Donders Graduate School for Cognitive Neuroscience
Master of Science Programme
MSc Thesis

Predicting movement intent in real-time: From brain to subjective experience

by

Anne Gerrits

Supervised by:

1. Dr. J. Farquhar
2. C. Verbaarschot, Msc

Radboud University Nijmegen

Date of final oral examination: 20th February 2017

Student name: Anne Gerrits
Student number: s4465652

Abstract

The readiness potential (RP) and the event-related desynchronization (ERD) are neural signals that build up over the motor cortex 1.5-2 seconds prior to movement onset. Bai et al. (2011) were amongst the first to reliably detect movement intent online based on these signals. Interestingly, these brain signals typically build up prior to the moment a person reports to consciously intend to act. However, how these subjective reports relate to these neural preparatory signals remains unclear. To investigate this, we developed a brain-computer interface (BCI), based on the Bai study, that predicts movement intent based on these brain signals and then feeds this prediction back by means of functional electrical stimulation (FES). Three experiments were conducted. In the first experiment we successfully replicated the Bai study offline. We found we could predict movement intent offline based on the ERD (-0.7±0.17s) and the RP (-0.43±0.84s) before movement onset. In the second experiment we investigated the effect of FES stimulation on EEG data. We showed FES stimulation mostly influences the EEG data on and after movement onset and was thus not an issue for our study. In our third experiment we used online classification to investigate if a person is aware of their intention to act when movement preparation is detected in the brain. The online classification did not work as expected due to a high false positive rate. Therefore, we could not answer the main question in this experiment. We believe the online classification was affected by an anticipation buildup over time. By using more time points for classifier training, building in trials that provide a measure of anticipation alone and creating more variance in action timing, we believe it will be possible to predict movement intent in real-time and investigate how the subjective experience of intending to act relate to the RP/ERD.

1. Introduction

On a daily basis, humans can spontaneously decide when to initiate a movement and when to commit to a certain course of action to accomplish a task (Haggard, 2008; Lew, Chavarriaga, Silvoni, & Millán, 2012). In a recent review article, Shibasaki (2012) described the main electrophysiological activities associated with self-paced voluntary movements, highlighting the readiness potential (RP) and the event-related desynchronization (ERD) over the (pre) motor cortex as factors commonly associated with movement preparation.

The RP is a slow negative cortical potential that is known to build up over the motor cortex as early as 2 seconds prior to movement onset. It can be divided into two components: the early RP, characterized by a slow negative slope, starting from 2 seconds before voluntary movement onset in the pre-supplementary motor area and the late RP, characterized by a steeper negative slope lateralized over the primary motor area which starts around 400 milliseconds before voluntary movement (Kornhuber & Deecke, 1965; Shibasaki & Hallett, 2006).

The ERD in the alpha (8-12 Hz) and beta (13-30 Hz) frequency bands also occurs around 1.5 - 2 seconds prior to movement onset and starts bilaterally over the motor cortex (Pfurtscheller & Aranibar, 1979). The beta ERD has been found to be contralateral prior to dominant hand movement as opposed to bilateral prior to non-dominant hand movement (Bai, Mari, Vorbach, & Hallett, 2005). These neural signals are typically not seen before movement onset when the movement was passive or involuntary (Müller et al., 2003; Shibasaki & Hallett, 2006).

Both the RP and ERD have been used to predict movement onset on a single trial level. Bai et al. (2011) were amongst the first to reliably detect the intention to move before movement onset based on EEG signals in real-time. They demonstrated that voluntary movement could be predicted on average 620 ± 250 milliseconds before its onset. Schneider, Houdayer, Bai and Hallett (2013), using the same methods as Bai et al. (2011), also showed they could predict movement intent before its onset in real-time with a low false positive rate. Based on these and other studies (Blankertz et al., 2006; Fried, Mukamel, & Kreiman, 2011; Lew et al., 2012) it should be possible to predict the intention to move based on the RP and alpha/beta ERD in real-time before movement onset. If prediction in real-time is possible, it provides a means to feed this prediction back to a person in order to investigate their subjective experience of intending to act at that moment.

Interestingly, these brain signals typically build up prior to the moment a person reports to consciously intend to act, suggesting the brain starts preparing an act before a person is aware of their intention (Libet, Gleason, Wright, & Pearl, 1983). However, how these subjective reports relate to these neural preparatory signals remains unclear.

In the current study we developed a brain-computer interface (BCI), based on the methods used by Bai et al. (2011), that tries to predict when someone has the intention to move based on the RP and/or ERD and then feeds this prediction back to the person by means of functional electrical stimulation (FES). We conducted three experiments. In our first experiment we replicated Bai et al. (2011) offline. In experiment 2 we investigated the effects of FES stimulation on EEG signals. Finally, in our third experiment we used online classification to investigate how the subjective reports of intending to act relate to the RP and/or ERD.

A BCI is a real-time system that can translate brain activity into control signals of external devices (Soekadar, Birbaumer, Slutzky, & Cohen, 2015). Neurophysiological signals from the brain are measured and used to make a direct online connection between the brain and a device such as a computer or a prosthetic device (van Gerven et al., 2009). This system can be used to replace, restore, enhance or improve neurophysiological activity in the brain (Wolpaw & Wolpaw, 2012).

Functional electrical stimulation (FES) is a technique that can be used to artificially activate the sensorimotor system by sending short pulses of electrical charge to the muscles (Popović, 2014). By placing electrodes near the motor point of the muscle and by applying short, constant-current pulses, the potential of the nerve is depolarized which leads to a contraction of the innervated muscle fibers. This suggests FES could potentially be used to feed the predictions back to the participant by performing a forced movement. By carefully selecting the right individual stimulation current and position of the FES electrodes, FES movements can be made very similar to a voluntary movement.

With this setup, we investigated if a person is aware of their intention to act when movement preparation is detected in the brain. To study this, participants were asked to report on their subjective experience after each movement. A movement could be initiated by the participant's muscle activity, by their brain (classifier) or randomly. Every type of movement ended with FES stimulation, making it ambiguous by what source the movement was initiated. This setup provided us with a way to compare the ERD and RP signals prior to movements that were reported as intended, unintended and movements they were not sure of.

Based on previous literature (Bai et al., 2011; Fried et al., 2011; Lew et al., 2012; Schneider et al., 2013a), we hypothesized that we can reliably predict movement onset in real-time based on the neural preparatory activity over the motor cortex. The RP is generally only visible after averaging over 40–50 voluntary movements. In contrast to the RP, the ERD can sometimes be observed on a single-trial level (Bai et al., 2011). Hence, we expected the ERD to be a better feature for online prediction as the RP is more difficult to detect on a single trial level.

Since the RP/ERD signals are only seen before the onset of voluntary movements (Müller et al., 2003; Shibasaki & Hallett, 2006) we expected these signals to be significantly different prior to a movement that was reported to be unintended compared to one that was reported as intended. Specifically, we expected only ERD or RP activity present before movement onset when a movement was reported as intended. No ERD or RP signal was expected to be present before a movement that was reported as unintended.

Next to providing us with a better understanding of how a person's subjective experience of intending to act relates to the neural preparatory activity for action, successful prediction of movement intent may provide us with faster and more convenient ways to control prosthetic devices or wheelchairs. In current brain-computer interface systems, users can control external devices by means of volitional or conscious control. Reliable prediction of voluntary movement may lead to a more effective BCI system that does not rely on attention, as sustained attention may tire the user (Sellers & Donchin, 2006). Additionally, 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 (Muralidharan, Chae, & Taylor, 2011). Finally, if early detection of movement intent is possible, this could be used to develop a BCI system for certain patients that can intervene with and inhibit an upcoming movement (Bai et al., 2011).

2. Experiment 1

Bai et al. (2011) were amongst the first to reliably detect movement intent before movement onset based on the ongoing EEG signals in real-time. In this first experiment we replicated Bai et al. (2011) offline to confirm we could predict movement intent before its onset based on the RP and alpha/beta ERD.

2.1 Methods

2.1.1. Participants

Four healthy participants took part in this experiment, all participated voluntarily and gave written informed consent. The average age of the participants was 24 (SD=3.1) ranging from 21 to 28 years old. Two participants were right-handed, two were left-handed and all had normal or corrected to normal vision.

2.1.2. Data acquisition

The experiment was designed and run in PsychoPy (Peirce, 2007). EEG was recorded using the 'Biosemi ActiveTwo' system with 64 Ag/AgCl active electrodes. They were placed according to the International 10/20 System (Klem, Lüders, Jasper, & Elger, 1999). Four EOG (electrooculography) electrodes were placed around the eyes to measure eye movements and eye blinks. One was placed above the left eye, one below the left eye and two on the sides of each eye. Two electrodes behind the ears (mastoids) were used as reference electrodes. Additionally, two EMG (electromyogram) electrodes were placed on the arm (extensor carpi ulnaris muscle) and wrist bone to record muscular activity from arm movements. Participants were asked to make tapping movements on a tap pad. The tap pad recorded the audio signal that was generated by the tapping and was used as a measure of movement onset.

