)$$

With B standing for bits per selection, N standing for number of possible selection-targets and P standing for the estimated probability of a correct classification, based on the online-selection-accuracy. To calculate the ITR, B was multiplied with the SPM, using the following formula (2), with S as the number of stimulus-selection-sequences that was chosen for each participant at each session and taking into account the duration of each stimulus-selection-sequence (3.75 s) as well as the post-stimulus-selection-sequence break (4 s):

$$SPM = \frac{60\ \text{s}}{S * 3.75\ \text{s} + 4\ \text{s}} \qquad (2)$$

RESULTS

EEG ERP data averages for each participant are depicted in Figure 2. Visual inspection suggests that the participants AT-1, AT-2, AT-3 and TA-2 were able to produce a solid P300 ERP pattern, with AT-2 slightly increasing and TA-2 slightly decreasing between t1, t2 and t3, while AT-1 and AT-3 remained stable between t1, t2 and t3. AT-1 is the only participant that increased to t4, while AT-2, AT-3 and TA-2 show a lesser to equally pronounced P300 ERP at t4. TA-1 and TA-3 do not show a solid P300 pattern over t1, t2, t3 and t4.

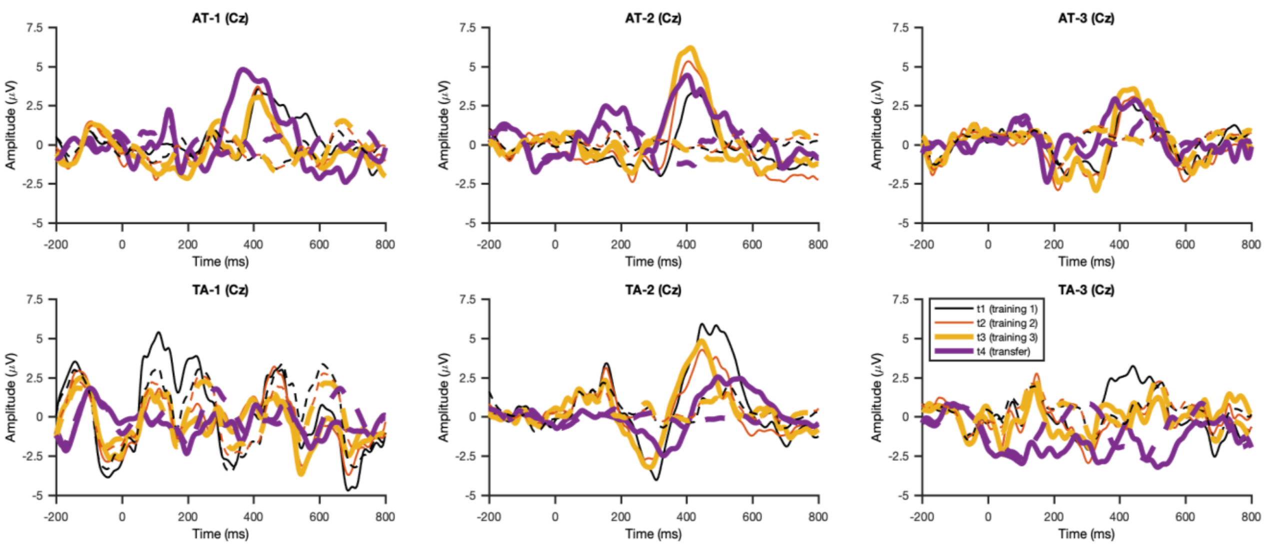

Figure 2: Average EEG ERP patterns for each participant over sessions t1, t2, t3 and t4, indicated by color, see legend. ERPs following target stimuli occurring at 0 ms are indicated by continuous lines, ERPs following non-target stimuli are indicated by dotted lines. Since earlier studies found the most pronounced P300 ERP pattern at electrode position Cz (e.g. [3, 4]), current analysis was focused on Cz.

 DOI: 10.3217/978-3-85125-682-6-46

Table 1 shows the development of P over t1, t2, t3 and t4 for each participant. While AT-1 shows a steadily increasing value of P that is above the BCI-inefficiency-threshold of 70% at t1, t2, t3 and t4, AT-2 shows a value of P above the BCI-inefficiency-threshold at t1, t2 and t3, but below the BCI-inefficiency-threshold of 70% selection-accuracy at t4. AT-3 shows values slightly above 70% at t1 and t3, the lowest value of P of all participants at t2 (27.08%), but the highest value of P of all participants at t4 (97.92%). TA-1 and TA-3 show values of P above the 70%-criterion at session t2 and t4 (TA-1) and t1, t2 and t3 (TA-3). TA-2 shows the highest P of all participants at t1 (89.58%), but values below the 70%-threshold at t2, t3 and t4.

