Proceedings of the
8th Graz Brain-Computer Interface Conference 2019 DOI: 10.3217/978-3-85125-682-6-47

FLIP-THAT-BUCKET
A FUN EEG-BCI GAME ON GOOEY MOVEMENT INTENTIONS

C.S. Verbaarschot[1], A.B.W. Gerrits[1], W.F.G. Haselager[1], J.D.R. Farquhar[1]

[1] Centre for Cognition, Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, the Netherlands

E-mail: c.verbaarschot@donders.ru.nl

ABSTRACT: "Flip-that-Bucket" is an open source, portable and enjoyable BCI game suitable for investigating or demonstrating movement intentions using scientific experiments or educational demonstrations in noisy environments. In the game, a sneaky virtual robot aims to predict a player's intentions to act based on their action history, muscle activity, or brain activity. The game can be used to assess (1) the accuracy of brain-based movement predictions, (2) the timing of these predictions relative to movement onset, (3) the potential benefit of brain-based over behavior-based predictions, and (4) the correlation between certain brain signals (e.g. the readiness potential and event-related desynchronization between 8-30Hz over the pre- and primary motor areas) and the experience of an intention to move. Answering each of these questions may greatly benefit future applications in prosthetics and motor rehabilitation. Flip-that-Bucket is made as an extension to the open-source buffer_bci toolbox, encouraging further development. Here, we demonstrate the idea, its practical implementation, the 'fun factor' and a first analysis of experimental results.

INTRODUCTION

When we move, imagine movement or observe movement, specific parts of our brain activate: the premotor, supplementary and primary motor cortices (termed 'motor cortex' in the remainder of this paper). When we perform self-paced voluntary movements, we typically see a readiness potential (RP) and event-related desynchronization (ERD) at 8-30Hz across the motor cortex [3, 5, 6, 8, 11]. Interestingly, these brain signals are visible on average around 1.5s before a person reports a conscious feeling of wanting to move. This suggests that the brain starts preparing a movement before a person reports a conscious intention to move [7]. Other studies show that using a real-time probing method, awareness of an intention to move can be reported up to 2s prior to movement onset [12]. How exactly a conscious intention to move relates to these neural signals remains unclear.
In this paper, we present "Flip-that-Bucket": a BCI game that serves as a tool to investigate or demonstrate the relation between the neural preparation and reported awareness of movement intentions (see Figure 1). In contrast to previous research [1, 2, 3, 4, 5, 6, 7, 8], the game provides an engaging real-time set-up to measure spontaneous self-paced right hand movements in an intuitive way. In the game, players try to beat a virtual robot opponent in a slime-bucket challenge. Across multiple rounds, both the participant and robot try to gather as much green slime as possible to throw it over their opponent's head. Both the player and robot have