2.1.3. Experimental procedure

The experiment took place in an electrically shielded room. Participants were seated in a chair in front of a computer screen at a distance of approximately 70 centimeters. Their right forearm was resting on a pillow in such a way they could relax their arm. They were asked to tap on the tap pad in front of them by making a self-paced voluntary wrist extension with their right hand. They were specifically asked not to plan their movements but to be as spontaneous as possible.

Each trial started with a green square in the middle of the screen and participants were instructed that the goal was to keep this square green. If they moved too fast (< 5 seconds), the green square turned red for 3 seconds, if they moved too slow (>15 seconds) the square turned blue for 3 seconds. This was done to make sure we got a good baseline for analyzing the EEG data before the participant intended to act and to keep the experiment going by making sure the participants didn't wait too long before acting. When they moved between 5-15 seconds the square stayed green. Similar to the red and blue feedback, there was a 3 second period before they could move again. However, now this was a "silent" feedback phase as the square stayed green. In 20% of the trials the square turned red or blue even though they moved in the right time range. This was done to keep the movements spontaneous. A fixation cross appeared in the middle of the square for 0.5 seconds to indicate a tap was registered and a new trial would start. The experiment consisted of 21 blocks of 10 trials. After each block participants received feedback on how many trials they got correct (i.e. kept the square green). The whole procedure, including cap fitting, took about 1.5 hours.

2.1.4. Data analysis

Data was analyzed in MATLAB (MathWorks Inc., Natick, MA, USA). Offline classification was performed for both the ERD and RP. Similar to Bai et al. (2011), EEG data from 1.5 seconds before tap onset to tap onset was labeled as 'movement state', where movement related activity was expected to occur. Data within 1,5 seconds from trial start was labeled as 'non-movement state', where no movement related activity was expected. Since we expected only relevant activities over the motor cortex, only the central electrodes (C1, C2, C3, C4, Cz, FC1, FC2, FC3, FC4, FCz, CP1, CP2, CP3, CP4, CPz) were used for classification. The time-frequency data was temporally filtered; the Welch method with a Hanning window was applied to estimate power spectral density (PSD). A 4 Hz frequency resolution (256/4= 64 segment length) was used to estimate the PSD. Temporal filtering was done to reduce spectral leakage (Smith, 1997). Relevant activities were expected in the alpha and/or beta bands, so only the 8-30Hz frequency bands were taken into account. Surface Laplacian derivation (SLD) was performed for spatial filtering, meaning the EEG signal from each electrode was referenced to the averaged potentials from the nearest four orthogonal electrodes (Hjorth, 1975). By using SLD, the local EEG potentials were enhanced by increasing spatial specificity (Bai et al., 2007; Pfurtscheller, 1988). For the RP the data was low-pass filtered at 10Hz and SLD was performed for spatial filtering. Finally, bad trials were removed when the channel or trial power deviated more than 3.5 standard deviations from the mean. This preprocessed data was used to train a linear classifier. The classifier was applied offline with 1.5 second time windows in steps of 100 milliseconds. This was repeated 10 times for cross-validation. For each run, the classifier was trained on a random 2/3 of the data and applied on the other 1/3 of the data. To calculate when the movement and non-movement class significantly differed from each other, a binomial test was performed to the determine the earliest point where the classification performance was higher than the 95% confidence interval for chance performance (Billinger et al., 2012).

2.2 Results

One participant was excluded from offline classification as the data contained a lot of electrical noise. The classifier training performance for the ERD for each participant was 74%, 68% and 89%. For the RP the training performances for each subject were 67%, 70% and 65%. Table 1 shows the results of the binomial test for each subject. It shows the first time points a significant class difference was found for each participant for the ERD and RP and the time point when movement intent was expected to be predicted. For instance, for participant 1, the first time a significant difference between 'move' and 'non-movement' classes for the ERD was found at -2.3 seconds. Since the classifier was applied with 1.5 second time windows, the actual prediction was made 1.5 seconds later. Thus, for subject 1 movement intent was expected to be predicted 0.8 seconds before movement onset. Predictions based on the ERD were found on average 0.7 (SD=0.17) seconds before movement intent. Predictions based on the RP were found on average 0.43 (SD=0.84) seconds before movement

Table 1. Results offline classification based on the ERD and RP for each subject. First columns show the first time point a difference was found for the ERD and RP classifiers. The last two columns show the time points the expected predictions were made for each classifier.

Participant	**Start window significant difference ERD (s)**	**Start window significant difference RP (s)**	**Prediction movement intent ERD (s)**	**Prediction movement intent RP (s)**
1	-2.3	-1.4	-0.8	0.1
2	-2.0	-2.9	-0.5	-1.4
3	-2.3	-1.5	-0.8	0

As shown in the table, predictions based on the ERD were made as early as 0.8 seconds before movement onset, whereas predictions based on the RP (with the exception of participant 2) were made on or a little after movement onset. Based on these results, we expected it should be possible to predict movement intent in real-time prior to movement onset based on the RP and/or ERD over the motor cortex. We also expected the ERD to be a better feature for online prediction than the RP as the predictions are generally made earlier.

3. Experiment 2

As mentioned briefly in the introduction, FES could provide a means to feed the predictions back to the participants. Qiu et al. (2016) quantitatively compared ERD patterns during active, passive and FES-induced movement. They found that beta ERD values induced by FES movements significantly correlated with beta ERD values from voluntary movements while no significant correlation was found between ERD values of FES-induced movements and passive movements, as well as voluntary and passive movement. Müller et al. (2003) found that the main difference between movement elicited by FES and self-paced or passive movements was that there was no ERD detected prior to movement onset. This suggests that sensorimotor processing during FES involves some of the same processes as voluntary movement. In both voluntary movement and movement by FES, efferent nerves of the muscles were stimulated, either by active motor commands or by FES, and in both cases afferent input could be anticipated.

We want to make the forced movements as similar as possible to voluntary movements. FES stimulation could be a good technique for that. One possible issue with using FES is the possibly large effect of FES stimulation on the ongoing EEG signals, which could make analysis difficult. In this second experiment we aimed to investigate the effect of FES stimulation on EEG data, specifically on the RP and alpha/beta ERD.

3.1 Methods

3.1.1. Participants

Three healthy participants took part in this experiment. All participated voluntarily and gave written informed consent. The average age was 30 (SD=10.3) ranging from 21 to 41. All participants had normal or corrected to normal vision and all were right-handed.

3.1.2. Data acquisition

The Motionstim 8 stimulator (Krauth & Timmermann, Germany) was used to perform functional electrical stimulation. Two oval FES electrodes (4x6cm) were placed around the extensor carpi ulnaris muscle in such a way that the movement made by FES was similar to the participant’s voluntary wrist extension. Stimulation frequency was set to 20 Hz with a 300 milliseconds pulse duration. Stimulation current was determined for each subject individually. For all three participants this was set to 20 microVolts. The same set-up for measuring EEG, EMG and audio data was used as in the first experiment.

3.1.3. Experimental procedure

After capfitting, the FES stimulation was set up. Starting with a low current (6 microVolts), the participant was introduced to the stimulation. The stimulation current was slowly increased until the stimulated movement was most similar to the participant’s own movement. The experimental task was the same as in experiment 1. However, now there were two conditions. At the beginning of the trial either the words ‘FES trial’ or ‘TAP trial’ were presented on the screen. In case of a ‘FES trial’, participants were asked to wait until the FES device would make a forced movement. The FES stimulation was made randomly between 5 and 15 seconds. In case of a ‘TAP trial’, the participant was asked to make a self-paced voluntary wrist extension with their right hand while performing the same task as described in experiment 1. The two conditions were randomized and the experiment consisted 21 blocks of 10 trials. The whole procedure, including cap fitting and setting up de FES, took about 2 hours.

3.1.4. Data analysis

Data was analyzed in MATLAB (MathWorks Inc., Natick, MA, USA). For each subject a time-frequency analysis and an ERP analysis was performed. The data was sliced in trials of -10 to 10 seconds around the recorded taps. All trials with movements before 5 seconds were deleted. The data was then split up in ‘TAP trials’ and ‘FES trials’ and the two conditions were preprocessed separately. The data was detrended and re-referenced to the average of the mastoids. After this the data was time-locked to EMG onset, as we found this to be a more accurate indicator of movement onset. The EMG channel on the wrist bone was subtracted from the channel on the forearm muscle. The signal was band-pass filtered (51-250 Hz), rectified, and a first order low-pass Butterworth filter was applied with a normalized cut-off frequency of 16/128 Hz (cut-off frequency / (sampling frequency/2)). EMG onset was determined as the mean plus 5 times the standard deviation of the EMG signal. Eye blinks and eye movements were subtracted. Bad channels and bad trials that deviated more than 2.5 standard deviations from the mean were removed. The data was visually inspected afterwards and if clear artifacts were still seen, extra trials/channels were deleted manually. The data was then low-pass filtered at 47 Hz and high-pass filtered 0.2 Hz. For the ERP analysis, preprocessed data was baseline corrected using a baseline window of 5 to 4 seconds before movement onset and a timelock analysis using FieldTrip (Oostenveld, Fries, Maris, & Schoffelen, 2011) was performed. For the time-frequency analysis, we used a (multi)taper approach implemented in FieldTrip. A Hanning window of 500 milliseconds was applied to frequencies between 2 and 30 Hz. This was done in steps of 100 milliseconds, from 5 seconds before tap to 3 seconds after the tap. Next to the individual analysis, a time-frequency analysis and ERP analysis was performed over the grand average of all participants.

