

8-2014

Complexity and familiarity enhance single-trial detectability of imagined movements with electroencephalography

Raechelle M. Gibson
Western University, rgibso5@uwo.ca

Srivas Chennu
University of Cambridge

Adrian M. Owen
Western University, dpavich@uwo.ca

Damian Cruse
Western University

Follow this and additional works at: <https://ir.lib.uwo.ca/anatomypub>

 Part of the [Anatomy Commons](#), and the [Cell and Developmental Biology Commons](#)

Citation of this paper:

Gibson, Raechelle M.; Chennu, Srivas; Owen, Adrian M.; and Cruse, Damian, "Complexity and familiarity enhance single-trial detectability of imagined movements with electroencephalography" (2014). *Anatomy and Cell Biology Publications*. 90.
<https://ir.lib.uwo.ca/anatomypub/90>

Title:

Complexity and familiarity enhance single-trial detectability of imagined movements with electroencephalography

Authors:

Raechelle M. Gibson^{*1}, Srivas Chennu², Adrian M. Owen¹, and Damian Cruse¹

Affiliations:

¹The Brain and Mind Institute, Department of Psychology, Western University, London, Ontario, CAN

²Department of Clinical Neurosciences, University of Cambridge, Cambridge, UK

Keywords:

Motor imagery; sensorimotor rhythm; movement complexity; electroencephalography; Vegetative State

Highlights:

- We use machine-learning to classify electrical brain activity during imagined hand movements in samples of experienced pianists, experienced ice hockey players, and age-matched controls.
- Imagery of complex actions yields more robust brain responses than imagery of relatively simpler actions, and imagery of familiar actions can result in more robust brain responses in some cases.
- Implications for future work with brain-computer interfaces for Vegetative State patients are discussed.

***Corresponding author at:**

The Brain and Mind Institute, Department of Psychology
Western University, London, ON, CAN

Tel.: +1 519 661 2111 x84672

Fax: +1 519 661 3613

E-mail: rgibso5@uwo.ca

Acknowledgments:

The authors thank Dr. J. Grahn and Dr. C. Higgins (Western University, London, Ontario, CAN) for their assistance in recruiting musicians and athletes for this work. This work was supported by grants to AMO from the James S. McDonnell Foundation and the Canada Excellence Research Chair (CERC) program. RMG was funded by a graduate scholarship from the Natural Sciences and Engineering Research Council of Canada (NSERC), and DC was funded by a fellowship from the Canadian Institutes of Health Research (CIHR).

Abstract

Objective: We sought to determine whether the sensorimotor rhythms (SMR) elicited during motor imagery (MI) of complex and familiar actions could be more reliably detected with electroencephalography (EEG), and subsequently classified on a single-trial basis, than those elicited during relatively simpler imagined actions.

Methods: Groups of healthy volunteers, including experienced pianists and ice hockey players, performed MI of varying complexity and familiarity. Their electroencephalograms were recorded and compared using brain-computer interface (BCI) approaches and spectral analyses.

Results: Relative to simple MI, significantly more participants produced classifiable SMR for at least one of the more complex imagined actions. During MI of performance of a complex musical piece, the EEG of the experienced pianists was classified significantly more accurately than during MI of performance of a simpler musical piece. The accuracy of EEG classification was also significantly more sustained during complex MI.

Conclusion: MI of complex actions results in more robust EEG responses than MI of relatively simpler actions, and familiarity with actions can enhance this response in some cases.

Significance: The accuracy of SMR-based BCIs in non-communicative patients may be improved by employing familiar and complex actions. Increased sensitivity to MI may also improve diagnostic accuracy for severely brain-injured patients in a Vegetative State.

1. Introduction

Patients with disorders of consciousness (DOC) are behaviourally characterized by varying levels of arousal and awareness measured primarily by their ability to exhibit reliable responses to external stimulation (Jennett, 2002; Bernat, 2006; Owen, 2008). Of the various conditions included in the DOC (*e.g.*, coma, the Minimally Conscious State, *etc.*), the Vegetative State (VS) is one of the most poorly understood (Jennett, 2002; Owen, 2008). After emerging from coma, VS patients retain cycles of eye opening and closing similar to the sleep-wake cycles of fully awake and aware individuals (Multi-Society Task Force on PVS, 1994*a, b*; Royal College of Physicians Working Group, 1996; *cf.* Cruse *et al.*, 2013). Critically, despite producing spontaneous movements, VS patients are unable to exhibit any purposeful outward responses to verbal commands, and are thereby diagnosed as ‘unaware’ (Jennett, 2002; Owen, 2008). Many VS patients have diffuse brain injury that may include insult to the peripheral motor system; these circumstances could lead to an inaccurate diagnosis of VS in a patient who retains awareness and cognitive function, but lacks the ability to respond purposefully in a behavioural assessment (Owen, 2008). In fact, researchers have reported that some patients who are diagnosed as VS can follow (*e.g.*, Owen *et al.*, 2006; Monti *et al.*, 2010; Bardin *et al.*, 2011; Cruse *et al.*, 2011; Goldfine *et al.*, 2011; Naci & Owen, 2013), or attempt to follow (*e.g.*, Bekinschtein *et al.*, 2011; Cruse *et al.*, 2012), commands by modulating their brain activity, despite being unable to follow commands with their external behaviour. These findings raise the possibility that assistive devices known as brain-computer interfaces (BCIs) could improve diagnostic accuracy in this group by detecting ‘covert’ signs of awareness, as well as by potentially offering the patient a means of communication (*e.g.*, Monti *et al.*, 2010; Lulé *et al.*, 2013).

BCIs are devices that can allow a person (the ‘user’) to operate a computer without producing a motor output. Using machine-learning techniques, subject-specific patterns of brain activity can be learned by a computer and subsequently classified into a predefined communicative output. For example, the computer may output the response “yes” when the user produces brain activity pattern A, and output the response “no” when the user produces brain activity pattern B (*e.g.*, Mason & Birch, 2003; Sorger *et al.*, 2009; Lulé *et al.*, 2013; Naci *et al.*, 2013). The computer algorithm must be trained on a series of trials in which the desired output from the user is known (the ‘training phase’ of machine-learning classification), and then tested on trials in which the desired output from the user is not known (the ‘testing phase’ of machine-learning classification) based on predefined features of the data (*e.g.*, power in a given frequency band of the electroencephalogram, EEG). From the testing phase of classification, one can obtain an accuracy value based on the number of successfully identified brain responses and, by extension, correctly executed communicative outputs from the BCI. Crucially, from a clinical perspective, when classification accuracy is significantly above chance, the individual is demonstrably capable of producing consistent and appropriate patterns of brain activity in response to commands, thus providing a means to identify covert command-following in the absence of a behavioural response (Cruse *et al.*, 2011; Owen, 2013). Since classification must be both accurate and reliable for successful communication and other BCI output functions, such as computer mouse cursor control, classification accuracy and task sensitivity are two of the most important measurements of any BCI.

A particular EEG signal called the sensorimotor-rhythm (SMR) is a practical option for BCIs intended for use by VS patients (Chatelle *et al.*, 2012; Naci *et al.*, 2012; Grosse-Wentrup & Schölkopf, 2013). Using as few as four surface electrodes placed on the head over the sensorimotor cortical areas (sites CP3, CP4, FC3, and FC4 from the modified international 10-20 system; Sharbrough *et al.*, 1991), one can acquire the SMR as a person kinesthetically imagines moving a body part. Power decreases known as event-related desynchronizations (ERDs) and power increases known as event-related synchronizations (ERSs) in the mu (7-13 Hz) and beta (13-30 Hz) frequency bands are typically used as the signal features for classification with SMR-based BCIs (Pfurtscheller & Neuper, 1997; Neuper & Pfurtscheller, 2001; Neuper *et al.*, 2009). Unlike other EEG-based BCI paradigms (*e.g.*, the P300 speller described in Farwell & Donchin, 1988), the imagination tasks used with SMR-based BCIs impose low sensory demands on the user. Furthermore, of particular importance for patients diagnosed as VS who, by definition, are unable to fixate their eyes, SMR BCIs need not involve visual stimulation (Chatelle *et al.*, 2012; Naci *et al.*, 2012; Grosse-Wentrup & Schölkopf, 2013).

