Mechanisms of neurofeedback: a computation-theoretic approach

Eddy J. Davelaar

Department of Psychological Sciences, Birkbeck, University of London
Malet Street, WC1E 7HX London, United Kingdom
e.davelaar@bbk.ac.uk

In Press in Neuroscience: https://doi.org/10.1016/j.neuroscience.2017.05.052

Neurofeedback training is a form of brain training in which information about a neural measure is fed back to the trainee who is instructed to increase or decrease the value of that particular measure. This paper focuses on electroencephalography (EEG) neurofeedback in which the neural measures of interest are the brain oscillations. To date, the neural mechanisms that underlie successful neurofeedback training are still unexplained. Such an understanding would benefit researchers, funding agencies, clinicians, regulatory bodies, and insurance firms. Based on recent empirical work, an emerging theory couched firmly within computational neuroscience is proposed that advocates a critical role of the striatum in modulating EEG frequencies. The theory is implemented as a computer simulation of peak alpha upregulation, but in principle any frequency band at one or more electrode sites could be addressed. The simulation successfully learns to increase its peak alpha frequency and demonstrates the influence of threshold setting - the threshold that determines whether positive or negative feedback is provided. Analyses of the model suggest that neurofeedback can be likened to a search process that uses importance sampling to estimate the posterior probability distribution over striatal representational space, with each representation being associated with a distribution of values of the target EEG band. The model provides an important proof of concept to address pertinent methodological questions about how to understand and improve EEG neurofeedback success.

Keywords: Neurofeedback, electroencephalography, computational neuroscience, computer model

Introduction

There are an increasing number of media reports about controlling electronic devices through brainwaves, whether it is for therapeutic reasons or for pure entertainment. These brain-computer interfaces utilise the ability of people to learn to voluntarily change their brain rhythms when provided with corrective feedback. This process is called neurofeedback and understanding how it works is the topic of this paper. Neurofeedback is a goal-directed process of modulating one's own neural dynamics by means of feedback-induced learning. The feedback that is obtained can either be internal phenomenological experiences or external provided stimulation of which visual and auditory stimulations are the most common modalities. The neural dynamics that are being influenced can be measured through electroencephalography (EEG), magnetoencephalography (MEG), blood-oxygen-level-dependent (BOLD) responses or any other direct or indirect methods.

This paper focuses on EEG neurofeedback. The aim here is to provide a proof of concept that using computational methods from neuroscience can pave the way for understanding how neurofeedback works. In several papers and books authored by leading practitioners and researchers, a common call is expressed to develop a theoretical

understanding of neurofeedback (e.g., Budzynski et al., 2009; Evans and Abarbanel, 1999). Although the neural circuitry involved in some brain rhythms are understood at a descriptive level, how these rhythms are influenced through feedback-based learning at a mechanistic level is still unclear.

Why would we need to seek to understand neurofeedback? Knowledge of the mechanisms underlying neurofeedback at the neural level provides a critical foundation for (1) interpreting findings (from the lab and clinic), (2) guiding research efforts, (3) developing new protocols, (4) improving existing protocols, (5) quality assurance, (6) risk assessment and management, and (7) approval of protocols. Currently, the research on neurofeedback is making great steps in validating the efficacy of neurofeedback training (see for recent special issues Gruzelier et al., 2014ab, van Boxtel and Gruzelier, 2014). However, there is no validated standardised methodology or a standard way of reporting the methods that have been used, leading to high study-exclusion rates in systematic reviews or meta-analyses (e.g., Emmert et al., 2016). To circumvent this, computational methods can be used to test whether certain choices, such as threshold settings, integrating neural activity across electrodes, the time window over which the neural signal is calculated, and the maximum feedback rate, have an effect of neurofeedback success and if so what the optimal parameters are. Given that properly conducted experimental studies are costly in terms of investment of money, time, and labour, dry-testing a protocol using a computational model can make the research efforts more efficient. At present, such an opportunity does not exist.

To facilitate the process, this paper puts forward an initial step towards a computational theory that explicates the neural mechanisms underlying neurofeedback training. By way of illustration, a computational model using spiking neurons is implemented that shows successful neurofeedback training and allows an analysis of the learning process. To demonstrate its benefit, the influence of threshold setting on learning is addressed in a mathematical abstraction of the model. Future work will address how each of the aforementioned points can be supported by using a computational modelling approach.

There are many detailed descriptions of the neuroanatomical circuitry involved in the generation of brain frequencies (e.g., SCP: Birbaumer et al., 1990; SMR: Hughes and Crunelli, 2005; Ritter et al., 2009; Sterman, 1996; theta: Gruzelier, 2009; Hsieh and Ranganath, 2014), but only some reports provide suggestions of what neurofeedback might be doing to this circuitry at the neurophysiological level (e.g., Koralek et al., 2012; Sterman, 1996). Here, a general theory of neurofeedback learning is proposed that is articulated at the neurophysiological level and addresses the contributions of the striatum and the thalamus. The theory is assumed to be applicable to any neurofeedback modality (EEG, MEG, BOLD) and breaks neurofeedback learning down into three stages, of which the first stage is then implemented in a computer simulation model. In order to address how the first stage unfolds a second simulation model is analysed in which the threshold setting for positive feedback is systematically varied. The results support the view that during stage 1, the striatum performs a search through representational space using importance sampling, i.e., maintaining only those sampled representations that lead to reward.

The importance of the striatum

Several studies employing a variety of methodologies confirm the critical contribution of the striatum in neurofeedback learning. Neuroimaging studies have shown the involvement of the entire striatum in neurofeedback (ventral striatum: Johnston et al., 2010; putamen: Hinterberger et al., 2005; caudate: Levesque et al., 2006). Johnston et al. (2010) trained participants using a real-time (rt) fMRI-neurofeedback to increase the activation in the emotion network, as defined by the collection of brain regions that was maximally responsive

to negative versus neutral stimuli. One of the non-target areas that was activated during the learning process was the ventral striatum. Hinterberger et al. (2005) demonstrated the involvement of the putamen and thalamus in regulating the slow cortical potential (SCP) over Cz. Levesque et al. (2006) observed increased functional activation of the caudate in ADHD children after a theta/SMR/beta1 protocol (20 sessions SMR increase with theta decrease followed by 20 sessions beta1 increase with theta decrease) over Cz. Furthermore, in a structural MRI study, Ghaziri et al. (2013) found increased white matter density in the anterior limb of the internal capsule (ALIC) after increasing beta1 at F4 and P4 with EEG neurofeedback. Increases in fractional anisotropy in the left ALIC, which includes cortico-striatal as well as frontal cortico-thalamic fibers, was correlated with enhanced visual attention. Finally, in a critical experiment involving rats, Koralek et al. (2012) measured the activity of motor neurons and transformed this activity into an auditory signal. They showed that rats lacking cortico-striatal plasticity could not learn to control the auditory pitch. In a series of experiments they also demonstrated that cortico-striatal plasticity is necessary for neuroprostethic control to occur. These studies provide strong support for the view that the entire striatum is involved in neurofeedback learning with lasting functional and structural consequences. Whether there is specificity in the recruitment of striatal subregions in relation to the EEG frequency, learning direction, and electrode site is yet unclear.

This is not to say that it is impossible for EEG spectrum modification to occur through synaptic changes at cortical sites only in the absence of a striatal contribution, but the current literature provides compelling evidence for a striatal theory of neurofeedback learning. In a recent meta-analysis of 12 rt-fMRI studies, Emmert et al. (2016) observed that the striatum and the anterior insula were non-target regions that were consistently activated during the neurofeedback learning. They suggested the existence of a "regulating network" of which the striatum and the anterior insula contribute through their involvement in reward-based learning and self-awareness processes, respectively. These findings are critical building blocks of the proposed theory to which we turn next.

A multi-stage theory of neurofeedback learning

What happens in the brain of a person from the first training session to demonstrable voluntary control over EEG brainwaves? In the theory advanced here three stages are discerned (see figure 1). In the first stage (indicated by the red parts in figure 1), trainees may try different things, such as the strategies provided by the trainer, strategies read from the internet, or idiosyncratic strategies that come to mind during the training session. Examples of strategies are trying to relax, focus on breathing, counting numbers, thinking back of positive events, trying to become angry or positive, staring at a point on the computer screen, and many more. This stage is the problem solving or exploration stage and involves performing various mental acts and evaluate their consequences on the feedback signal. It is expected that frontal brain areas are critically involved in this stage, as it requires retrieval, creation, and maintenance of goal representations (i.e., possible strategies), execution of these strategies as response plans, and evaluating the consequences of the mental actions. The evaluation is driven by the feedback and it is assumed that the frontal-striatal connectivity is the locus of the learning (cf. Koralek et al., 2012), with positive feedback strengthening the connections between the (frontal) goal representation and the striatal response plan.

In the second stage (indicated by the orange parts in figure 1), trainees may have found a strategy that seems to elicit a high rate of positive feedback signals. The trainee then enters the consolidation phase. Two neural processes are operational. The goal representation gets associated through reinforcement learning with the striatal representation, sharpening it in the process. This leads to changes to the neural architecture from the striatum to cortex via

the thalamus and can be demonstrated by various functional (baseline EEG recordings, EEG coherence) and structural (white/grey matter density; see Ghaziri et al., 2013) brain measurements. In particular, it is assumed that the resulting striatal representation influences the activation in the thalamic nucleii either through the striatal-thalamic, the striatal-pallidal-thalamic, or the striatal-nigral-thalamic pathways. Synaptic plasticity occurs in all these pathways and between the specific thalamic nucleii and the reticular thalamic nucleus. These synaptic modifications at the thalamic level may lead to changes in baseline recordings across sessions. Using SMR-upregulation as an example, Sterman (1996) suggested that neurofeedback alters inhibitory mechanisms in the thalamus by increasing GABA synthesis or up-regulating GABA receptors. Other protocols might lead to changes within different thalamic pathways. It is assumed here that the precise striatal-thalamic and thalamo-cortical pathways that undergo synaptic modification will govern the topographical and frequency specificity of the learning.