3.2 Results

As expected, we found clear FES artifacts when a forced movement was performed. Figure 1 shows an example of the effect of FES stimulation in channel Cz (central channel on the motor cortex). As shown in this figure, both RP and ERD signals returned to baseline within 500 milliseconds. Furthermore, the FES artifacts are found mostly on and after movement onset. Since we are interested in the EEG signals before movement onset, we don't expect the FES stimulation to be an issue for analyzing the neural preparatory activity.

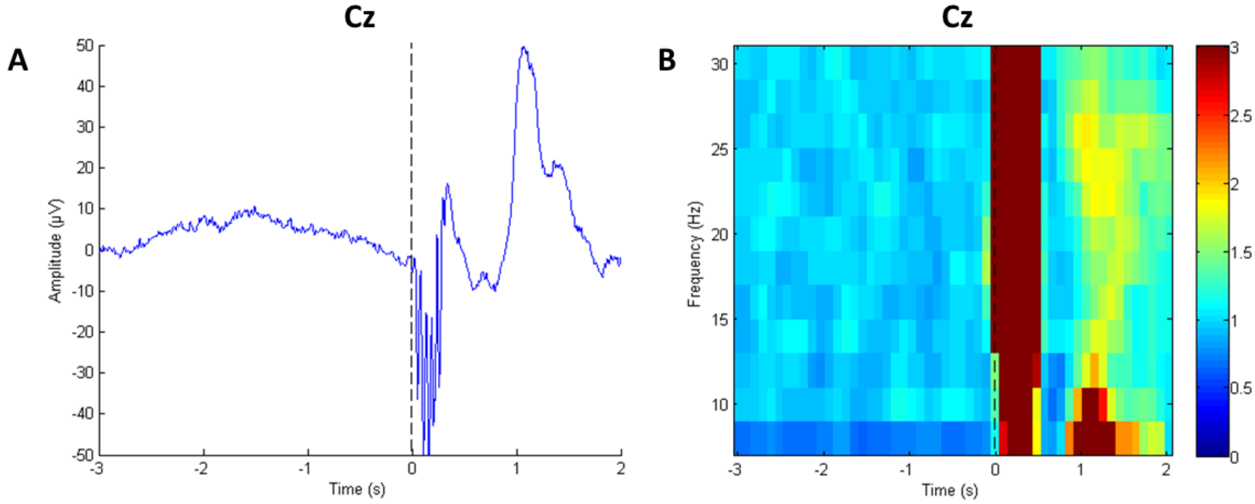

Figure 1. **A**. Grand average of the RP for channel Cz. **B**. Grand average for time-frequency representation for channel Cz. The colorbar shows the relative power change. Time 0 (marked by the dashed line) is the point of movement onset based on the EMG signal.

4. Experiment 3

In the final experiment we combined the two previous experiments and developed a brain-computer interface that predicts when someone has the intention to move. As soon as movement intent was detected, the person received feedback through a forced movement by means of FES. With this setup, we aimed to investigate whether we can reliably predict movement onset in real-time based on the ongoing EEG activity, as well as how the subjective experience of intending to act relates to the neural preparatory signals.

Investigating the subjective experience of intending to act is quite complex. There are many ways to ask about a person's intention and many ways a person can interpret what an intention is (Verbaarschot, Farquhar, & Haselager, 2015). Furthermore, internal and external influences can affect a person's report of their intention to act. For instance, the amount of effort you have to invest in your movement influences whether or not you attribute a movement to yourself. Moreover, a temporal delay of an action's consequence leads to attributing the consequences to others (Minohara, Wen, Hamasaki, Maeda, & Kato, 2016). Another well-known example is the "self-serving bias", meaning that we tend to attribute positive outcomes to oneself and negative outcomes to external factors (Gentsch & Synofzik, 2014; Greenberg, Pyszczynski, & Solomon, 1982). This suggests positive or negative feedback during the experiment could influence the participant's subjective report. In our experiment we explained the difference between wanting to move or not wanting to move as the difference between something you do (or wanted to do) and something that happened to you.

4.1 Methods

4.1.1 Participants

Eight healthy participants took part in this study. Participants received course credits or money for their participation. The average age of the participants was 27.4 (SD=7.8) ranging from 19 to 43 years old. All participants were right-handed and had normal or corrected to normal vision. All gave written informed consent. One participant was excluded before the start of the experiment since we could not invoke a movement with FES that was similar to the participant's own movement. Three participants were excluded after stage 1 of the experiment: one was excluded because the classifier training performance was too low (<50%) and two participants were excluded due to technical difficulties with online classification (further explained in 4.4). This leaves four participants who completed the whole experiment.

4.1.2 Data acquisition

The experiment was designed and run in PsychoPy (Peirce, 2007) and MATLAB (MathWorks Inc., Natick, MA, USA). The same set-up for measuring EEG and EMG was used as in the first and second experiment. The tap pad was no longer used to record movement onset, since the recorded taps did not provide an accurate indicator of movement onset in experiment 1 and 2. EMG electrodes were placed on the arm (extensor carpi ulnaris muscle) and the wrist bone to record muscular activity from arm movements and to continuously detect movement onset. The Motionstim 8 stimulator (Krauth & Timmermann, Germany) was again used to perform functional electrical stimulation. Two oval FES electrodes (4x6cm) were placed around the extensor carpi ulnaris muscle and stimulation frequency was set to 20 Hz with a 300 milliseconds pulse duration. Stimulation current was determined for each subject individually and ranged from 18 to 22 microVolts.

4.1.3 Experimental procedure

The experiment took place in an electrically shielded room. Participants were seated in a chair in front of a computer screen at a distance of approximately 70 centimeters. Both forearms were resting on the table in front of them and a button box was placed by their left hand. After capfitting and setting up the FES stimulation, a quick EMG training was performed to continuously decode muscle movement. Participants were informed that this training served as a check to see if we could correctly detect their muscle activity. Next, participants received the experimental instructions.

The experiment was divided into two identical parts, with a questionnaire in between to give participants a short break from the stimulation. Both parts took about 35 minutes (depending on the participant's reaction time). Another questionnaire was given at the end of the experiment. Including capfitting, setting up the FES, filling out the questionnaires and debriefing the experiment took approximately 2,5 hours.

Figure 2. The experimental set-up.

4.1.4 Experimental paradigm

The task participants performed was similar to that of experiment 1 and 2: self-paced voluntary movements while keeping the green square on the screen. However, this third experiment consisted of two stages. The first stage consisted of voluntary movements and randomly triggered FES movements and served to gather training data for our classifier. The second stage consisted of voluntary movements and FES triggered movements in response to real-time predictions made by the classifier. At the start of each stage, participants received identical instructions. They were informed that we were trying to develop a new brain-computer interface that predicts when they were intending to move and that the FES triggers were based on their brain data (their intentions). This was done to create uncertainty about the decision of who made the movement: their own motor system, their brain or something random.

In the first stage FES triggers were sent randomly. Participants were asked to make one tap movement each trial by making a self-paced voluntary wrist extension with their right hand. During each trial a FES trigger would be sent at random. After 10 trials, the running average of the participant's reaction time was calculated and random FES triggers were sent within a range of -2 to 2 seconds around that average. However, if EMG activity was detected prior to the random trigger, a FES trigger was sent right away to initiate a forced movement. This was done in order to make the voluntary movements as similar as possible to the forced FES movements. After each trial the following question was presented: "Did you want to make this movement?". Participants could answer with "Yes", "No" or "Don't know" using the button box. Since participants received FES stimulation during both voluntary and forced movements, they could not answer this question based on the presence of FES stimulation alone. After they answered the question, they received feedback on whether their movement was made in the right time range (red, green or blue square). The feedback was given after the question to avoid the effect of negative or positive feedback on the participant's report. This first stage consisted of 11 blocks of 10 trials. The first block was intended to familiarize them with the experiment. After each block the participant received feedback on how many trials they kept the green square on the screen and there was time for a short break. After stage 1, participants filled in a questionnaire that went into more detail about their subjective experience during the first part of the experiment. While the participants were filling in the questionnaire, a classifier was trained based on trials from the first stage. If the training performance of the classifier was less than 50%, the subject was excluded from the second stage of the experiment.