Despite the potential benefits of bedside EEG-based BCIs for patients diagnosed as VS and their families, there is substantial intra- and inter-subject variability in BCI performance (Wolpaw *et al.*, 2002; Pfurtscheller *et al.*, 2006; Naci *et al.*, 2012; Grosse-Wentrup & Schölkopf, 2013). In many studies of healthy volunteers and people with severe motor impairments, some individuals are simply unable to reliably regulate the brain signals necessary to operate a BCI without training (*e.g.*, Guger *et al.*, 2003; Wolpaw & McFarland, 2004; Cruse *et al.*, 2011; Hammer *et al.*, 2012). In the current work, we propose modifications to the traditional SMR-based BCI design that may optimize BCI performance for behaviourally non-responsive patients in particular. These modifications apply to the nature of the task used to generate the SMR and the nature of the comparisons made during signal classification (see also Curran & Stokes, 2003; Curran *et al.*, 2004).

In published SMR-based BCI research to date, users are typically instructed to imagine moving their hands, feet, or tongue to generate an SMR (*e.g.*, Neuper & Pfurtscheller, 2001; Kübler *et al.*, 2005; Cruse *et al.*, 2011). With only a few exceptions, users are asked to imagine very simple actions, such as repeatedly squeezing one of their hands into a fist. However, actions that are more complex could result in a more robust and consistent SMR (Curran & Stokes, 2003; Curran *et al.*, 2004). Indeed, there is evidence that complex imagined actions are associated with more robust brain responses than simpler imagined actions. For instance, there is converging evidence from functional magnetic resonance imaging (fMRI), functional near-infrared spectroscopy (fNIRS), and transcranial magnetic stimulation (TMS) studies that complex motor imagery is associated with greater hemodynamic change and higher amplitude motor-evoked potentials than simple motor imagery (*e.g.*, Kuhtz-Buschbeck *et al.*, 2003; Roosink & Zijdwind, 2010; Holper & Wolf, 2011). In the latter studies and in this paper, “complex” motor imagery is defined as tasks that involve sequences of movements (*e.g.*, Kuhtz-Buschbeck *et al.*, 2003; Roosink & Zijdwind, 2010; Holper & Wolf, 2011), and/or more than one body part (*e.g.*, Holper & Wolf, 2011). Based on these findings, we hypothesized that actions involving more than one body part and more complex action sequences would result in more robust SMRs and, consequently, higher classification accuracy than traditional SMR-based BCI imagery tasks.

Additionally, it has been proposed in previous work that asking users to imagine actions which they are familiar with could improve SMR classification (Curran & Stokes, 2003; Curran *et al.*, 2004). In this paper, we chose to explore the role of action familiarity in modulation of the SMR by drawing from samples of experienced athletes and musicians, given that the effects of long-term motor learning have been studied extensively in these groups already (see Münte *et al.*, 2002, and Nakata *et al.*, 2010, for reviews). While imagining actions involving the sport or instrument of their expertise, experienced athletes and musicians produce more focused patterns of brain activation (*e.g.*, Lotze *et al.*, 2003; Milton *et al.*, 2007; Olsson *et al.*, 2008; Wei & Luo, 2010) and report more objectively accurate imagery than novices (*e.g.*, Louis *et al.*, 2012; see also Rieger, 2012). Based on the latter findings, it was expected that familiar imagery would result in a more reliable SMR, and thus higher classification accuracy, than traditional SMR-based BCI imagery. If supported, this hypothesis could be extended to future work with brain-injured patients by selecting imagery tasks based on the skills and hobbies that the patient had prior to their injury. Furthermore, based on the findings reviewed previously regarding the influence of action complexity on brain responses to motor imagery, it was also hypothesized that actions that were both familiar and complex would further enhance the SMR and increase classification accuracy - thereby improving the ability to detect covert awareness in future work with behaviourally non-responsive patients.

Another convention in SMR-based BCI research to date is that comparisons are almost exclusively made between the SMRs generated for various imagined movements (*e.g.*, left hand versus right hand, *etc.*; Neuper & Pfurtscheller, 2001; Guger *et al.*, 2003; Kübler *et al.*, 2005; Pfurtscheller, Brunner, *et al.*, 2006; Cruse *et al.*, 2011). Although these types of comparisons render acceptable classification accuracy in most healthy people (*e.g.*, Guger *et al.*, 2003; Bai *et al.*, 2008), as discussed in Cruse *et al.* (2012), these types of comparisons may not be appropriate for VS patients. Indeed, as part of the standard behavioural assessment tool for VS patients, the Coma Recovery Scale - Revised (CRS-R; Kalmar & Giacino, 2005), awareness is assessed by a patient's ability to produce one type of action following a command. If, on three out of four occasions, the patient is able to successfully follow that command, the patient is diagnosed as, at least, minimally conscious (Kalmar & Giacino, 2005; see also Cruse *et al.*, 2012). In this work, we therefore performed comparisons that are more similar to behavioural assessments of awareness than previous SMR-based BCI paradigms. Specifically, contrasts were made between one imagined action of interest and periods of rest (mind-wandering), rather than between two distinct imagined actions. In addition to providing a potential BCI-based assessment of awareness that is more similar to the standard behavioural method than previous BCIs, we propose that this technique is more practical for behaviourally non-responsive patients because maintaining more than one imagined action in working memory may impose excessive cognitive demands on some patients.

As a final consideration, it is important to note that behavioural assessments of awareness such as the CRS-R (Kalmar & Giacino, 2005) are based on the *reliability* of a patient's ability to follow commands. Patients are evaluated across multiple assessment sessions and multiple attempts to follow the same command in each session in order to ensure that an accurate assessment of awareness is obtained. If neuroimaging-based methods are to be used in clinical assessments of awareness, these methods should maximize the likelihood of detecting a reliable command-following response to reduce the risk of misdiagnosing awareness. SMR-based BCIs

naturally provide an additional measure of reliability that is not automatically available with standard behavioural assessments, *i.e.*, the statistical significance of the classification of the brain response. As described above, when two brain states produced by a patient can be differentiated with statistical significance, it may be inferred that the patient possesses a covert ability to follow commands. Thus, increasing the accuracy of BCI classification not only increases the accuracy of potential communication devices, but also increases the number of patients in whom covert command-following may be detected when it is present.

The research question in this work was whether having users perform motor imagery tasks involving more complex and familiar movements than previous investigations would improve classification accuracy and result in more users (and, in future work, patients diagnosed as VS) with SMRs that could be detected reliably. This question was addressed in three studies of healthy, young adults with an experimental set-up suitable for future clinical work with VS and other non-communicative patients. In Study 1 (Complexity), participants imagined simple hand actions (squeezes) of the sort typically used with SMR-based BCIs alongside other more complex bimanual actions not commonly used with BCIs. It was predicted that, in accordance with prior evidence of increased brain activity during complex motor imagery (Kuhtz-Buschbeck *et al.*, 2003; Roosink & Zijdwind, 2010; Holper & Wolf, 2011), classification accuracy (versus rest) would be higher when the participant was imagining complex actions than when imagining relatively simpler actions. In Study 2 (Familiarity), groups of experienced pianists, experienced ice hockey players, and age-matched controls were instructed to imagine completing hand squeezes and actions from hockey and piano. It was predicted that classification accuracy between rest and imagery would be highest for the athletes and musicians in the action with which they were most familiar (*i.e.*, pianists imagining playing the piano, and hockey players imagining playing hockey; see Lotze *et al.*, 2003; Fourkas *et al.*, 2008; Olsson *et al.*, 2008; Wei & Luo, 2010). In Study 3 (Complexity and Familiarity), the experienced pianists from Study 2 imagined playing one simple piece of music and one relatively more complex piece of music on the piano. It was expected that classification accuracy would be highest for the complex piece versus rest comparison. Implications and modifications for future BCI work with VS patients are discussed.

2. Methods

2.1. Ethics Statement

All participants gave informed written consent. The Psychology Research Ethics Board of the Western University (London, ON, CAN) provided ethical approval for the studies.

