

# Usage of drip drops as stimuli in an auditory P300 BCI paradigm

Minqiang Huang<sup>1</sup> · Jing Jin<sup>1</sup> · Yu Zhang<sup>1</sup> · Dewen Hu<sup>2</sup> · Xingyu Wang<sup>1</sup>

Received: 17 January 2017/Revised: 17 July 2017/Accepted: 10 October 2017/Published online: 16 November 2017  
© Springer Science+Business Media B.V. 2017

**Abstract** Recently, many auditory BCIs are using beeps as auditory stimuli, while beeps sound unnatural and unpleasant for some people. It is proved that natural sounds make people feel comfortable, decrease fatigue, and improve the performance of auditory BCI systems. Drip drop is a kind of natural sounds that makes humans feel relaxed and comfortable. In this work, three kinds of drip drops were used as stimuli in an auditory-based BCI system to improve the user-friendliness of the system. This study explored whether drip drops could be used as stimuli in the auditory BCI system. The auditory BCI paradigm with drip-drop stimuli, which was called the drip-drop paradigm (DP), was compared with the auditory paradigm with beep stimuli, also known as the beep paradigm (BP), in items of event-related potential amplitudes, online accuracies and scores on the likability and difficulty to demonstrate the advantages of DP. DP obtained significantly higher online accuracy and information transfer rate than the BP ( $p < 0.05$ , Wilcoxon signed test;  $p < 0.05$ , Wilcoxon signed test). Besides, DP obtained higher scores on the likability with no significant difference on the difficulty ( $p < 0.05$ , Wilcoxon signed test). The results

showed that the drip drops were reliable acoustic materials as stimuli in an auditory BCI system.

**Keywords** P300 · Auditory BCI · Drip drops · Online accuracy · User-friendness

## Introduction

The first P300 Speller system was presented by Farwell and Donchin nearly 30 years ago (Farwell and Donchin 1988). Although visual-based P300 brain-computer interfaces (BCIs) can obtain high classification accuracy and information-transfer rate (ITR) (Jin et al. 2014, 2015; Martinez et al. 2007; Pan et al. 2013; Xu et al. 2016; Zhang et al. 2016; Zhu et al. 2010), visual-based P300 BCIs cannot be used by people with visual disabilities or disorder of consciousness (DOC) while the effects of auditory-based P300 BCIs on these people were validated (Başar et al. 2012; Lulé et al. 2013; Puanhvuan et al. 2017; Yin et al. 2013).

The P300 potential is a large positive-going potential, occurring at ~ 300–500 ms after stimulus onset (Farwell 2012; Farwell et al. 2013; Park et al. 2016), and is firstly discovered by Sutton et al. (1965). The oddball paradigm is a classical paradigm used to evoke the P300 potential (Donchin et al. 2000; Monica et al. 1995). The efficacies of a four-choice auditory P300 BCI between amyotrophic lateral sclerosis (ALS) patients and health subjects were compared in Sellers and Donchin's work firstly and it showed that the auditory P300 BCI could evoke recognizable ERPs on the ALS patients as well as the healthy people. This study showed the potential application of auditory P300 BCI on the ALS groups (Sellers and Donchin 2006). In the proposed auditory oddball

---

✉ Jing Jin  
jinjingat@gmail.com

✉ Xingyu Wang  
xywang@ecust.edu.cn

<sup>1</sup> Key Laboratory of Advanced Control and Optimization for Chemical Processes, Ministry of Education, East China University of Science and Technology, Shanghai, People's Republic of China

<sup>2</sup> College of Mechatronics and Automation, National University of Defense Technology, Changsha, Hunan 410073, People's Republic of China

paradigms of BCI systems, different acoustic signals are utilized as stimuli.

Beeps with different frequencies are widely used in auditory event-related potential (ERP) researches (Güntekin and Başar 2010). Hill et al. (2005) proposed an auditory paradigm that contained both left and right synchronous auditory streams. These auditory streams consisted of beeps at different frequencies: 800 Hz (non-target) and 880 Hz (target) for the left; and 1500 Hz (non-target) and 1650 Hz (target) for the right. This paradigm presented an encouraging result for its high classification accuracy. Halder et al. (2010) further studied the beeps in auditory-based BCIs and focused on three effects of beep stimuli: pitch, loudness, and direction effects. Pitch and direction can improve the performance of auditory-based BCIs better than the loudness. Schreuder et al. (2010) developed an eight-class auditory paradigm with eight pitches from eight directions and analyzed the performances of different inter-stimulus interval (ISI) conditions (1000, 300, and 175 ms). The highest average ITR was 175 ms ISI, which achieved 17.39 bits/min. The speed of auditory BCIs could be comparable to the visual-based BCIs based on covert attention (Schreuder et al. 2011).

Recently, different auditory stimuli have been found to exert various effects on auditory BCI users. A proper auditory stimulus would help to increase the performance of the BCI system and decrease the fatigue of users after long-term use. Several studies reported that natural sounds would be a good choice for auditory-based BCIs (Theunissen and Elie 2014). The special properties of natural sounds make them easy to be perceived, and the auditory systems of humans can also adapt well to these properties (Theunissen and Elie 2014). Natural sounds include a relatively wide variety of sounds, and different natural sounds are selected to develop auditory paradigms. Klobassa et al. (2009) presented a high-throughput auditory BCI system with five different sounds (bell, bass, ring, chord, and buzz). The findings indicated that this auditory paradigm which utilized natural sounds is feasible and reliable with stable classification accuracy for users with vision limitations. The vocalization of numbers was also investigated in the research on auditory-based BCIs. Furdea et al. (2009) used Arabic numerals (labeled 1–10) on each row and column of a  $5 \times 5$  speller, and the participants can select their targets by listening to the vocalization of the labeled number of the desired row or column. High classification accuracies demonstrated the availability of this speller on the majority of the participants (9/13). Kübler et al. (2009) tested this paradigm on four locked-in patients, but the results were disappointing. Two letter streams and two natural speech streams read by humans were used as stimuli in Lopez's study (Lopez-Gordo et al. 2012). Natural speeches obtained high classification. Besides the

human voice, sounds of duck, singing bird, frog, seagull, and dove are used in auditory BCI systems (Baykara et al. 2016; Halder et al. 2016; Simon et al. 2015). Höhne (2012) verified that natural sound stimuli perform better than beep stimuli by comparing the beeps with spoken and sung syllables; both of which obtained high classification accuracies and sufficient positive feedback from the participants. Appropriate natural sounds were shown as not only effective stimuli to evoke distinguishable ERPs but can also improve auditory BCIs with respect to ergonomics and performance.

