A Brain-Computer Interface Based on Abstract Visual and Auditory Imagery: Evidence for an Effect of Artistic Training

Kiret Dhindsa[1(✉)], Dean Carcone[2], and Suzanna Becker[2]

[1] School of Computational Science and Engineering,
McMaster University, Hamilton, ON, Canada
dhindsj@mcmaster.ca
[2] Department of Psychology, Neuroscience, and Behaviour,
McMaster University, Hamilton, ON, Canada
becker@mcmaster.ca

Abstract. Various kinds of mental imagery have been employed in controlling a brain-computer interface (BCI). BCIs based on mental imagery are typically designed for certain kinds of mental imagery, e.g., motor imagery, which have known neurophysiological correlates. This is a sensible approach because it is much simpler to extract relevant features for classifying brain signals if the expected neurophysiological correlates are known beforehand. However, there is significant variance across individuals in the ability to control different neurophysiological signals, and insufficient empirical data is available in order to determine whether different individuals have better BCI performance with different types of mental imagery. Moreover, there is growing interest in the use of new kinds of mental imagery which might be more suitable for different kinds of applications, including in the arts.

This study presents a BCI in which the participants determined their own specific mental commands based on motor imagery, abstract visual imagery, and abstract auditory imagery. We found that different participants performed best in different sensory modalities, despite there being no differences in the signal processing or machine learning methods used for any of the three tasks. Furthermore, there was a significant effect of background domain expertise on BCI performance, such that musicians had higher accuracy with auditory imagery, and visual artists had higher accuracy with visual imagery.

These results shed light on the individual factors which impact BCI performance. Taking into account domain expertise and allowing for a more personalized method of control in BCI design may have significant long-term implications for user training and BCI applications, particularly those with an artistic or musical focus.

Keywords: Brain-computer interface · Mental imagery · Individual differences · Performance predictors · Domain expertise · User-centred design · Auditory imagery · Visual imagery

1 Introduction

Brain-computer interfaces (BCIs) allow a user to control a computerized device using their brain activity directly [53]. This is achieved by interpreting user intentions or reactions from brain recordings in real time. BCIs based on mental imagery are particularly flexible because they potentially allow for a high number of inputs, or mental commands, and because they can be implemented such that the user may issue his/her mental commands at will, rather than as a reaction to a stimulus (see for example [30,37]). Thus, mental imagery BCIs can be categorized as spontaneous BCIs (also called asynchronous BCIs) [7,13,24,31].

Further advances to mental imagery BCIs may bring a more conscious, creative, and free interactive BCI experience in the future. As signal processing and machine learning algorithms become more reliable and generalizable in translating mental commands recorded in electroencephalography (EEG) and other brain recording technologies into BCI outputs, BCI users will be able to interact with BCIs in more varied and personalized ways. However, current BCIs are much more restrictive than this. At present, BCIs are capable of recognizing only a few predetermined mental commands reliably, and users are asked to learn how to modulate specific neurophysiological signals using mental imagery which is narrowly defined by the design of the BCI itself.

Mental imagery BCIs are restricted to a few predefined mental commands because doing so simplifies the problem of translating brain activity into BCI outputs. If users are instructed to use mental images which have well-characterized neurological correlates, then the BCI will know what changes in brain activity to look for. By far the most common form of mental imagery used in BCI is motor imagery [37,49], in which the user imagines performing a specific action involving one or more parts of their body. Motor imagery is convenient in the BCI context because it is known to modulate the sensorimotor rhythm (SMR), an oscillation pattern typically in the 8–12 Hz frequency band over sensorimotor cortex (also known as the μ rhythm) [41], in a similar fashion as real motor actions [33,42]. Furthermore, different motor images can be localized spatially. For example, real and imagined left versus right hand movements result in a suppression of the SMR in a localized region on the opposite hemisphere of the brain [40]. Therefore, motor imagery lends itself to create a relatively simple mental imagery BCI.

Despite the advancement it has brought to the field, the current reliance on motor imagery to drive the development of mental imagery BCI methods and applications may be limited in the long term. Individuals vary significantly in their ability to voluntarily modulate their SMR [45,46], and the ability to modulate the sensorimotor rhythm is correlated with cognitive profile and past experience outside of the BCI context [1,9,18,19,47,52]. This may explain why an estimated 15%–25% of individuals are unable to control a BCI with motor imagery [4,22].

It has been suggested previously that making mental imagery BCIs more reliable for the general user may require more than merely training unsuccessful users to use different kinds of motor imagery or to modulate their SMR in differ-

ent ways. Instead, the solution might be to allow different users to use different kinds of neurophysiological signals altogether [2]. In this study, we ask whether it is possible to use different kinds of mental imagery with a BCI designed for generalizability and to allow different users to use different specific mental imagery (we call this an Open-Ended BCI [14]). Furthermore, given that successful modulation of the SMR and successful use of BCIs based on motor imagery is at least partially dependent on individual factors, we ask whether it is also the case that success with different kinds of mental imagery depends on background experience relevant to the sensory modality used when controlling the BCI. In particular, we compare motor imagery to abstract visual imagery and abstract auditory imagery and ask whether success with any of these modalities is related to artistic or athletic background. The results of this study have potentially profound implications for BCI design and training, especially in the context of creative or artistic BCI applications.

2   Methods

Thirteen undergraduate and graduate participants practiced controlling an EEG-based BCI using three different kinds of mental imagery (data from three participants were excluded due to poor signal quality, so only data from ten participants are reported here). Visual imagery was used to change the size of a circle, auditory imagery was used to control the pitch of a tone, and motor imagery, used for comparison, was used to control the position of a circle on a computer display. Three 30-minute sessions were completed for one type of mental imagery over the course of one week (with some variation to accommodate the schedules of each participant) before moving to the next type of mental imagery. The order in which the three different types of mental imagery tasks were completed was counter-balanced across participants. The experiment was approved by the McMaster Research Ethics Board.

Participants were free to choose their own particular mental commands within each sensory modality. However, each participant was asked to make sure that their mental commands were very distinct and invoked rich and salient sensory imagery. Furthermore, since it was very difficult for participants to employ only one type of sensory imagery at the complete exclusion of others (e.g., as known from previous studies, it is difficult to engage in purely kinesthetic motor imagery without any accompanying visual imagery [36,48]), the requirement was only that the appropriate sensory modality was the most dominant and salient feature of each mental command. The mental commands chosen by each participant for each task are summarized in Table 2.