Stage 2 was similar to stage 1 except now the FES triggers were not random, but based on classifier predictions. This stage consisted of 10 blocks of 10 trials. Afterwards, a second questionnaire consisting of questions about their subjective experience in this stage, as well as compared with the first, was filled out. We hypothesized that EMG-triggered movements would be reported as "intended" close to 100% of the time in both stages as we expected muscle activity to occur only when the participant intended to make a movement. For the random-triggered movements it is more difficult. Random FES triggers were sent within a range of -2 to 2 seconds around the running average of the participant's reaction time and a movement was classified as a random-triggered movement only if the random trigger was sent before muscle onset. We assume that the participant can be aware of their intention when probed 2 seconds before muscle activity (Verbaarschot, Haselager, & Farquhar, 2016). Moreover, if we assume that both the distribution of actions and the distribution of random FES triggers are uniform and 4 seconds wide, the chance the FES trigger is sent in the 2 second window prior to the action, given that the trigger happened before the action, was estimated to be around 75%[1]. Thus the random-triggered trials were expected to be reported as intended 75% of the time. Finally, if the RP and alpha/beta ERD are predictive of the awareness of intending to act, we expected that the brain-triggered trials (the trials with classifier predictions) of stage 2 would be reported as 'intended' close to 100% of the time.

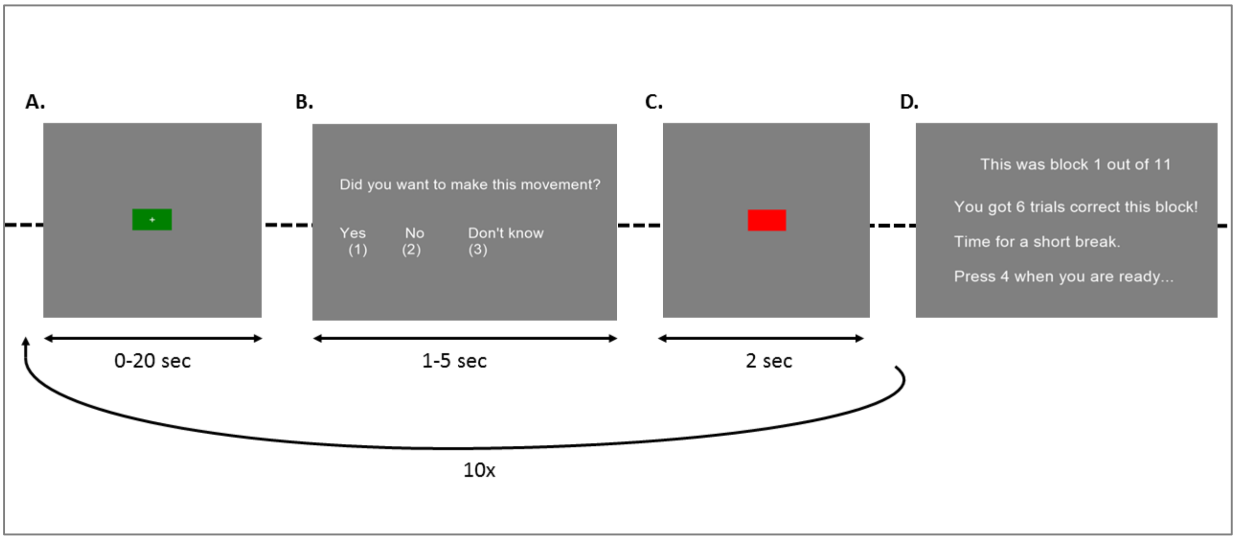

Figure 3. The experimental design. **A.** The start of the trial is indicated with a green square and fixation cross in the middle of the screen. **B.** Question with corresponding buttons to press. **C.** In this example the feedback was red (indicating they moved too fast), but it could also be a blue or green square. **D.** This shows an example of feedback at the end of each block.

4.2 Real-time prediction

4.2.1 EMG detection

For training the EMG detector, participants performed a short task. In the task, either the word "relax" or "right hand" was shown on the computer screen for 3 seconds. Participants were asked to rest their right arm as long as the word "relax" was on the screen and make wrist

[1] In a 4 second window, we will always find 100% intention for the first 2 seconds, assuming the subject is aware of their intention 2 seconds before movement. If the action happens later in time, the time window up until the point of the action gets bigger, but the intention window stays the same length (2 sec). This then decreases the chance the trigger is presented in the intention window. This chance decreases as a function of movement time (2/4 when action happened at 4 sec) and we estimate this to be 75%.

extensions as long as the word "right hand" was on the screen. Six "relax" and six "right hand" trials were performed in alternating order.

The EMG channel on the wrist bone was subtracted from the channel on the forearm muscle. The signal was band-pass filtered (70-250 Hz), rectified, and a first order low-pass Butterworth filter was applied with a normalized cut-off frequency of 16/128 Hz. Three seconds of data from the start of the "right hand" trials were labeled as movement data. The mean and standard deviation of the movement data was calculated and the threshold for movement was set as the mean plus 3 times the standard deviation. EMG input was continuously decoded every 50 milliseconds and whenever the input was detected to be higher than the threshold, a trigger was sent to the FES device to make a forced movement.

4.2.2 EEG classifier

EMG-triggered trials from stage 1 of the experiment were used to train a linear classifier based on the RP and ERD. EMG onset was determined as the point where the EMG signal was higher than the threshold (as described in 4.2.1). The data from 500 milliseconds before EMG onset to EMG onset was labeled as 'movement' data and the data from -5000 to -4500 milliseconds before EMG onset was labeled as 'non-movement' data.

Surface Laplacian derivation was performed for spatial filtering. For the ERD, data was temporally filtered using the welch method with a Hanning window, with a 4 Hz frequency resolution (256/4= 64 segment length) applied to estimate PSD. We only expected relevant information in the alpha and beta frequency bands, so only the 8-30 Hz bands were taken into account. RP data was spatially filtered by performing SLD and low-pass filtered at 10Hz. Relevant EEG activity was assumed to occur over the motor cortex. However, all 64 EEG electrodes were used for training the classifier. This might introduce more noise, but if there are artifacts we expect them to occur in the whole brain, thus the other electrodes provide us with useful information.

4.2.3 Determining the optimal threshold for prediction

Initially, cross-validated classifier predictions from the training data were used to determine a threshold for prediction. The median for the positive predictions for the RP and ERD was computed and used as the RP and ERD threshold. The classifier was continuously applied to the data every 500 milliseconds, with steps of 100 milliseconds. Every step, the last 10 predictions were saved and a prediction event was sent to activate the FES device whenever 8 out of 10 ERD prediction values were higher than the ERD threshold and 4 out of 10 RP prediction values were higher than the RP threshold. A larger amount of ERD predictions was chosen as we expected the ERD to be a better feature for online prediction and thus to predict movement intent more accurately than the RP.

This method was tested with two pilot subjects. During testing it was found that for one participant almost no predictions were sent and for the other participant predictions were sent all the time. The data was analyzed offline and we found the RP did not contribute much to early detection of movement intent. Based on this, the results of our first pilot and previous literature (Bai et al., 2011) we decided to only use the ERD signal for online classification. As our thresholds for the pilot subjects were not reliable, we used another approach to determine a threshold for online prediction in the experiment.

Based on the data of the trained ERD classifier, receiver operation characteristic (ROC) curves were made for each participant showing the true positive rate (sensitivity) against the false positive rate (1-specificity) at different threshold settings for the ERD data. In an ideal situation, predictions would be made with a low false positive rate (predicted movement, when there was no movement) and a high true positive rate (predicted movement when there was movement). However, there is a tradeoff between the two. For our experiment we decided it was more important to have a low false positive rate than to have a high true positive rate, so that when a prediction was made, the movement would likely follow. As the ERD and RP are known to build up over the motor cortex up to 2 seconds before movement onset (Kornhuber & Deecke, 1965; Pfurtscheller & Aranibar, 1979), positive predictions that occur in that time window were considered to be true positives. Positive predictions that were made prior to that time window were considered to be false positives. To ensure a minimal false positive rate and to predict movement intent before its onset, we tried several different thresholds for predicting movement intent online. These will be described in the paragraphs below.

We wanted to allow one false positive prediction per trial. The average trial length was expected to be approximately 7 seconds. The classifier was applied every 500 milliseconds to the data, with steps of 100 milliseconds, so every 100 milliseconds a prediction was made. Under the assumption that these predictions were independent of each other, this means that every second we get 10 predictions. Thus, the false positive rate was set at 1/70, allowing one false prediction per 70 predictions. The ROC curve was used to determine the ERD threshold matching the desired false positive rate. Two participants were tested in both stages and we found the false positive rate (FPR) to be much higher than expected. Offline analysis showed there was a large negative baseline shift and more extreme prediction values in the second part of the experiment compared to the first part. This difference in prediction values between the first and second stage is shown in figure 4. This means determining a threshold for classification based on the first part of the experiment alone is not sufficient. Instead, we need an adaptive threshold that continuously adapts to hit our target FPR.