In the field of acoustics, it had been proved that water sounds were the optimal sounds among natural sounds to mask the noises in the environment and make people relaxed (Jeon et al. 2010). However, there were no studies that focused on the performance of water sound stimuli used in auditory-based BCI system. An auditory BCI evoked by drip drops is proposed and named as the drip-drop paradigm (DP) in this study as the drip drops were one kind of water sounds. This work aimed to verify whether the proposed auditory paradigm can make users feel comfortable and ease their tiredness, as well as not deteriorate the classification accuracy performance of the auditory BCI. A traditional beep paradigm (BP) was compared with DP to verify the performance of the latter. We hypothesized that DP was superior to BP in terms of classification accuracy, decreasing fatigue, and long-term performance.

## Methods

### Participants

Eleven healthy subjects (mean age: 23.4 years old, SD 1.29, range 21–26 years old) participated in this study. Five of the participants had no previous BCI experience. Each participant had no neurological diseases and presented normal audition. Each participant was given information about the experiments without exposing intension. All the participants signed written consent forms before the experiments, and they received 100 RMB each.

### Auditory stimuli

The DP presented three different drip-drop clips which were intercepted from a music work named *Fragile Hope* (WANDER/WONDER, Balam Acab). The three drip-drop clips were selected carefully to ensure that each clip could be distinguished from the others, and no clip sounded more salient than the others. The three drip-drop clips presented clear and different orientations (left/middle/right). The “first drip-drop” stimulus was played in the right

headphone, the “second drip-drop” stimulus was played through both headphones to sound as if it came from the middle, and the “third drip-drop” stimulus was played in the left headphone. Three beeps with different frequencies were used for BP. The frequencies of beeps were 800, 1000, and 1200 Hz. The “1200 Hz” stimulus was played in the right headphone, the “1000 Hz” stimulus was played through both headphones to sound as if it came from the middle, and the “800 Hz” stimulus was played in the left headphone. Details of all the stimuli were shown in Fig. 1.

The stimulus on time for each pattern was 200 ms, and the stimulus onset asynchrony (SOA) time was 550 ms. The same stimulus in each pattern was not played successively to avoid the “double-stimulus” effect and increase the target to target interval (TTI).

### Experimental design

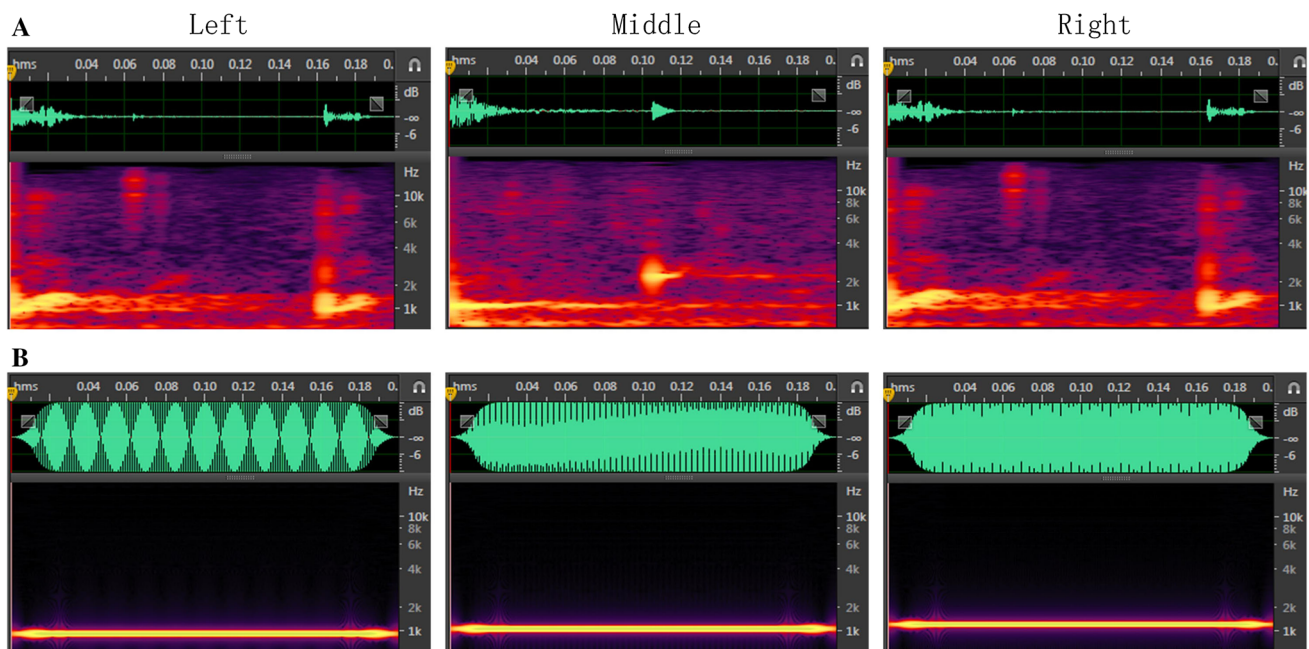
EEG data were recorded with active electrodes (Ag–AgCl) in a 64-channel “g.EEGcap” EEG cap (Guger Technologies, Graz, Austria), of which 15 channels (F3, Fz, F4, T7, C3, Cz, C4, T8, CP3, CPz, CP4, P3, Pz, P4, and Oz) were used. The electrodes were placed in accordance with the international 10–20 system. FPz was set as the ground, and the right earlobe was set as the reference. The EEG was measured with a 16-channel “g.USBamp” amplifier (Guger Technologies, Graz, Austria), band-pass filtered between 0.1 and 100 Hz, notch-filtered at 50 Hz, and

digitized at a rate of 512 Hz. Impedances of all electrodes were kept below 10 K $\Omega$ .