2.1   EEG Hardware

The Emotiv Epoc [16] was used to record EEG. The Epoc is a consumer-grade EEG headset previously shown to provide useful EEG but with poor signal

quality compared to research-grade devices [3,15,25]. However, successful BCI studies have been conducted using this device in the past [10,28].

The Emotiv Epoc is equipped with 14 saline-based electrodes with additional channels for Common Mode Sense (CMS) and Drven Right Leg (DLR) located at P3 and P4 according to the International 10–20 system (these are used for referencing and noise reduction). EEG is recorded with a sampling rate of 128 Hz and a 0.2–45 Hz bandpass filter along with 50 Hz and 60 Hz notch filters are implemented in the hardware. The electrode configuration is shown in Fig. 1.
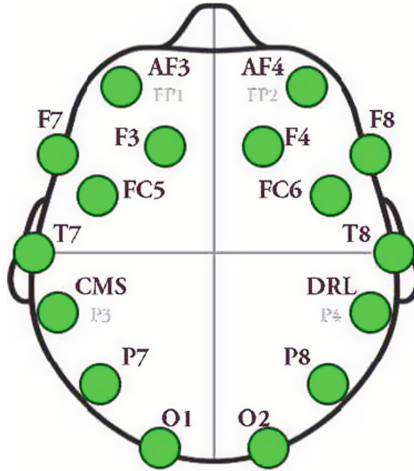

Fig. 1. Emotiv Epoc electrode layout. Symmetrically on each hemisphere is one electrode on visual cortex, one on parietal cortex, one on temporal cortex (with one near the border of temporal and frontal cortices), and three on frontal cortex.

2.2 Experimental Procedure

Before beginning the experiment, participants were asked to complete a brief questionnaire examining their background experience in the arts and in athletic activities. The questionnaire asked participants to indicate how many years of practice, how many hours per week they practice, and for their self-rated expertise in visual arts, music, and athletics/sports. The questions and responses are given in Table 3. The order of imagery tasks was then determined by counterbalancing with previous participants.

At the beginning of each session, an experimenter fit the EEG headset to the participant. Since the Emotiv Epoc does not allow for direct measurements of impedance, impedance was estimated using the proprietary toolbox that accompanies the device. In this toolbox, a colour-coded display indicates the signal quality at each electrode site. Electrodes were readjusted and saline solution was reapplied until all 14 sites showed "good" signal quality according to the

proprietary software. In cases where good signal quality was especially difficult to achieve, at most two electrodes were allowed to show less than "good" signal quality.

Data collection was completed with Matlab 2013b [32], Simulink, and Psych-toolbox [8]. At the start of each session, on-screen text reiterated the description of the experiment and all necessary instructions, including instructions to avoid blinking, head/eye movements, jaw clenches, and any other muscular activity during the mental imagery period. Each session included 10 blocks of 20 trials, where each trial spanned approximately nine seconds (the structure of each trial is given in Fig. 2). The first block of every session was used for pretraining. Therefore, no classification was performed and no feedback was provided to the participant. These twenty trials were used to construct models with which to classify trials in the next block. The models were updated at the end of every block, and the newly updated models were used to classify trials in the next block.
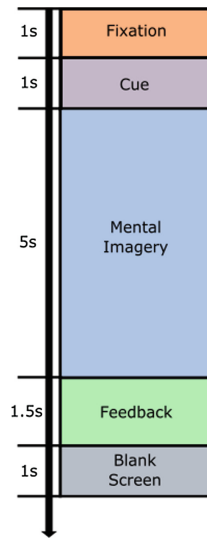

Fig. 2. The structure of each trial. A white fixation cross appeared for 1 s over a black background to indicate the start of a new trial. A textual cue (e.g., "low note", "shrink", "left", etc.) then appeared in white font in place of the fixation cross and persisted for 1 s. This cue was replaced by the fixation cross for 5 s, marking the mental imagery period. The feedback stimulus was then presented for 1.5 s corresponding to the classification confidence level. At the end of the trial, the screen was left blank for 1 s.

After each session, participants completed a questionnaire asking them to describe the specific mental commands used and to rate their level of interest in the task. The mental commands used by participants are summarized in Table 2. Correlational analyses comparing task interest and the accuracy of the BCI are given in Sect. 3.1.

2.3   EEG Processing Pipeline

Each BCI used the same processing pipeline so that performance across types of mental imagery could be fairly compared. Common spatial patterns (CSP) [34, 44] and power spectral density estimation (PSD) were used to extract features. Minimum-Redundancy Maximum-Relevance (MRMR) [39] was used for feature selection. Finally, a linear Support Vector Machine (SVM) [12] was used for binary classification.

Feature Extraction: Common Spatial Patterns and Power Spectral Density Estimation. CSP is a PCA-based supervised spatial filter typically used for motor imagery classification for EEG-based BCIs [34, 44], but various extensions of CSP have also been used to classify other types of mental imagery in EEG (e.g., emotional imagery [21]). CSP is a supervised method that aims to construct a spatial filter which yields components (linear combinations of EEG channels) whose difference in variance between two classes is maximized.

The CSP filter W is constructed with respect to two $N \times S_1$ and $N \times S_2$ EEG data matrices X_1 and X_2, where N is the number of EEG channels and S_1 and S_2 are the total number of samples belonging to class one and class two respectively. The normalized spatial covariance matrices of X_1 and X_2 are then computed as follows:

$$R_1 = \frac{X_1 X_1^T}{trace(X_1 X_1^T)} \qquad R_2 = \frac{X_2 X_2^T}{trace(X_2 X_2^T)}, \qquad (1)$$

where T denotes the transpose operator. The composite covariance matrix is then taken using

$$R_c = R_1 + R_2. \qquad (2)$$

The eigendecomposition of R_c

$$R_c = V \lambda V^T \qquad (3)$$

can be taken to obtain the matrix of eigenvectors V and the diagonal matrix of eigenvalues in descending order λ. The whitening transform

$$Q = V \sqrt{\lambda^{-1}} \qquad (4)$$

is then computed so that $Q R_c Q^T$ has all variances (diagonal elements) equal to one. Because Q is computed using the composite covariance matrix in Eq. 2,