Simultaneously, the subjective experience associated with the particular brain state can become more salient to the trainee and available for introspection. For example, the trainee might be able to say whether "being in the zone" means feeling relaxed or tensed or even feeling disconnected from the body. Phenomenological analyses have shown that different neurofeedback protocols are associated with different subjective experiences (Edge and Lancaster, 2004) and that differences in learning success may in part be due to differences in subjective experiences (Davelaar et al., 2016). This subjective feeling is a higher level interoceptive representation that consists of an integrated percept of the state of the body (see Ceunen et al., 2016) and that is of importance in the final stage. Based on the meta-analysis by Emmert et al. (2016), the anterior insula might be crucial in the genesis of this subjective experience.

In the final stage (indicated by the blue parts in figure 1), the subjective or interoceptive representation, when of sufficient distinctiveness, can function as a secondary reinforcer, as it is always paired with the feedback reward. Therefore, the system is more likely to want to remain in that interoceptive state. Losing that sensation would lead to activating the goal representation to return to that state. In this way, the desired interoceptive state forms a new homeostatic set point that aids in neurofeedback learning At present, it is speculated that this stage of interoceptive-homeostasis relies on the interplay between the insula and dopaminergic midbrain areas (Geisler et al., 2007; Watabe-Uchida et al., 2012; Arsenault et al., 2014).

Although the multi-stage theory as outlined here introduces the stages as operating sequentially. The sequentiality of stages comes from the order in which the processes come online and it is assumed that interaction occurs thereafter. For example, the (slower) synaptic changes within the thalamus (stage 2) will affect the magnitude of the target neural signal and thus the probability of reward, which in turn influences the reward-modulated synaptic change at the striatum (stage 1). In addition, in those situations where a distinctive subjective representation exists (stage 3), the explicit effort of aiming to reach that subjective feeling will also influence the striatal learning. Both interactive routes lead to bootstrapping the neurofeedback learning and demonstrate that multiple processes can be operational simultaneously. We turn to a computational analysis of what happens during neurofeedback learning after the multi-stage theory is compared to a number of other views in the literature regarding how neurofeedback works.

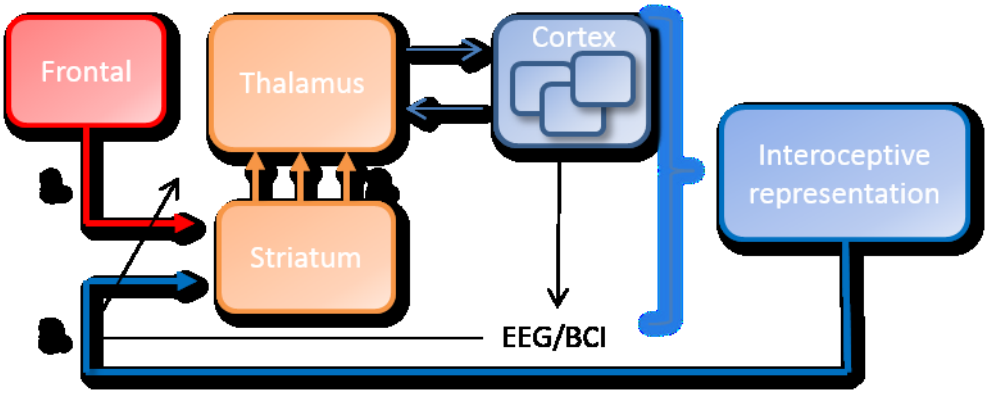

Figure 1. Overview of the theory. There are three stages. In stage 1, the striatal-exploration stage, the frontal or executive system generates representations that create an activation pattern in the striatum. These frontal representations are goals that activate associated response plans. The striatal representation or response plan will modulate the activation of the thalamus and thereby the EEG signal recorded in the cortex (three smaller blue squares). Stage 1 involves modifying the synaptic connections between the frontal and striatal representations, using reward as the primary reinforcer. In stage 2, the thalamic-consolidation stage, a relatively stable frontal representation is present and the synaptic modifcations occur between the striatum and the thalamic nucleii. Modifications in stage 2 may lead to further stage 1 learning. In the final stage 3, the interoceptic-homeostasis stage, which is not obligatory, a distinctive internal representation of the consequences of the target brain state becomes associated with the reward and in turn a secondary reinforcer.

Related accounts of neurofeedback learning

Beyond the general clinical view that neurofeedback normalises the brain, a more neurocognitive perspective adds that it increases the neural efficiency or brings the brain closer to self-organised criticality that is conducive for maximal information transfer (e.g., Ros et al., 2014). Current theorising on the mechanisms underlying neurofeedback takes a high level view (Niv, 2013; Ros et al., 2014) in that neurofeedback promotes synaptic modifications, leading to strengthening of neural circuitry. These modifications alter the dynamics of brain networks (Ros et al., 2013; Kluetsch et al., 2014) such as the default mode network which is then assumed to be associated with changes in cognitive performance and therapeutic benefit. However, these explanations do not address the detailed neural mechanism underlying neurofeedback learning, but there are some notable exceptions that largely overlap with the aforementioned theoretical view.

In a review paper, Birbaumer et al. (2013) largely agree with the first two stages described above, but disagree in two details. Consistent with the multi-stage theory, their view is that a response plan that produces the desired brain state is learned in the same way as other skills are learned and that there is an initial stage of trial-and-error followed by consolidation of the appropriate response plan. Whereas there is also agreement that reinforcement learning underlies neurofeedback learning, they equate the response plan with the strategy or goal representation and in effect do not distinguish between the frontal and striatal representations. Separating these representations allows for a single goal representation to exist, while a striatal response plan is created via reward-based learning. When such a striatal response plan fails to increase reward frequency, the frontal representation will be purged and a new representation is activated. The two representations form a hierarchy with a slower time-course at the frontal compared to the striatal level.

Another point of difference is that although Birbaumer et al. (2013) highlight the importance of motivational signals to stabilise the desired response plan through primary and secondary reinforcers, subjective experience is not seen as a potential secondary reinforcer. In

fact, Birbaumer et al. (2013) view neurofeedback learning as an implicit learning process that does not require consciousness (see also Chang et al., 2014). This, however, is inconsistent with the methods in animal studies of making the animals hungry or thirsty before training and the observation that humans are actively trying to find a suitable strategy. The learning seems explicit and goal-driven. They acknowledge that unsuccessful studies might have failed due to the lack of putting the subject in a state of needing the reward signal (food or juice). Finally, their view that the "response plan is modified based on the difference between the anticipated effect and the actual effect" (p. 297), does hint at a mental representation of the anticipated effect (i.e., an internal model), but they did not explicate what that representation is. Instead of assuming only implicit or explicit learning processes, the multi-stage theory allows for both components. The results of activation in the anterior insula (Emmert et al., 2016) and its role in interoception (Craig, 2002) could suggest a subjective mental representation based on proxies of bodily sensations. This representation could emerge during stage 2 and consolidated in stage 3 and have an influence on striatal learning when it functions as a secondary reinforcer.

A computational approach was adopted by Legenstein and colleagues (Legenstein et al., 2010; see also Legenstein et al., 2008 for work at the single neuron level). They focused on a study by Jarosiewicz et al. (2008) in which monkeys were trained to control a cursor in a 3D virtual reality environment by modifying the activity in the motor cortex. Legenstein et al.'s computational model did not include a division between cortical and striatal areas, but consisted of two functional pools of neurons, where neurons in the first pool send input to neurons in the second pool. Legenstein et al. did not commit to a neural substrate and allowed for the possibility that the reward-based learning could operate in extra-cortical and non-motor cortical areas, or even in the motor cortex itself. The multi-stage theory advanced here makes the stronger assumption that the learning mechanism is more likely to operate in the cortico-striatal projections. Activation of the simulated neurons was then used to control the cursor movement. The global feedback signal that was provided for adjusting the synaptic connectivity was calculated from the angular match between the actual and desired cursor direction. Although Legenstein et al. addressed a neurofeedback protocol, their focus was on the learning rule that underlies the synaptic plasticity between the two neuronal pools. Critically, they showed that a reward-modulated Hebbian learning rule can solve the credit-assignment problem inherent in the task by making synaptic weight changes contingent on the correlation between deviations in the reward signal and in the postsynaptic potential compared to the longer-term average of the signal and potential. Adopting Legenstein et al.'s (2008) findings, the multi-stage theory assumes that the reward-modulated Hebbian learning rule operates at the level of the striatal medium spiny neurons, thus providing a physiologically plausible solution to the credit-assignment problem – finding which neurons out of many simultaneously active ones are responsible for the probabilistic reward signal – present during stage-1 neurofeedback learning.

Ros et al. (2014) agrees with Legenstein et al.'s (2010) view that synaptic plasticity is governed by a reward-modulated Hebbian learning rule. They further suggest that the neurofeedback system essentially augments the sensory repertoire, allowing the brain to "sense" the neuroelectrical patterns and thus make them amenable to control in a homeostatic manner. In addition, they argue that feedback signals need to be presented above the threshold for awareness in order to be effective. Although they do not provide explicit mechanisms by which neurofeedback operates, they use dynamical systems theory to argue that neurofeedback functions to bring a disordered brain into a regime where there is maximal information transfer, which in turn explains the benefits on cognitive and creative abilities of neurofeedback training (see for reviews Gruzelier, 2014ab).