Before the experiments, participants were trained for the auditory tasks. Before each target task, a female voice would tell the subjects what the target sound was. Afterward, the subjects only needed to count the number of times they heard the targets. In the offline experiment, three different sound stimuli composed one trial and sixteen trials composed one run. Each session comprised five target tasks (five runs). Three sessions were established in the offline experiment for each pattern. The subjects were allowed to rest for 2–3 min after each session. The order of the paradigms in the experiments was counterbalanced for each participant. The offline training of each paradigm was separated into three sessions and the recorded data was used to train the classifier model.

In the online experiments, subjects were required to finish 36 target copy selections in each session. The number of trials used in each run was selected by an adaptive strategy (Jin et al. 2011). Each online trial block takes about 10 min. Compared with the previous work Huang et al. (2016), the amount of target selection tasks was increased from 24 to 36, which will help to show the performance of this auditory-based BCI when more tasks were done by users.

The EEG data were filtered with a third-order Butterworth band-pass filter between 0.1 and 30 Hz. The EEG was down sampled by selecting every eighth sample from



**Fig. 1** Oscillograms and spectrograms of DP and BP are presented. In the subplot **a**, the oscillograms and spectrograms of drip drops in DP are presented. The direction of each stimulus is marked on top of

the figures. Subplot **b** presents the oscillograms and spectrograms of beeps in BP. The frequencies of beeps from left to right are 800, 1000, and 1200 Hz, respectively

the EEG. The first 1000 ms of EEG data after each stimulus presentation was used for feature extraction.

Bayesian linear discriminant analysis (BLDA) was used as the classification algorithm for its properties that prevent overfitting in high-dimension data and its better classification accuracies than those of Fisher's linear discriminant analysis (FLDA). The details of BLDA can be found in Hoffmann et al. (2008).

### Online accuracy and speed

The online accuracy of each participant was calculated as

$$Acc_{\text{online}} = \frac{T}{F + T} \quad (1)$$

$F$  is the number of false selections, and  $T$  is the number of true selections in the whole online experiment. The value of  $F + T$  is equal to 36 (the number of total online targets). Online accuracy is an important index which reflects the true performance of the BCI system.

The  $r$ -square value was computed to evaluate the classification discriminant degree. The  $r$ -square value was  $r^2(x)$ , and the  $r(x)$  was calculated as

$$r(x) = \frac{\sqrt{N_1 N_2} \text{mean}\{x_i | l_i = 1\} - \text{mean}\{x_i | l_i = 2\}}{N_1 + N_2 \text{std}\{x_i | l_i = 1, 2\}} \quad (2)$$

where  $N_1$  and  $N_2$  are the number of variables in the target and non-target groups, respectively.  $x_i$  is the  $i$ th variable, and  $l_i$  is the class label of the  $i$ th variable. Variable is the feature vector extracted from the EEG epoch in the targets and non-targets.

ITR is one of the important indexes to demonstrate the property of BCI paradigms. ITR is calculated as the following formula:

$$B = b^* \left( \log_2 N + P \log_2 P + (1 - P) \log_2 \left[ \frac{(1 - P)}{(N - 1)} \right] \right) \quad (3)$$

$$b = \frac{60}{t * AVT}$$

$B$  is the value of ITR.  $N$  indicates the number of possible targets and  $P$  is the possibility that the desired target is selected.  $AVT$  is the number of trials to output a target and  $t$  (1.65 s) is the time of each trial (SOA is 550 ms and each trial contains 3 stimuli).

### Feedback

After completing the last run of each session, each subject was asked two questions about each condition. Each question can be answered on a 1–5 rating scale indicating strong disagreement, moderate disagreement, neutrality, moderate agreement, or strong agreement. All questions were asked in Chinese. The two questions were

**Fig. 2** Grand average ERP waveforms across 11 participants of DP and BP. Thick solid red lines indicate the average target ERP responses of DP, and thick solid blue lines depict the average target responses of BP. Thin red dashed lines indicate the average non-target ERP responses of DP, and thin blue dashed lines represent the average non-target ERP responses of BP. (Color figure online)

1. Do you prefer this pattern?
2. Is this pattern difficult?

### ERP analyses

To survey the performance of DP and BP, the grand averaged amplitudes and latencies of ERPs was presented. The grand averaged amplitudes of targets/non-targets for each electrode were obtained by averaging the 1 s time window EEG of all target stimuli/non-target stimuli and the start point of the time windows is set according to the start time point of each stimulus. The peaks of ERPs were the peak points from 75 to 180 ms for the N1, from 150 to 250 ms for the P2, from 250 to 400 ms for the P300 and from 380 to 600 ms for the N400-like component. The time point of each peak was defined as the peak latency.