$$R_1^* = Q R_1 Q^T \qquad and \qquad R_2^* = Q R_2 Q^T \qquad (5)$$

have a common matrix of eigenvectors V^* such that

$$R_1^* = V^* \lambda_1 V^{*T}, \qquad R_2^* = V^* \lambda_2 V^{*T}, \qquad and \qquad \lambda_1 + \lambda_2 = I, \qquad (6)$$

where I is the identity matrix. Hence, the largest eigenvalues for R_1^* are the smallest eigenvalues for R_2^* and vice versa. Since R_1^* and R_2^* are whitened spatial

covariance matrices for X_1 and X_2, the first and last eigenvectors of V^*, which correspond to the largest and smallest eigenvalues in λ_1, define the coefficients for two linear combination of EEG channels which maximize the difference in variance between both classes. Given this result, the CSP filter W is constructed with

$$W = (V^{*T}Q)^T \tag{7}$$

and is used to decompose EEG trials into CSP components like any other linear spatial filter:

$$C = WX_{EEG}. \tag{8}$$

For classification, W can be constructed using only the top M and bottom M eigenvectors from V^*, where $M \in \{1, 2, \ldots \lfloor N/2 \rfloor\}$ is a parameter that must be chosen, or alternatively, only the top M and bottom M rows of C can be used for feature extraction. Assuming the latter (i.e., that W was constructed using all eigenvectors in V^*), then features f_j, $j = 1, \ldots 2M$, are extracted by taking the log of the normalized variance for each of the $2M$ components in $Z = \{1, \ldots M, N - M + 1, \ldots N\}$:

$$f_j = \log \left[\frac{var(C_m)}{\sum_{i \in Z} var(C_i)} \right], \tag{9}$$

where $m \in Z$. These $2M$ features can then be used for classification.

Because CSP is a supervised spatial filter, it also allows for the estimation and visualization of the discriminative EEG spatial patterns corresponding to each class. In particular, the columns of W^{-1} can be interpreted as time-invariant EEG source distributions, and are called the common spatial patterns [5, 44].

This study involved three particular challenges with respect to the mental commands used by our participants: (1) a wide variety of mental commands were used between participants and between the three sensory modalities, (2) many of these mental commands were abstract and atypical for BCI use, and (3) the mental commands used by participants were not known a priori. Therefore, the EEG processing pipeline needed to cast a wide net in order to attempt to classify trials in the presence of these extra sources of variability. To do this, CSP models were computed from EEG after applying an 8–30 Hz 4th order Butterworth bandpass filter. We pre-selected $M = 2$, resulting in four CSP components and therefore four CSP features per trial. In addition to CSP features, the power of each CSP component was computed in non-overlapping 1 Hz bins, resulting in an additional 88 features per trial with which to attempt to find an optimally discriminative subset.

A total of 92 features per trial is too many for reliable classification given only a maximum of 180 trials for training, and only a small subset of these features were expected to have discriminative value. However, we could not know in advance which features would be useful because the choice of mental commands was left to the participants. In fact, it was expected that different features would be important for different types of mental imagery and for different participants, hence the need for feature selection.

Feature Selection: Minimum-Redundancy Maximum-Relevance.
MRMR is a supervised feature selection method based on mutual information
[39]. Its objective is to find a subset of features Z which has maximum mutual
information with the true class labels (maximum relevance) while at the same
time minimizing the mutual information between the selected features them-
selves (minimum redundancy). MRMR was chosen for this study because its
approach makes it particularly effective when the candidate features are highly
correlated and where only a small subset contribute distinct discriminative infor-
mation.

MRMR selects K features, where K is a chosen integer less than the total
number of features. Features are selected from the list of candidate features
sequentially. The first selected feature, z_1, is chosen by finding the candidate
feature which has the highest mutual information with the class labels in a
training set:

$$z_1 = \max_i I(F = \{f_i, i = 1, \ldots, N\}; Y), \tag{10}$$

where $f_i \in F$ are the individual candidate feature vectors in the candidate feature
matrix of the training set F, N is the total number of candidate features, Y are
the true class labels in the training set, and I is the mutual information function.
Each subsequent k^{th} selected feature for $k = 2, \ldots K$ is chosen by maximizing
the difference between relevance and redundancy, $D - R$, where

$$D = I(Z_k = \{z_i, i = 1, \ldots, k\}; Y), \tag{11}$$

which is estimated by

$$\bar{D} = \frac{1}{k} \sum_{z_i \in Z_k} I(z_i; Y) \tag{12}$$

in order to avoid computing potentially intractable joint probability densities,
and

$$R = I(Z_k, Z_k), \tag{13}$$

which is estimated by

$$\bar{R} = \frac{1}{k^2} \sum_{z_i, z_j \in Z_k} I(z_i; z_j). \tag{14}$$

During model construction and model updates (i.e., after every block of 20
trials within each session), we test a classifier with $K = 5, 10, \ldots, 40$ and choose
the model with the highest classification accuracy.

Classification: Linear Support Vector Machine. The linear SVM imple-
mentation from the libSVM Matlab toolbox was used [11]. In order to minimize
the time between blocks, we did not optimize the SVM parameters C and G
during model construction or model updates. The classifier, along with the CSP
filter and list of selected features, was updated after every block of trials to
incorporate all trials performed within that session (e.g., at the end of block 5,
the models were recomputed using all of the 100 trials completed during that

session). Each session was independent from previous sessions, even within the same sensory modality. New models were initialized and trained after the first block of every session without any reference to the models or trials obtained in previous sessions.

2.4   BCI Outputs and Feedback

Feedback was provided to participants after each trial according to the parameters given in Table 1. The feedback given was proportional to classifier confidence, where classifier confidence was the estimated probability of belonging to each class using a parametric model to fit posterior densities (see [11, 26, 27, 43, 54]). Using these probability estimates, weighted feedback could be presented between the two binary extremes for each type of mental imagery. For example, a classification decision in favour of a high tone in the auditory imagery case would result in a feedback tone with a frequency closer to the highest possible tone than the lowest possible tone. In contrast to using only binary feedback, participants were instructed to aim for maximally high tones or maximally low tones, thus training to improve classification confidence rather than just training to improve classification accuracy alone.

Table 1. The features of the BCI outputs for each type of mental imagery. The extreme outputs were shown during the pretraining block. In subsequent blocks, the feedback provided was somewhere in between the low and high extremes, based on classifier confidence, in the direction of the classifier's decision. Classifier confidence greater than 0.8 for either class also resulted in the extreme output.