In the following section, a computational model is presented that implements the first stage of the multi-stage theory of neurofeedback learning. The basic components for demonstrating neurofeedback learning are (1) stochastic activation of striatal neurons, with the activation probability contingent on the changing fronto-striatal synaptic weights that in turn (2) influence the thalamic input to a cortical pool of neurons that produce an EEG frequency, that (3) is analysed in real-time to provide feedback. Although the model focuses on stage 1 of the theory outlined in figure 1, the current computational implementation does not feature (1) the precise reward-modulated Hebbian learning rule (only an abstraction of it), (2) detailed modeling of the basal ganglia and thalamus, (3) synaptic modifications at the thalamic level (stage 2), and (4) the emergence of subjective experience as a secondary reinforcer (stage 3). All these aspects augment the core part of the theory and their direct implementations are left for the future.

The task

The task that will be used here comes from a paper by Zoefel and colleagues (2011). They were interested in investigating the effect of training the peak alpha frequency (PAF) as defined by the peak amplitude in the 8-12 Hz range. To increase the PAF, the averaged upper alpha frequency (UAF) band over P3, Pz, P4, O1, and O2 was rewarded, where UAF = [PAF, PAF+2Hz]. Participants were trained over five sessions, each session consisting of seven 5-minutes blocks. The first and last blocks were measurement-only blocks during which the participants sat quietly. The first block was used to estimate the individual PAF for that day. The following five blocks were training blocks during which participants had to increase the spectral power in the UAF range. To accomplish this, participants were provided with visual feedback regarding whether the UAF value of the preceding second was higher (red colour) or lower (blue colour) than the average UAF measured in the baseline block, with colour saturation indicating the UA amplitude. PAF has been found to be associated with attention (Klimesch et al., 1998), short-term memory (Nan et al., 2012), working memory capacity (Moran et al., 2010), problem solving (Haarmann et al., 2012), and mind wandering (Ros et al., 2013) among others. In general, higher PAF is associated with better cognitive performance (Klimesch, 1999; Hanslmayr et al., 2005). As the alpha rhythm amplitude fluctuates (Palva and Palva, 2007; Omata et al., 2013), the EEG measurements are essentially samples from a distribution of amplitudes. This distribution can be regarded as the within-subject variability of the alpha rhythm, which can be compared across sessions.

Simulation study 1: a spiking neuron model of upper alpha upregulation

Here, a computational model is presented that mimics the results of Zoefel et al. (2011). The model consists of an EEG-generator and a feedback-based learning mechanism that based on the feedback it receives adjusts the neural configuration through which the EEG-generator is more likely to produce the target frequency.

There are many computational models in the literature that address how brain rhythms are generated (see for a review, Van Drongelen, 2013). For the purposes of demonstrating a proof of concept, the details are not too important, other than that the model should be able to produce a useful proxy for the EEG signal. The choice for a spiking neuron EEG generator over a mean-field model is that it demonstrates the utility of models at the neurophysiological level, which is then inherited by mean-field abstractions. It also implies that the current EEG generator could be replaced by models with increasing neurophysiological realism and detail, allowing future investigations to consider for instance the interactions between neurofeedback and neuroactive drugs. The simplest EEG generator chosen for this purpose is

the model by Izhikevich (2003) and was adopted without any changes to the code. The reader is referred to Izhikevich (2003) for further details. By using a simple EEG generator, the model's scope and focus is on the striatal learning process instead of on the mechanisms underlying alpha oscillations or alpha/theta training protocols. The model as presented by Izhikevich (2003) happens to produce only a clear oscillation in the alpha frequency band without the need for an explicit implementation of the thalamus and its reticular nucleus. When those components are implemented, the model produces a more realistic frequency spectrum, including theta rhythms (Izhikevich and Edelman, 2008). Such a model will be explored in future research efforts.

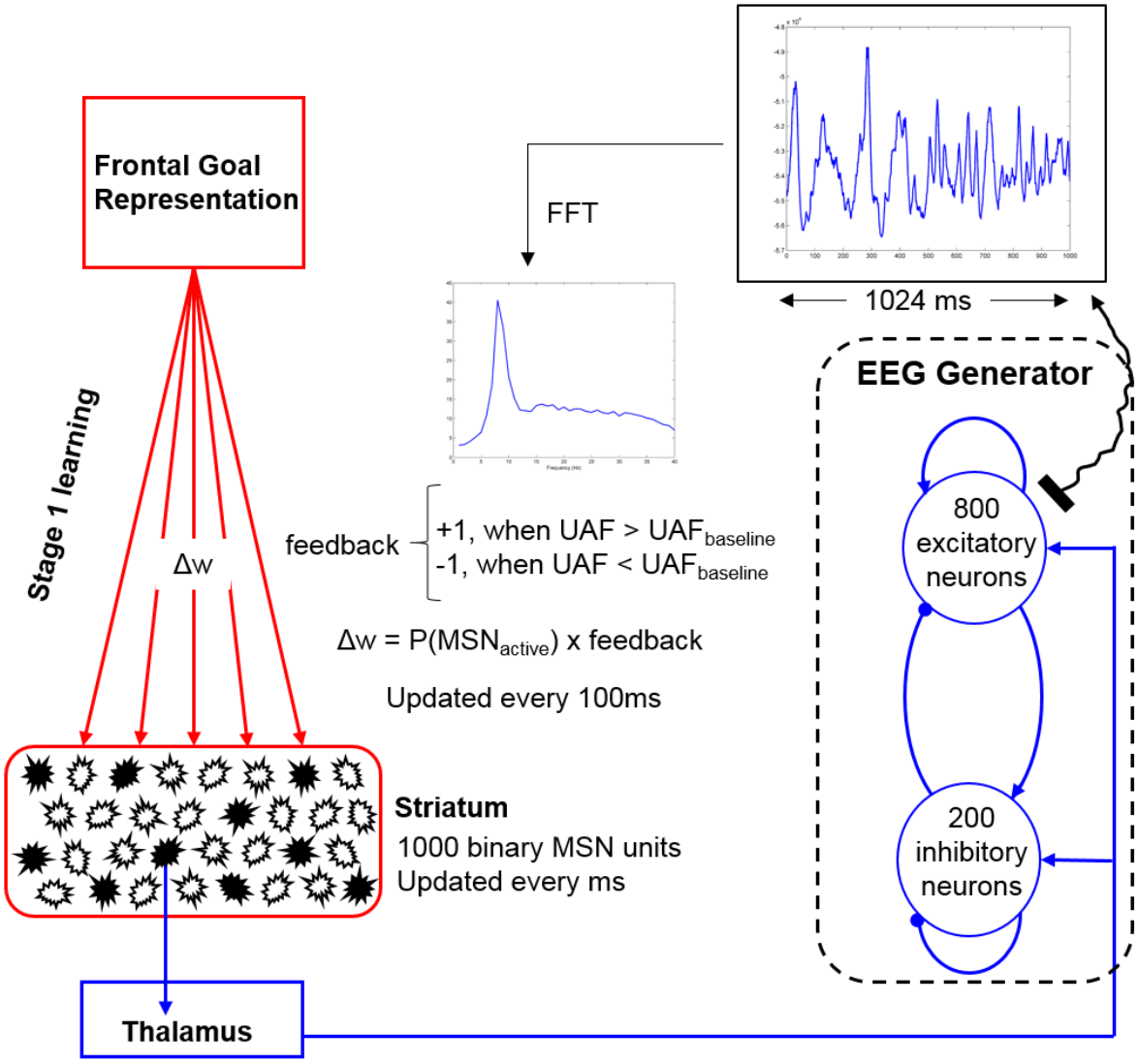

Figure 2. Architecture of the computational model of stage 1 learning: the striatal-exploration stage. The frontal goal representation is continuously activated during the simulation run and influences via the modifiable fronto-striatal projections the activity the striatal units. The striatum consists of 1000 binary units that are updated every ms. Whereas only one target unit is connected to the thalamus (for simplicity) at any one time approximately 10 MSN units are active. The model needs to identify and strenghten the connection to the target unit. The thalamus sends noisy input to the EEG generator with a higher input when the target MSN unit is active. The EEG generator is identical to the model by Izhikevich (2003). The feedback is based on the spectral analysis over a 1024 ms window of EEG and updated every 100 ms. Only the red fronto-striatal connections are modified. The blue connections remain fixed for this simulation, although are modified during the thalamic-consolidation stage.

Methods

The computational model addresses stage 1 in the multi-stage theory and its structure is shown in Figure 2. A single frontal representation is continuously active during a simulation run and represents the goal. It is connected via modifiable weights to the striatal part. The configuration of active striatal units forms the response plan. The task that the model addresses is to find the distribution of active striatal units that maximises the frequency of reward feedback. This is accomplished by changing the synaptic strengths of the fronto-striatal connections. The output of the response plan is the activation (or inhibition) of the thalamus and thalamic reticular nucleus, which in turn influence the frequency spectrum at the cortical level. To simplify the model as much as possible, the output of the correct response plan, here a single target unit, is used to activate the EEG generator, in essence assuming that the thalamus is only relaying the neural activation. Only the fronto-striatal connections undergo synaptic change in this simulation. All other connections, including those of the EEG generator, are fixed throughout a simulation run.