### Statistic analyses

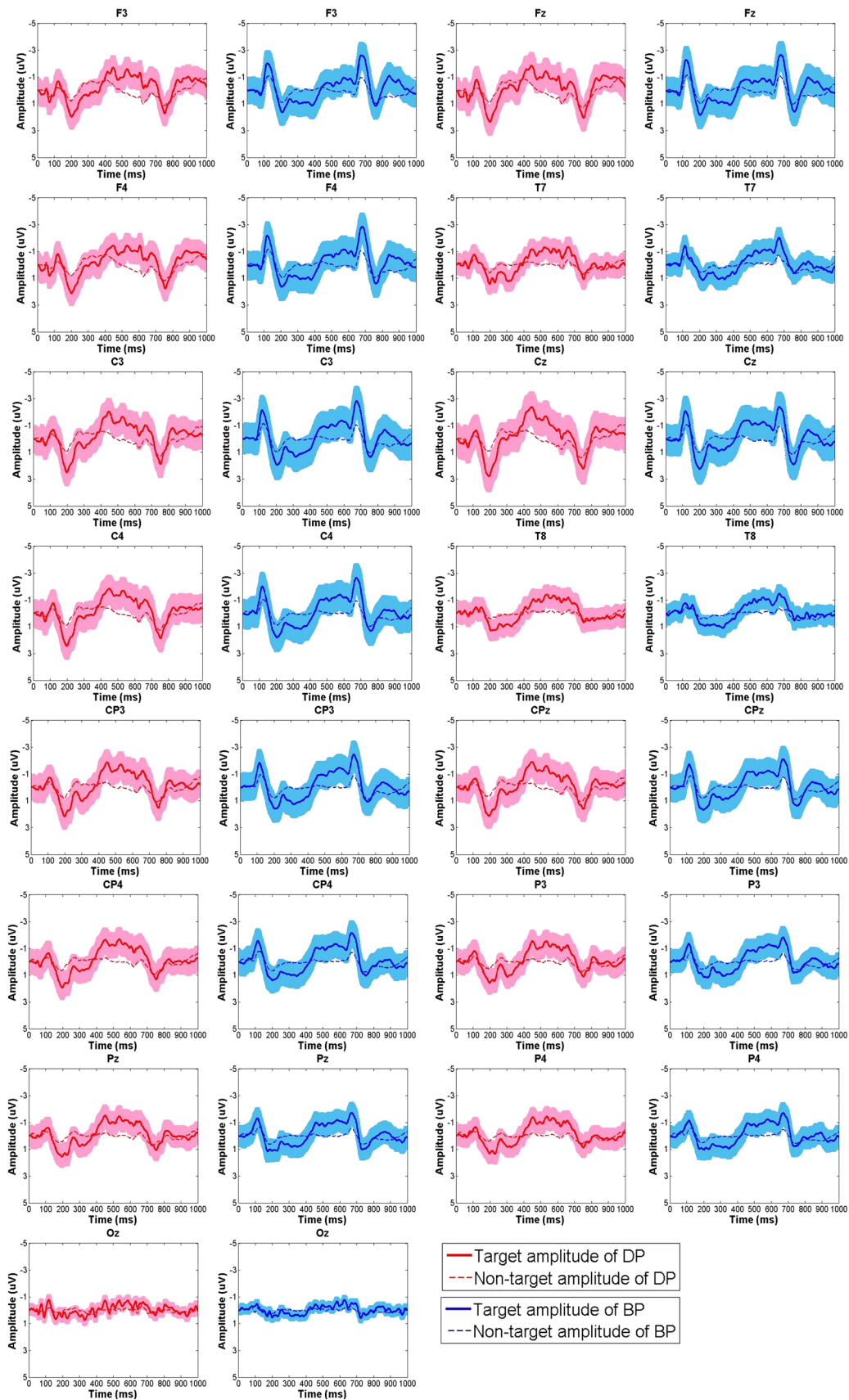
In this study, the online accuracy, average amplitudes of ERPs and scores collected from feedback were the interested items to analysis. A paired  $t$  test was used to show the statistical differences of the amplitudes of ERPs across all participants between two paradigms. Since the classification accuracy and user feedback scores were not meet the normal distribution, a Wilcoxon signed rank test was used to do the statistical analysis.

## Results

### ERPs

The grand averaged amplitudes of the target and non-target of two paradigms are shown in Fig. 2. Clear N1-P2 components could be observed in two paradigms. The amplitudes of N1 of the BP was significantly higher than that of DP at C3 ( $p < 0.01$ ,  $t = -4.12$ , paired  $t$  test). The P2 amplitude of DP was significantly higher than the BP at T8 ( $p < 0.05$ ,  $t = 2.24$ , paired  $t$  test). It was found that significant differences existed on the P300 amplitude between DP and BP at Fz ( $p < 0.05$ ,  $t = -3.12$ , paired  $t$  test). Following the P300 potential, a negative component began at 400 ms. DP had a higher average peak value of the





negative component than that of BP at Cz ( $p < 0.01$ ,  $t = -4.27$ , paired  $t$  test).

To evaluate the classification contributions of the components, r-square values were calculated (Fig. 3), which showed that P300 and N400-like contributed more for classification accuracy compared to other ERPs.

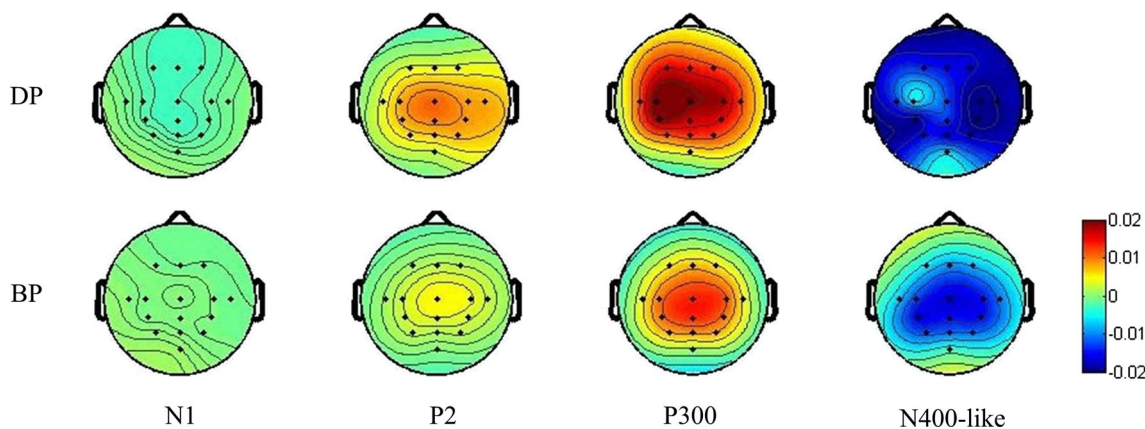
**Offline accuracy**

The offline accuracies of 11 participants for each pattern are shown in Fig. 4. All participants achieved higher accuracies of both patterns than the chance level (33.3%),