Imagery	Output type	Low extreme	Midpoint	High extreme
Motor	100 pixel diameter white circle	−500 pixel shift	0 pixel shift	500 pixel shift
Visual	150 pixel diameter white circle	100 pixel diameter decrease	0 pixel change	100 pixel diameter increase
Auditory	Pure tones	220 Hz (A3)	440 Hz (A4)	880 Hz (A5)

2.5   Offline Analysis

Offline analysis using the Fieldtrip toolbox's [38] statistical thresholding-based artifact rejection was performed to remove trials contaminated by artifacts and reduce the risk that BCI performance could be explained by muscular activity. Visual inspection was performed after automatic artifact rejection in order to remove any trials which were not free of artifacts with high confidence. For each session, 15-fold cross-validation was performed where on each iteration all artifact-free trials belonging to that session were randomly partitioned into a

training and test set (the test set contained 25% of the trials), feature extraction was performed using the same method as in online analysis (CSP filters were trained using only the training set), and a linear SVM was used.

3   Results

3.1   BCI Performance

BCI performance varied considerably across participants. Figure 3 shows the average classification accuracy of the last three blocks of each session (the last three blocks were used as an estimate of final model performance). Similarly, the results of offline analysis are shown in Fig. 4. There was no specific effect of
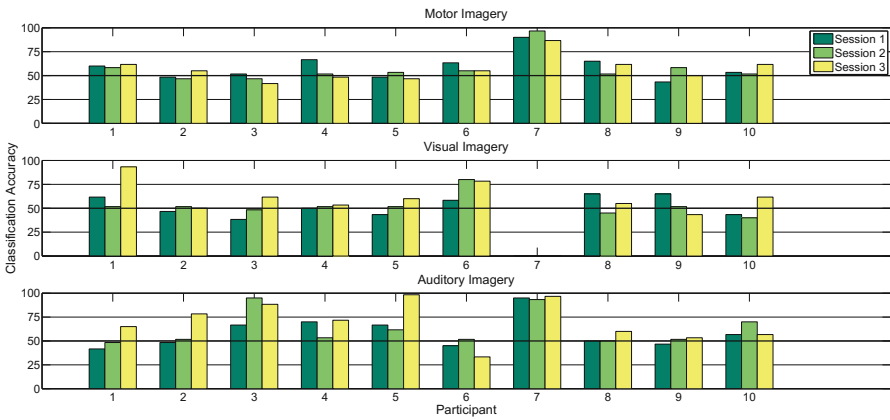

Fig. 3. Online classification accuracy over the final three blocks of each session. All participants and types of mental imagery are shown.
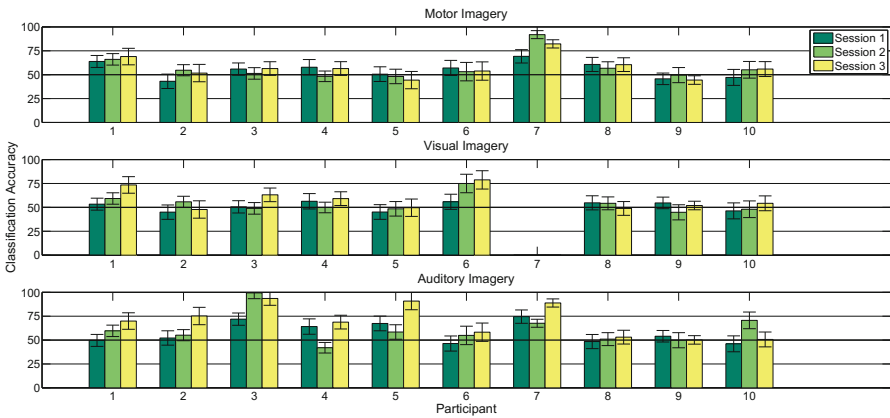

Fig. 4. Offline cross-validation accuracy for each session. Error bars show the standard deviation computed from all 15 cross-validation iterations.

task order ($F_{2,84} = 1.22$, $p = 0.30$) or sensory modality ($F_{2,84} = 2.39$, $p = 0.10$). However, there was weak but significant positive correlation between reported interest in the task and performance ($r = 0.28$, $p < 0.05$).

The specific mental commands performed by each participant are given in Table 2. The corresponding common spatial patterns are shown in Fig. 5.

Table 2. A summary of mental commands chosen by each participant for each training session. "Feedback stimulus" means the participant imagined the BCI outputs directly.

Imagery	Motor	Visual	Auditory
P1	Sweeping right arm/Sweeping left leg	Feedback stimulus	Piercing high note/Muffled low note
P2	N/A	N/A	N/A
P3	N/A	N/A	Opera singer/Chanting monks
P4	Guitar chord with left hand/Slapping with right hand	Growing blue circle/Shrinking marble	Buzzy kazoo/Leonard Cohen singing
P5	Boxing with right hand/Guarding with left hand	Self expanding/Shrinking a ball in hands	Jazz trumpet/Heavy metal vocals
P6	Retracting hand from hot stove/Painting with brush	Moon getting closer/Car driving away	Opera singer/Chanting om
P7	Punching with right hand/Stretching right arm to the left	*Withdrew from study*	Singing or playing high notes/Singing or playing low notes
P8	Right hand and left hand actions (not described)	Feedback stimulus	Feedback stimulus
P9	Lifting a dumbell/Dribbling a basketball	Car driving away/Balloon expanding	Bell ringing/Foghorn
P10	Turning a car right/Turning a car left	Inflating a balloon/Deflating a balloon	Screeching chalkboard/Growling lion