2.2. Participants and Stimuli

2.2.1. Study 1 (Complexity)

Sixteen healthy, right-handed young adults participated in the complexity study (five men; age range=17-20 years; median age of 18 years). For the simple imagined movement phase of the complexity study (Study 1), the participants were instructed to imagine repeatedly squeezing their left hand, right hand, or both hands following the auditory cues of “left”, “right”, and “both”, respectively. For the complex imagined movements phase of Study 1, the participants were instructed to imagine either playing the guitar, clapping their hands, or juggling

using both hands. These tasks were cued with the words “guitar”, “clap”, and “juggle”, respectively. In each task phase, participants were also asked to cease the previously-cued mental imagery and mind-wander following the cue “relax”. The order of the simple and complex imagined movement phases were counter-balanced across participants. All auditory instruction stimuli were 1 second in length.

2.2.2. Study 2 (Familiarity)

Forty-eight healthy, right-handed young adults participated in the familiarity study. Sixteen participants were experienced ice hockey players (seven men; age range=18-29 years; median age of 20 years); sixteen participants were experienced pianists (six men; age range=18-29 years; median age of 20 years); and sixteen participants had either limited or no experience playing the piano or hockey (eight men; age range=18-28 years; median age of 18 years). All hockey players had played regular, competitive ice hockey for at least ten years, and all pianists had formal musical training and had played and practiced piano regularly for at least ten years. There were no significant differences in mean age of first play experience, mean years of total play experience, or mean self-reported hours of regular play per week between the groups of athletes and musicians, pairwise $ps > .51$ (Bonferroni correction; see Table 1). As shown in Table 1, the three groups also did not differ significantly in mean age, sex, handedness (Oldfield, 1971), or imagery ability (Gregg *et al.*, 2010; $ps > .34$). All participants were instructed to imagine making a slap shot (a bimanual action from hockey), playing a musical piece on the piano using both hands, or squeezing their right hand into a fist following the auditory cues of “hockey”, “piano”, and “right hand”, respectively. As in Study 1, participants were asked to mind-wander following the cue “relax”, and all instructions were 1 second in length.

2.2.3. Study 3 (Complexity and Familiarity)

The experienced pianists ($n=16$) from Study 2 completed Study 3 in the same EEG recording session. In Study 3, the pianists were instructed to imagine playing ascending and descending C-major scales and B-major arpeggios over two octaves using both hands following the auditory cues of “scale” and “arpeggio”. The pieces of music were selected based on the curriculum of the Royal Conservatory of Music (RCM), which is a prominent music education institution in Canada. In the RCM curriculum, piano students are evaluated on scales and the key of C-major from the 1st grade level, arpeggios from the 4th grade level, and the key of B-major from the 7th grade level (Royal Conservatory of Music, 2008). Given the different grade levels at which the C-major scale and B-major arpeggio are evaluated in the RCM curriculum, one can conclude that the B-major arpeggio represents a more difficult, *i.e.*, complex, action than the C-major scale. It is also important to note that all the pianists in this study reported high familiarity with both pieces and were able to recall both pieces from memory. Finally, the pianists were also instructed to mind-wander following the auditory cue of “relax”, and all instructions in Study 3 were 1 second in length. The Study 3 procedure was always conducted following the Study 2 procedure in order to prevent pianists from selecting the particular musical pieces from Study 3 for the piano action in Study 2.

2.3. Procedure

Before the EEG recording session, each participant completed a series of short questionnaires. All participants completed the Edinburgh Handedness Inventory (Oldfield, 1971) and the Movement Imagery Questionnaire - Revised Second Version (Gregg *et al.*, 2010).

Participants in Studies 2 and 3 also completed a questionnaire regarding their experiences playing hockey, piano, and other sports and instruments. At the conclusion of Studies 2 and 3, participants rated the vividness of their imagined actions using a 5-point Likert scale (see Table 1).

All auditory cues were pre-recorded by one female speaker and presented to the participant using ER-1 insert earphones (Etymotic Research Inc., Elk Grove Village, IL). Each trial began with an auditory cue and was followed by 5 to 8 seconds of silence before the onset of the next auditory cue. The duration of the silent interval was selected randomly from a uniform distribution on each trial. Studies 1 and 2 were completed in four blocks of 48 trials (12 trials of each instruction per block); Study 3 was completed in three blocks of 48 trials (16 trials of each instruction per block) as there were only three (rather than four) trial types in the latter task; and each block of 48 trials was approximately six minutes in duration. All trials were presented in a pseudorandom order so that no more than two cues of the same type were presented consecutively, and the first trial of each block was always an imagined action trial (rather than a 'relax' trial). Participants were provided with short breaks between blocks in order to reduce fatigue. Participants were also instructed to imagine completing each action repeatedly from the offset of the auditory cue to the onset of the next auditory cue in order to account for potential differences in the duration of the imagined actions. To reduce ocular artefacts, participants were instructed to keep their eyes closed throughout the tasks.

2.4. Electrophysiological Data Acquisition and Pre-Processing

In all three studies, EEG was recorded using the g.Gamma active electrode system (g.tec Medical Engineering GmbH, Austria). In Study 1, EEG was recorded with a four-channel montage housed in an electrode cap; the electrodes were placed at sites CP3, FC3, CP4, and FC4 (Sharbrough *et al.*, 1991). In Studies 2 and 3, EEG was recorded from the same four scalp sites as in Study 1, and additional electrodes were placed at sites TP7, FT7, CPz, FCz, TP8, and FT8 (Sharbrough *et al.*, 1991). The reported analyses for Studies 2 and 3 consist of data from only the four electrodes used in Study 1, as this montage has been previously shown to provide robust SMR recordings (Guger *et al.*, 2003; Cruse *et al.*, 2012). In all three studies, bipolar surface electromyographic (EMG) recordings were obtained from both forearms on the ventral surface (placed over the *flexor digitorum profundus*) in order to detect overt movements. Online, the EEG data were filtered from 0.5 to 60 Hz with a 60 Hz notch filter using an infinite impulse response (IIR) digital Butterworth filter set using the g.USBamp graphical user interface (GUI). The EMG data were filtered from 5 to 250 Hz with a 60 Hz notch filter using the g.USBamp GUI. The EEG recordings were referenced to the right earlobe with a forehead (Fpz) ground, and the right elbow (*olecranon*) was used for the EMG ground. The EMG and EEG data were sampled at 600 Hz with impedances kept below 5 k Ω for the EEG recording.

Offline, the EEG data were down-sampled to 100 Hz, filtered between 0.5 and 40 Hz using the EEGLAB function 'pop_eegfilt', and segmented into 6-second epochs time-locked to the onset of the auditory cue. The EEGLAB filter function consisted of a two-step least-squares finite impulse response (FIR) filter; in the first step, data were filtered with a high-pass cut-off of 0.5 Hz, and in the second step, data were filtered with a low-pass cut-off of 40 Hz. The EMG data were rectified and then filtered with a 10 Hz high pass filter using the same EEGLAB least-squares FIR filter function described previously. Trials containing physiological artefacts,

including overt hand movements as evident from the EMG, were identified by visual inspection and removed. After artefact rejection, the median number of trials included in each imagery and rest condition per participant was: Study 1 - 40 (range: 29-48); Study 2 - 43 (range: 27-48); and Study 3 - 43 (range: 28-48). Finally, the EEG data were re-referenced offline to form two bipolar channels (FC3-CP3, FC4-CP4) that are subsequently identified as C3' and C4', respectively; this bipolar approach is known to detect changes in mu and beta power with high accuracy across many people (Cruse *et al.*, 2012).

2.5. EEG Single-Trial Classification Procedure

A machine-learning algorithm was used for single-trial classification of the EEG data as described in Cruse *et al.* (2011, 2012). For these analyses, the log bandpower values of four frequency bands at electrodes C3' and C4' were the classification features. Based on previous work (Cruse *et al.*, 2011, 2012), the frequency bands were 7-13 Hz (mu), 13-19 Hz (low-beta), 19-25 Hz (mid-beta), and 25-30 Hz (high-beta), for a total of eight features per classification analysis (two electrodes x four frequency bands). For the single-trial analyses, the spectral power in each band was estimated with a sliding window of 1-second (as recommended by Pfurtscheller & Lopes da Silva, 1999) moving in 50 ms steps using a short-time Fourier transform (MATLAB function 'spectrogram').