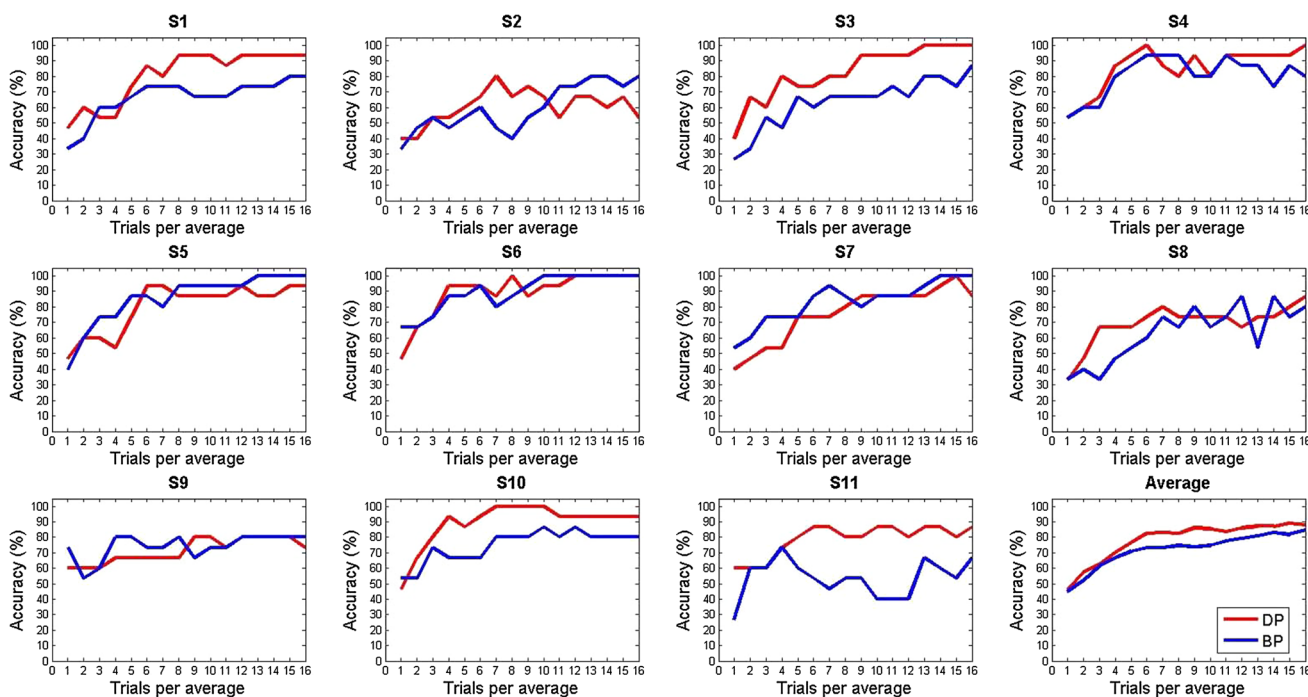
showing that the DP, which presented natural stimuli, could evoke distinguishable ERPs for classifier. No significant difference was found on the single-trial offline accuracy between DP and BP ( $p > 0.05$ , Wilcoxon signed test).

**Online accuracy and ITR**

The online accuracies and ITRs of each participant in the two patterns are listed in Table 1. The average online accuracies of DP and BP across 11 participants were 73.48 and 65.91%, respectively. The online accuracies and ITRs



**Fig. 3** Topographic maps of r-square values across 11 participants of DP and BP



**Fig. 4** Average accuracy of each participant and grand average accuracy across all participants for DP and BP. The solid red lines refer to the accuracy of DP, and the solid blue lines refer to the accuracy of BP. The

pictures from S1 to S11 represent the accuracies of participants from S1 to S11 and the last picture shows the grand average accuracy across all participants of DP and BP. (Color figure online)

**Table 1** Online accuracy of each participant and each paradigm, as well as group average accuracy of each paradigm

| Participant | Online accuracy (%) |              | ITR (bit/min) |             |
|-------------|---------------------|--------------|---------------|-------------|
|             | DP                  | BP           | DP            | BP          |
| S1          | 75.00               | 55.56        | 2.86          | 0.82        |
| S2          | 63.89               | 63.89        | 1.53          | 1.53        |
| S3          | 80.56               | 72.22        | 3.71          | 2.48        |
| S4          | 69.44               | 61.11        | 2.13          | 1.26        |
| S5          | 80.56               | 83.33        | 3.70          | 4.19        |
| S6          | 88.89               | 77.78        | 5.29          | 3.26        |
| S7          | 66.67               | 61.11        | 1.81          | 1.26        |
| S8          | 69.44               | 69.44        | 2.13          | 2.13        |
| S9          | 69.44               | 61.11        | 2.13          | 1.26        |
| S10         | 69.44               | 66.67        | 2.13          | 1.82        |
| FS11        | 75                  | 52.78        | 2.86          | 0.63        |
| Avg         | 73.48 ± 7.39        | 65.91 ± 9.22 | 2.75 ± 1.10   | 1.88 ± 1.08 |

of DP were significantly higher than that of BP ( $p < 0.05$ , Wilcoxon signed test;  $p < 0.05$ , Wilcoxon signed test). Three participants obtained accuracy levels surpassing 80% in the DP, but only one participant's accuracy was beyond 80% in the BP. The average ITRs of DP and BP were 2.75 bit/min and 1.88 bit/min (Table 1).