3.2 Effect of Background Experience

A significant effect of background expertise was found, as evaluated by our background experience questionnaire ($F_{2,80} = 14.0$, $p < 0.0001$, with variance explained $\omega^2 = 0.22$). Self-reported expertise in athletics, visual arts, or music was also significantly correlated with BCI performance ($r = 0.46$, $p < 0.0001$).
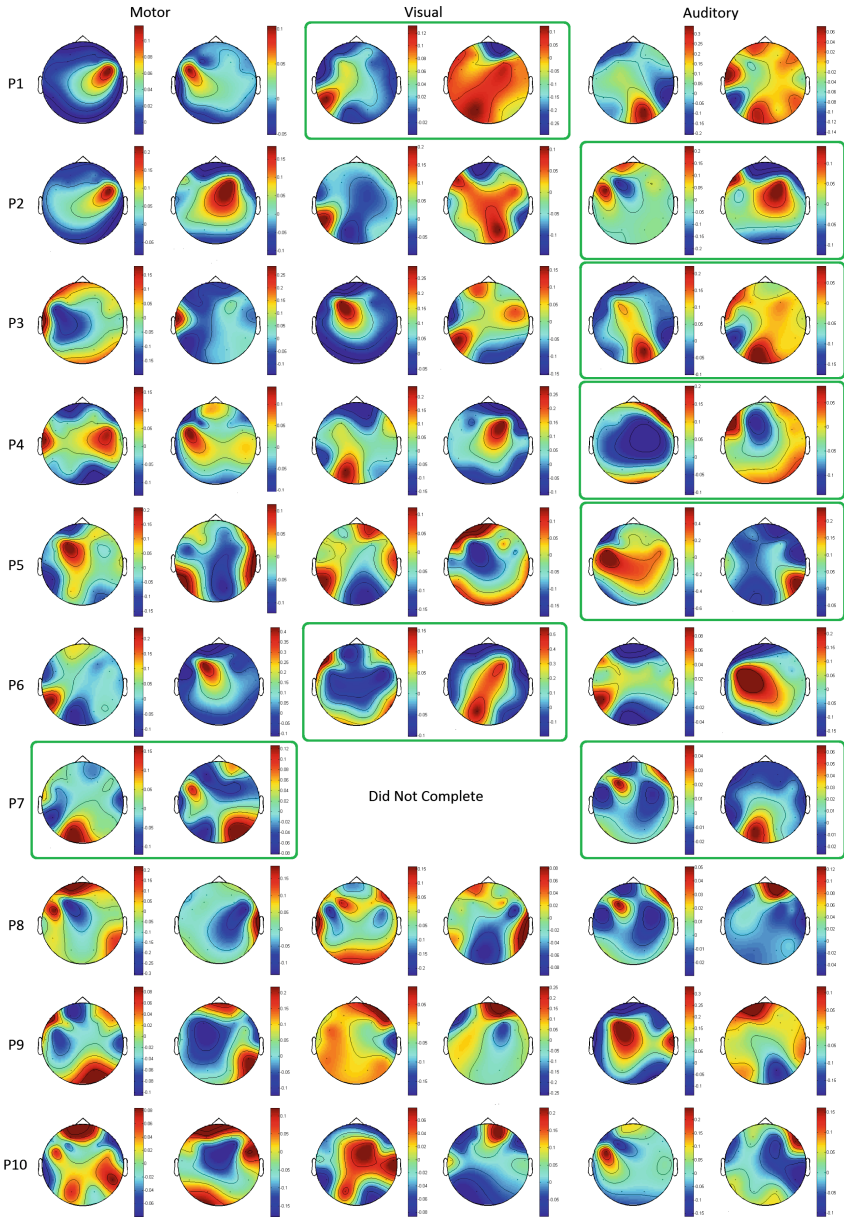
Fig. 5. The common spatial patterns (i.e., the first column and last column of W^{-1}) for the last session of each sensory modality for each participant. All trials within a session were used to construct the common spatial patterns shown here. The left pattern of each pair corresponds to the negative class (i.e., left shift, shrink, and low tone), and the right pattern corresponds to the positive class (i.e., right shift, grow, high tone). Sessions during which more than 70% classification accuracy was achieved are boxed in green. (Color figure online)
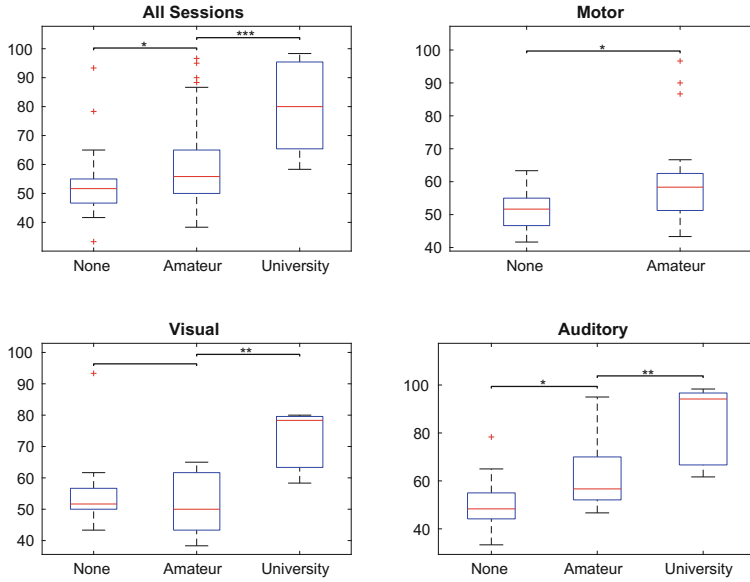
Fig. 6. BCI performance across sessions at varying levels of self-reported expertise in a relevant domain (athletics for motor imagery, visual art for visual imagery, and music for auditory imagery). $*p < 0.05$, $**p < 0.01$, $***p < 0.001$

BCI performance in all sessions organized by self-reported expertise in the corresponding domain is shown in Fig. 6.

4   Discussion and Conclusions

In this study we present two important findings which may impact how mental commands are chosen for BCI control. First, we found that by using a broad feature extraction approach, it was possible to enable user control over a BCI with abstract visual and auditory imagery, even when the specific mental commands were not known a priori. Second, it was found that participants were able to control a BCI using only one or two of the available types of mental imagery, and that this result may be related to the participant's artistic background.

4.1   Brain-Computer Interfacing with Abstract Mental Imagery

From Fig. 3, it can be seen that nine of out ten participants were able to achieve above chance level performance with at least one type of mental imagery on at least one session. Furthermore, eight out of ten participants achieved their best performance with 70% classification accuracy or above, where 70% is considered the minimum threshold for a communication device such as a BCI [23].

Table 3. Responses to the background experience questionnaire. The questionnaire asks participants to state the number of hours/week of practice, the number of years spent practicing, and a self-rating of their overall proficiency of performance level. Note that P1, P2, and P3 completed an earlier version of this questionnaire, so only their level of expertise is available.