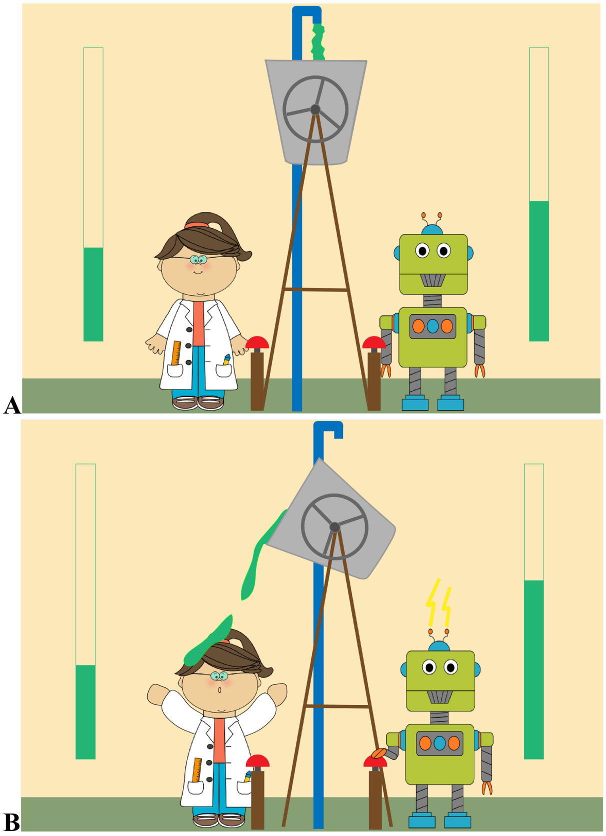

Figure 1. Screenshots of Flip-that-Bucket. The game consists of several rounds during which both the player (the scientist) and the robot can gather as much gooey green slime as possible. At the start of each round, an empty bucket is displayed for 2s. (A) This bucket fills with green slime over time, although the exact content of the bucket remains hidden. (B) The player and robot can flip over the bucket and empty its contents onto their opponent any time they want. Whoever flips the bucket first, will get points for the amount of slime they threw over their opponent. Whoever has been slimed the least wins the game.

 DOI: 10.3217/978-3-85125-682-6-47

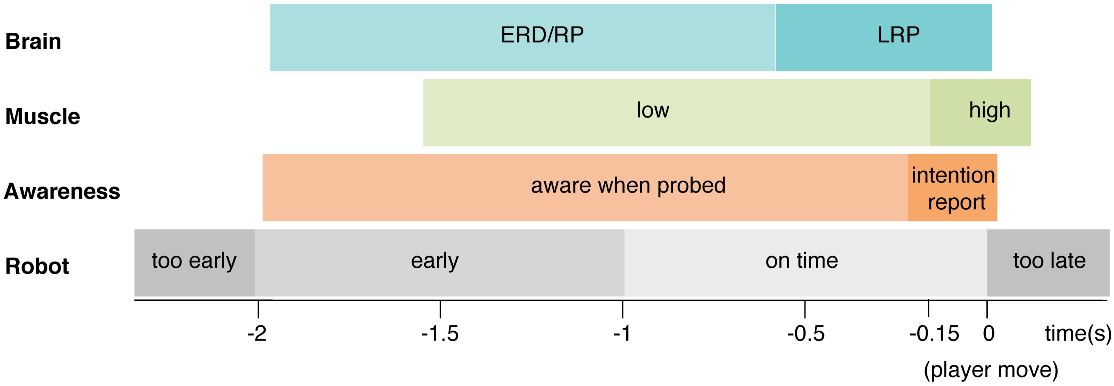

Figure 2. Schematic visualization of the possible moments in time at which a robot could predict a player's intention to move (movement onset is at time 0s). We distinguish four possible categories of robot predictions. (1) Too early: no visible ERD/RP, no visible muscle activity, only negative intention reports. (2) Early: no or weak ERD/RP, no visible muscle activity, more negative/unknown intention reports than positive ones. (3) On time: medium to strong ERD/RP (possibly including an LRP), low to high muscle activity, more positive intention reports than negative/unknown ones. (4) Too late: strong (lateralized) ERD/RP, high muscle activity, no intention reports collected.

access to the same bucket, which can flip over only once each round to spill out its gooey content. The robot is sneaky: he tries to predict a player's intention to move using their action history, muscle activity, or brain activity. As soon as the robot detects a player's intention to move, he will flip over the bucket of slime. Because the robot uses different types of predictions (based on action history, muscle or brain activity), the game serves as a thought-provoking means to explain to a general public how movement intent travels from the brain to the muscles in order to perform a voluntary movement.

In a first experiment, we use robot predictions based on action history. We collect EEG data to assess whether there is a strong correlation between the RP and/or ERD and the awareness of an intention. To do so, participants are asked to report whether they experienced an intention to move at the moment when a prediction is made (i.e. when the robot moves). This ante-hoc probing strategy measures awareness of movement intent prior to movement performance, which avoids the potential confound of movement execution on post-hoc awareness reports [4]. We expect to see a clear RP and ERD prior to a movement intent prediction that is reported as "intended" and no (or a weak) RP and ERD prior to a movement intent prediction that was reported as "unintended" (see Figure 2).

Flip-that-Bucket is implemented as an extension to an existing open-source BCI development toolbox called "buffer_bci",[1] encouraging further development of our project ideas by both academics and the general developer community. Anyone with access to an EEG system can try the game. It can be used at home, during public events or in scientific experiments.