Classification of each imagined action (*e.g.*, squeezing the right hand in Study 1, *etc.*) and the corresponding rest condition was performed using a naïve Bayes classifier (MATLAB's 'naivebayes' object). Each classification analysis was conducted using ten-fold cross-validation. For the cross-validation procedure, each participant's trials for one type of imagined action and the rest condition from the same experiment were separated into ten approximately equal groups. The naïve Bayes classifier was trained on the features of nine of these groups ('training'), and then the class of each trial in the tenth group was predicted in order to calculate the classifier's accuracy ('testing'). Specifically, during training, the naïve Bayes classifier estimated the parameters of a probability distribution per training feature per class; the parameters were the mean and standard deviation of a normal distribution, the training features were bandpower per frequency band at each electrode, and the two classes were the rest and imagery trial types. Using Bayes' Theorem during testing, the features of the test trials were used to calculate the posterior probabilities for each class, and then each test trial was placed in the class with the highest posterior probability (for more information regarding naïve Bayes classification, see Jiang *et al.*, 2007). The classification procedure was repeated ten times so that each trial served as a test trial in exactly one of the ten cross-validation folds. The average classification accuracy across the ten folds was then calculated at each time-point, and the time-course of the cross-validated classification accuracy was smoothed with a sliding-window of 500 ms to control for outliers (*cf.* Cruse *et al.*, 2012).

To determine the statistical significance of the classification accuracy, a permutation test with 1000 repetitions that controlled for familywise error was used (Maris, 2004; see also Cruse *et al.*, 2012). For each permutation, the class labels of imagery or rest were randomly shuffled across trials, and the cross-validated classification procedure described previously was repeated. The maximum smoothed accuracies across all time-points from each of the 1000 repetitions were used to form a distribution representing the expected classification results if the classifier were operating at chance (the null hypothesis). The classification accuracy obtained for the

participant's original data (*i.e.*, the data with the correct trial labels) was then evaluated against this distribution to calculate a familywise error-corrected significance value for the original classification results at each time-point. Finally, to control for the multiple comparisons of bandpower (*i.e.*, one comparison for each time point of imagery versus rest), a control of False Discovery Rate (FDR) approach was used (implemented via MATLAB's 'fdr' function; Benjamini & Hochberg, 1995; Verhoeven *et al.*, 2005). The control of FDR approach is known to reduce the risk of Type I error without requiring as stringent reductions in power as Bonferroni procedures (see Verhoeven *et al.*, 2005).

2.6. EEG Spectral Analyses

In addition to the single-trial classification analyses of the data, the EEG data from all three studies were analyzed using the same spectral analysis procedure reported in Cruse *et al.* (2012). For each time-point at C3' and C4', spectral power estimates were calculated using a Hanning window (1-second) time-frequency transformation via the 'ft_freqstatistics' function of the open-source MATLAB toolbox, FieldTrip (Oostenveld *et al.*, 2011). The time-frequency data at both electrodes were then compared between the imagined movements and rest using cluster-based permutation testing (*cf.* Maris & Oostenveld, 2007; Cruse *et al.*, 2012) implemented via FieldTrip. For the cluster-based testing, the time-frequency data for a given imagery condition and rest (or another imagery condition; see section 2.7 for a description of the comparisons for each study) were log-transformed and then compared at each data point by a paired-samples *t*-test. All significant data points ($p < .025$) were then arranged into groups, *i.e.*, clusters, based on their temporal and spectral proximity to each other, and the sum of the *t* values was calculated for each cluster. To determine the familywise error-corrected significance value for each summed *t*-value cluster, a Monte Carlo randomization test that controlled for familywise error was used. In the randomization test, the condition labels were randomly permuted to remove task-related differences, and the clustering procedure was repeated 1000 times. The maximum summed *t*-value clusters from each repetition were used to form a distribution, and this distribution was then used to test the null hypothesis that the original summed *t*-value cluster (*i.e.*, the summed *t*-value cluster computed from the data with the correct trial labels) occurred by chance.

2.7. Group-level Statistical Analyses

For the single-trial analyses of the EEG data (see Table 2 and Figure 1), all group-level statistical analyses were conducted using IBM SPSS Statistics version 21.0, and where applicable, the Dunn-Sidak correction was used for the follow-up tests. For the spectral analyses of the EEG data (see Figures 2-4), all statistical analyses were conducted using the cluster-based permutation testing described previously (*cf.* Maris & Oostenveld, 2007; Cruse *et al.*, 2012) via custom MATLAB script using the open-source toolbox, FieldTrip (Oostenveld *et al.*, 2011).

For the group comparisons of single-trial analyses of the Study 1 and Study 3 data, several parametric and non-parametric repeated-measures statistical tests were used. Two paired-samples *t*-tests were used to compare maximum classification accuracy and the time at which maximum classification accuracy occurred relative to the onset of the instruction in both Studies 1 and 3. Wilcoxon Signed Rank Tests were used to compare the number of time-points for which a significant classification was obtained in both studies, and this test was also used to compare the self-reported vividness ratings of the imagined actions from the pianists between Study 2 and

Study 3. An exact (rather than asymptotic) calculation of the p -value was used with the test of time-points in Study 1 to account for the positive skew of the count data (given that many participants had zero significant time-points, especially in the simple complexity condition). Finally, the number of trials included in each complexity condition of each study was compared using the Friedman test.

To further illustrate the difference in terms of the significance of the single-trial analyses between the complexity levels in Studies 1 and 3, participants were assigned to a binary category based on whether or not at least one imagined movement in each complexity condition was classified significantly from rest for the participant (0 = no significant classifications). The number of participants with at least one significant classification and the number of participants with no significant classifications in each complexity level were then compared using Fisher's Exact Test.

For Study 2, three 3 (Group: Pianist, Hockey, Control) x 3 (Action: Play Piano, Slap-shot, Squeeze) mixed analyses of variance (ANOVAs) were used to compare the averaged maximum classification accuracies, the time at which the maximum accuracy occurred, and the total number of trials included in each condition (*N.B.*, the trial numbers were rank-transformed to meet the statistical assumptions of the ANOVA). Additionally, a Kruskal-Wallis test was performed to compare the self-reported imagery vividness ratings between the groups. Given the lack of significant differences in classification accuracy in Study 2, no comparisons were made for the number of significant time points in the imagery versus rest comparisons.

The group spectral analyses were conducted with the time-frequency data averaged across all trials in each condition per participant. For Study 1, the time-frequency data for each participant were averaged across all of the imagined actions in each complexity level, and the cluster-based permutation testing was conducted between each complexity level and rest, and between the two complexity levels (Figure 2). For Study 2, comparisons were made separately for each familiarity group between each imagery condition and rest (Figure 3). Finally, for Study 3, comparisons were made between each of the two imagery conditions and rest, and between the two imagery conditions (Figure 4). Although we were primarily concerned with the imagery versus rest comparisons because we posit that these comparisons are more practical for future BCI work with brain-injured patients, we conducted *post-hoc* comparisons between the imagery conditions in the spectral analyses of the Studies 1 and 3 data to attempt to identify the neurophysiological correlates of the complexity effect observed in the single-trial analyses of Study 3. We did not make comparisons between the familiar and control imagery conditions in Study 2 because there was no evidence of an effect of action familiarity in the single-trial analyses.

3. Results

3.1. Study 1 (Complexity)

In terms of the single-trial analyses of the EEG data, there was a trend for classification accuracy (imagery versus rest) to be higher for the complex imagined actions, $t(15)=-1.963$, $p=.068$, $d=0.49$ (Simple: $M=60.68\%$, $SE=0.74\%$; Complex: $M=62.74\%$, $SE=0.93\%$). There was no significant difference for the time at which the maximum classification accuracy occurred

between the two complexity conditions, $p=0.29$. From the familywise permutation tests, there were significantly more time-points at which significant classifications were obtained in the complex condition than in the simple condition, $Z=-2.197$, exact $p=.026$, $r=.55$ (Simple: median of 0 significant time-points [range: 0-26]; Complex: median of 14.5 significant time-points [range: 0-31]). It is also worth noting that there was no significant difference in the number of trials in each complexity level, $p=.29$.