## Feedback

The feedback about the likability and difficulty of each paradigm is shown in Table 2. The scores in Table 2 indicated that most participants liked the DP better than the BP, and they thought the DP was less difficult than the BP. The group mean scores of likability were  $3.73 \pm 0.79$  for the DP and  $3.00 \pm 0.89$  for the BP. It showed that

**Table 2** Scores of each paradigm of each participant in items of likability and difficulty

| Participant | Likability  |             | Difficulty  |             |
|-------------|-------------|-------------|-------------|-------------|
|             | DP          | BP          | DP          | BP          |
| S1          | 4           | 3           | 4           | 3           |
| S2          | 4           | 3           | 3           | 4           |
| S3          | 3           | 3           | 3           | 4           |
| S4          | 2           | 4           | 4           | 3           |
| S5          | 3           | 5           | 4           | 3           |
| S6          | 4           | 3           | 2           | 3           |
| S7          | 4           | 2           | 3           | 4           |
| S8          | 5           | 3           | 3           | 5           |
| S9          | 4           | 2           | 4           | 3           |
| S10         | 4           | 3           | 3           | 4           |
| S11         | 4           | 2           | 3           | 3           |
| Avg         | 3.73 ± 0.79 | 3.00 ± 0.89 | 3.27 ± 0.65 | 3.55 ± 0.69 |

The score ranged from 1 to 5. Higher scores mean deeper degree on likability or difficulty

participants like the DP much more than the BP ( $p < 0.05$ , Wilcoxon signed test). The group mean scores on the difficulty were  $3.27 \pm 0.65$  for the DP and  $3.55 \pm 0.69$  for the BP. For the DP, two participants (S1 and S9) gave higher scores on both likability and difficulty, and two participants (S4 and S5) gave lower scores on likability and higher scores on difficulty, compared with those for the BP.

## Discussion

An auditory paradigm that used drip drops as stimuli was proposed in this study. Drip drops are natural sounds exhibiting special structure compared with artificial sounds. Therefore, this study investigated whether an auditory paradigm with drip drops performed better than the paradigm with beeps.

## Sound comparison

The structures of the beeps were different from drip drops, as shown by the spectrograms in Fig. 1. The drip drops presented wider frequency range and more complex structure than the beeps. Some neurophysiological investigations showed that the structures of natural sounds affect neural coding, such as higher information rates measured in the auditory system of the brain, indicating that human auditory systems are sensitive to natural sounds (Altmann et al. 2007; Cummings et al. 2006; Theunissen and Elie 2014; Wang and Chang 2008).

## ERPs

In this study, clear N1-P2 and P300 components were evoked in DP and BP. BP evoked higher amplitude of N1 than DP, but DP obtained higher P2 than BP. As mentioned

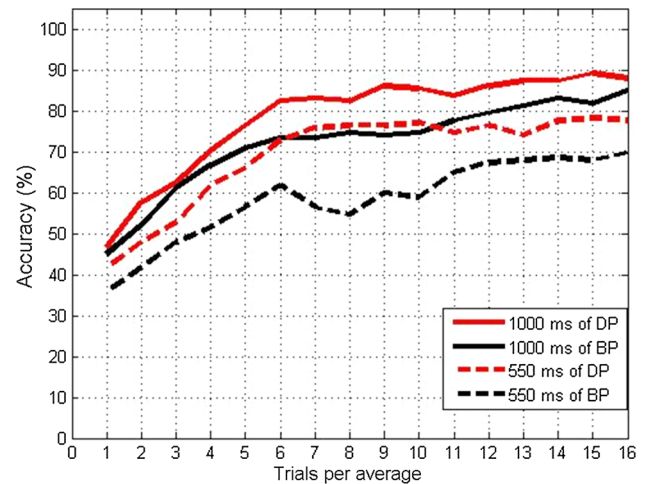
before, the drip drops had more complex structure than the beep, and related research demonstrated that the sounds with complex spectral could enhance the amplitude of P2 but not N1 (Del Cul et al. 2007). The sounds with wide spectrum were easily to be perceived than the sounds with single frequency (Tervaniemi et al. 2000) and this might make BCI users recognize the sounds in the DP more easily compared to the sounds in the BP. The contributions for classification of P300 components of DP and BP were showed in the Fig. 3. Obviously, the P300 component of DP contributed more for classification than that of BP ( $p < 0.05$ ,  $t = 2.33$ , paired  $t$  test, at C4). Although neither DP nor BP met the conditions to evoke a standard N400 component, this N400-like effect also could be found in other researches (Nijboer et al. 2008; Hill and Schölkopf 2012; Zhou et al. 2016). There was a hypothesis that the occurrence of this N400-like effect might be related to the amount of stimuli (Nijboer et al. 2008; Hill and Schölkopf 2012; Zhou et al. 2016).

Compared DP and BP, a clear difference was a negative component around 700 ms which was not shown in the DP (see Fig. 2). Martens et al. (2009) found that the short TTI or target-to-non-target-interval (TNI) would affect the morphologies of ERPs in a visual BCI system. In this study, this negative component around 700 ms in the target waveform from the BP might be an early N1 component which was evoked by the next non-target stimuli. N1 is an exogenous component and it also could be elicited by the non-target stimuli (Amenedo and Diaz 1998). The amplitude N1 of non-targets in the BP was significantly higher than that of DP ( $p < 0.01$ ,  $t = -4.74$ , paired  $t$  test). It indicated that the subject was affected by the non-target stimuli more in the BP compared to the DP.

### Accuracy and ITR

The hypothesis implying that the DP obtained higher accuracy than the BP was verified. Online accuracies of the DP were higher than the BP. The mean online accuracy of the DP was 73.48%. Compared with other auditory BCI systems with natural sounds, the classification accuracy of DP was above average. The average online accuracy determined by Klobassa et al. (2009) was 59.38%; the highest classification accuracy achieved by Höhne (2012) was below 70%; Lopez-Gordo et al. (2012) obtained an average classification accuracy of approximately 73% across all participants; the online classification accuracy of Simon et al. (2015) was 76.73%; Zhou et al. (2016) obtained an average online classification accuracy of approximately 74%. These studies showed that the classification accuracy of DP was acceptable.

Taking account of the overlap phenomenon, we limited the length of EEG data to 550 ms to train the classifier and



**Fig. 5** The comparison of offline accuracies using 1000 ms data and 500 ms data of DP and BP. The red solid line indicates the 1000 ms data of DP and the black solid line indicates the 500 ms data of BP. The dash red line indicates the 550 ms data of DP and the dash black line indicates the 550 ms data of BP. (Color figure online)

compared the offline accuracies of two conditions (1000 and 550 ms). It was found that the classification accuracies of both two patterns decreased a lot when 550 ms EEG data was used (Fig. 5). Because of this, 1000 ms EEG data was better than the 550 ms EEG data.