The EEG generator

The EEG generator consists of 800 excitatory neurons and 200 inhibitory spiking neurons that are fully interconnected and receive noisy thalamic input (Gaussian noise with mean = 5 or 2 for excitatory and inhibitory neurons, respectively with sd = 1). The neurons obey the following dynamics:

$v' = 0.04v^2 + 5v + 140 - u + I$
$u' = a(bv - u)$
if $v \geq 30$ mV, then $\{v = c$ and $u = u + d\}$

I stands for input, which comes from the thalamus and from all other excitatory and inhibitory neurons. *V* represents the membrane potential and *u* the membrane recovery variable. After *v* reaches 30 mV, *v* and *u* are reset to values *c* and *u* + *d*, respectively and activation is propagated over the synaptic connections The two pools of neurons differ from each other in the type of spiking behaviour, which is governed by the parameters, *a*, *b*, *c*, and *d*. For the excitatory neurons: $a = 0.02$, $b = 0.2$, $c = -65 + 15*r_e^2$, $d = 8 - 6*r_e^2$. For the inhibitory neurons: $a = 0.02 + 0.08*r_i$, $b = 0.25 - 0.05*r_i$, $c = -65$, $d = 2$. At the beginning of each simulation, r_e and r_i are drawn from a uniform distribution between 0 and 1, for each neuron. This produces parameter ranges of a, b, c, and d of {0.02, 0.2, [-65, -50], [2, 8]} for excitatory and {[0.02, 0.1], [0.20, 0.25], -65, 2} for inhibitory neurons. As shown by Izhikevich (2003) this scheme produces a pool of excitatory neurons that consists predominantly of regular spiking neurons and a pool of low-threshold and fast spiking inhibitory neurons. Figure 3 shows traces of membrane potentials for four neurons over a 2000 ms time interval. The top two and bottom two traces represent excitatory and inhibitory neurons, respectively. As can be seen, the parameter values greatly affect the neural dynamics. To obtain an EEG signal, the sum of the membrane potentials of all excitatory neurons, Σv_e, was low-pass filtered using: $EEG(t) = 0.9*EEG(t-1) + 0.1*\Sigma v_e(t)$.

Striatal learning

The feedback-based learning mechanism is assumed to be located in the striatum. The model simplifies the striatal part by assuming 1000 binary medium spiny neuron (MSN) units that have a probability $p < 1$ of being active (i.e., set to 1). These units receive input from a frontal goal representation (set equal to 1) which is multiplied by the modifiable synaptic weight. For simplicity, no intrastriatal connectivity is used here, other than the normalisation of activation as a proxy for the high level of inhibitory connectivity. Only one of those units is able to push

the EEG-generator to produce a higher PAF by disinhibiting the thalamus to produce a higher mean value. Although it is anticipated that the appropriate striatal response plan consists of more than one neuron, the choice for one unit here is to demonstrate the nontriviality of finding the one target unit among 999 other units that can be active simultaneously. Recent computational work has shown that increasing the mean of the noisy thalamic input increases the actual peak frequency (Cohen, 2014). This forms a network version of the input-frequency curve for single neurons. To speed up the simulations, the target unit increases the mean thalamic input to the cortex by 1. At first, the normalised probability is 1/100 for every unit (i.e., the expected number of active units at any one time is 10), but with positive feedback, the probability of the active units increases. Thus, over the course of a successful neurofeedback session, the likelihood of the target unit becoming active increases, whereas the likelihood of the other units decreases. This is not a trivial learning task, as the units are updated every millisecond, whereas the feedback about the UAF in the last 1024 samples is given every 100ms (see below). This creates a credit assignment problem – identifying which MSN unit among many competing ones to strengthen based on the feedback that is delayed and nonspecific – that is solved here by matching the activation probability of the units with the thalamic activation. This solution guarantees that the model finds the correct target unit, although the time to find it varies greatly across simulations. Recent work by Legenstein et al. (2010) shows how the credit assignment problem can be solved in a physiological plausible way by making fronto-striatal synaptic modification contingent on the covariance between the change in postsynaptic activity and the change in reward signal.

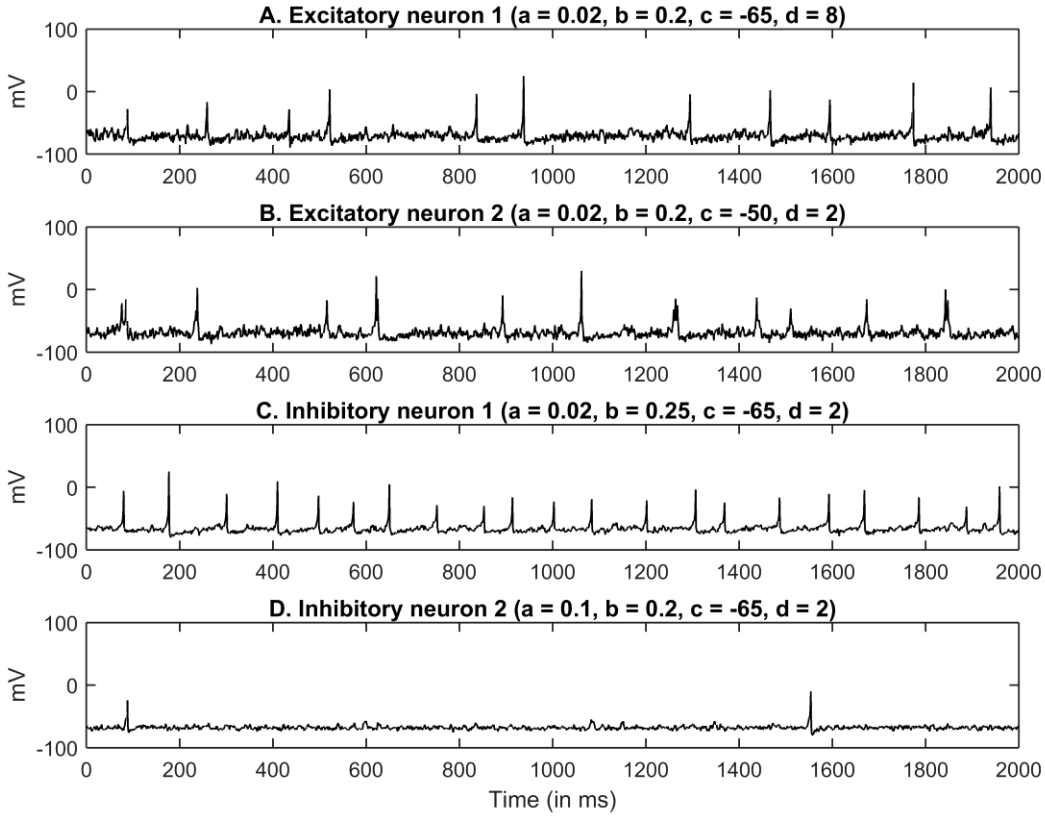

Figure 3. Example of simulated traces of membrane potentials of four neurons of the EEG generator. The traces were recorded for the first 2000 ms of the EEG generator with all 1000 neurons. All other neurons had their parameter values initialised using the random function. The neurons in these panels were fixed to the boundary values. A. Trace of an excitatory neuron initialised with $r_e = 0$. B. Trace of an excitatory neuron initialised with $r_e = 1$. C. Trace of an inhibitory neuron initialised with $r_e = 0$. D. Trace of an inhibitory neuron initialised with $r_i = 1$.

The methods used for the EEG analysis are the same as used by Zoefel et al. (2011). In short, the sampling frequency is 1 kHz and a moving window consisting of 1024 samples is updated every 100ms. The Hamming-windowed EEG is subjected to fast Fourier transform from which the average amplitude in the UAF band is extracted. This measured UAF is then compared to the mean UAF of the baseline block and information is send to the feedback-based learning mechanism.

The learning mechanism updates the activation probabilities of the MSN units by increasing or decreasing them when feedback is presented or absent, respectively. When positive feedback is given, the fronto-striatal weights to all the MSN units are increased based on the number of iterations each unit was active during the preceding 1024 iterations (min = 0, max = 1024). When feedback is absent, the weights are decreased based on the mean activation of the units in the preceding 1024 iterations. The weights are normalised and multiplied by 10 to obtain activation probabilities for the next 100 iterations after which the next feedback signal is given.

The training sequence for the model consisted of 5 minutes worth of baseline recording, which is used to estimate the PAF. This is followed by one minute worth of training and then 5 minutes worth of post-training baseline measurement. Several runs were conducted to verify reliability in performance.

Results and discussion

Figure 4A presents representative windows of simulated EEG before and after the neurofeedback training. It can be seen that the training led to more alpha rhythm. The spectral density plots in Figure 4B confirm that the training did increase PAF. In other words, the neurofeedback training was successful.