MATERIALS AND METHODS

Participants: 41 healthy volunteers were tested at the InScience festival in Nijmegen, the Netherlands.[2] The experiment was conducted in accordance with the ethical standards provided by the 1964 Declaration of Helsinki. The study protocol was approved by the local Ethics Committee of Faculty of Social Sciences of the Radboud University Nijmegen. All participants gave their written informed consent. Ten participants were excluded from analysis because they did not follow instructions correctly or would not stop talking or moving during the experiment.

Task: Participants play 4 blocks of Flip-that-Bucket[3] (see Figure 1): a practice block of 3 trials, a training block of 60 trials, a hidden validation block of 15 trials and a test block of 60 trials. This experimental structure is used to accommodate future experimentation including brain-based robot predictions. In between the training and validation block is a self-paced break and a short questionnaire. Participants are informed that we are testing a new BCI that attempts to predict the moment at which they intend to move and asked to report whether or not they wanted to move when a prediction (i.e. robot move) was made. During the training block, predictions of movement intent are made based on the participant's action history. Based on [12], we expected awareness of intending no earlier than 2s prior to action onset. Each trial, a minimum cost function takes the current mean and standard deviation of the player's action times (relative to trial start) and calculates a distribution of possible robot action times such that: 1/5 of robot acts are performed earlier than 2s prior to the average scientist move, 3/5 of robot acts between 2s and 0s prior to the average scientist move, and 1/5 of robot acts are performed after the scientist moves (i.e. the robot loses). The robot action time for a

[1] www.github.com/jadref/buffer_bci

[2] www.insciencefestival.nl

[3] Flip-that-Bucket can be found at: www.github.com/jadref/buffer_bci/tree/master/matlab/movementBCIgame

 DOI: 10.3217/978-3-85125-682-6-47

given trial was drawn randomly from this distribution. After the training block, a combined features classifier (incorporating both RP and 8-30Hz ERD features of pre-move and non-move data) was trained on the collected labeled EEG data and used to provide brain-based predictions during the test block. Unfortunately, due to technical errors (i.e. accidentally including post-move data in the training set and switching class labels during real-time prediction) the brain-based predictions were not executed properly and the predictions made in the second block were roughly random.
At the end of the train and test block, participants fill in a short questionnaire asking them (1) what they thought about the game on a scale of 1 (boring) to 5 (fun), (2) whether they felt free to do what they want (Yes/No), (3) how difficult it was to win on a scale of 1 (easy) to 5 (difficult), and (4) how accurate the robot was in predicting their actions on a scale of 1 (inaccurate) to 5 (accurate). At the end of the test block, they were asked how good they thought the robot predictions were in the second block compared to the first on a scale of 1 (worse) to 5 (better). For additional motivation, a high-score list across all players and robots is maintained. In total, the experiment took 24 minutes (excluding cap fitting).

Data acquisition: The experiment was run in Matlab.[4] Instructions and visual stimuli were displayed using a 17 inch TFT screen with a resolution of 800 by 600 pixels and a refresh rate of 60Hz that was placed roughly at 70cm directly in front of the participant. To flip the bucket, participants press SPACE with their right hand on a regular keyboard. EEG data was collected using the TMSi Porti system,[5] with water-based electrodes sampled at 512 Hz placed at Fp1, Fp2 F3, Fz, F4, C4, Cz, C4, P3, Pz, P4, POz, TP9 and TP10 (according to the International 10/20 system). In addition, muscle activity was recorded using two EEG electrodes in a bipolar pair on the wrist and the center of the right forearm (flexor pollicis longus).