The average ITR of DP was 2.75 bit/min and apparently the ITR level of DP could not match the highest ITR (17.39 bit/min) of the auditory BCI proposed by Schreuder et al. (2009). The ITRs of auditory BCIs proposed in recent years ranged from 5.26 bit/min. For DP was a three-class paradigm in this study, the low-class would result in a low ITR. Another factor also needed to be considered was that the classification algorithms applied in this study was basic BLDA and it implied that the classification accuracy of DP still could be improved further (Long et al. 2011).

### Feedback about likability and difficulty

Likability and difficulty were common indices used to evaluate the friendliness of the BCI systems in many related works (Höhne 2012; Simon et al. 2015; Zhou et al. 2016). Participants preferred the DP for its euphonious stimuli ( $p < 0.05$ , Wilcoxon signed test). Although two participants (S1 and S9) reported that the DP was more difficult than the BP, they rated higher scores on likability. The mean score on the difficulty of the DP was lower than the BP. These indexes showed the DP was a reliable auditory paradigm with favorable user-experience.

### Limitations

Two limitations existed in this work. First, the criterion for designing the workload of online tasks remains unclear and



would be explored in further research. Second, the paradigm was tested on healthy participants, but its performance on patients was not evaluated. The current findings proved the feasibility of using drip drops as stimuli in an auditory BCI system, but further studies should be conducted to verify the performance of this strategy on patients.

## Conclusion

In this work, the usage of drip drops as stimuli in an auditory P300 paradigm was explored by comparing with the paradigm using beeps as stimuli. The drip-drop paradigm obtained better accuracy and was more user-friendly than the beep paradigm. Hence, the auditory BCI system can be a pleasant communication method for the users.

**Acknowledgements** This work was supported in part by the Grant National Natural Science Foundation of China, under Grant Nos. 91420302, 61573142, and 61703407. This work was also supported by the programme of Introducing Talents of Discipline to Universities (the 111 Project) under Grant B17017 and Shanghai Chenguang Program under Grant 14CG31, and the Foundation of Key Laboratory of Science and Technology for National Defense (No. 6142222030301).

## References

- Altmann CF, Bledowski C, Wibral M, Kaiser J (2007) Processing of location and pattern changes of natural sounds in the human auditory cortex. *Neuroimage* 35(3):1192–1200
- Amenedo E, Diaz F (1998) Aging-related changes in processing of non-target and target stimuli during an auditory oddball task. *Biol Psychol* 48(3):235–267
- Başar E, Güntekin B, Atagün İ, Gölbaşı BT, Tülay E, Özerdem A (2012) Brain's alpha activity is highly reduced in euthymic bipolar disorder patients. *Cogn Neurodyn* 6(1):11–20
- Baykara E et al (2016) Effects of training and motivation on auditory P300 brain–computer interface performance. *Clin Neurophysiol* 127(1):379–387
- Cummings A, Čeponienė R, Koyama A, Saygin AP, Townsend J, Dick F (2006) Auditory semantic networks for words and natural sounds. *Brain Res* 1115(1):92–107
- Del Cul A, Baillet S, Dehaene S (2007) Brain dynamics underlying the nonlinear threshold for access to consciousness. *PLoS Biol* 5(10):e260
- Donchin E, Spencer KM, Wijesinghe R (2000) The mental prosthesis: assessing the speed of 39 P300-based brain–computer interface. *IEEE Trans Rehabil Eng Publ IEEE Eng Med Biol Soc* 8(2):174–179
- Farwell LA (2012) Brain fingerprinting: a comprehensive tutorial review of detection of concealed information with event-related brain potentials. *Cogn Neurodyn* 6(2):115
- Farwell LA, Donchin E (1988) Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalogr Clin Neurophysiol* 70(6):510–523
- Farwell LA, Richardson DC, Richardson GM (2013) Brain fingerprinting field studies comparing P300-MERMER and P300 brainwave responses in the detection of concealed information. *Cogn Neurodyn* 7(4):263
- Furdea A, Halder S, Krusienski DJ, Bross D, Nijboer F, Birbaumer N, Kübler A (2009) An auditory oddball (P300) spelling system for brain–computer interfaces. *Psychophysiology* 46(3):617–625
- Güntekin B, Başar E (2010) A new interpretation of P300 responses upon analysis of coherences. *Cogn Neurodyn* 4(2):107–118
- Halder S et al (2010) An auditory oddball brain–computer interface for binary choices. *Clin Neurophysiol Off J Int Fed Clin Neurophysiol* 121(4):516–523
- Halder S, Käthner I, Kübler A (2016) Training leads to increased auditory brain–computer interface performance of end-users with motor impairments. *Clin Neurophysiol* 127(2):1288–1296
- Hill N, Schölkopf B (2012) An online brain–computer interface based on shifting attention to concurrent streams of auditory stimuli. *J Neural Eng* 9(2):026011
- Hill NJ, Lal TN, Bierig K, Birbaumer N (2005) Attention modulation of auditory event-related potentials in a brain–computer interface. In: *IEEE international workshop on biomedical circuits and systems*, pp S3/5/INV–S3/17–20
- Hoffmann U, Vesin JM, Ebrahimi T, Diserens K (2008) An efficient P300-based brain–computer interface for disabled subjects. *J Neurosci Methods* 167(1):115–125
- Höhne J (2012) Natural stimuli improve auditory BCIs with respect to ergonomics and performance. *J Neural Eng* 9(4):2099–2102
- Huang M, Daly I, Jin J, Zhang Y, Wang X, Cichocki A (2016) An exploration of spatial auditory BCI paradigms with different sounds: music notes versus beeps. *Cogn Neurodyn* 10(3):1–9
- Jeon JY, Lee PJ, You J, Kang J (2010) Perceptual assessment of quality of urban soundscapes with combined noise sources and water sounds. *J Acoust Soc Am* 127(3):1357–1366
- Jin J, Allison BZ, Sellers EW, Brunner C, Horki P, Wang X, Neuper C (2011) An adaptive P300-based control system. *J Neural Eng* 8(3):292–301
- Jin J, Allison BZ, Zhang Y, Wang X, Cichocki A (2014) An ERP-based BCI using an oddball paradigm with different faces and reduced errors in critical functions. *Int J Neural Syst* 24(8):1450027–1450027
- Jin J, Sellers EW, Zhou S, Zhang Y, Wang X, Cichocki A (2015) A P300 brain–computer interface based on a modification of the mismatch negativity paradigm. *Int J Neural Syst* 25(3):595–599
- Klobassa DS, Vaughan TP, Schwartz NE, Wolpaw JR, Neuper C, Sellers EW (2009) Toward a high-throughput auditory P300-based brain–computer interface. *Clin Neurophysiol Off J Int Fed Clin Neurophysiol* 120(7):1252–1261
- Kübler A, Furdea A, Halder S, Hammer EM, Nijboer F, Kotchoubev B (2009) A brain–computer interface controlled auditory event-related potential (p300) spelling system for locked-in patients. *Ann N Y Acad Sci* 1157(1):90–100
- Long J, Gu Z, Li Y, Yu T, Li F, Fu M (2011) Semi-supervised joint spatio-temporal feature selection for P300-based BCI speller. *Cogn Neurodyn* 5(4):387–398
- Lopez-Gordo MA, Fernandez E, Romero S, Pelayo F, Prieto A (2012) An auditory brain–computer interface evoked by natural speech. *J Neural Eng* 9(3):408–417
- Lulé D, Noirhomme Q, Kleih SC (2013) Probing command following in patients with disorders of consciousness using a brain–computer interface. *Clin Neurophysiol Off J Int Fed Clin Neurophysiol* 124(1):101–106
- Martens S, Hill N, Farquhar J (2009) Overlap and refractory effects in a brain? computer interface speller based on the visual P300 event-related potential. *J Neural Eng* 6(2):026003
- Martinez P, Bakardjian H, Cichocki A (2007) Fully online multi-command brain–computer interface with visual neurofeedback using SSVEP paradigm. *Comput Intell Neurosci* 2007:94561. <https://doi.org/10.1155/2007/94561>