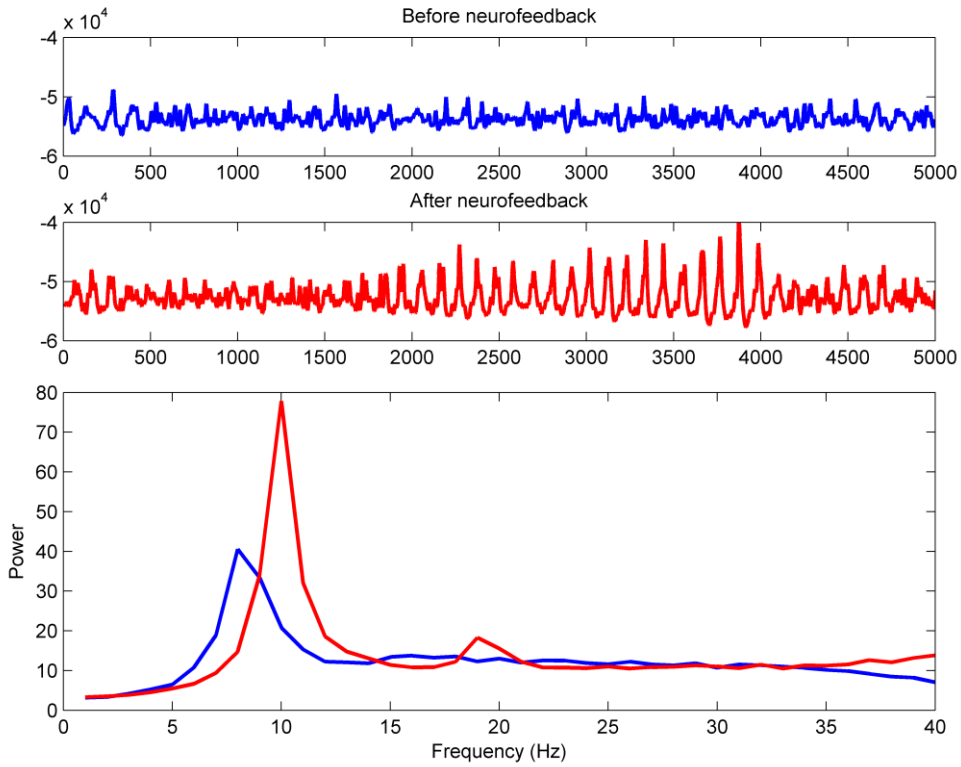

Figure 4. Results of simulation study 1. The top panels show representative EEG samples before and after simulated neurofeedback training. The bottom panel shows the spectral density plots for before and after simulated neurofeedback training of the upper alpha band.

The model shows the ability to learn to enhance the alpha frequency, but what exactly happens in the model as a consequence of the learning? One way to address this question is by comparing what happens to the distribution of UAF values before and after training. Figure 5A shows that training causes the distribution to increase in variance while retaining the lower limit. In other words, the stage 1 learning model as implemented here leads to higher UAF values than expected with the model without the target MSN unit. The distribution after neurofeedback training is therefore a weighted combination of two parent distributions, $D_{(UAF|P(target\_MSN) = 0)}$, which is the same as $D_{(UAF|before)}$, and $D_{(UAF|P(target\_MSN) = 1)}$, with $P(target\_MSN)$ as the weighting parameter.

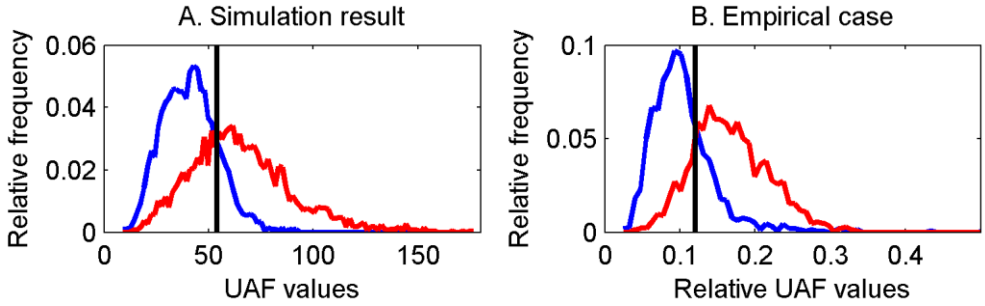

Figure 5. Distributions of upper alpha (UAF) values before (blue) and after (red) neurofeedback training. A. Simulated results. B. Empirical case study of a person trained on alpha upregulation at Pz for 10 months. Note that in both panels, the distributions share the minimum values and that the post-training distribution is stretched to encompass higher values. The vertical black line optimally separates the two distributions.

Figure 5B presents an example of an empirical case study of a person trained on alpha over Pz for 10 months. The sampling frequency was 256 Hz, using a filter-based approach. Further technical details are unavailable. The empirical data files were obtained after the local ethics committee approved analysing this dataset as secondary data which was done after the simulation results were known. The resemblance between the two figure panels is striking.

In the model, even though the desired state exists in the repertoire of the simulated participant before training, its probability of occurring is very low and the probability of it remaining active long enough to influence neural dynamics is negligible. The model equates the neurofeedback learning to a search task in which a particular neural state is searched for that enhances the likelihood of a positive reward. When the target MSN unit is active, the minimum UAF value remains the same and therefore $min(D_{(UAF|P(target\_MSN)})$ will be constant for any value of $P(target\_MSN)$, resulting in a stretching of the distribution. Only when the striatal representation produces a higher mean value for a target frequency will the entire distribution shift, which could happen in the thalamic-consolidation stage (stage 2, not modelled here).

The conceptualisation of neurofeedback training changing the distribution of values of the target frequency suggests that the placement of the threshold is critical. Set too high, the model will decrease all weights and decrease of the target frequency will occur. Set too low and the model will not learn at all as it increases the weight for all MSN units, which are then subsequently normalised. In figure 5, the vertical line represents the level which optimally separates the two distributions. In order to systematically investigate the effect of threshold setting on the learning, simulations are needed that run the neurofeedback training at each threshold setting. We now turn to such a simulation study.

Simulation study 2: EEG distribution sampling

The minimalist spiking neuron model showed that a computation-theoretic approach to neurofeedback can predict and capture real data on distributions. Whether other data patterns

can be accounted for is an important question for future studies. However, the aim of the current modelling efforts is to understand what happens during stage 1, the lowest level of the learning process, i.e., how does the system find the right striatal representation? To address this question, an abstraction of the previous model is used based on the results that the model changes the distribution of UAF values. This allows for varying the threshold for every simulation run. The learning trajectories are then examined at each threshold level.

Method

Two parent distributions, $D_{(UAF|P(target\_MSN) = 0)}$ and $D_{(UAF|P(target\_MSN) = 1)}$, consisting of 100,000 UAF values each, were generated with the previous model (see figure 6C). The current model consisted of 1,000 binary MSN units receiving an activation of 1 multiplied by the normalised synaptic weight to each MSN unit. This value functioned as the probability of the unit being active (i.e., output activation = 1). The initial synaptic weight was 1, making the initial activation probability, including the target unit, equal to 0.001. On every iteration, 10 units were active. When the target unit was part of the 10 active units, a UAF value from $D_{(UAF|P(target\_MSN) = 1)}$ was selected, otherwise a UAF value from $D_{(UAF|P(target\_MSN) = 0)}$ was selected. This value was compared against the threshold, T, for that simulation. If the value was larger than T, the weights to all (10) active MSN units was increased by 0.1, otherwise their weights were decreased by 0.1. For the next iteration, the weights were normalised (i.e., divided by the sum of weights) in order to obtain a new set of activation probabilities.

Each simulation ran for 10,000 iterations. There were 100 simulation trials for each of 137 threshold levels. The threshold levels were run from the minimum and maximum of $D_{(UAF|P(target\_MSN) = 0)}$, which were 15 and 151, respectively. The mean of $D_{(UAF|P(target\_MSN) = 0)}$ was 64.98 (blue vertical line in figure 6C) and of $D_{(UAF|P(target\_MSN) = 1)}$ was 91.00. The threshold setting that optimally separates the two distributions is $T = 78$ (green vertical line in figure 6C).

Results and discussion

The results are shown in figure 6. Figures 6A and 6B show the activation probability of the target MSN unit for different levels of T over 100,000 iterations or for time = 100,000, respectively. Figures 6A and 6B demonstrate that the model is learning to increase the probabilities, but that there are two limits. First, there are asymptotic levels that depend on the threshold setting. The closer the threshold setting is to the optimal setting (T = 78), the faster is the learning (green trajectory in figure 6A) and the higher the asymptotic value (on the vertical green line in figure 6B). Figure 6B indicates that with a high threshold setting the asymptotic level drops. Figure 6D shows the percentage of trials (out of 100) in which the model learns, i.e., the target MSN unit increases in synaptic weight. Thus, the second limit is that when the threshold is set higher than the optimal level, the model increasingly does *not* learn. When the threshold is set too high, negative weight changes dominate, leading to unlearning of all connections. Due to the probabilistic nature of the sampling, a dominating negative weight change can only occur at thresholds higher than the mean of the initial distribution, i.e., T > 64.98. This simulation shows that the dominance occurs at thresholds higher than the optimal setting. Due to the balance of learning and unlearning, the number of participating MSN units varies. The result is a small collection of MSN units (out of 1000, including the target MSN unit) that form a stable response plan for the protocol. This resonates with Birbaumer et al. (2013) assertion that learning sharpens the response plan.