Analysis: All analyses were performed on the 31 participants who followed instructions correctly. To assess the correlation between the RP and/or ERD and a reported intention to move, epochs of -15 to 15s around a player and robot act are analyzed. Epochs around a robot act are further subdivided in epochs in which the robot acted when the participant did or did not want to move. To ensure a decent baseline period, only epochs in which the player or robot acted slower than 4s after trial start are kept for analysis. Linear trends are removed from the data. Subsequently, the data is re-referenced by subtracting the average signal from all outer channels (Fp1, Fp2, F3, P3, POz, F4, P4, TP9, TP10) from each individual channel. This was done to subtract as much noise as possible without subtracting the signals of interest. Since most recorded channels cover the motor cortex, a full common average reference would subtract much of the signal of interest along with the noise (leading to a decrease in RP amplitude). Eye-artifacts are removed using linear de-correlation of channels Fp1 and Fp2 with respect to the other EEG channels. Only the central channels (F3, C3, P3, Pz, Fz, F4, Cz, C4 and P4) are kept for further analysis. Channels that differ more than 2 times the standard deviation in power from the median are removed. If necessary, spherical spline interpolation is used to reconstruct missing central channels. The data is band-pass filtered between 0.2 and 35Hz. Bad trials that differ more than two times the standard deviation in power from the median are removed. For the ERD, frequencies of interest are defined from 4 to 30Hz using 2Hz bins. A flexible Hanning window is used such that it includes at least 7 cycles of each frequency of interest. The baseline activity is defined per electrode, frequency and trial as the median power between 3.5 and 2.5s prior to action. A relative baseline (where a value of 1 means no signal change compared to baseline) is subtracted from the data. The ERD is calculated per participant by taking the median power across trials for each electrode, frequency and trial.

Offline classification: For each participant, a linear classifier is trained using 10-fold cross-validation to distinguish baseline data from pre-movement data. Baseline data is extracted from the last 500ms of the baseline period. Pre-movement data is extracted from the last 500ms prior to a player move. All training data is extracted from the training block. The data is pre-processed using the same steps as described in the *Analysis* section. The RP features consist of 257 time points for each channel and epoch. The ERD features consist of the average power of each of the 14 frequencies (4,6,…,12,14Hz) for each channel and epoch. The classifier is trained on both RP and ERD features, giving a total of 2439 features per epoch. Data from the validation block is used to set an optimal threshold for classification. This threshold is set such that the number of *on time* (between 1 and 0s prior to a player move) predictions is maximized whereas the number of *too early* (more than 2s prior to a player move) predictions is minimized (see Figure 2). The accuracy of the classifier is assessed using data from the

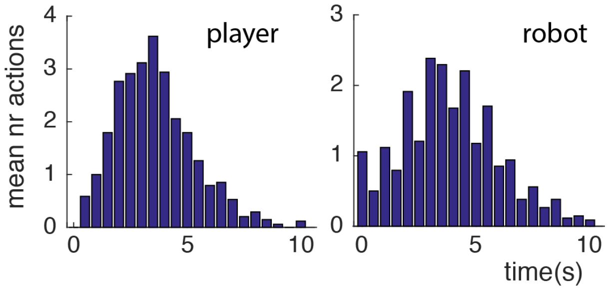

Figure 3. Distribution of player and robot moves relative to trial start (0s) across all participants during the training block.

[4] www.mathworks.com
[5] www.tmsi.com/products/porti/

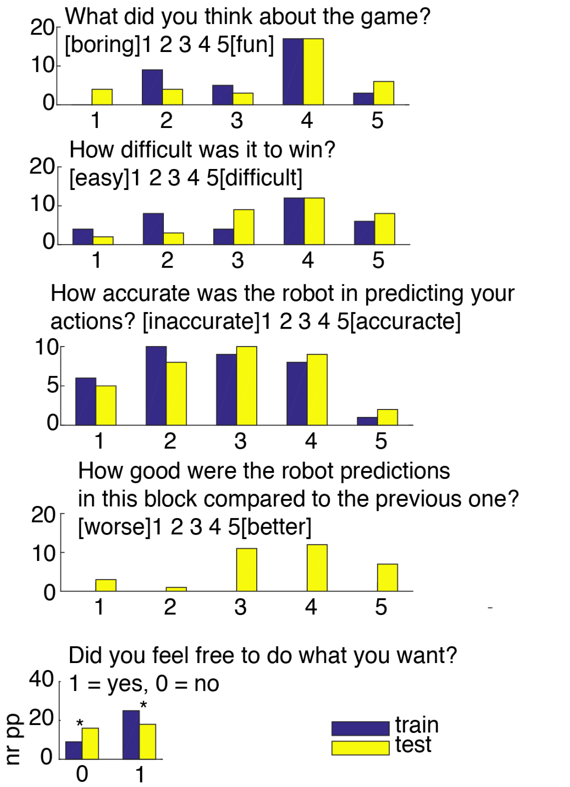

Figure 4. Questionnaire results. Participants felt significantly less free to do what they wanted during the test compared to the train block. In question 4, "this block" refers to the test block and "the previous one" refers to the training block.

test block. The trained classifier is applied to non-overlapping 500ms epochs starting from the start of a trial until a player or robot move. For each trial that includes an intention to move (the player moved or reported an intention to move at the time of a robot move), the first classifier prediction that exceeds the optimal threshold for motor intention detection is selected. The timing of each first motor intention prediction is determined relative to the corresponding robot or player move and categorized as *too early*, *early*, *on time* or *too late* (see Figure 2).