There was some variability between and within subjects for the single-trial analyses of the Study 1 EEG data. While at least one simple imagined action type was classified significantly from rest for only four of the 16 participants (25%), at least one complex imagined action type was classified significantly from rest for significantly more participants (11 of 16, or 69%), Fisher's exact $p=.032$ (two-tailed). Of the eleven participants who produced significant responses for the complex imagery versus rest comparisons, three participants produced significant responses for all three of the complex actions; five participants produced significant responses for two of the three complex actions; and three participants produced a significant response for only one of the complex actions. In the simple imagery comparisons, two of the four participants produced significant responses for all three of the simple actions; one participant produced significant responses for two of the simple actions; and one participant produced a significant response for only one of the simple actions. Furthermore, there were some participants in the sample who did not produce any significant brain responses for any of the imagined actions (four of the sixteen participants, or 25% of the sample). Only one participant in the sample produced significant brain responses for every imagined action in the simple and complex task phases. The inter- and intra-subject variability in classification accuracy is summarized in Table 2 and detailed in the supplementary data tables (Table A1).

From the group spectral analyses of the EEG data (Figure 2), there were statistically significant ERDs over the left hemisphere in the low-beta band in the complex imagery versus rest comparison ($ps<.014$). In the simple imagery versus rest comparison, there was an ERD over the right hemisphere in the mid-beta band that approached statistical significance ($p=.050$). Additionally, there were no significant clusters in the simple imagery versus complex imagery comparisons ($ps>.10$).

Even though the role of action familiarity was not explicitly examined in Study 1, there was one interesting finding in this experiment that emphasized the importance of this factor and its potential influence on single-subject performance. In Figure 1, we depict the time-course of the single-trial classification accuracies for one of two experienced guitarists who participated in Study 1. In line with the group trends already reported, these participants did not produce significantly classifiable SMRs for any of the simple imagined actions; the averaged, maximum classification accuracy obtained for these participants on the simple imagined actions was 57.58% ($SE=1.01\%$), and the classification results for the simple imagery versus rest comparisons were not statistically significant at any time-point for either participant. Both participants produced SMRs for the instruction to imagine clapping (one of the actions in the complex imagery condition) that were significantly classifiable from rest for a short time; significant classification results occurred for 6-7 time-points, and the smoothed maximum classification accuracy for each participant was 64.79% ($SE=3.52\%$) and 64.47% ($SE=4.04\%$; shown in Figure 1). Most interestingly, however, both participants produced a markedly robust

SMR for the instruction to imagine playing the guitar that was significantly classifiable from rest for most of the epoch (68-69 time-points) with very high accuracy (maximum accuracies of 81.10%, $SE=3.44\%$, as shown in Figure 1, and 71.28%, $SE=3.97\%$, no figure provided).

3.2. Study 2 (Familiarity)

In terms of the single-trial analyses of the EEG data, there were no significant differences in accuracy for any of the imagined action versus rest comparisons, or for any group on any of the imagined action versus rest comparisons, $ps>.44$. The main effect of group on accuracy approached significance ($p=.054$), and this was driven by the relatively low overall classification accuracy of the hockey players ($M=59.70\%$, $SE=1.03\%$) compared to the control group ($M=63.52\%$, $SE=1.24\%$; pairwise $p=.054$). There were no significant differences in terms of the time at which maximum classification accuracy occurred for any imagined action or any group by imagined action type, $ps>.64$, and there was also no significant difference in the number of trials included in any of the imagined action types or rest conditions on average or by group, $ps>.41$. The three groups also did not differ in their self-reported vividness ratings of the imagined actions, $p=.34$, and in motor imagery ability as measured by the MIQ-RS ($ps>.545$; Gregg *et al.*, 2010; see Table 1).

From the group spectral analyses of the EEG data (Figure 3), there were significant ERDs for the familiar imagery versus rest comparisons for both the experienced pianists ($ps<0.015$) and the experienced hockey players ($ps<0.019$). Although the ERDs were significant bilaterally (rather than unilaterally) and over a longer period of time for the pianists, the significant ERDs were similar between the hockey players and pianists for the familiar imagery, in that both ERDs featured significant clusters in both the mu and low-beta bands. However, the pianists also had significant ERDs for the hockey imagery ($ps<.017$) and simple imagery ($ps<.017$) versus rest comparisons, and the hockey players had an ERD that approached statistical significance ($p=.036$) in the simple imagery versus rest comparison. Furthermore, the control group (which consisted of age-matched individuals without significant experience playing hockey or piano) also produced significant ERDs for both the piano ($p=.021$) and hockey imagery ($p=.002$) versus rest comparisons, though these ERD clusters were smaller in their spatiotemporal extent than those generated by pianists. There were no significant clusters in any of the other comparisons ($ps>0.08$). Thus, much like the single-trial analyses, the spectral analyses do not provide strong evidence that there was an advantage in terms of SMR detection for any group regardless of their familiarity (or lack thereof) with the imagined actions.

3.3. Study 3 (Complexity and Familiarity)

In terms of classification accuracy, there was a significant advantage for the complex imagined action (Complex: $M=69.60\%$, $SE=2.03\%$) compared to the simple imagined action for the pianists-only study (Simple: $M=66.34\%$, $SE=1.96\%$; $t(15)=-2.589$, $p=.021$, $d=0.65$). Furthermore, there was an advantage for the complex imagined action compared to the simple imagined action in that significantly more time-points were classified significantly from rest in the former comparison, $Z=-2.510$, $p=.009$ (two-tailed), $r=.63$ (Simple: median of 9.0 significant time-points [range:0-76]; Complex: median of 32.0 significant time-points [range:0-74]). There was no significant difference between the number of participants with significant imagery versus rest comparisons in the two complexity levels, Fisher's exact $p=0.25$. It is important to note that there was no significant difference in terms of when the maximum classification accuracy

occurred relative to the onset of the auditory cue for the two imagined actions, $p=.72$, and it is also worth noting that the differences between the complexity conditions were not driven by a difference in the number of trials in any condition, $p=.16$, or by a difference between the self-reported vividness of the imagined actions in Study 3 compared to Study 2, $p=.10$.

The results of the group spectral analyses are shown in Figure 4 for each comparison. Compared with rest, significant ERDs ($ps<.017$) occurred in the mu and low-beta bands over the left hemisphere and in the low-, mid-, and high- beta bands over the right hemisphere beginning approximately one second after the offset of the instruction to imagine playing the simple piece. For the complex imagery versus rest comparison, a similar response was observed with the same time-course, although the ERDs were significant ($ps<.011$) over both hemispheres throughout the mu band, and throughout the low- and mid-beta bands. In a comparison of the two imagery conditions, there was more of a desynchronization bilaterally for the complex imagery in the mu and low-beta bands that approached statistical significance ($ps=.044$ [C3'] and $.038$ [C4']).

4. Discussion

In this paper, we presented a series of three experiments in which movement complexity and familiarity with movements were manipulated in a SMR-based imagery paradigm. The purpose of this work was to increase the likelihood of detecting reliable and robust SMRs, and thereby improve how well covert command-following can be detected in behaviourally non-responsive patients. In future work, we will apply these manipulations to SMR paradigms used with BCIs intended for communication with severely brain-injured patients, including individuals diagnosed as VS.

In Study 1, imagery of a range of bimanual sequences of actions (“complex imagery”) resulted in SMRs that were classified from rest with similar accuracy as imagery of the simple hand squeezes typically used with SMR-based BCIs. There was a group trend that the complex actions used in this task were classified with higher accuracy than the simple hand squeezes typically used with SMR-based BCIs, although this result did not reach statistical significance ($p=.068$). Furthermore, we found a significant advantage for the complex actions in that the SMRs for these actions were classified significantly from rest for a longer period of time than the simple actions. We also found that significantly more single subjects produced significantly classifiable SMRs for at least one of the complex actions than for at least one of the simple actions. Overall, the findings of Study 1 align with our prediction based on previous work that there would be an enhancement of the brain response for the complex imagery (*e.g.*, Kuhtz-Buschbeck *et al.*, 2003; Roosink & Zijdwind, 2010; Holper & Wolf, 2011), and support our hypothesis that motor imagery that involves more than one body part and sequences of actions is likely to improve the ability to detect covert command-following or communication in future work.