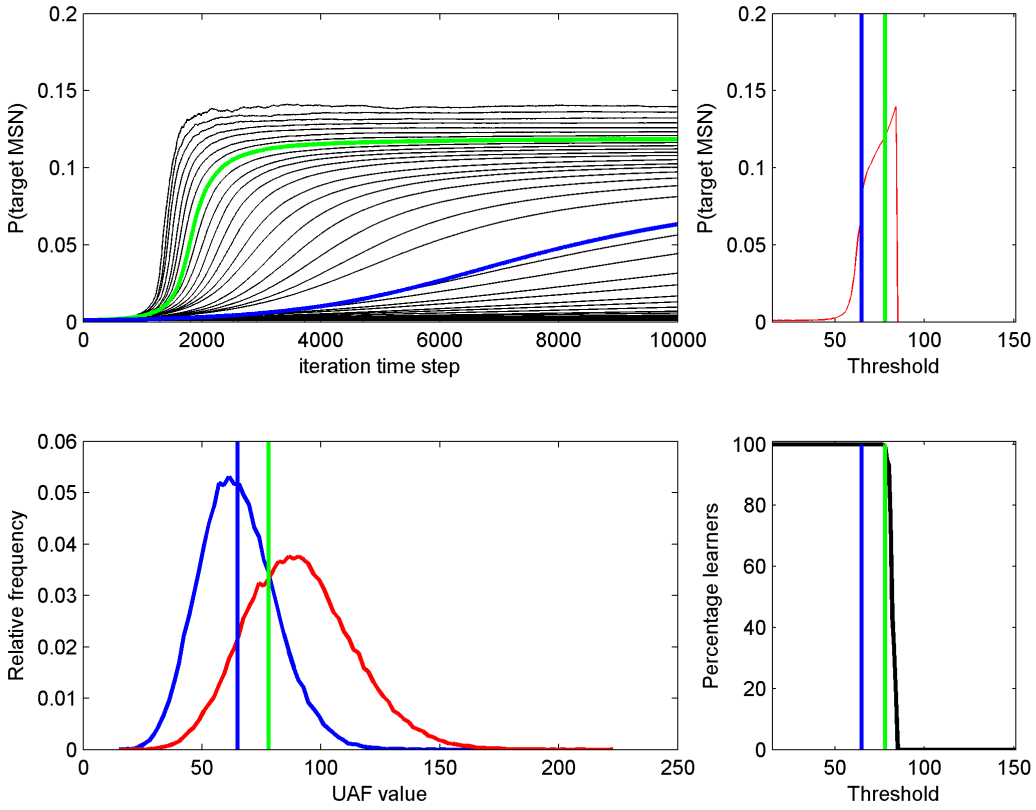

Figure 6. Results of simulation study 2. A. Trajectory of the averaged activation probability of the target MSN unit over 10,000 iterations for all simulated threshold levels. The blue trajectory corresponds to a threshold level that is the mean of the baseline distribution. The green trajectory is associated with the optimal threshold setting. B. The averaged activation probability of the target MSN unit at time = 10,000 iterations. Note the sharp rise and fall below and above the threshold corresponding to the mean of the baseline distribution and the optimal threshold, respectively. C. Relative frequency distributions used in the simulation study. Both the baseline (blue) and target (red) distributions consist of 100,000 samples simulated using the model from simulation study 1. Also shown are the mean of the baseline distribution (blue vertical line) and the optimal threshold setting (green vertical line). D. The percentage of simulation runs (100 runs in total per threshold level) that learned. Note the sharp drop in learners when the threshold is set higher than the optimal threshold (green vertical line).

There is a critical range of threshold settings that supports learning and within this range there is a narrow temporal window in which most of the learning takes place. What happens within these critical ranges? To address this question, the responses were classified as hits, misses, correct rejections, and false alarms and averaged for every 100 iterations, providing a time series of 100 epochs.

Figure 7A shows the evolution of the averaged UAF values in each epoch for different threshold levels: the learning curves. There is a clear separation between the lowest and highest values. Figures 7B and 7C show the distributions of UAF values across thresholds for the first and final epochs, respectively. As expected, the level of the UAF values at the beginning centre on the mean of the initial distribution, $D_{(UAF|P(target\_MSN) = 0)}$. Figures 7C shows the sharp divide between simulations in which learning takes place and those in which learning is absent. The sudden change in the UAF values over epochs (see figure 7A) suggests that reaching the right state leads to a self-promoting cycle akin to an attractor state. This can be understood as follows. When the target MSN unit is active, a UAF value is sampled from $D_{(UAF|P(target\_MSN) = 1)}$, which is more likely to lead to a positive

feedback than a value sampled from $D_{(UAF|P(target\_MSN) = 0)}$. The reward signal leads to an increase in probability of the target MSN unit to be active again on the next iteration, which leads again to sampling from $D_{(UAF|P(target\_MSN) = 1)}$, and so on. This attractor dynamics explains the suddenness of the change in averaged UAF values.

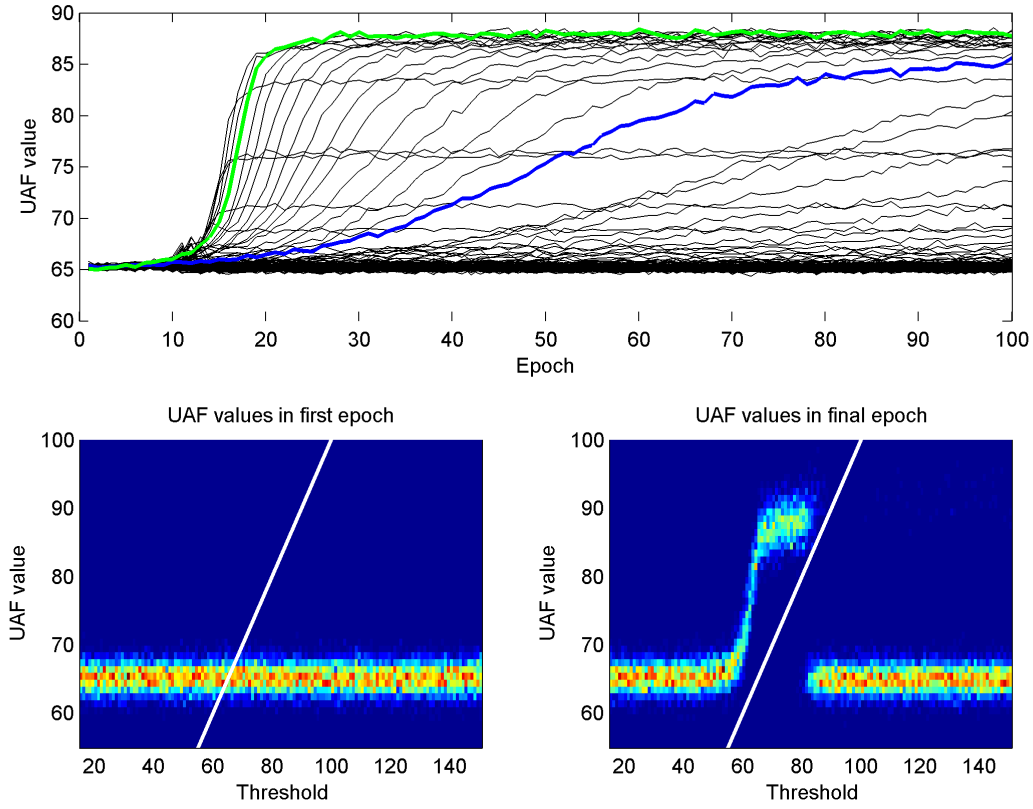

Figure 7. Results on the UAF values sampled in simulation study 2. A. Learning curves showing the average UAF value in each of 100 epochs (1 epoch = 100 iterations) for every threshold setting. B. Distributions of UAF values in the first epoch across threshold levels. As no learning has taken place, the distribution is identical tot he baseline distribution. C. Distributions of UAF values in the final epoch across threshold levels. Learning has taken place for some threshold levels. Note that within the narrow range, the distribution of samples is removed from the baseline distribution. The white diagonal line in 7B and 7C indicates when the threshold level equals the UAF value.

The initial probability of activating the target MSN unit is very low, but learning, or more precisely unlearning, is taking place prior to reaching the target attractor state. Figure 8A shows the difference in hits and false alarms as a function of epoch and threshold. The values are maximal in the narrow range of thresholds between the medians of $D_{(UAF|P(target\_MSN) = 0)}$ and $D_{(UAF|P(target\_MSN) = 1)}$. The maximum occurs earlier for higher thresholds within this range. Figure 8B shows that the probability of the target MSN unit being active in each epoch is close to binary, again indicating the attractor behaviour. To understand what drives this pattern, the probability of correctly rejecting a UAF value as coming from $D_{(UAF|P(target\_MSN) = 1)}$ (figure 8C) is compared against the hit rate (figure 8D). The attractor state is not entered because of increasing the weights to the target MSN unit, but by decreasing the weights to non-target units. The normalisation that follows increases the relative probability of activating the target MSN unit. Only when this happens, the attractor dynamics will quickly make the activation probability maximal. A threshold set lower than the median of $D_{(UAF|P(target\_MSN) = 0)}$ or higher than the median of $D_{(UAF|P(target\_MSN) = 1)}$ will lead to increasing

the likelihood of activating non-targets and decreasing the likelihood of activating the target MSN unit, respectively.

To conclude, this simulation demonstrates that learning in this abstracted version of neurofeedback training is due to unlearning of unsuitable representations followed by enhancing the likelihood of the suitable one.