RESULTS

An average of 69 (min:42, max:104) player moves, 20 (min:5, max:54) robot moves with player intent and 33 (min:9, max:67) robot moves without player intent are collected across the experiment. An overview of player and robot move times during training is provided in Figure 3.
The majority of participants reported the game as fun and slightly difficult to win (see Figure 4). Although opinions on the accuracy of the robot predictions varied greatly, participants judged the robot predictions of the second experimental block as more accurate than the first (even though these predictions were roughly random). Furthermore, a within-subject t-test on the questionnaire data of the first en second experimental blocks revealed that participants felt significantly ($p < .05$) less free to act during the second block of the experiment compared to the first. No significant differences were found between the number of robot actions that happened at a time when the player did or did not intend to move during the first and second experimental blocks.

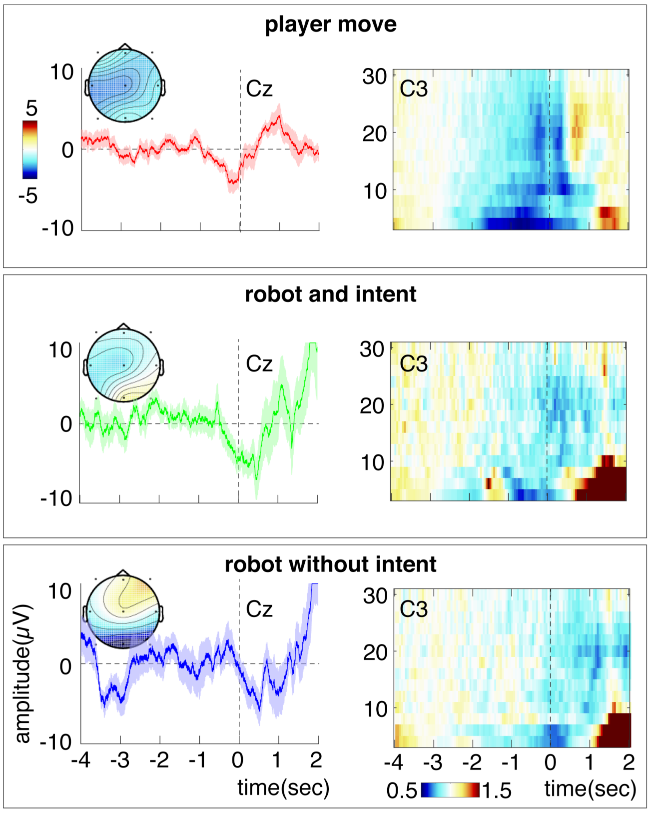

Figure 5. Grand average EEG data of Flip-that-Bucket. Left column: grand average ERPs including the standard error, and topoplots across the last 500ms prior to a player/robot move until action onset (0s). Right column: grand average spectrogram across 4 to 30Hz before player/robot action onset (at 0s). Baseline activity between 3.5 and 2.5s prior to action onset is subtracted from the ERD data. Although a clear RP and ERD is visible prior to a player move (top row), no RP and or ERD is visible prior to a robot act that happened when the player did not intend to move (bottom row). A weak RP and ERD is visible when the player did intend to move when the robot acted (middle row).

A clear RP and 8-30Hz ERD are visible in the grand average prior to a player move, starting 1 to 2s prior to movement onset respectively (see Figure 5). A weak RP and ERD is visible prior to a robot act that happened at a time when the player intended to move (i.e. a correct prediction), and no RP or ERD is visible prior to a robot act that happened when the player did not intend to move (i.e. an incorrect prediction). After a robot acts, a big positive response is visible in the recorded EEG. This response is due to the additional button press that is required to report whether the robot moved at a time when the player wanted to move. Any differences

 DOI: 10.3217/978-3-85125-682-6-47

between the post robot move responses may be due to the presence or absence of an intention to move, leading to an superimposed error potential in one case rather than the other. Moreover, a player may show different levels of surprise or frustration after the robot moved, depending on the presence or absence of an intention to move.