Another interesting finding from Study 1 was the between- and within-subject variability for the various imagined actions. Nearly half of the sample (seven or eight of the sixteen participants; see Table 2) produced a SMR that was significantly classifiable from rest for each of the complex imagined actions, while only one to four participants produced a SMR that was significantly classifiable from rest for each of the simple imagined actions. We posit that, at the

group-level, the advantage of the complex imagery was not statistically significant for classification accuracy because of this variability. In other words, there was no advantage for the complex imagery at the group-level because most participants (eight of the eleven participants with significant complex imagery versus rest comparisons) only produced significant responses for one or two of the complex imagined actions, rather than for all three of these actions. The latter observation is well-illustrated anecdotally by the classification results of two guitarists who participated in Study 1 (see section 3.1 and Figure 1). There was a marked improvement in classification accuracy when the guitarists imagined playing the guitar as part of the complex imagery phase of the experimental procedure (maximum accuracy: 81.10%, $SE=3.44\%$; see Figure 1 and section 3.1) compared with all of the other imagined actions (maximum accuracy range=54.27-64.79%). Indeed, both guitarists only produced a significantly classifiable brain response for one of the other two complex imagined actions, and for none of the simple imagined actions.

We designed Studies 2 and 3 to follow up on the group trend of the advantage for the complex actions in Study 1 and to explore the role of action familiarity in the imagery paradigm. Study 2, which included groups of experienced pianists, experienced ice hockey players, and age-matched controls, did not meet our expectations in that there was no advantage for any of the imagined actions for any group, regardless of their familiarity with the imagined actions (*e.g.*, pianists imagining playing piano, *etc.*). Compared to novices, experienced athletes and musicians utilize fewer regions of the brain when imagining actions involving the sport or instrument with which both groups have familiarity (*e.g.*, Lotze *et al.*, 2003; Fourkas *et al.*, 2008; Olsson *et al.*, 2008; Wei & Luo, 2010). Although expert brain responses to familiar imagery are fairly consistent within and across individuals (*e.g.*, Langheim *et al.*, 2002), these responses did not result in an enhancement of the SMR in Study 2.

Interestingly, in Study 3, we found an advantage in SMR classification for specific, familiar actions (performance of two musical pieces), such that classification accuracy was higher and significant over a longer period of time when the experienced pianists from Study 2 imagined playing the more complex musical piece. Moreover, the mu and beta ERDs from the spectral analyses of Study 3 were statistically significant for the longest period of time and associated with the largest log-ratio differences in power of all the imagery conditions with significant ERDs across all three experiments (see Figures 2-4). As an exploratory *post-hoc* test, we also compared the classification accuracy for the pianists on the instruction to imagine playing a musical piece of their choice on the piano (versus rest) in Study 2 with the classification accuracy for imagined performance of each of the two specific musical pieces (versus rest) in Study 3. This analysis resulted in a significant effect of movement type, $F(15)=16.016$, $p=.001$, $\eta_p^2=.361$, that was driven by the significantly lower classification accuracy for the piano imagery in Study 2 compared to the complex piano imagery in Study 3 ($p=.003$; Complex Musical Piece [Study 3]: $M=69.60\%$, $SE=2.03\%$; Musical Piece of Choice [Study 2]: $M=63.31\%$, $SE=1.49\%$; other pairwise $ps>.06$). This finding provides additional support for our claim that it is the imagery of a *specific*, *complex*, and *familiar* action that leads to an advantage in SMR classification, rather any of these three properties individually.

A combination of factors likely contributed to the finding that the specific, complex, familiar imagery from Study 3 resulted in the most robust brain responses in this work. The

finding that the complex imagery resulted in an enhancement of the SMR compared to the simple imagery within Study 3 aligns well with previous work (*e.g.*, Kuhtz-Buschbeck *et al.*, 2003; Roosink & Zijdewind, 2010; Holper & Wolf, 2011). Moreover, there was more potential variability in the brain responses between participants in the piano imagery for Study 2 versus Study 3 due to variations in the particular pieces that each person chose to imagine in Study 2; this variability likely resulted in less consistent and less robust brain responses between individuals, regardless of their familiarity with the piano. On a related note, it is highly unlikely that the specificity of the instructions alone underlies the advantage of the Study 3 imagery, given that all of the other imagery tasks in this work also involved specific instructions (*e.g.*, imagine squeezing your right hand, *etc.*). Playing the piano involves highly temporally and spatially complex movements (see Zatorre *et al.*, 2007, for a review), and analogous finger-sequencing actions that do not require prior knowledge of the piano are also associated with more robust brain responses than less temporally and spatially complex actions (*e.g.*, Bengtsson *et al.*, 2004). It is thus possible that the piano performance imagery was simply more conducive to an enhanced SMR than the other imagined actions in this work, but that this advantage was only evident when the actions were well-specified and performed by individuals who were highly familiar with the actions. Although the particular imagined actions from Study 3 are not appropriate for non-musicians, these findings nevertheless provide some criteria that can be generalized to other imagery tasks for future work with SMR-based BCIs.

In summary, the three experiments in this work provide three important findings regarding the roles of action familiarity and complexity in the EEG correlates of motor imagery. Firstly, allowing an individual to imagine performing a task that involves bimanual sequences of actions can result in more robust brain responses from some individuals, as illustrated in Figure 1, and these modified imagery tasks do not impair performance compared to the hand-squeeze imagery typically used with SMR-based BCIs. Secondly, a familiar action may not always lead to a more robust SMR during motor imagery than other actions (Study 2), but, thirdly and lastly, an action that is both familiar and involves sufficiently complex and well-specified actions can lead to an enhanced and more sustained SMR during motor imagery (Study 3). Furthermore, we speculate that the role of action familiarity in modulation of the SMR is that individuals must have some experienced executing an action in order to perform motor imagery of that action reliably (see Olsson & Nyberg, 2010), but other factors, such as the complexity of the particular imagined action, determine how robust the brain response will be for that action. In fact, in studies of short-term motor learning, brain responses to motor imagery are enhanced following overt practice of novel actions (*e.g.*, Lacourse *et al.*, 2005; Baeck *et al.*, 2012). For this reason, although this work only examined familiar imagery among highly experienced athletes and musicians, we propose that it is worthwhile to select an imagery task based on a person's skills and interests to attempt to ensure a reliable brain response during motor imagery, regardless of the person's level of expertise in executing the action. Most importantly, we posit that the subtle but important changes in task instructions proposed here may provide benefits to those individuals who are unable to control a conventional SMR-based BCI (*e.g.*, Figure 1), given the substantial variability between and within subjects in previous work (*e.g.*, Hammer *et al.*, 2012).

We therefore conclude that these results provide a framework for modifications of SMR-based BCI paradigms that may be used to detect covert command-following in, and even communicate with, behaviourally non-responsive patients. Based on the findings from the

healthy individuals used in this work, we propose that the imagery tasks used with patients diagnosed as VS in SMR-based BCI paradigms should be customized to each patient whenever possible. Indeed, during conventional behavioural assessment with the CRS-R, the type of command that a patient is asked to follow should be based on the patient's physical capacity in order to maximize the likelihood of detecting a response (Kalmar & Giacino, 2005). Based on the findings of Study 3 and the interesting results depicted in Figure 1, it may be useful to confer with care-givers in order to select an imagery task that involves more than one sensory modality (*e.g.*, auditory imagery) and a specific sequence of actions with which the patient has had some experience. Furthermore, the supplementary motor area (SMA) and other premotor cortical areas are active in both musicians and non-musicians when listening to complex rhythms (*e.g.*, Bengtsson *et al.*, 2009) and imagining familiar melodies (*e.g.*, Halpern & Zatorre, 1999; for a review, see Zatorre *et al.*, 2007). Thus, imagery that involves music and rhythm may also generalize well to non-musicians and behaviourally non-responsive patients (see also Schaefer *et al.*, 2011). In general, customizing the BCI imagery task to suit each patient when possible and utilizing other sensory modalities and sequences of actions in the imagery task should maximize the likelihood of the patient producing a robust, reliable SMR. By extension, this enhanced SMR may increase the chances of correctly identifying patients who are able to follow the imagery commands, and may even allow for some patients to exploit the imagery response for communication. With the appropriate reliability checks, successful use of a SMR-based BCI could become part of a neuroimaging-based assessment of awareness in VS patients, and thereby improve the currently low rate of diagnostic accuracy for this group (Childs *et al.*, 1993; Andrews *et al.*, 1996; Schnakers *et al.*, 2009).