Table 1: Development of P (% correct selections) over training sessions t1, t2 and t3 to transfer session t4 for the participants of group AT and group TA, values below the BCI-inefficiency-threshold of 70% selection-accuracy [16] are highlighted by grey background

Participant	*t1*	*t2*	*t3*	*t4*
AT-1	77.08	79.17	87.50	95.83
AT-2	75.00	93.75	81.25	64.58
AT-3	70.83	27.08	72.92	97.92
TA-1	56.25	87.50	64.58	87.50
TA-2	89.58	68.75	68.75	66.67
TA-3	75.00	81.25	83.33	58.33

ITR values over t1, t2, t3 and t4 for each participant are listed in Table 2. Since P contributes a substantial part to the calculation of the ITR, it is to be expected, that if participants had low values of P at a certain session, the ITR is relatively low as well. This pattern is evident in Table 2, especially if participants were not able to exceed the 70%-criterion.

Table 2: Development of ITR (bits/min) over training sessions t1, t2 and t3 to transfer session t4 for the participants of group AT and group TA, values affected by values of P below the BCI-inefficiency-threshold of 70% selection-accuracy [16] are highlighted by grey background

Participant	*t1*	*t2*	*t3*	*t4*
AT-1	0.51	0.69	1.20	1.98
AT-2	0.43	2.09	1.20	0.16
AT-3	0.29	0.00	0.31	1.93
TA-1	0.02	1.03	0.14	1.20
TA-2	1.17	0.21	0.27	0.16
TA-3	0.50	0.96	1.10	0.05

SPM is reported in Table 3 over t1, t2, t3 and t4 for each participant. Since the ITR is determined by P and the SPM, ITR scores are relatively high, if a relatively high value of P is achieved in combination with a relatively high SPM value, which lead to ITR values that are more than twice as high as the ITR at t1, for example at t2 (AT-2, TA-1, TA-3), t3 (AT-1, AT-2, TA-3) or t4 (AT-1, AT-3, TA-1).

Table 3: Development of SPM (1/min) over training sessions t1, t2 and t3 to transfer session t4 for the participants of group AT and group TA, numbers in brackets indicate the corresponding number of chosen stimulus-selection-sequences based on the offline-calibration-heuristic [3]

Participant	*t1*	*t2*	*t3*	*t4*
AT-1	2.26 (6)	2.64 (5)	2.64 (5)	2.64 (5)
AT-2	2.26 (6)	3.16 (4)	3.93 (3)	2.64 (5)
AT-3	2.26 (6)	1.98 (7)	1.98 (7)	2.26 (6)
TA-1	1.45 (10)	2.26 (6)	2.26 (6)	2.64 (5)
TA-2	2.26 (6)	1.98 (7)	2.64 (5)	1.98 (7)
TA-3	2.64 (5)	3.16 (4)	3.16 (4)	2.26 (6)

DISCUSSION

The descriptive analysis of the six participants revealed a high variety of inter-individual differences of BCI-usability, measured by EEG ERPs at electrode position Cz, P, ITR and SPM. The analysis of P, ITR and SPM revealed that five participants were able to use the BCI version they were first assigned to intuitively with above 70% accuracy in the first session (all except TA-1) and five participants profited from training at the second and/or third session (all except TA-2). Three participants profited from transfer to the BCI version they did not use before in the last session (AT-1, AT-3, TA-1), while three were not able to transfer successful BCI-use in the last session (AT-2, TA-2, TA-3). Even though BCI-inefficiency occurred in some sessions, all participants showed successful BCI-use in at least one session, using either the developed streaming-based P300 BCI version with auditory stimulation or the tactile version or in three cases with both versions, where also successful transfer of learning was possible (AT-1, AT-3, TA-1). Despite some participants partly showing BCI-inefficiency, there were no drop-outs and, defying potentially discouraging results, every participant completed all sessions, leading to success for example in the case of AT-3, who showed the lowest value of P of all participants at the second training session (27.08%), but eventually the highest value of P of all participants at the transfer session (97.92%), potentially facilitated by transfer of learning.

Yet, several limiting aspects have to be noted, indicating the need for further research. After showing the best results in the first training session, participant TA-2 failed to replicate these results in the following three sessions, even though TA-2 produced a relatively solid ERP pattern. TA-2 reported emerging monotony throughout the paradigm post-hoc after the last session, indicating that the used tactile BCI version might not be optimal. Furthermore, the P300 BCI is thought to be driven by P300 elicitation, but high values of P, ITR or SPM did not always correspond with a clear P300 pattern at Cz (e.g. TA-1 at t2 and t4, TA-3 at t2 and t3) or vice versa (e.g. AT-2 at t4, AT-3 at t2, TA-2 at t2, t3 and t4). This should be further examined, especially

 DOI: 10.3217/978-3-85125-682-6-46

since only one electrode position was considered in the current analysis.
Training lead to relatively high ITR increases, but ITR in general was relatively low in comparison to earlier studies (e.g. [3, 4]). Both used BCI versions should be optimized with regard to ITR maximization, for example by lowering the amount of time needed for target-selection (e.g. using less non-targets) and by updating the signal processing pipeline (e.g. using shrinkage LDA). Eventually, both BCI versions have to be tested by LIS users, for example in multiple training sessions or long-term independent home use [1, 3] and therefore include a completely non-visual task design.
For future research, a more exhaustive analysis on a higher number of participants is planned, taking into account data not yet reported in this paper, like questionnaire measures of motivation and of the user-centered design criteria efficiency and satisfaction, further ERP analysis (e.g. different electrode positions) as well as including inferential statistics.