Classifiers are trained on an average of 48 (min:44, max:50) pre-move and 47 (min:42, max:50) baseline epochs from the training block. A mean classifier performance of 74% across all participants is found on the training data. Although these results seem promising, the performance of the classifier on the sequential test data are rather poor. On average a mere 12% of all motor intentions is detected *on time* (within 1s prior to a player or robot move). The majority of motor intentions is detected *too early* (24%) or *too late* (54%).

PP	**Perf.**	**Too early**	**Early**	**On time**	**Too late**
1	76	7	0	0	93
2	92	0	8	13	79
3	64	6	8	8	78
4	90	17	0	12	71
5	78	15	6	17	62
6	90	18	0	20	63
7	79	45	15	10	30
8	61	64	31	6	0
9	72	21	38	29	13
10	60	41	14	5	41
11	69	0	0	0	100
12	83	56	12	9	23
13	88	4	0	2	93
14	66	29	3	12	56
15	60	28	26	16	30
16	74	70	17	4	9
17	75	85	5	5	5
18	61	21	3	18	59
19	65	2	2	2	93
20	84	0	0	0	100
21	76	21	50	25	4
22	59	35	13	15	37
23	98	19	2	7	71
24	71	17	22	26	35
25	68	0	0	0	100
26	80	24	10	19	48
27	68	6	2	10	82
28	60	32	6	3	58
29	82	17	9	20	54
30	70	23	12	37	28
31	65	35	7	11	48
mean	**74**	**24**	**10**	**12**	**54**

Table 1. Offline classification results. For each participant the cross-validated test performance of the classifier on the data of the training block is provided. Furthermore, for each trial in the test block that includes a motor intention, the first classifier prediction that exceeds the chosen threshold for motor intent is determined. These predictions are categorized as too early, early, on time or too late relative to the robot or player move. The percentage of predictions within each category compared to the total number of motor intentions present is provided in the remaining columns.

DISCUSSION

Moving your body at will seems trivially easy. You probably do it all the time. You can knock on a table, stomp with your feet and flap your arms whenever you want to. But what mechanism enables you to initiate these movements? Moreover, how does this mechanism relate to your conscious experience of wanting to move? Flip-that-Bucket is a fun and thought-provoking EEG-BCI game on movement intentions. In the game, a virtual robot tries to predict a player's intention to move using the player's history of action times, the onset of muscle activity or the neural preparation for movement. Flip-that-Bucket can serve as a means to educate a general public about the neural mechanisms that underlie our ability to perform voluntary movements: the connection between a participant's brain signals and their voluntary movements are made explicit to them by means of the robot opponent.

Here, we demonstrate that predictions based on action history are effective to create a competitive game. Furthermore, we demonstrate that both the RP and 8-30Hz ERD are clearly visible across the motor cortex prior to a player move, whereas it is not visible prior to an incorrect robot prediction (i.e. the robot acted when the participant did not experience an intention to move). In case of a correct robot prediction (i.e. the robot acts when the participant experienced an intention to move), a weak RP and 8-30Hz ERD are visible across the motor cortex. Along with previous research [1, 2, 9, 10], these results suggest that brain-based predictions of movement intent may be reasonably successful. If this is the case, brain-based predictions may be used to assess the relation between the RP and ERD in real-time in further experiments. It would be interesting to see whether brain-based predictions are more accurate compared to behavior-based predictions (i.e. the robot predictions are more often reported to be correct when using brain-based rather than behavior-based data).