For the adaptive threshold the time-constant of the adaptation. i.e. how fast the threshold was expected to change, was set at 50 predictions. As we wanted a minimum amount of false positive predictions we set the target percentile to 95%, meaning 95% of the predictions should be below the threshold (< 0). It adapts by computing a running estimate of the 95% point of the last N predictions. It then shifts the outputs such that predictions lower than the 95% percentile have values less than 0 and those higher having a value greater than 0. Whenever a prediction was higher than the threshold an event was sent, initiating a forced FES movement. The classifier was again applied every 500 milliseconds with steps of 100 milliseconds. Two participants completed the whole experiment with this adaptive threshold.

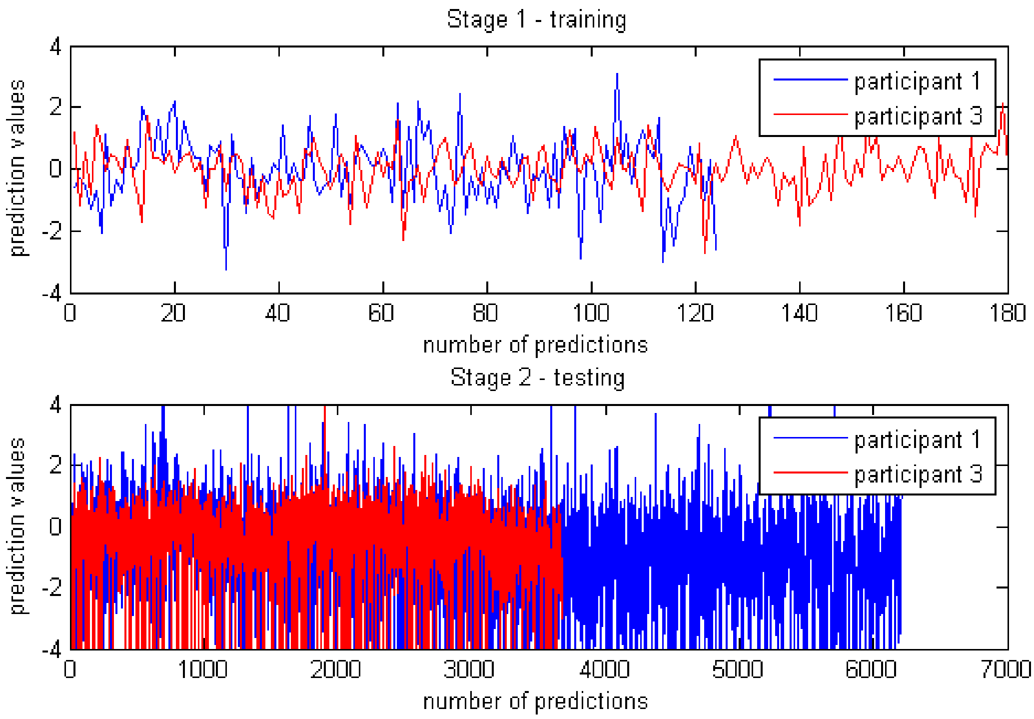

Figure 4. Difference between training and testing. The upper plot shows the prediction values for each prediction that was made during the training phase. The plot below shows the prediction values for each prediction in the testing phase.

4.3 Data analysis

Data was analyzed in MATLAB (MathWorks). For each participant a time-frequency analysis and an ERP analysis was performed. All trials with movements before 5 seconds were deleted. The data was then split up in 'intended trials' and 'unintended trials' based on the subjective reports. The conditions were preprocessed separately. The data was detrended and re-referenced to the average of the mastoids. Eye blinks and eye movements and muscle artifacts were subtracted. Bad channels were removed when they deviated more than 3 standard deviations from the mean. Bad trials were removed when they deviated more than 2 standard deviations from the mean. The data was visually inspected afterwards and if clear artifacts were still seen, extra trials/channels were deleted manually. Moreover, the data was low-pass filtered at 47 Hz and high-pass filtered 0.2 Hz. Within subject cluster-based permutation tests[2] were performed to investigate whether there was a significant difference in ERD and RP signals between intended and unintended trials. The response rate for EMG-triggered, random-triggered and brain-triggered trials was computed and for each subject and the questionnaires were analyzed.

4.4 Results

4.4.1. Questionnaire

The questionnaire after the first part of the experiment showed that 5/7 participants used a strategy to decide when to act. Two out of those five, reported to be counting during the beginning of the experiment (until the experimenter told them to be more spontaneous) in order to time their actions correctly. Other timing strategies included actively thinking of something else before acting (1 participant) or trying to be surprising (2 participants). These strategies were not considered a problem for the experiment as these result in more spontaneous movements. In the first stage FES triggers were random and EMG based. 4/7 Participants reported that the FES triggers were sometimes right, 2/7 reported the FES triggers to be accurate most of the time. 1/7 Participants reported the FES triggers were not accurate. 5/7 Participants felt free to make a movement whenever they wanted to. 4/7 Participants reported they really wanted move when the FES was activated, 3/7 reported they wanted to move most of the time. 3/7 Participants felt in control of their movements, 1/7 did not feel in control and the other 3/7 reported feeling in control only in trials the computer predicted accurately according to them, i.e. when the FES was activated based on the EMG.

The second part of the questionnaire was filled in by the 4 participants that completed both stages of the experiment. All 4 participants reported FES triggers, now brain and EMG based, were less accurate compared to the first stage. Differences that were noticed were that they got more red squares (indicating they moved too fast), that the FES device stimulated more often and that the computer seemed more sensitive. 1/4 Participants reported to feel free to make a movement whenever they wanted to, 2/4 sometimes felt free and 1/4 felt like he/she had

[2] See: www.fieldtriptoolbox.org/tutorial/cluster_permutation_timelock

to move before the FES trigger would be sent. Finally, 3/4 participants reported feeling in control of their movement when the FES was initiated by the EMG, not when the FES was initiated by the classifier. 1/4 Participants did not feel in control of their movements.

4.4.2. Response rate intended report

For each subject the EEG data was split up in 'intended trials', 'unintended trials' and 'don't know trials' based on the subjective reports. Table 3 shows an overview of the amount of 'intended', 'unintended' and 'don't know' reports for EMG-triggered and random-triggered trials in stage 1 of the experiment and EMG-triggered and brain-triggered trials in stage 2 of the experiment. For each participant we calculated the response rate for intended trials, e.g. what percentage of brain-triggered trials were reported as intended. This was calculated by dividing the number of intended trials by the total number of trials for each movement condition. Figure 5 shows these response rates.

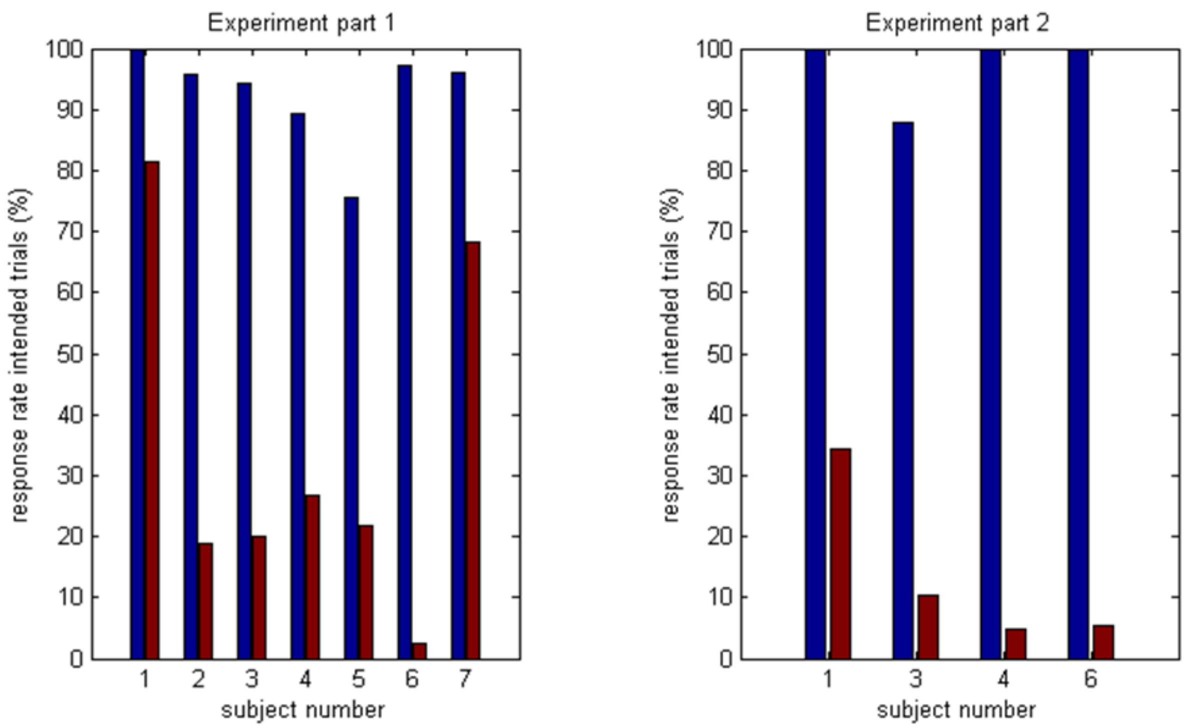

Figure 5. Response rate for intended trials for each condition. The blue colored bars are the EMG-triggered trials in both plots. The red colored bars in the left plot are the random-triggered trials and the red colored bars in the right plot are the brain-triggered trials. As an example, for subject one 100% of the EMG-triggered trials were reported as intended, and 81,58% of the random-triggered trials were reported as intended.