CONCLUSION

The presented results indicate that the streaming-based P300 BCI is a promising approach with auditory stimuli as well as with tactile stimuli and that the study design motivated the participants; all could achieve successful BCI-use in at least one session of the current study. It seems that there is no superior modality, but individual preferences. These findings should be considered for future research with healthy and LIS users.

REFERENCES

[1] Kübler, A., Müller, K.-R., & Guan, C. (2017). The P300 BCI: on its way to end-users? In Müller-Putz, G.R., Steyrl D., Wriessnegger S.C., & Scherer, R. (Eds.), *Proceedings of the 7th Graz Brain-Computer Interface Conference 2017*. Graz: Verlag der Technischen Universität Graz. doi: 10.3217/978-3-85125-533-1-00.

[2] Riccio, A., Mattia, D., Simione, L., Olivetti, M., & Cincotti, F. (2012). Eye-gaze independent EEG-based brain–computer interfaces for communication. *Journal of Neural Engineering, 9(4),* doi: 10.1088/1741-2560/9/4/045001.

[3] Halder, S., Käthner, I., & Kübler, A. (2016). Training leads to increased auditory brain–computer interface performance of end-users with motor impairments. *Clinical Neurophysiology, 127,* 1288-1296.

[4] Herweg, A., Gutzeit, J., Kleih, S., & Kübler, A. (2016). Wheelchair control by elderly participants in a virtual environment with a brain-computer interface (BCI) and tactile stimulation. *Biological Psychology, 121,* 117-124.

[5] Hill, N.J., Ricci, E., Haider, S., McCane, L.M., Heckman, S., …, & Vaughan, T.M. (2014). A Practical, Intuitive Brain-Computer Interface for Communicating “Yes” or “No” by Listening. *Journal of Neural Engineering, 11,* doi: 10.1088/1741-2560/11/3/035003.

[6] Chavarriaga, R., Fried-Oken, M., Kleih, S., Lotte, F., & Scherer, R. (2017). Heading for new shores! Overcoming pitfalls in BCI design. *Brain-Computer Interfaces, 4,* doi: 10.1080/2326263X.2016. 1263916.

[7] Kübler, A., Holz, E.M., Riccio, A., Zickler, C., Kaufmann, T., …, & Mattia, D. (2014). The user-centered design as novel perspective for evaluating the usability of BCI-controlled applications. *PLoS One, 9(12),* doi: 10.1371/journal.pone.0112392.

[8] Lotte, F., Larrue, F., & Mühl, C. (2013). Flaws in current human training protocols for spontaneous Brain-Computer Interfaces: lessons learned from instructional design. *Frontiers in Human Neuroscience, 7,* doi: 10.3389/fnhum.2013.00568.

[9] Marshall, D., Coyle, D., Wilson, S., & Callaghan, M. (2013). Games, gameplay, and BCI: the state of the art. IEEE Transactions on *Computational Intelligence and AI in Games, 5(2),* 82-99.

[10] World Medical Association. (2013). World Medical Association Declaration of Helsinki ethical principles for medical research involving human subjects. *JAMA: Journal of the American Medical Association, 310(20),* 2191-2194.

[11] Sharbrough, F., Chatrian, G. E., Lesser, R.P., Lüders, H., Nuwer, M., & Picton, T.W. (1991). American electroencephalographic society guidelines for standard electrode position nomenclature. *Journal of Clinical Neurophysiology, 8,* 200-202.

[12] Schalk, G., McFarland, D.J., Hinterberger, T., Birbaumer, N., & Wolpaw, J.R. (2004). BCI2000: A general-purpose brain-computer interface (BCI) system. *IEEE Transactions on Biomedical Engineering, 51,* 1034-1043.

[13] Delorme, A. & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods, 134(1),* 9-21.

[14] Kothe, C.A.E. & Jung, T.P. (2016). *U.S. Patent Application No. 14/895,440.* https://patentimages.storage.googleapis.com/c3/6b/dc/7dadfae33c0062/US20160113587A1.pdf

[15] Halder, S., Rea, M., Andreoni, R., Nijboer, F., Hammer, E.M., ..., & Kübler, A. (2010). An auditory oddball brain–computer interface for binary choices. *Clinical Neurophysiology, 121(4),* 516-523.

[16] Kübler, A., Neumann, N., Kaiser, J., Kotchoubey, B., Hinterberger, T., & Birbaumer, N.P. (2001b). Brain-computer communication: self-regulation of slow cortical potentials for verbal communication. *Archives of Physical Medicine and Rehabilitation, 82(11),* 1533-1539.

[17] Wolpaw, J.R., Birbaumer, N., McFarland, D.J., Pfurtscheller, G., & Vaughan, T.M. (2002). Brain-computer interfaces for communication and control. *Clinical Neurophysiology, 113,* 767-791.