- Monica F, Demetrios K, Emanuel D (1995) P300 and recall in an incidental memory paradigm. Humana Press, New York
- Nijboer F, Furdea A, Gunst I, Mellinger J, Mcfarland DJ, Birbaumer N, Kübler A (2008) An auditory brain–computer interface (BCI). *J Neurosci Methods* 167(1):43–50
- Pan J, Li Y, Gu Z, Yu Z (2013) A comparison study of two P300 speller paradigms for brain–computer interface. *Cogn Neurodyn* 7(6):523–529
- Park M et al (2016) Dysfunctional information processing during an auditory event-related potential task in individuals with Internet gaming disorder. *Transl Psychiatry* 6:e721. <https://doi.org/10.1038/tp.2015.215>
- Puanhvuan D, Khemmachotikun S, Wechakarn P, Wijarn B, Wong-sawat Y (2017) Navigation-synchronized multimodal control wheelchair from brain to alternative assistive technologies for persons with severe disabilities. *Cogn Neurodyn* 11(2):1–18
- Schreuder M, Tangermann M, Blankertz B (2009) Initial results of a high-speed spatial auditory BCI. *Int J Bioelectromagn* 11(2):105–109
- Schreuder M, Blankertz B, Tangermann M (2010) A new auditory multi-class brain–computer interface paradigm: spatial hearing as an informative cue. *PLoS ONE* 5(3):e9813
- Schreuder M, Rost T, Tangermann M (2011) Listen, you are writing! Speeding up online spelling with a dynamic auditory BCI. *Front Neurosci* 5:112. <https://doi.org/10.3389/fnins.2011.00112>
- Sellers EW, Donchin E (2006) A P300-based brain–computer interface: initial tests by ALS patients. *Clin Neurophysiol* 117(3):538–548
- Simon N, Käthner I, Ruf CA, Pasqualotto E, Kübler A, Halder S (2015) An auditory multiclass brain–computer interface with natural stimuli: usability evaluation with healthy participants and a motor impaired end user. *Front Hum Neurosci* 8:1039. <https://doi.org/10.3389/fnhum.2014.01039>
- Sutton S, Braren M, Zubin J, John ER (1965) Evoked-potential correlates of stimulus uncertainty. *Science* 150(3700):1187–1188
- Tervaniemi M, Schröger E, Saher M, Näätänen R (2000) Effects of spectral complexity and sound duration on automatic complex-sound pitch processing in humans—a mismatch negativity study. *Neurosci Lett* 290(1):66–70
- Theunissen FE, Elie JE (2014) Neural processing of natural sounds. *Nat Rev Neurosci* 15(6):355–366
- Wang D, Chang P (2008) An oscillatory correlation model of auditory streaming. *Cogn Neurodyn* 2(1):7–19
- Xu M et al (2016) Use of a steady-state baseline to address evoked vs. oscillation models of visual evoked potential origin. *Neuroimage* 134:204–212
- Yin E, Zhou Z, Jiang J, Chen F, Liu Y, Hu D (2013) A novel hybrid BCI speller based on the incorporation of SSVEP into the P300 paradigm. *J Neural Eng* 10(2):026012
- Zhang Y, Guo D, Xu P, Zhang Y, Yao D (2016) Robust frequency recognition for SSVEP-based BCI with temporally local multivariate synchronization index. *Cogn Neurodyn* 10(6):505–511
- Zhou S, Allison BZ, Kübler A, Cichocki A, Wang X, Jin J (2016) Effects of background music on objective and subjective performance measures in an auditory BCI. *Front Comput Neurosci* 10:105. <https://doi.org/10.3389/fncom.2016.00105>
- Zhu D, Bieger J, Molina GG, Aarts RM (2010) A survey of stimulation methods used in SSVEP-based BCIs. *Comput Intell Neurosci*. <https://doi.org/10.1155/2010/702357>