Imagery	Motor			Visual			Auditory		
	Hrs/Wk	#Years	Level	Hrs/Wk	#Years	Level	Hrs/Wk	#Years	Level
P1	N/A	N/A	2	N/A	N/A	1	N/A	N/A	1
P2	N/A	N/A	1	N/A	N/A	1	N/A	N/A	1
P3	N/A	N/A	1	N/A	N/A	2	N/A	N/A	2
P4	3	2	2	1	1	1	4	4	2
P5	2	3	2	1	1	1	1	6	3
P6	2	1	1	3	4	3	1	1	1
P7	1	6	2	1	1	2	4	4	3
P8	4	4	2	2	3	2	1	1	2
P9	4	2	2	2	7	2	1	5	2
P10	4	3	2	1	3	2	1	5	2

Coding:
- Hours/Week (*1 = less than 1 h, 2 = 1–2 h, 3 = 3–4 h, 4 = 5–10 h, 6 = 10+ h*).
- #Years (*1 = less than 1 year, 2 = 1–2 years, 3 = 3–4 years, 4 = 5–6 years, 5 = 7–8 years, 6 = 9–10 years, 6 = 10+ years*)
- Level (*1 = Do not practice/perform, 2 = Amateur, 3 = Varsity/University level, 4 = Professional*)

The results obtained through offline analysis validate the BCI performance levels achieved during the online training experiment. In several cases, classification accuracy was higher in offline analysis than in online analysis. The main differences between the two analyses were that in offline analysis trials contaminated by artifacts were removed. In addition, all trials were shuffled during offline analysis before partitioning training and test sets, resulting in the training data containing a mix of trials from different blocks of each session. These two differences together may have made offline analysis more robust than its online counterpart, but the offline analyses do suggest that BCI performance was not substantially driven by artifacts. However, it is important to note that it is possible that subvocal muscle activity, micro eye movements, or micro muscle activations impacted performance. This possibility is discussed in greater detail in Sect. 4.3.

4.2   Evidence for an Effect of Background Experience

It is interesting that most participants performed much better with one sensory modality compared to others. Participants most often performed best with auditory or visual imagery rather than motor imagery, even though motor imagery

is usually considered simpler to classify. This might be explained by the electrode configuration of the Emotiv Epoc headset (this point is discussed further in Sect. 4.3). However, we also note that most of our participants reported having greater expertise in visual arts or music rather than in athletic performance or sports.

The differences in performance were not related to task order, perceived accuracy, or interest in each task (see Sect. 3.1). However, it was found that performance varied with background experience. Specifically, it was found that self-reported expertise or performance level in athletics/sports, visual arts, and music had an effect on BCI performance with different sensory modalities (see Sect. 3.2). While performance is also correlated with interest, which itself related to domain expertise, the effect of domain expertise specifically was stronger than the effect of interest in each task. Therefore, we conclude that domain expertise had a specific significant effect on BCI performance.

We suggest that results with a larger sample size and replication in other contexts is needed before these results should be incorporated into training humans to use a BCI. However, these results may have a profound impact on how BCI training is done. In particular, BCIs designed for artistic or creative applications, or BCIs designed to allow mental commands involving abstract visual or auditory imagery, may need to take into consideration the artistic background of its users during training. Likewise, BCIs intended as assistive or rehabilitative tools might benefit from designing for types of mental imagarey associated with any domain expertise acquired by the patient pre-injury. This observed effect of domain expertise may also have implications for BCI training more generally.

If it is indeed the case that artistic background or domain expertise more broadly has a significant impact on BCI performance, the suggestion to design BCIs which enable different users to employ different mental commands, even if different neurophysiological signals are used [14,17,35], must be examined more closely. Achieving this, however, requires the BCI community to meet the challenge of creating a truly generalizable BCI which does not need to know the kinds of mental imagery that will be used a priori.

There is also growing attention being brought to the need for improved BCI training and neurofeedback for humans [29]. While we do not explicitly present methods for this here, the results of this study may be relevant. In addition to improving methods for BCI training, the effect of background experience on BCI performance seen in the present study suggests that we should also consider which mental commands should be trained with which individuals in the first place.

The co-adaptive BCI approach is a good example of an advancement in the direction of individual-based mental command selection [50,51]. However, because it attempts to find an optimal subset of mental commands from a predefined set of choices, it cannot fully take advantage of individual factors influencing the best choice of mental commands. In order to do so with this approach would require an expontentially increasing number of combinations of mental commands to test. The BCI presented here takes a different approach

to reach a similar goal. Rather than trying to find an optimal subset of mental commands from a list of mental commands, we left the choice of mental command open to the user and aimed to find an optimal set of features from a list of candidate features.

4.3   Limitations and Future Work

BCI performance and direct comparisons in performance between different sensory modalities or specific kinds of mental commands are limited in this study by the Emotiv Epoc hardware. The unchangeable electrode configuration of the headset is less optimal for some types of mental commands than others. In particular, no sensors are placed over locations C3 or C4, which are most commonly used to detect the sensorimotor rhythm which is modulated during motor imagery. Similarly, only two electodes are available over the occipital cortices, which might have otherwise played a more central role in detecting visual imagery. Instead, the Emotiv Epoc relies most heavily on the frontal cortices, which may in part determine which specific mental commands were most successful. For example, perhaps mental commands with different emotional content or with differing degrees of cognitive load would be more successful with this electrode configuration, but it is not clear from descriptions of the mental commands used whether this was an explanatory factor in differences in BCI performance in this study.

The Emotiv Epoc is also known to have significantly lower a signal-to-noise ratio compared to research-grade devices [3,15,25] and to result in lower BCI performance (e.g.,[6,15], or comparing [28] and [20]). However, our aim here was not to achieve state of the art BCI performance, but rather to assess BCI performance in the context of abstract user-defined visual and auditory imagery and to compare this performance to relevant domain expertise.

It is possible that artifacts could have interfered with BCI performance in a significant way. We conducted an offline classification analysis using the same feature extraction methods and classifier as in the online experiment but included standard artifact rejection software included in the Fieldtrip toolbox [38]. We saw a slight improvement in classification accuracy, suggesting that at least common artifacts, such as eye blinks and jaw clenches, were not driving BCI performance. However, there remains the possibility that very subtle muscle activity, such as subvocal laryngeal contractions influenced BCI performance. These would require electromyography (EMG) electrodes to detect, and therefore we cannot confirm whether these significantly affected performance. We would not expect, however, that the tendency to perform subvocal laryngeal contractions or other types of muscle activity would be so highly related to domain expertise, especially given the variety of mental commands used in this study (many of which did not correspond to the actual skill participants had specific training in). Therefore, we do not expect that BCI performance was driven mainly by such subtle muscle contractions.