As explained in section 4.1.4, EMG-triggered trials and brain-triggered trials were expected to be reported as intended close to 100% of the time. Random-triggered trials were expected to be reported as intended 75% of the time. To test whether these hypotheses were true, a Wilson confidence interval was performed with an α of 0.05 (Wilson, 1927). The confidence interval tells us that, at a given level of uncertainty (0.05), if our null-hypothesis is correct, the true value for the population will likely be in that range. Table 2 shows the results of this binomial test. In stage 1, the majority of EMG-triggered trials were significantly different from our null hypothesis, meaning that the EMG-triggered trials were not reported as intended 100% of the time for those subjects. Random-triggered trials were significantly different for five subjects in stage 1. For two subjects the null-hypothesis was not rejected, meaning the response rate was likely around 75%. All confidence intervals can be found in table 2.

Table 2. Response rates and their confidence intervals for EMG-triggered, random-triggered trials and brain-triggered trials for each subject.

Subject	**Response rate EMG trials**	**Wilson confidence interval (H0 = 1.0)**	**Response rate random trials**	**Wilson confidence interval (H0 = .75)**
Stage 1				
1	100%	[0.9417 – 1.0000]	81.58%	[0.6658 – 0.9078]
2	95.74%	[0.8956 – 0.9833] *	18.75%	[0.0659 – 0.4301] *
3	94.44%	[0.8765 – 0.9760] *	20%	[0.0807 – 0.4160] *
4	89.47%	[0.8170 – 0.9418] *	26.67%	[0.1090 – 0.5195] *
5	75.64%	[0.6506 – 0.8381] *	21.88%	[0.1102 – 0.3875] *
6	97.1%	[0.9003 – 0.9920] *	2.44%	[0.0043 – 0.1260] *
7	96%	[0.8654 – 0.9890] *	68.33%	[0.5577 – 0.7869]
Subject	**Response rate EMG trials**	**Wilson confidence interval (H0 = 1.0)**	**Response rate brain trials**	**Wilson confidence interval (H0 = 1.0)**
Stage 2				
1	100%	[0.9442 – 1.0000]	34.29%	[0.2083 – 0.5085] *
3	87.88%	[0.7267 – 0.9518] *	10.45%	[0.0515 – 0.2003] *
4	100%	[0.9536 – 1.0000]	4.76%	[0.0085 – 0.2267] *
6	100%	[0.8794 – 1.0000]	5.56%	[0.0218 – 0.1343] *

* = significantly different from null-hypothesis

Table 3. Overview of all trials in stage 1 and 2 for each subject *before* removal of bad trials. This table shows the amount of intended, unintended and don't know reports for EMG-triggered, random-triggered movements in stage 1 and EMG-triggered and brain-triggered movement in stage 2.

Report / Participant		**Intended**	**Unintended**	**Don't know**	**Total**
Stage 1					
1	EMG	62	0	0	62
	Random	31	7	0	38
2	EMG	90	4	0	94
	Random	3	13	0	16
3	EMG	85	5	0	90
	Random	4	16	0	20
4	EMG	85	10	0	95
	Random	4	11	0	15
5	EMG	59	12	7	78
	Random	7	24	1	32
6	EMG	67	1	1	69
	Random	1	32	8	41
7	EMG	48	2	0	50
	Random	41	15	4	60
Stage 2					
1	EMG	65	0	0	65
	Brain	12	23	0	35
3	EMG	29	4	0	33
	Brain	7	59	1	67
4	EMG	79	0	0	79
	Brain	1	20	0	21
6	EMG	28	0	0	28
	Brain	4	56	12	72

4.4.3. EEG data

As can be seen in table 3, there were only a few trials reported as 'don't know', moreover for these trials it would be unclear what to look for when analyzing the data. Therefore, we chose not to analyze the 'don't know' trials. Furthermore, there were much less trials that were reported as unintended compared to ones that were reported as intended. Important to note is that table 3 shows the amount of trials before bad trial removal. As a consequence of the unbalanced number of intended and unintended trials, we chose to not perform cluster-based permutation tests to compare intended versus unintended trials for each subject separately. Instead, within-subject cluster-based permutation tests were performed to see if there were significant differences over all subjects between intended movements and unintended movements. The tests were performed per stage and for the RP and ERD separately (4 tests in total). Only the central electrodes were included (C1, C2, C3, C4, Cz, FC1, FC2, FC3, FC4, FCz, CP1, CP2, CP3, CP4, CPz) since we only expected relevant activity over the motor cortex. For all four tests, no significant clusters were found. These results show that, over all subjects, there was no difference found between intended and unintended movements for both the ERD and RP signals in both stage 1 and stage 2. The grand averages over all subjects for intended vs unintended trials for the RP and ERD were calculated and plotted in figure 6.

4.4.4. Online predictions

Participant 1 and 3 were tested with the original threshold based on the ROC curve as explained in 4.2.3. Participant 4 and 6 were tested with the adaptive threshold. During testing it was found that the adaptive threshold did not work well for all participants. For two participants the online predictions were constantly sent within 1 second after trial start, hence it was decided to end the experiment. Histograms were made offline to check the timing of the first predictions in each trial. Figure 7 shows examples of two cases: subject 1, where the online classification worked quite well, but not a lot of predictions were made and subject 6, where there were a lot of predictions, but these were often made early.

These histograms show that the online classification did not work as expected. Either predictions were not made often (participant 1) or the false positive rate was much higher than expected (participant 3, 4 and 6). The clear trade-off between false positive predictions and true positive predictions is visual in these plots. To achieve a low false positive rate, we must have a low true positive rate, i.e. not many predictions. To achieve a reasonable true positive rate, we get many false predictions. Finding a good trade-off between the two is challenging.