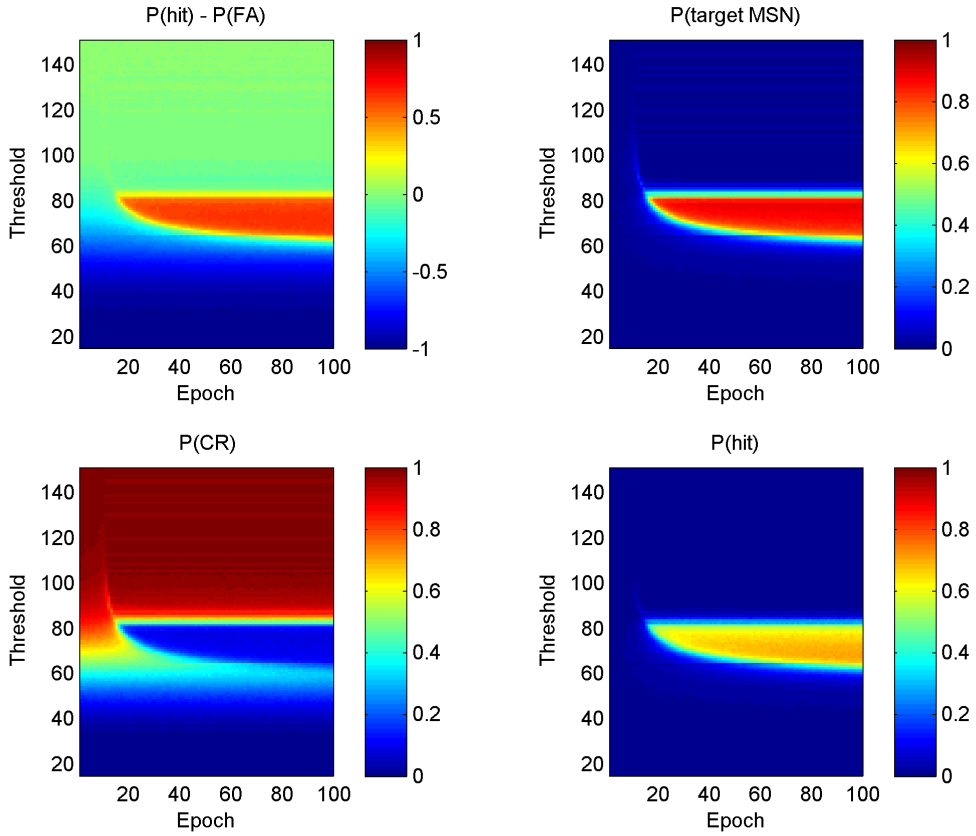

Figure 8. Signal-detection theoretic analyses of simulation study 2. A. Difference between hit rate and false alarm rate as a function of epoch and threshold setting. The results show a narrow range of threshold levels and a temporal window in which the samples are correctly identified as coming from the target distribution. B. Actual probability of the target MSN being active as a function of epoch and threshold setting. This figure demonstrates that the system is almost completey binary in activating the target MSN unit, which supports the attractor-style dynamics discussed in the text. C. Probability of correctly rejecting samples as coming from the target distribution as a function of epoch and threshold setting. This reveals that the transition observed in 8B requires levels of correct rejections between 0.4 and 0.8. D. Probability of correctly accepting samples as coming from the target distribution as a function of epoch and threshold setting. The figure shows that hit rate is not driving the almost sudden transition observed in 8B.

General discussion

This paper addressed the neural mechanisms involved in learning to change the EEG spectral profile at a given electrode site. The multi-stage theory of neurofeedback learning consists of three stages, striatal-exploration, thalamic-consolidation, and interoceptive-homeostasis (or insular-homeostasis). The early period of learning is dominated by frontal brain areas generating different representations and maintaining those that produce a positive feedback signal. The middle period is characterised by activating the winning frontal representation and modifying the synaptic connections to and within the thalamus. The final phase involves

awareness of the bodily state associated with the target brain state and use the subjective experiential representation as secondary reinforcer for closing the homeostatic loop.

Although the theory involves several stages, these stages are not ordered in terms of one stage finishing before the next one starts, instead the stages are ordered in sequence of coming online during the overall neurofeedback learning process. This means that after stage 2 or 3 has got some momentum they alter the neural signal and thus the distribution of reward signals, which in turn impacts on the stage 1 fronto-striatal learning. In order to assess the interactivity of these stages, measurements at multiple timescales are needed. In addition to interactivity, the stages decrease in essentiality to neurofeedback learning. Whereas reward-based learning is the core process in stage 1, neurofeedback learning does not need thalamic-level of synaptic modifications. Within-session improvements may simply reflect the results of stage 1 learning. Thalamic synaptic changes may be reflected in improvements over sessions and in particular in the baseline blocks across sessions. Finally, interoceptive awareness may not occur at all or when it occurs may not lead to noticeable improvements. However, evidence for such a stage is reflected in the differential subjective experiences associated with different protocols (Edge and Lancaster, 2004).

Related to topic of subjective experiences is the notion that neurofeedback does not require consciousness at all (Birbaumer et al., 2013; Chang et al., 2014). Although the multi-stage theory allows for explicit processes to influence learning, the striatal-exploration stage is unconscious to extent that the actual response plan may not be amenable to introspection. In such a situation, different trainees may provide different answers to the question of what they did to increase reward frequency. Note that the assessment of subjective experience requires more than a question such as "What did you do as a strategy?". Instead, we (Davelaar et al. 2016) advocate that a time is spent interviewing the trainee to get a detailed narrative that is then subjected to a phenomenological analysis. On the other side, the interoceptive representation that emerges during stage 3 need not be verbalisable. For example, just being able to recognise the "gut feeling" is sufficient to be able to say that there exists a subjective representation. Such a representation can become a secondary reinforcer during the neurofeedback training sessions and can be used as the only reinforcer when the trainee is not connected to the EEG equipment. As a result, further training using only subjective representations becomes very much like meditation practises.

The aim of this paper was two-fold: introducing the multi-stage theory of neurofeedback learning and demonstrating how a methodology, computational modelling, in neuroscience can contribute to neurofeedback research. The computational model was not targeted to implement the entire theory, but instead focused on the very first stage of neurofeedback learning. It was shown that an off-the-shelf spiking neuron model of alpha generation augmented with a striatal learning component was able to demonstrate the unfolding of the striatal-exploration phase. From analyses of an abstracted version based on distributions of UAF values it was found that the neurofeedback model learns initially by correctly rejecting low UAF values as coming from the target state, which increases the likelihood of the target state, from which point the model continues to learn through correctly accepting high UAF values as coming from the target state. The end result is that the model samples from an asymptotic density function that is close to the target distribution.

Although the simulations are necessarily oversimplifications, a few limitations are worthy of mention to be taken into account in future extensive analyses. First, the sampling of UAF values from the distribution is independent of the sampling history. This means that the model does not capture the autocorrelation present in actual data. High levels of autocorrelation are associated with decreased information content. Thus, if an analysis is needed to address the system-level dynamics, the model needs to account for autocorrelation,

either through sampling changes in UAF values from a parent distribution or by using the type of model used in simulation study 1.

Another limitation is the focus on a single "electrode site", which assumes that neurofeedback using one electrode will not influence the spectral profile observed at another electrode. Several empirical reports have documented the nonspecificity of EEG neurofeedback, with changes in power occurring in non-target frequencies and/or at non-target electrodes (e.g., Egner et al., 2004; Haarmann et al., 2012; Ros et al., 2013) and structural changes occurring in areas removed from the EEG electrode sites (Ghaziri et al., 2013). These observations support the argument for a network-level instead of localisation approach to neurofeedback (Gruzelier, 2014c) which is further strengthened by the finding that neurofeedback reconfigures functional connectivity in brain networks (e.g., Haller et al., 2013; Ros et al., 2013). The theory accommodates these effects through the provision of modification of synaptic strengths at the reticular thalamic nucleus that originate from the basal ganglia (e.g., Gasca-Martinez et al., 2010) or prefrontal cortex (e.g., Zikopoulos and Barbas, 2006). Interregional oscillations may occur when two regions share the same specific thalamic nucleus and/or the same reticular projection zone (for a computational account see Victor et al., 2011).

A major utility of a model is that it can also be used to test various proposed indices on which to base the feedback or quantify learning. For example, Dempster and Vernon (2009) trained participants to increase alpha (8-12 Hz) over Pz. Feedback was analogue and based on the alpha amplitude above the average baseline amplitude. They compared amplitude, percent time, and integrated alpha as indices of learning and found that the amplitude measure was superior in showing learning across sessions both with and without baseline correction. Discussion on the appropriate index has not converged and more extensive work is needed, but such an endeavour is very time-consuming (10 weeks/participant in Dempster and Vernon, 2009). Without a measure of effect size to calculate the number of participants needed to obtain a reasonable level of statistical power, such a study could easily suffer from underpowered designs. A model could run the suggested designs and provide estimates of effect size that can be used in planning the studies.

The theory also makes testable predictions that could further the field. For example, it predicts that dopaminergic signals are necessary for learning to occur, which provides a pathway to increase or decrease neurofeedback efficacy or understand why certain populations are non-trainable. It also predicts that training on the same protocol can follow different learning pathways dependent on the dynamic status of the cortico-striatal-thalamic circuitry and thus produce different patterns of synaptic change. Dry-testing predictions through computational implementation enhances the cost-effectiveness of empirical studies and can guide future development of research methodology and data-analytics.

The work presented here paves the way for more detailed computational efforts. Two areas for further developments are (1) using more neurophysiologically realistic neural circuitries, and (2) applying the model to test commonly used methods. Regarding the former, directions for research include developing models with laminar organisation that have explicit cortico-thalamic and cortico-basal ganglia-thalamic loops, which can produce multiple brain rhythms and incorporate synaptic plasticity that are sensitive to neuromodulators involved in reinforcement learning. Importantly, the theory is not limited to EEG, but is in principle applicable to rt-fMRI or rt-NIRS. Regarding the latter, various choices regarding threshold setting, windowing, Fourier versus filtering, feedback type and feedback modality, (and more) can be systematically investigated without running lengthy and costly studies. The system can be subjected to mathematical analysis to address optimisation in high-dimensional search.

Conclusion

This paper presented a theoretical account of mechanisms underlying neurofeedback learning and analysed a minimal model of the neural learning dynamics underlying successful EEG neurofeedback. The view is that models developed in the field of computational neuroscience can be applied to address the thorny question in the field of neurofeedback: "How does neurofeedback work?". Model analyses show that neurofeedback can be likened to a search process that first rejects representations that produce low levels of the desired EEG profile, which increases the likelihood of activating a representation that produces the desired levels. The success of this minimal model provides the necessary proof of concept that neurofeedback research can be facilitated by incorporating computational approaches developed in other disciplines to address theoretical and practical questions. This paper demonstrated that such an endeavour has merit.