References

- Andrews K, Murphy L, Munday R, Littlewood C. Misdiagnosis of the vegetative state: Retrospective study in a rehabilitation unit. *BMJ* 1996;313:13-16.
- Baek J-S, Kim Y-T, Seo J-H, Ryeom H-K, Lee J, Choi S-M, *et al.* Brain activation patterns of motor imagery reflect plastic changes associated with intensive shooting training. *Behav Brain Res* 2012;234:26-32.
- Bai O, Lin P, Vorbach S, Floeter MK, Hattori N, Hallett M. A high performance sensorimotor beta rhythm-based brain-computer interface associated with human natural motor behavior. *J Neural Eng* 2008;5:24-35.
- Bardin JC, Fins JJ, Katz DI, Hersh J, Heier L a, Tabelow K, *et al.* Dissociations between behavioural and functional magnetic resonance imaging-based evaluations of cognitive function after brain injury. *Brain* 2011;134:769-782.
- Bekinschtein TA, Manes FF, Villarreal M, Owen AM, Della-Maggiore V. Functional imaging reveals movement preparatory activity in the vegetative state. *Front Hum Neurosci* 2011;5:1-5.
- Bengtsson SL, Ehrsson HH, Forssberg H, Ullén F. Dissociating brain regions controlling the temporal and ordinal structure of learned movement sequences. *Eur J Neurosci* 2004;19:2591-2602.
- Bengtsson SL, Ullén F, Ehrsson HH, Hashimoto T, Kito T, Naito E, *et al.* Listening to rhythms activates motor and premotor cortices. *Cortex* 2009;45:62-71.
- Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J Roy Stat Soc B Met* 1995;57:289-300.
- Bernat JL. Chronic disorders of consciousness. *Lancet* 2006;367:1181-1192.
- Chatelle C, Chennu S, Noirhomme Q, Cruse D, Owen AM, Laureys S. Brain-computer interfacing in disorders of consciousness. *Brain Injury* 2012;26:1510-1522.
- Childs NL, Mercer WN, Childs HW. Accuracy of diagnosis of persistent vegetative state. *Neurology* 1993;43:1465-1467.
- Cruse D, Chennu S, Chatelle C, Bekinschtein TA, Fernández-Espejo D, Pickard JD, *et al.* Bedside detection of awareness in the vegetative state: A cohort study. *Lancet* 2011;378:2088-2094.
- Cruse D, Chennu S, Fernández-Espejo D, Payne WL, Young GB, Owen AM. Detecting awareness in the vegetative state: Electroencephalographic evidence for attempted movements to command. *PLoS One* 2012;7:e49933.

- Cruse D, Thibaut A, Demertzi A, Nantes JC, Bruno M-A, Gosseries O, *et al.* Actigraphy assessments of circadian sleep-wake cycles in the Vegetative and Minimally Conscious States. *BMC Med* 2013;11:18.
- Curran E, Stokes MJ. Learning to control brain activity: A review of the production and control of EEG components for driving brain-computer interface (BCI) systems. *Brain Cognition* 2003;51:326-336.
- Curran E, Sykacek P, Stokes MJ, Roberts SJ, Penny W, Johnsrude I, *et al.* Cognitive tasks for driving a brain-computer interfacing system: A pilot study. *IEEE T Neur Sys Reh* 2004;12:48-54.
- Farwell LA, Donchin E. Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials. *Electroen Clin Neuro* 1988;70:510-523.
- Fourkas AD, Bonavolontà V, Avenanti A, Aglioti SM. Kinesthetic imagery and tool-specific modulation of corticospinal representations in expert tennis players. *Cereb Cortex* 2008;18:2382-2390.
- Goldfine AM, Victor JD, Conte MM, Bardin JC, Schiff ND. Determination of awareness in patients with severe brain injury using EEG power spectral analysis. *Clin Neurophysiol* 2011;122:2157-2168.
- Gregg M, Hall C, Butler A. The MIQ-RS: A suitable option for examining movement imagery ability. *Evid-Based Compl Alt* 2010;7:249-257.
- Grosse-Wentrup M, Schölkopf B. A review of performance variations in SMR-based Brain-Computer Interfaces (BCIs). In: Guger C, Allison BZ, Edlinger G, editors. *Brain-Computer Interface Research Berlin, Heidelberg: Springer; 2013. p. 39-51.*
- Guger C, Edlinger G, Harkam W, Niedermayer I, Pfurtscheller G. How many people are able to operate an EEG-based brain-computer interface (BCI)? *IEEE T Neur Sys Reh* 2003;11:145-147.
- Halpern AR, Zatorre RJ. When that tune runs through your head: A PET investigation of auditory imagery for familiar melodies. *Cereb Cortex* 1999;9:697-704.
- Hammer EM, Halder S, Blankertz B, Sannelli C, Dickhaus T, Kleih SC, *et al.* Psychological predictors of SMR-BCI performance. *Biol Psychol* 2012;89:80-86.
- Holper L, Wolf M. Single-trial classification of motor imagery differing in task complexity: A functional near-infrared spectroscopy study. *J Neuroeng Rehabil* 2011;8:34.
- Jennett B. *The Vegetative State. Medical Facts, Ethical and Legal Dilemmas.* Cambridge: Cambridge University Press; 2002.

- Jiang L, Wang D, Cai Z, Yan X. Survey of Improving Naive Bayes for Classification. In: Alhajj R, Gao H, Li X, Li J, Zaiane OR, editors. Proceedings of the Third International Conference of Advanced Data Mining and Applications Harbin, China: Springer; 2007. p. 134-145.
- Kalmar K, Giacino JT. The JFK Coma Recovery Scale — Revised. *Neuropsychol Rehabil* 2005;15:454-460.
- Kübler A, Nijboer F, Mellinger J, Vaughan TM, Pawelzik H, Schalk G, *et al.* Patients with ALS can use sensorimotor rhythms to operate a brain-computer interface. *Neurology* 2005;64:1775-1777.
- Kuhtz-Buschbeck JP, Mahnkopf C, Holzknecht C, Siebner H, Ulmer S, Jansen O. Effector-independent representations of simple and complex imagined finger movements: A combined fMRI and TMS study. *Eur J Neurosci* 2003;18:3375-3387.
- Lacourse MG, Orr ELR, Cramer SC, Cohen MJ. Brain activation during execution and motor imagery of novel and skilled sequential hand movements. *NeuroImage* 2005;27:505-519.
- Langheim FJP, Callicot JH, Mattay VS, Duyn JH, Weinberger DR. Cortical systems associated with covert music rehearsal. *NeuroImage* 2002;16:901-908.
- Lotze M, Scheler G, Tan H-R., Braun C, Birbaumer N. The musician's brain: Functional imaging of amateurs and professionals during performance and imagery. *NeuroImage* 2003;20:1817-1829.
- Louis M, Collet C, Champely S, Guillot A. Differences in motor imagery time when predicting task duration in alpine skiers and equestrian riders. *Res Q Exercise Sport* 2012;83:86-93.
- Lulé D, Noirhomme Q, Kleih SC, Chatelle C, Halder S, Demertzi A, *et al.* Probing command following in patients with disorders of consciousness using a brain-computer interface. *Clin Neurophysiol* 2013;124:101-106.
- Maris E. Randomization tests for ERP topographies and whole spatiotemporal data matrices. *Psychophysiology* 2004;41:142-151.
- Maris E, Oostenveld R. Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Meth* 2007;164:177-190.
- Mason SG, Birch GE. A general framework for brain-computer interface design. *IEEE T Neur Sys Reh* 2003;11:70-85.
- Milton J, Solodkin A, Hlustík P, Small SL. The mind of expert motor performance is cool and focused. *NeuroImage* 2007;35:804-813.