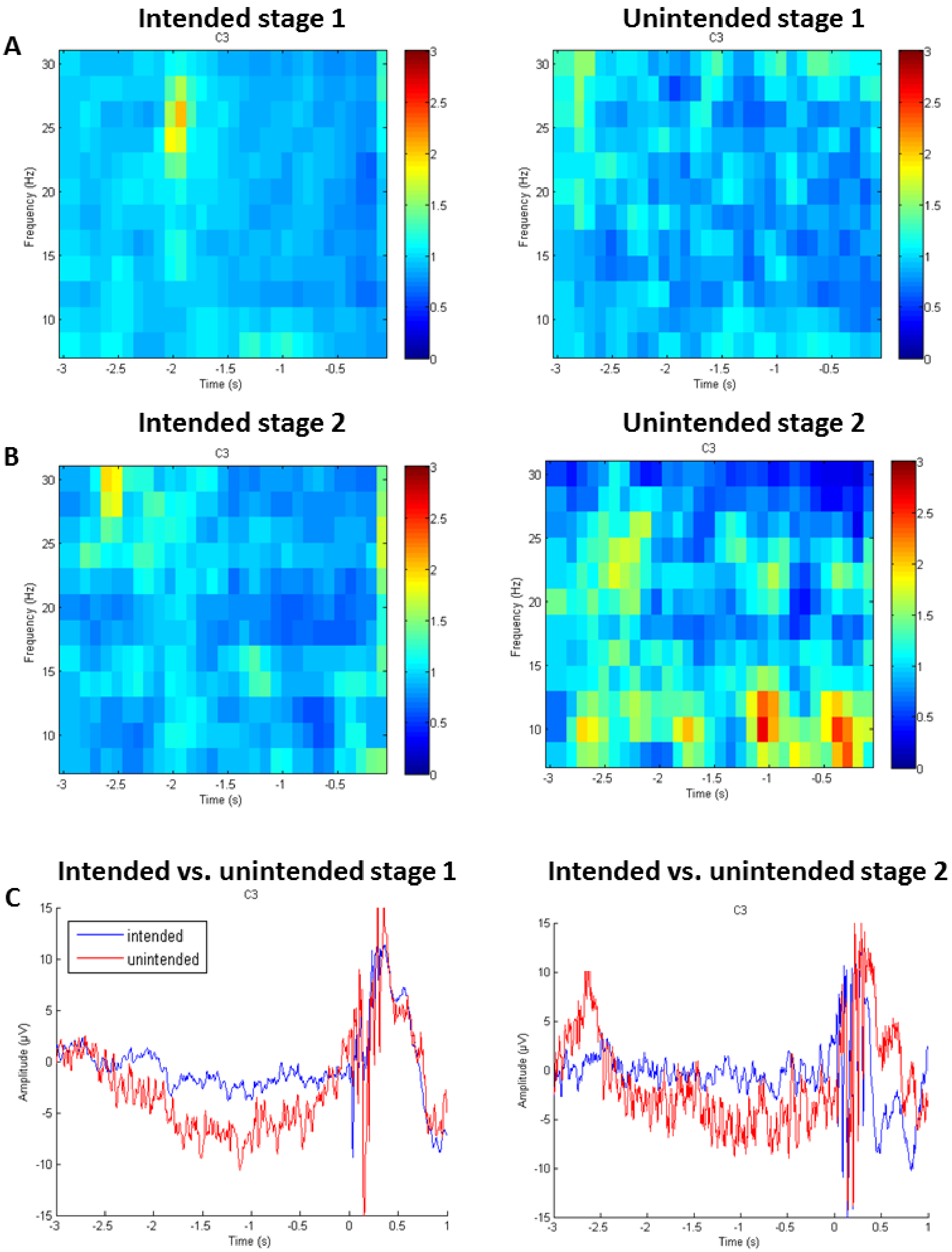

Figure 6. Grand averages over all subjects. Stage 1 included 461 intended vs. 129 unintended trials. Stage 2 included 169 intended vs. 23 unintended trials. **A** shows the grand averages of the ERD of channel C3 for intended data vs unintended data. **B.** shows the same as A but now for stage 2. **C.** shows the grand averages for the RP in both stages. The left plot shows the intended (blue line) vs. unintended (red line) in stage 1, whereas the right plot shows this for stage 2. The color bar in A and B shows the relative power change compared to the baseline (-4 to -3 seconds).

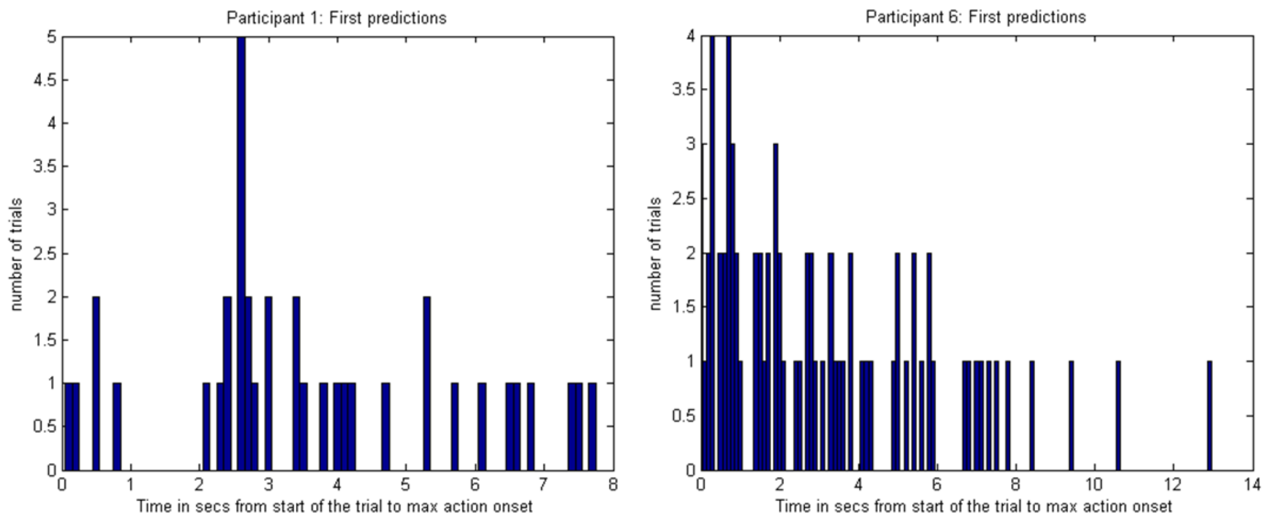

Figure 7. histogram of online predictions for participant 1 and 6. For each trial we looked when the first prediction was made relative to trial start. This was done in windows of 100 milliseconds. A prediction was always followed by an action. The histograms show all predictions made from trial start (0) to the maximum action onset in seconds for the participant.

5. Discussion

We conducted three experiments to investigate if we can predict movement intent in real-time based on the RP and/or ERD and to find out how the subjective reports of intending to act relate to these signals. We hypothesized we could predict movement intent in real-time based on the ongoing neural preparatory signals. Furthermore, we expected these signals to be significantly different prior to an electrically stimulated movement that was reported to be unintended compared to one that was reported as intended.

With our first experiment we successfully replicated Bai et al. (2011) offline. We showed we could predict movement intent offline before its onset and that movement intent was detected earlier based on the ERD (-0.7±0.17 s) than the RP (-0.43±0.84 s). The prediction onsets we found were in line with the prediction onsets that Bai et al. found (-0.62±0.25 s). Moreover, predictions based on the ERD were more consistent than predictions based the RP over all subjects (as shown by the standard deviations). Since the ERD predictions were made earlier and more consistently than the RP predictions, the ERD seems to be a better and more reliable feature for online prediction of movement intent than the RP.

In the second experiment we showed FES stimulation mostly influences the data on and after movement onset. In this experiment, we asked participants to just wait until the FES device made a forced movement. When analyzing the data for ‘FES trials’, no ERD or RP signals were found before movement onset (see figure 1), providing more evidence the RP and ERD signals are not seen before the onset of involuntary movements (Müller et al., 2003; Shibasaki & Hallett, 2006).

Before interpreting the results of the third experiment, it is important to note that the sample size of 7 participants for this experiment was very low. Even fewer participants were tested in the second stage of the experiment. Furthermore, there were not enough trials in all conditions to make good comparisons for each subject separately, as can be seen in table 3 in section 4.4.2. As a consequence, it was chosen to perform a cluster-based permutation test on the grand averages over all participants. No significant differences were found between electrically stimulated movements that were reported as intended compared to ones that were reported as unintended. This was the same for the RP and ERD and both experiment stages. Here, it is important to mention that there were many fewer unintended than intended trials (e.g. Stage 1: 129 unintended trials vs. 461 intended trials over all subjects and for stage 2: 23 unintended trials vs. 129 intended trials over all subjects). Few participants and few trials make it difficult to interpret our results as this leads to a low statistical power (Button et al., 2013). Moreover, the online classification did not work as expected and predictions were often made more than 2 seconds before movement onset, thus resulting in a lot of false predictions.

For these reasons, we could not answer one of our main questions of whether a person was aware of their intention to act when ERD or RP activity was detected in the brain. In the next section possible explanations and solutions will be provided.

For 5 out of 7 participants random-triggered trials were reported as intended much less often than expected. Looking at the questionnaire, 3 out of those 5 reported to only feel in control of their own movements, not of the FES-initiated movements. This suggests random-triggered movements were easily distinguishable from the EMG-triggered ones and that these participants were not aware of their intention to act when the random-triggers happened. One of the 5 participants reported only 2.4% of random-triggered trials as intended in stage 1. This participant reported to not feel free to make a movement whenever he/she wanted to and did not feel in control of their movement at all, confirming the low percentage of intended random trials.

Even though the random-triggered trials were reported as intended much less often than expected, there were a lot less unintended reports for each participant compared to intended reports. This could be explained by having only a small amount of random-triggered trials per subject (as can be seen in table 3). Random triggers were only sent when they occurred before muscle onset was detected. Thus, if the subject consistently moved earlier than expected few random triggers were sent. A possible solution to ensure that more randomly-triggered movements were experienced would be to weight the random-trigger distribution. By making the triggers in the period before the mean reaction time weigh more heavily, more random triggers could be made before the participants moved themselves. Moreover, the brain-triggered trials in stage 2 were not reported as intended close to 100% of the time. The classifier made a lot of false predictions. This means the FES triggers that were based on the brain were often made very early. For these false predictions, no intended reports were expected.

The high false positive rate was one of the reasons our classification did not work well. In order to say something about the relation between the RP and/or ERD and the subjective experience, we need enough true positives, i.e. trials that happen in the participant's 'intention awareness' window. However, since there is a trade-off between the true and false positive rate, this results in either many false predictions of movement when there was no movement intention, or in very few true predictions when there was a movement intention. Finding a good trade-off between the two is difficult.

During the development of this experiment we identified trial-time as a significant potential confound in this experiment. Even when participants were asked to be spontaneous, they tended to move at particular points in time after trial start. Since we wanted to identify brain signatures associated with movement, this is a potential problem as it is difficult to disentangle movement-related brain signals from brain signals associated with time since trial start. For instance, if a participant always moves at 8s after movement start, and they generate two main brain signals, one movement related signal in the 2 seconds before the movement, and one trial-related signal which only occurs 6 seconds after trial start, then with this experimental design we cannot separate these two signals.