The exact reasons background expertise may impact BCI performance with abstract mental imagery remains unknown. It is possible that someone who is

musically trained or merely innately musically talented is able to generate more salient, consistent, and rich auditory imagery than others. It is also possible that individuals who are able to produce such auditory imagery are also drawn to practising music. In addition to investigating the effect of background experience on BCI performance more broadly and with a larger sample of participants, it would be of great benefit to separate the effect of the quality (e.g., including saliency, consistency, and richness) of the mental commands themselves to see if these are highly correlated with background experience and if these factors are the primary drivers affecting the differences in BCI performance seen in this study.

References

1. Allison, B., Luth, T., Valbuena, D., Teymourian, A., Volosyak, I., Graser, A.: BCI demographics: how many (and what kinds of) people can use an SSVEP BCI? IEEE Trans. Neural Syst. Rehabil. Eng. **18**(2), 107–116 (2010)
2. Allison, B.Z., Neuper, C.: Could anyone use a BCI? In: Tan, D.S., Nijholt, A. (eds.) Brain-Computer Interfaces, pp. 35–54. Springer, Heidelberg (2010)
3. Badcock, N.A., Mousikou, P., Mahajan, Y., de Lissa, P., Thie, J., McArthur, G.: Validation of the Emotiv EPOC EEG gaming system for measuring research quality auditory ERPs. PeerJ **1**, 2 (2013)
4. Blankertz, B., Sannelli, C., Halder, S., Hammer, E.M., Kübler, A., Müller, K.R., Curio, G., Dickhaus, T.: Neurophysiological predictor of SMR-based BCI performance. NeuroImage **51**(4), 1303–1309 (2010)
5. Blankertz, B., Tomioka, R., Lemm, S., Kawanabe, M., Muller, K.-R.: Optimizing spatial filters for robust EEG single-trial analysis. IEEE Signal Proc. Mag. **25**(1), 41–56 (2008)
6. Bobrov, P., Frolov, A., Cantor, C., Fedulova, I., Bakhnyan, M., Zhavoronkov, A.: Brain-computer interface based on generation of visual images. PLoS One **6**(6), e20674 (2011)
7. Borisoff, J.F., Mason, S.G., Bashashati, A., Birch, G.E.: Brain-computer interface design for asynchronous control applications: improvements to the LF-ASD asynchronous brain switch. IEEE Trans. Biomed. Eng. **51**(6), 985–992 (2004)
8. Brainard, D.H.: The psychophysics toolbox. Spat. Vis. **10**, 433–436 (1997)
9. Burde, W., Blankertz, B.: Is the locus of control of reinforcement a predictor of brain-computer interface performance? na (2006)
10. Carrino, F., Dumoulin, J., Mugellini, E., Khaled, O.A., Ingold, R.: A self-paced BCI system to control an electric wheelchair: evaluation of a commercial, low-cost EEG device. In: 2012 ISSNIP Biosignals and Biorobotics Conference (BRC), pp. 1–6, January 2012
11. Chang, C.C., Lin, C.J.: LIBSVM: a library for support vector machines. ACM Trans. Intell. Syst. Technol. **2**, 27:1–27:27 (2011). http://www.csie.ntu.edu.tw/cjlin/libsvm
12. Cortes, C., Vapnik, V.: Support-vector networks. Mach. Learn. **20**(3), 273–297 (1995)

13. del Millan, J.R., Mouriño, J.: Asynchronous BCI and local neural classifiers: an overview of the adaptive brain interface project. IEEE Trans. Neural Syst. Rehabil. Eng. **11**(2), 159–161 (2003)
14. Dhindsa, K., Carcone, D., Becker, S.: An open-ended approach to BCI: embracing individual differences by allowing for user-defined mental commands. In: Conference Abstract: German-Japanese Adaptive BCI Workshopp (2015). Front. Comput. Neurosci
15. Duvinage, M., Castermans, T., Petieau, M., Hoellinger, T., Cheron, G., Dutoit, T.: Performance of the Emotiv Epoc headset for P300-based applications. Biomed. Eng. Online **12**, 56 (2013)
16. Emotiv Systems. Emotiv - brain computer interface technology, May 2011. http://www.emotiv.com
17. Friedrich, E.V., Scherer, R., Neuper, C.: The effect of distinct mental strategies on classification performance for brain-computer interfaces. Int. J. Psychophysiol. **84**(1), 86–94 (2012)
18. Hammer, E.M., Halder, S., Blankertz, B., Sannelli, C., Dickhaus, T., Kleih, S., Müller, K.R., Kübler, A.: Psychological predictors of SMR-BCI performance. Biol. Psychol. **89**(1), 80–86 (2012)
19. Jeunet, C., NKaoua, B., Subramanian, S., Hachet, M., Lotte, F.: Predicting mental imagery-based BCI performance from personality, cognitive profile and neurophysiological patterns. PloS One **10**(12), e0143962 (2015)
20. Kindermans, P.-J., Verschore, H., Verstraeten, D., Schrauwen, B.: A p300 BCI for the masses: prior information enables instant unsupervised spelling. In: Advances in Neural Information Processing Systems, pp. 710–718 (2012)
21. Kothe, C.A., Makeig, S., Onton, J.A.: Emotion recognition from EEG during self-paced emotional imagery. In: Proceedings - 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction, ACII, pp. 855–858 (2013)
22. Kübler, A., Müller, K.R.: An introduction to brain computer interfacing. In: Dornhege, G., del Millán, J.R., Hinterberger, T., McFarland, D., Müller, K.R. (eds.) Toward Brain-Computer Interfacing. MIT Press, Cambridge (2007)
23. Kübler, A., Neumann, N., Kaiser, J., Kotchoubey, B., Hinterberger, T., Birbaumer, N.P.: Brain-computer communication: self-regulation of slow cortical potentials for verbal communication. Arch. Phys. Med. Rehabil. **82**(11), 1533–1539 (2001)
24. Kus, R., Valbuena, D., Zygierewicz, J., Malechka, T., Graeser, A., Durka, P.: Asynchronous BCI based on motor imagery with automated calibration and neurofeedback training. IEEE Trans. Neural Syst. Rehabil. Eng.: Publ. IEEE Eng. Med. Biol. Soc. **20**(6), 823–835 (2012)
25. Lievesley, R., Wozencroft, M., Ewins, D., Lievesley, M., Wozencroft, R.: The Emotiv EPOC neuroheadset: an inexpensive method of controlling assistive technologies using facial expressions and thoughts? J. Assist. Technol. **5**(2), 67–82 (2011)
26. Lin, C.J., Weng, R.C., et al.: Simple Probabilistic Predictions for Support Vector Regression. National Taiwan University, Taipei (2004)
27. Lin, H.-T., Lin, C.-J., Weng, R.C.: A note on Platts probabilistic outputs for support vector machines. Mach. Learn. **68**(3), 267–276 (2007)
28. Liu, Y., Jiang, X., Cao, T., Wan, F., Mak, P.U., Mak, P.I., Vai, M.I.: Implementation of SSVEP based BCI with Emotiv EPOC. In: Proceedings of IEEE International Conference on Virtual Environments, Human-Computer Interfaces, and Measurement Systems, VECIMS, pp. 34–37 (2012)
29. Lotte, F., Larrue, F., Mühl, C.: Flaws in current human training protocols for spontaneous Brain-Computer Interfaces: lessons learned from instructional design. Front. Hum. Neurosci. **7**(September), 568 (2013)

