NeuroImage 134 (2016) 142–152

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Reinforcement learning of self-regulated sensorimotor β-oscillations improves motor performance

G. Naros [a,*], I. Naros [a,b], F. Grimm [a,b], U. Ziemann [c,d], A. Gharabaghi [a,b,*]

[a] Division of Functional and Restorative Neurosurgery, University of Tuebingen, Germany
[b] Centre for Integrative Neuroscience, University of Tuebingen, Germany
[c] Department of Neurology and Stroke, University of Tuebingen, Germany
[d] Hertie Institute for Clinical Brain Research, University of Tuebingen, Germany

ARTICLE INFO

Article history:
Received 28 October 2015
Accepted 7 March 2016
Available online 2 April 2016

Keywords:
Operant conditioning
Beta oscillations
Brain–computer interface
Brain–machine interface
Brain–robot interface
Motor learning
Neurofeedback
Hand function
Functional restoration
Stroke

ABSTRACT

Self-regulation of sensorimotor oscillations is currently researched in neurorehabilitation, e.g. for priming subsequent physiotherapy in stroke patients, and may be modulated by neurofeedback or transcranial brain stimulation. It has still to be demonstrated, however, whether and under which training conditions such brain self-regulation could also result in motor gains.

Thirty-two right-handed, healthy subjects participated in a three-day intervention during which they performed 462 trials of kinesthetic motor-imagery while a brain–robot interface (BRI) turned event-related β-band desynchronization of the left sensorimotor cortex into the opening of the right hand by a robotic orthosis. Different training conditions were compared in a parallel-group design: (i) adaptive classifier thresholding and contingent feedback, (ii) adaptive classifier thresholding and non-contingent feedback, (iii) non-adaptive classifier thresholding and contingent feedback, and (iv) non-adaptive classifier thresholding and non-contingent feedback. We studied the task-related cortical physiology with electroencephalography and the behavioral performance in a subsequent isometric motor task.

Contingent neurofeedback and adaptive classifier thresholding were critical for learning brain self-regulation which, in turn, led to behavioral gains after the intervention. The acquired skill for sustained sensorimotor β-desynchronization correlated significantly with subsequent motor improvement. Operant learning of brain self-regulation with a BRI may offer a therapeutic perspective for severely affected stroke patients lacking residual hand function.

Introduction

Lost motor function following brain lesions may limit the relearning of movements when physical practice is no longer possible (Doyon and Benali, 2005; Halsband and Lange, 2006). In such cases, motor imagery (MI) might be an alternative for physical practice (Boe et al., 2014; Halsband and Lange, 2006) since it activates the sensorimotor system without any overt movement (Gao et al., 2011; Szameitat et al., 2012; Vukelić and Gharabaghi, 2015a). This volitional modulation of oscillatory activity during MI can be supported by providing visual and/or proprioceptive feedback about the user's current brain state to facilitate operant learning of oscillatory patterns that are considered beneficial to recovery (Boe et al., 2014; Dobkin, 2004; Gomez-Rodriguez et al., 2011; Vukelić and Gharabaghi, 2015a). When applied together with robotic rehabilitation technology, these neurofeedback tools are also referred to as brain–robot interfaces (BRI; Bauer et al., 2015; Vukelić and Gharabaghi, 2015a). First studies applying neurofeedback in stroke rehabilitation are promising, with the largest clinical gains in the subacute patient population (Pichiorri et al., 2015). There is, however, only very little evidence up to now that these MI-based interventions achieve effects beyond a general *priming* of subsequent physiotherapy, i.e. that they indeed induce *operant conditioning* of the targeted brain states to facilitate task-specific functional gains (Naros and Gharabaghi, 2015; Zoefel et al., 2011).