- Monti MM, Vanhaudenhuyse A, Coleman MR, Boly M, Pickard JD, Tshibanda L, *et al.* Willful modulation of brain activity in disorders of consciousness. *New Engl J Med* 2010;362:579-589.
- Multi-Society Task Force on PVS. Medical aspects of the persistent vegetative state (first part). *New Engl J Med* 1994;330:1499-1508.
- Multi-Society Task Force on PVS. Medical aspects of the persistent vegetative state (second part). *New Engl J Med* 1994;330:1572-1579.
- Münte TF, Altenmüller E, Jäncke L. The musician's brain as a model of neuroplasticity. *Nat Rev Neurosci* 2002;3:473-478.
- Naci L, Cusack R, Jia VZ, Owen AM. The brain's silent messenger: Using selective attention to decode human thought for brain-based communication. *J Neurosci* 2013;33:9385-9393.
- Naci L, Monti MM, Cruse D, Kübler A, Sorger B, Goebel R, *et al.* Brain computer interfaces for communication with non-responsive patients. *Ann Neurol* 2012;72:312-323.
- Naci L, Owen AM. Making every word count for nonresponsive patients. *JAMA Neurology* 2013;1-7.
- Nakata H, Yoshie M, Miura A, Kudo K. Characteristics of the athletes' brain: Evidence from neurophysiology and neuroimaging. *Brain Res Rev* 2010;62:197-211.
- Neuper C, Pfurtscheller G. Event-related dynamics of cortical rhythms: Frequency-specific features and functional correlates. *Int J Psychophysiol* 2001;43:41-58.
- Neuper C, Scherer R, Wriessnegger S, Pfurtscheller G. Motor imagery and action observation: Modulation of sensorimotor brain rhythms during mental control of a brain-computer interface. *Clin Neurophysiol* 2009;120:239-247.
- Oldfield RC. The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia* 1971;9:97-113.
- Olsson C-J, Jonsson B, Larsson A, Nyberg L. Motor representations and practice affect brain systems underlying imagery: An fMRI study of internal imagery in novices and active high jumpers. *Open Neuroimag J* 2008;2:5-13.
- Olsson C-J, Nyberg L. Motor imagery: If you can't do it, you won't think it. *Scand J Med Sci Spor* 2010;20:711-715.
- Oostenveld R, Fries P, Maris E, Schoffelen J-M. FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci* 2011;2011:1-9.

- Owen AM. Disorders of consciousness. *Ann Ny Acad Sci* 2008;1124:225-238.
- Owen AM. Detecting consciousness: A unique role for neuroimaging. *Annu Rev Psychol* 2013;64:109-133.
- Owen AM, Coleman MR, Boly M, Davis MH, Laureys S, Pickard JD. Detecting awareness in the vegetative state. *Science* 2006;313:1402.
- Pfurtscheller G, Brunner C, Schlögl A, Lopes da Silva FH. Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks. *NeuroImage* 2006;31:153-159.
- Pfurtscheller G, Lopes da Silva FH. Event-related EEG/MEG synchronization and desynchronization: Basic principles. *Clin Neurophysiol* 1999;110:1842-1857.
- Pfurtscheller G, Neuper C. Motor imagery activates primary sensorimotor area in humans. *Neurosci Lett* 1997;239:65-68.
- Rieger M. Motor imagery in typing: Effects of typing style and action familiarity. *Psychon B Rev* 2012;19:101-107.
- Roosink M, Zijdwind I. Corticospinal excitability during observation and imagery of simple and complex hand tasks: Implications for motor rehabilitation. *Behav Brain Res* 2010;213:35-41.
- Royal College of Physicians Working Group. The permanent vegetative state. *J Roy Coll Phys Lond* 1996;30:119-121.
- Royal Conservatory of Music. Piano Syllabus: The Royal Conservatory of Music official examination syllabus. Mississauga, ON: Frederick Harris Music Company, Limited; 2008. p. 1-142.
- Schaefer RS, Farquhar JDR, Blokland Y, Sadakata M, Desain P. Name that tune: Decoding music from the listening brain. *NeuroImage* 2011;56:843-849.
- Schnakers C, Vanhaudenhuyse A, Giacino J, Ventura M, Boly M, Majerus S, *et al.* Diagnostic accuracy of the vegetative and minimally conscious state: Clinical consensus versus standardized neurobehavioral assessment. *BMC Neurol* 2009;9:35.
- Sharbrough F, Chatrian G-E, Lesser RP, Luders H, Nuwer M, Picton TW. American electroencephalographic society guidelines for standard electrode position nomenclature. *J Clin Neurophysiol* 1991;8:200-202.
- Sorger B, Dahmen B, Reithler J, Gosseries O, Maudoux A, Laureys S, *et al.* Another kind of “BOLD Response”: Answering multiple-choice questions via online decoded single-trial brain signals. *Prog Brain Res* 2009;177:275-292.

Verhoeven KJF, Simonsen KL, McIntyre LM. Implementing false discovery rate control: Increasing your power. *Oikos* 2005;108:643-647.

Wei G, Luo J. Sport expert's motor imagery: Functional imaging of professional motor skills and simple motor skills. *Brain Res* 2010;1341:52-62.

Wolpaw JR, Birbaumer N, McFarland DJ, Pfurtscheller G, Vaughan TM. Brain-computer interfaces for communication and control. *Clin Neurophysiol* 2002;113:767-791.

Wolpaw JR, McFarland DJ. Control of a two-dimensional movement signal by a noninvasive brain-computer interface in humans. *P Natl Acad Sci Usa* 2004;101:17849-17854.

Zatorre RJ, Chen JL, Penhune VB. When the brain plays music: Auditory-motor interactions in music perception and production. *Nat Rev Neurosci* 2007;8:547-558.

Figure Legends

Figure 1. Mean smoothed, cross-validated classification accuracy from the EEG single-trial analyses across time for one subject (an experienced guitar player) from Study 1 by imagery versus rest comparison. Time is measured relative to the offset of the auditory instruction cue. Shaded regions depict ± 1 standard error of the mean (smoothed), and stars denote time-points with statistically significant classification results for the corresponding accuracy time course.

Figure 2. Averaged, group ($n=16$) time-frequency plots from the spectral analyses of the EEG data for Study 1 (Complexity) averaged across the three imagined actions in each imagery condition. The range of power values (log ratio difference) that are plotted is ± 0.6121 . Significant clusters ($p < .014$) are outlined with solid lines; dashed lines highlight a cluster with $p = .050$. Plots on the left and right reflect the left- and right-hemisphere EEG channels (C3' and C4', respectively), as indicated. Frequency (Hz) is indicated on the vertical axis, and time is measured relative to the offset of the instruction.

Figure 3. Averaged, group time-frequency plots from the spectral analyses of the EEG data for Study 2 (Familiarity) by imagery versus rest comparison per familiarity group ($n=16$ per group). The range of power values (log ratio difference) that are plotted is ± 0.6121 . Significant clusters ($p < .021$) are outlined with solid lines; dashed lines highlight a cluster with $p = .036$ (C3' for the simple imagery minus rest comparison for the hockey players). Plots on the left and right in each pair reflect the left- and right-hemisphere EEG channels (C3' and C4', respectively), as indicated. Frequency (Hz) is indicated on the vertical axis, and time is measured relative to the offset of the instruction.

Figure 4. Averaged, group ($n=16$) time-frequency plots from the spectral analyses of the EEG data from Study 3 (Complexity and Familiarity) by comparison. The range of power values (log ratio difference) that are plotted is ± 0.6121 . Significant clusters ($p < .017$) are outlined with solid lines; dashed lines highlight clusters with $p = .044$ (C3') and $p = .038$ (C4'). Plots on the left and right reflect the left- and right-hemisphere EEG channels (C3' and C4', respectively), as indicated. Frequency (Hz) is indicated on the vertical axis, and time is measured relative to the offset of the instruction.