Flip-that-Bucket can also facilitate scientific investigations on the neural preparation for a voluntary movement. Although some studies aimed to predict movement onset in real-time prior to movement performance based on the RP and 8-30Hz ERD across the motor cortex, e.g. [1, 2, 9, 10], the accuracy and timing of single-trial predictions remains difficult to assess since often only averages across participants are reported. Flip-that-Bucket can be used to assess the accuracy and timing of brain-based predictions on single-trial continuous data. A correct prediction of movement intent would happen at a time when a participant reports that they want to move, their muscles are active in preparation of the upcoming movement or when it coincides with movement performance. Furthermore, predictions that happen very early (more than 2s) prior to movement onset would be considered incorrect (false positive), whereas those happening shortly (within 1s) prior to movement onset would be correct (true positive). A first offline analysis of these prediction results suggests that our combined features

 DOI: 10.3217/978-3-85125-682-6-47

classifier would perform rather poorly, detecting only about 12% of all motor intentions. The current discrepancy between the performance on the training and test data, may be due to a potential expectation effect that builds up during a trial: the longer a trial develops, the more likely a player or the robot is to act. This expectation effect may induce additional brain responses that are maximally different between baseline and pre-move training data, but less so between subsequent epochs of test data. Based on previous research [1, 2, 9, 10], we expect that better results should be possible. Possible improvements could be made by (1) including more instances of "non-move" data during classifier training, (2) implementing a more sophisticated method to set an optimal threshold for detecting a motor intention or (3) using different brain signals that are indicative of movement preparation.

Assessing the accuracy and timing of real-time predictions of movement intent on continuous data may greatly benefit future applications in prosthetics and motor rehabilitation. Delays in activating an assistive device could be minimized by detecting movement intent early on, which potentially increases the therapeutic benefit by minimizing the time between motor planning in the cortex and the execution of that plan with the assistive device.

Flip-that-Bucket is made as an extension to the existing open-source buffer_bci development toolbox, encouraging further development of the game by both academics and the general developer community. This extended toolbox is available to anyone with access to an EEG system. It can be used at home, during public events or in scientific experiments.

REFERENCES

[1] Bai, O., Rathi, V., Lin, P., Huang, D., Battapady, H., Fei, D. Y., ... & Hallett, M. (2011). Prediction of human voluntary movement before it occurs. *Clinical Neurophysiology*, *122*(2), 364-372.

[2] Blankertz, B., Dornhege, G., Lemm, S., Krauledat, M., Curio, G., & Müller, K. R. (2006). The Berlin brain-computer interface: Machine learning based detection of user specific brain states. *J. UCS*, *12*(6), 581-607.

[3] Fried, I., Mukamel, R., & Kreiman, G. (2011). Internally generated preactivation of single neurons in human medial frontal cortex predicts volition. *Neuron*, *69*(3), 548-562.

[4] Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nature neuroscience*, *5*(4), 382.

[5] Kornhuber, H. H., & Deecke, L. (2016). Brain potential changes in voluntary and passive movements in humans: readiness potential and reafferent potentials. *Pflügers Archiv-European Journal of Physiology*, *468*(7), 1115-1124.

[6] Lew, E., Chavarriaga, R., Silvoni, S., & Millán, J. D. R. (2012). Detection of self-paced reaching movement intention from EEG signals. *Frontiers in neuroengineering*, *5*, 13.

[7] Libet, B., Gleason, C. A., Wright, E. W., & Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential) the unconscious initiation of a freely voluntary act. *Brain*, *106*(3), 623-642.

[8]Pfurtscheller, G., & Aranibar, A. (1979). Evaluation of event-related desynchronization (ERD) preceding and following voluntary self-paced movement. *Electroencephalography and clinical neurophysiology*, *46*(2), 138-146.

[9] Schneider, L., Houdayer, E., Bai, O., & Hallett, M. (2013). What we think before a voluntary movement. *Journal of cognitive neuroscience*, *25*(6), 822-829.

[10]Schultze-Kraft, M., Birman, D., Rusconi, M., Allefeld, C., Görgen, K., Dähne, S., ... & Haynes, J. D. (2016). The point of no return in vetoing self-initiated movements. *Proceedings of the National Academy of Sciences*, *113*(4), 1080-1085.

[11]Shibasaki, H., & Hallett, M. (2006). What is the Bereitschaftspotential?. *Clinical neurophysiology*, *117*(11), 2341-2356.

[12]Verbaarschot, C., Haselager, P., & Farquhar, J. (2016). Detecting traces of consciousness in the process of intending to act. *Experimental brain research*, *234*(7), 1945-1956.