30. Mak, J.N., Arbel, Y., Minett, J.W., McCane, L.M., Yuksel, B., Ryan, D., Thompson, D., Bianchi, L., Erdogmus, D.: Optimizing the P300-based brain-computer interface: current status, limitations and future directions. J. Neural Eng. **8**(2), 025003 (2011)
31. Mason, S.G., Birch, G.E.: A brain-controlled switch for asynchronous control applications. IEEE Trans. Biomed. Eng. **47**(10), 1297–1307 (2000)
32. MATLAB. Version 8.2.0 (R2013b). The MathWorks Inc., Natick, Massachusetts (2013)
33. McFarland, D.J., Miner, L.A., Vaughan, T.M., Wolpaw, J.R.: Mu and beta rhythm topographies during motor imagery and actual movements. Brain Topogr. **12**(3), 177–186 (2000)
34. Müller-Gerking, J., Pfurtscheller, G., Flyvbjerg, H.: Designing optimal spatial filters for single-trial EEG classification in a movement task. Clin. Neurophysiol. **110**(5), 787–798 (1999)
35. Neuper, C., Pfurtscheller, G.: Neurofeedback training for BCI control. In: Graimann, B., Pfurtscheller, G., Allison, B. (eds.) Brain-Computer Interfaces, pp. 65–78. Springer, Heidelberg (2010)
36. Neuper, C., Scherer, R., Reiner, M., Pfurtscheller, G.: Imagery of motor actions: differential effects of kinesthetic and visual-motor mode of imagery in single-trial EEG. Cogn. Brain. Res. **25**(3), 668–677 (2005)
37. Nicolas-Alonso, L.F., Gomez-Gil, J.: Brain computer interfaces, a review. Sensors **12**(2), 1211–1279 (2012)
38. Oostenveld, R., Fries, P., Maris, E., Schoffelen, J.M.: FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput. Intell. Neurosci. **2011**, 1–9 (2011)
39. Peng, H.C.: Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. IEEE Trans. Pattern Anal. Mach. Intell. **27**, 1226–1238 (2005)
40. Pfurtscheller, G., Neuper, C., Flotzinger, D., Pregenzer, M.: EEG-based discrimination between imagination of right and left hand movement. Electroencephalogr. Clin. Neurophysiol. **103**(6), 642–651 (1997)
41. Pfurtscheller, G., Lopes da Silva, F.H.: EEG event-related desynchronization (ERD), event-related synchronization (ERS). Electroencephalogr.: Basic Princ. Clin. Appl. Relat. Fields **4**, 958–967 (1999)
42. Pfurtscheller, G., Neuper, C.: Motor imagery activates primary sensorimotor area in humans. Neurosci. Lett. **239**(2), 65–68 (1997)
43. Platt, J., et al.: Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. Adv. Large Margin Classif. **10**(3), 61–74 (1999)
44. Ramoser, H., Müller-Gerking, J., Pfurtscheller, G.: Optimal spatial filtering of single trial EEG during imagined hand movement. IEEE Trans. Rehabil. Eng. **8**(4), 441–446 (2000)
45. Randolph, A.B.: Not all created equal: individual-technology fit of brain-computer interfaces. In: Proceedings of the Annual Hawaii International Conference on System Sciences, pp. 572–578 (2011)
46. Randolph, A.B., Jackson, M.M., Karmakar, S.: Individual characteristics and their effect on predicting mu rhythm modulation. Int. J. Hum.-Comput. Interact. **27**(1), 24–37 (2010)
47. Scherer, R., Faller, J., Friedrich, E.V., Opisso, E., Costa, U., Kübler, A., Müller-Putz, G.R.: Individually adapted imagery improves brain-computer interface performance in end-users with disability. PloS One **10**(5), e0123727 (2015)

48. Stinear, C.M., Byblow, W.D., Steyvers, M., Levin, O., Swinnen, S.P.: Kinesthetic, but not visual, motor imagery modulates corticomotor excitability. Exp. Brain Res. **168**(1–2), 157–164 (2006)
49. Thomas, E., Dyson, M., Clerc, M.: An analysis of performance evaluation for motor-imagery based BCI. J. Neural Eng. **10**(3), 031001 (2013)
50. Vidaurre, C., Sannelli, C., Müller, K.R., Blankertz, B.: Co-adaptive calibration to improve BCI efficiency. J. Neural Eng. **8**(2), 025009 (2011)
51. Vidaurre, C., Sannelli, C., Müller, K.-R., Blankertz, B.: Machine-learning-based coadaptive calibration for brain-computer interfaces. Neural Comput. **23**(3), 791–816 (2011)
52. Vuckovic, A., Osuagwu, B.A.: Using a motor imagery questionnaire to estimate the performance of a brain-computer interface based on object oriented motor imagery. Clin. Neurophysiol. **124**(8), 1586–1595 (2013)
53. Wolpaw, J.R., Birbaumer, N., McFarland, D.J., Pfurtscheller, G., Vaughan, T.M.: Brain-computer interfaces for communication and control. Clin. Neurophysiol. Official J. Int. Fed. Clin. Neurophysiol. **113**(6), 767–791 (2002)
54. Ting-Fan, W., Lin, C.-J., Weng, R.C.: Probability estimates for multi-class classification by pairwise coupling. J. Mach. Learn. Res. **5**(Aug), 975–1005 (2004)