This ambiguity may be related to various methodological differences of previous studies with regard to the cortical frequency-band trained by event-related desynchronization (ERD), the applied modality and type of feedback, the adaptation strategy and/or application or lack of additional brain stimulation, e.g. transcranial direct current stimulation (tDCS):

Feedback target

Sensorimotor α- (8–12 Hz) and β-frequency (15–35 Hz) bands are both modulated during actual and imagined movements and show a

* Corresponding authors at: Division of Functional and Restorative Neurosurgery, Eberhard Karls University, Otfried-Mueller-Str. 45, 72076 Tuebingen, Germany.
E-mail addresses: georgios.naros@med.uni-tuebingen.de (G. Naros), alireza.gharabaghi@uni-tuebingen.de (A. Gharabaghi).

http://dx.doi.org/10.1016/j.neuroimage.2016.03.016
1053-8119/

highly correlated pattern while serving distinct functional mechanisms (Brinkman et al., 2014; Kilavik et al., 2013; van Wijk et al., 2012). Previous studies implemented feedback of the α-band or of individual frequency bands with optimal classification properties, i.e. the bands that separated the rest and the task condition most effectively (Ang et al., 2011, 2014b; Buch et al., 2008, 2012; Ramos-Murguialday et al., 2013; Shindo et al., 2011; Pichiorri et al., 2015). Alternative approaches applied β-ERD feedback (Naros and Gharabaghi, 2015; Bauer et al., 2015; Brauchle et al., 2015; Vukelić and Gharabaghi, 2015a,b; Kraus et al., 2016a) on account of its physiological role in disinhibiting the sensorimotor cortex and in mediating coherent interaction with the muscles (Aumann and Prut, 2014; Kilavik et al., 2013; Kristeva et al., 2007; Mima et al., 2000; van Wijk et al., 2012). In the present study, we chose the latter approach in a bid to achieve behavioral gains through improved cortico-muscular communication (Kraus et al., 2016a,) and not to maximize the classification accuracy of the training device (Spüler et al., 2014), which might be better achieved by the former approach. However, the frequencies of cortico-muscular beta-band interactions can vary individually. Their distribution (Rossiter et al., 2012) and peak frequency (Fang et al., 2009) are known to be altered (or even absent) in stroke patients in comparison to healthy subjects. Operant conditioning of individually determined markers in stroke patients with a persistent deficit (the patient population which we intend to address in future) might therefore reinforce a pathological (or at least non-physiological) pattern. We have thus decided to explore the feasibility of training predefined biomarkers, i.e. fixed frequencies in the beta-band (17–23 Hz), in this study with healthy subjects. Albeit reduced following stroke (Rossiter et al., 2014), this frequency band is known to be physiologically relevant during unimpaired movement, i.e. by mediating the disinhibition of the sensorimotor cortex and the coherent interaction with the muscles (Mima et al., 2000; Kristeva et al., 2007; van Wijk et al., 2012; Kilavik et al., 2013; Aumann and Prut, 2014), thereby acting as a suitable biomarker for motor learning.

Feedback modality

Providing visual feedback of MI-associated β-oscillations was recently shown to increase the laterality at the targeted brain regions (Boe et al., 2014) and the movement-associated desynchronization of the β-frequency band (Bai et al., 2014). Proprioceptive feedback of MI-associated β-oscillations activated a distributed cortical network (Vukelić et al., 2014). Furthermore proprioceptive feedback, and not motor imagery alone, has been shown to address the abilities and cortical networks related to motor execution, thereby providing a potential backdoor for training the motor system in those patients for whom physical practice is no longer an option due to lost motor function (Bauer et al., 2015). A direct comparison of these two feedback modalities revealed that proprioceptive input facilitated decoding of MI-induced brain states (Gomez-Rodriguez et al., 2011), self-regulation of β-oscillations and entrainment of cortical motor networks (Vukelić and Gharabaghi, 2015a). We therefore applied proprioceptive feedback with a BRI in the present study to approximate actual physical practice as closely as possible.