Due to the random triggers used in our experiment, we expect an anticipation signal to build up over time, from trial start as the participant anticipates the upcoming stimulation. A well-known signal that is associated with anticipation is the contingent negative variation (CNV). The CNV is a slow negative cortical potential that can be observed during response anticipation over the motor cortex and typically depends on the contingency of two successive stimuli (Gangadhar, Chavarriaga, & Millan, 2009; Tecce, 1972; Walter, Cooper, Aldridge, McCallum, & Winter, 1964). Since participants knew trial start was always followed by FES stimulation, there was likely a buildup of the CNV. Thus, we indeed have a potential confound of brain signals associated with trial-time (via the CNV) and those associated with movement intention. This confound is an issue when training a classifier to identify movement intention from non-movement intention. This means the classifier can 'cheat' by using the CNV to make movement intention predictions indirectly based on time since trial start, instead of using movement related brain signals. There are a few ways to address this confound issue. When training the classifier, we need to weaken the connection between trial-time and the labelling of the data as ‘movement’ or ‘non-movement’ data. One way to do this is to broaden the range of data which is labeled as ‘movement’ or ‘non-movement’. For instance, if we label all data from 1500 milliseconds before movement onset until the movement as ‘movement’ and all data from trial start until 3500 milliseconds before the movement as ‘non-movement’, the classifier is less able to ‘cheat’ as it must find a brain signal which generalizes over a range of trial-times.

Furthermore, inducing more variation in movement time would help a lot, as then, for example, 4 seconds after trial start would sometimes be a classified as a ‘movement’ and sometimes as a ‘non-movement’, again forcing the classifier to identify a brain signal that’s not time-based to identify movement intentions.

Moreover, it would be good to include trials with no self-initiated movement, but only movement by FES, so that we have a measure of expectancy alone without an intention process in there. These trials can then also be used to train the classifier by using them as a negative ‘non-movement’ class. Finally, it would be good to ensure enough time for the data to return to baseline, as otherwise the classifier could respond to other events (e.g. visual feedback). This could for instance be done, by making the classifier wait 1 second after trial start before listening to data.

We suspect a lot of experiments that investigate the intention to act deal with this anticipation build up over time. In future research we need take this into account when designing our experiments. By building in trials where there is anticipation but no intention process and by forcing the classifier to identify movement-related brain signals we think this issue can be addressed. In the future, we would like to implement these possible solutions, use a bigger sample size and provide a more convincing argument to whether or not a person is aware of their intention act when the neural preparatory signals are predicted in the brain. We believe that, if we implement these solutions we will be able to control the false positive rate, and that we can achieve similar results as Bai et al. (2011), Lew et al. (2012) and Schneider et al. (2013) in real-time. With these changes, we hope we can contribute to finding out how a person’s subjective experience relates to the neural preparatory activity.

References

Bai, O., Lin, P., Vorbach, S., Li, J., Furlani, S., & Hallett, M. (2007). Exploration of computational methods for classification of movement intention during human voluntary movement from single trial EEG. *Clinical Neurophysiology*, *118*(12), 2637–2655.

Bai, O., Mari, Z., Vorbach, S., & Hallett, M. (2005). Asymmetric spatiotemporal patterns of event-related desynchronization preceding voluntary sequential finger movements: a high-resolution EEG study. *Clinical Neurophysiology*, *116*(5), 1213–1221.

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.

Billinger, M., Daly, I., Kaiser, V., Jin, J., Allison, B. Z., Müller-Putz, G. R., & Brunner, C. (2012). Is It Significant? Guidelines for Reporting BCI Performance (pp. 333–354).

Blankertz, B., Dornhege, G., Krauledat, M., Kunzmann, V., Losch, F., Curio, G., & Müller, K.-R. (2006). The Berlin Brain-Computer Interface: Machine learning based detection of user specific brain states. *Computer*, *12*(6), 581–607.

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, *14*(5), 365–376.

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.

Gangadhar, G., Chavarriaga, R., & Millan, J. del R. (2009). Anticipation based Brain-Computer Interfacing (aBCI). In *2009 4th International IEEE/EMBS Conference on Neural Engineering*, 459–462.

Gentsch, A., & Synofzik, M. (2014). Affective coding: the emotional dimension of agency. *Frontiers in Human Neuroscience*, *8*, 608.

Greenberg, J., Pyszczynski, T., & Solomon, S. (1982). The self-serving attributional bias: Beyond self-presentation. *Journal of Experimental Social Psychology*, *18*(1), 56–67.

Haggard, P. (2008). Human volition: towards a neuroscience of will. *Nature Reviews. Neuroscience*, *9*(12), 934–46.

Hjorth, B. (1975). An on-line transformation of EEG scalp potentials into orthogonal source derivations. *Electroencephalography and Clinical Neurophysiology*, *39*(5), 526–30.

Klem, G. H., Lüders, H. O., Jasper, H. H., & Elger, C. (1999). The ten-twenty electrode system of the International Federation. The International Federation of Clinical Neurophysiology. *Electroencephalography and Clinical Neurophysiology. Supplement*, *52*, 3–6.

Kornhuber, H. H., & Deecke, L. (1965). Changes in the brain potential in voluntary movements and passive movements in man: Readiness potential and reafferent potentials. *Pflugers Archiv Fur Die Gesamte Physiologie Des Menschen Und Der Tiere*, *284*, 1–17.

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

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). *Brain*, *106*(3), 623–642.

Minohara, R., Wen, W., Hamasaki, S., Maeda, T., & Kato, M. (2016). Strength of Intentional Effort Enhances the Sense of Agency. *Frontiers in Psychology*, *7*, 1–5.

Müller, G. R., Neuper, C., Rupp, R., Keinrath, C., Gerner, H. J., & Pfurtscheller, G. (2003). Event-related beta EEG changes during wrist movements induced by functional electrical stimulation of forearm muscles in man. *Neuroscience Letters*, *340*(2), 143–147.

Muralidharan, A., Chae, J., & Taylor, D. M. (2011). Extracting attempted hand movements from eegs in people with complete hand paralysis following stroke. *Frontiers in Neuroscience*, *5*.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Computational Intelligence and Neuroscience,* 1–9.

Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1), 8–13.

Pfurtscheller, G. (1988). Mapping of event-related desynchronization and type of derivation. *Electroencephalography and Clinical Neurophysiology*, *70*(2), 190–3.

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.

Popović, D. B. (2014). Advances in functional electrical stimulation (FES). *Journal of Electromyography and Kinesiology*, *24*(6), 795–802.

Qiu, S., Yi, W., Xu, J., Qi, H., Du, J., Wang, C., … Ming, D. (2016). Event-Related Beta EEG Changes During Active, Passive Movement and Functional Electrical Stimulation of the Lower Limb. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *24*(2), 283–290.

Schneider, L., Houdayer, E., Bai, O., & Hallett, M. (2013a). What We Think before a Voluntary Movement. *Journal of Cognitive Neuroscience*, *25*(6), 822–829.

Schneider, L., Houdayer, E., Bai, O., & Hallett, M. (2013b). What we think before a voluntary movement. *Journal of Cognitive Neuroscience*, *25*(6), 822–9.

Sellers, E. W., & Donchin, E. (2006). A P300-based brain-computer interface: Initial tests by ALS patients. *Clinical Neurophysiology*, *117*(3), 538–548.

Shibasaki, H. (2012). Cortical activities associated with voluntary movements and involuntary movements. *Clinical Neurophysiology*.

Shibasaki, H., & Hallett, M. (2006). What is the Bereitschaftspotential? *Clinical Neurophysiology*.

Smith, S. W. (1997). *The scientist and engineer's guide to digital signal processing*. California Technical Pub.

Soekadar, S. R., Birbaumer, N., Slutzky, M. W., & Cohen, L. G. (2015). Brain–machine interfaces in neurorehabilitation of stroke. *Neurobiology of Disease*, *83*, 172–179.

Tecce, J. J. (1972). Contingent negative variation (CNV) and psychological processes in man. *Psychological Bulletin*, *77*(2), 73–108.

van Gerven, M., Farquhar, J., Schaefer, R., Vlek, R., Geuze, J., Nijholt, A., … Desain, P. (2009). The brain-computer interface cycle. *Journal of Neural Engineering*, *6*(4), 41001.

Verbaarschot, C., Farquhar, J., & Haselager, P. (2015). Lost in time...The search for intentions and Readiness Potentials. *Consciousness and Cognition*, *33*, 300–315.

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.

Walter, W. G., Cooper, R., Aldridge, V. J., McCallum, W. C., & Winter, A. L. (1964). Contingent negative variation: An electric sign of sensorimotor association and expectancy in the human brain. *Nature*, *203*, 380–384.

Wilson, E. B. (1927). Probable Inference, the Law of Succession, and Statistical Inference. *Journal of the American Statistical Association*, *22*(158), 209.

Wolpaw, J. R., & Wolpaw, E. W. (2012). *Brain-computer interfaces : principles and practice*. Oxford University Press.