Transcranial brain stimulation

Anodal transcranial direct current stimulation (tDCS) of the motor cortex increased in a polarity-specific way both α- and β-band ERD of the motor cortical network in a subsequent motor task (Notturno et al., 2014). Similarly, when examining the α-band only, anodal tDCS – not cathodal or sham stimulation – increased sensorimotor ERD in a subsequent motor *imagery* task (Matsumoto et al., 2010). When a BRI task similar to the present one was performed over several days and preceded by tDCS, anodal stimulation improved the self-regulation of α-band ERD in a polarity-specific way from training day 3 onwards, thus suggesting that there are cumulative effects of transcranial stimulation (Soekadar et al., 2014). We therefore also applied anodal tDCS before the BRI task to research additive effects on learning β-band self-regulation.

Feedback strategy

Previous BRI studies provided contingent feedback to successful brain self-regulation, i.e. the participants were rewarded with, for example, robotic opening of the hand, when the predefined brain state was achieved and sustained; whenever the respective ERD was insufficient, the robotic movement ceased but could be restarted to continue hand opening when the ERD threshold was reached again (Bauer et al., 2015; Ramos-Murguialday et al., 2013; Vukelić and Gharabaghi, 2015a; Vukelić et al., 2014). However, this practice of contingent feedback to sustained ERD is challenging and is characterized by a strong sensation of frustration on the part of the participants (Fels et al., 2015). In a bid to reduce this frustration, we explored an alternative, i.e. non-contingent, feedback strategy. In this approach, as soon as the predefined threshold was reached for the first time, feedback began and continued until the end of the task.

Threshold adaptation

During BRI neurofeedback, different performance measures provide information about the subject's ability for brain self-regulation as well as an indirect measure of his/her cognitive resources for coping with the mental load that occurs during a misalignment between ability and difficulty (Allal and Pelgrims Ducrey, 2000; Bauer and Gharabaghi, 2015a; Schnotz and Kürschner, 2007; Naros and Gharabaghi, 2015). Such an alignment could, for example, be achieved by altering the sensitivity and specificity of the BRI classifier (Bauer and Gharabaghi, 2015b; Naros and Gharabaghi, 2015; Thompson et al., 2013). The selection of this difficulty threshold is currently determined by the intent to maximize the classification accuracy and results, usually in a fixed threshold throughout the sessions (Thomas et al., 2013; Thompson et al., 2013). However, mathematical neurofeedback modeling on the basis of Bayesian simulations indicates that brain self-regulation can be improved when an adaptation strategy for threshold selection is applied in the course of the training (Bauer and Gharabaghi, 2015b). Such an adaptation would require an alteration of the classifier threshold, i.e. the level of difficulty of the feedback device, during the training to challenge the participant while preserving his/her motivation (Naros and Gharabaghi, 2015). We therefore directly compared adaptive and non-adaptive, i.e. standard, classifier thresholding in this study.

In the present study, we reasoned that brain self-regulation would result in motor gains when applying a comprehensive training strategy for operant conditioning of oscillatory states optimized on the basis of knowledge gained from previous studies. In a three-day intervention, we therefore studied proprioceptive feedback of sensorimotor β-ERD following anodal tDCS. In a parallel group design, we compared four different combinations of feedback and thresholding strategies: (i) adaptive classifier thresholding with contingent feedback, (ii) adaptive classifier thresholding with non-contingent feedback, (iii) non-adaptive classifier thresholding with contingent feedback, and (iv) non-adaptive classifier thresholding with non-contingent feedback. We hypothesized that reinforcement learning is achieved best by adapting the BRI task to the performance and learning experience of the participant. Finally, we sought to unravel the brain–behavior specificity of the intervention by analyzing the oscillatory cortical activity that mediates the expected motor gains.

Materials & methods

Having given their informed consent, 32 BRI-naïve healthy subjects (age: 25.9 ± 0.5 years [mean ± SEM], 16 females) were enrolled in this study which had been approved by the local ethics committee