Acknowledgments

I thank Mervyn Etienne for providing the EEG data shown in figure 5B and Henk Haarmann for detailed comments and suggestions. This research did not receive any specific grant from funding agencies in the public, commercial, or non-for-profit sectors.

Disclosure Policy

The author declares that there is no conflict of interest regarding the publication of this paper.

Arsenault JT, Rima S, Stemmann H, Vanduffel W (2014) Role of the primate ventral tegmental area in reinforcement and motivation. Curr Biol 24:1347–1353.

Birbaumer N, Elbert T, Canavan A, Rockstroh B (1990) Slow potentials of the cerebral cortex and behavior. Physiol Rev 70:1-41.

Birbaumer N, Ruiz S, Sitaram R. (2013) Learned regulation of brain metabolism. Trends Cogn Sci 17:295-302.

Budzynski TH, Budzynski HK, Evans JR. Abarbanel A. (2009) Introduction to quantitative EEG and neurofeedback. Academic Press: San Diego.

Chang M, Iizuka H, Naruse Y, Ando H, Maeda T (2014) Unconscious learning of auditory discrimination using mismatch negativity (MMN) neurofeedback. Scientific Reports 4:6729.

Ceunen E, Vlaeyen JWS, Van Diest I (2016) On the origin of interoception. Front Psychol 7:743.

Cohen MX (2014) Fluctuations in oscillation frequency control spike timing and coordinate neural networks. J Neurosci 34:8988–8998.

Craig AD (2002) How do you feel? Interoception: the sense of the physiological condition of the body. Nat Rev Neurosci 3:655-666.

Davelaar EJ, Barnby JM, Almasi S, Eatough V (2016) Neurophenomenology and neurofeedback: a pilot study. Front Hum Neurosci Conference Abstract: SAN2016 Meeting.

Dempster T, Vernon D (2009) Identifying indices of learning for alpha neurofeedback training. Appl Psychophys Biof 34:309-318.

Edge J, Lancaster L (2004) Phenomenological analysis of superior musical performance facilitated by neurofeedback: enhancing musical performance through neurofeedback: playing the tune of life. Transpers Psychol Rev 8:23-35.

Egner T, Zech TF, Gruzelier JH (2004) The effects of neurofeedback training on the spectral topography of the electroencephalogram. Clin Neurophysiol 115:2452-2460.

Emmert K, Kopel R, Sulzer J, Brühl AB, Berman BD, Linden DEJ, Horovitz SG, Breimhorst M, Caria A, Frank S, Johnston, S., Long Z, Paret C, Robineau F, Veit R, Bartsch A, Beckmann CF, Van De Ville D, Haller S (2016) Meta-analysis of real-time fMRI neurofeedback studies using individual participant data: how is brain regulation mediated? Neuroimage 124:806-812.

Evans JR, Abarbanel A (1999) Introduction to quantitative EEG and neurofeedback. Academic Press: San Diego.

Gasca-Martinez D, Hernandez A, Sierra A, Valdiosera R, Anaya-Martinez V, Floran B, Erlij D, Aceves J (2010) Dopamine inhibits GABA transmission from the globus pallidus to the thalamic reticular nucleus via presynaptic D4 receptors. Neuroscience 169:1672-1681.

Geisler S, Derst C, Veh RW, Zahm DS (2007) Glutamatergic afferents of the ventral tegmental area in the rat. J Neurosci 27:5730-5743.

Ghaziri J, Tucholka A, Larue V, Blanchette-Sylvestre M, Reyburn G, Gilbert, G, Lévesque J, Beauregard M (2013) Neurofeedback training induces changes in white and gray matter. Clin EEG Neurosci 44:265–272.

Gruzelier JH (2009) A theory of alpha/theta neurofeedback, creative performance enhancement, long distance functional connectivity and psychological integration. Cogn Process 10:S101-109.

Gruzelier JH (2014a) EEG-neurofeedback for optimising performance. I: a review of cognitive and affective outcome in healthy participants. Neurosci Biobehav R 44:124-141.

Gruzelier JH (2014b) EEG-neurofeedback for optimising performance. II: creativity, the performing arts and ecological validity. Neurosci Biobehav R 44:142-158.

Gruzelier JH (2014b) EEG-neurofeedback for optimising performance. III: a review of methodological and theoretical considerations. Neurosci Biobehav R 44:159-182.

Gruzelier JH, Bamidis P, Babiloni F, de Ridder D (2014a) Applied neuroscience: models, methods, theories, reviews. Neurosci Biobehav R 44:1-3.

Gruzelier JH, Bamidis P, Pagiani L, Reiner M, Ros, T (2014b) Applied neuroscience. Int J Psychophysiol, 93: ix-xii.

Haarmann HJ, George T, Smaliy A, Dien J (2012) Remote associates test and alpha brain waves. J Probl Solv 4:5.

Haller S, Kopel R, Jhooti P, Haas T, Scharnowski F, Lovblad K-O, Scheffler K, Van de Ville D (2013) Dynamic reconfiguration of human brain functional networks through neurofeedback. NeuroImage 81:243-252.

Hanslmayr S, Sauseng P, Doppelmayr M, Schabus M, Klimesch W (2005) Increasing individual upper alpha power by neurofeedback improves cognitive performance in human subjects. Appl Psychophys Biof 30:1-10.

Hinterberger T, Veit R, Wilhelm B, Weiskopf N, Vatine J-J, Birbaumer N (2005) Neuronal mechanisms underlying control of a brain-computer interface. Eur J Neurosci 21:3169-3181.

Hsieh LT, Ranganath C (2014) Frontal midline theta oscillations during working memory maintenance and episodic encoding and retrieval. Neuroimage 85:721-729.

Hughes SW, Crunelli V (2005) Thalamic mechanisms of EEG alpha rhythms and their pathological implications. Neuroscientist 11:357-372.

Izhikevich EM (2003) Simple model of spiking neurons. IEEE T Neural Networ 14:1569–1572.

Izhikevich EM, Edelman GM (2008) Large-scale model of mammalian thalamocortical systems. P Natl Acad Sci USA 105:3593-3598.

Jarosiewicz B, Chase SM, Fraser GW, Velliste M, Kass RE, Schwartz AB (2008) Functional network reorganization during learning in a brain-computer interface paradigm. P Natl Acad Sci USA 105:19486–19491.

Johnston SJ, Boehm SG, Healy D, Goebel R, Linden DEJ (2010) Neurofeedback: a promising tool for the self-regulation of emotion networks. NeuroImage 49:1066-1072.

Klimesch W (1999) EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. Brain Res Rev 29:169-195.

Klimesch W, Doppelmayr M, Russegger H, Pachinger T, Schwaiger J (1998) Induced alpha band power changes in the human EEG and attention. Neurosci Lett 244:73-76.

Kluetsch RC, Ros T, Théberge J, Frewen PA, Calhoun VD, Schmahl C, Jetly R, Lanius R A (2014) Plastic modulation of PTSD resting-state networks and subjective wellbeing by EEG neurofeedback. Acta Psychiat Scand 130:123-136.

Koralek AC, Jin X, John D. Long JD, Costa RM, Carmena JM (2012) Corticostriatal plasticity is necessary for learning intentional neuroprosthetic skills. Nature 483:331-335.

Legenstein R, Chase SM, Schwartz AB, Maass W (2010) A reward-modulated hebbian learning rule can explain experimentally observed network reorganization in a brain control task. J Neurosci 30:8400-8410.

Legenstein R, Pecevski D, Maass W (2008) A learning theory for reward-modulated spike-timing-dependent plasticity with application to biofeedback. PLoS Comput Biol 4:e1000180.

Levesque J, Beauregard M, Mensour, B (2006) Effect of neurofeedback training on the neural substrates of selective attention in children with attention-deficit/hyperactivity disorder: A functional magnetic resonance imaging study. Neurosci Lett 394:216-221.

Moran RJ, Campo P, Maestu F, Reilly RB, Dolan RJ, Strange BA (2010) Peak frequency in the theta and alpha bands correlates with human working memory capacity. Front Hum Neurosci 11:200.

Nan W, Rodrigues JP, Ma J, Qu X, Wan F, Mak P, Mak PU, Vai MI, Rosa A (2012) Individual alpha neurofeedback training effect on short term memory. Int J Psychophysiol 86:83-87.

Niv S (2013) Clinical efficacy and potential mechanisms of neurofeedback. Pers Ind Diff 54:676–686.

Omata K, Hanakawa T, Morimoto M, Honda M (2013) Spontaneous slow fluctuation of EEG alpha rhythm reflects activity in deep-brain

structures: a simultaneous EEG-fMRI study. PLoS ONE 8:e66869.

Palva S, Palva JM (2007) New vistas for alpha-frequency band oscillations. Trends Neurosci 30:150-158.

Ritter P, Moosmann M, Villringer A (2009) Rolandic alpha and beta EEG rhythms' strengths are inversely related to fMRI-BOLD signal in primary somatosensory and motor cortex. Hum Brain Mapp 30:1168-1187.

Ros T, Baars BJ., Lanius RA, Vuilleumier P (2014) Tuning pathological brain oscillations with neurofeedback: a systems neuroscience framework. Front Hum Neurosci. 8:1008

Ros T, Theberge J, Frewen PA, Kluetsch R, Densmore M, Calhoun VD, Lanius RA (2013) Mind over chatter: plastic up-regulation of the fMRI salience network directly after EEG neurofeedback. Neuroimage 65:324-335.

Sterman MB (1996) Physiological origins and functional correlates of EEG rhythmic activities: implications for self-regulation. Biofeedback Self-Reg 21:3-33.

van Boxtel GJM, Gruzelier JH (2014) Neurofeedback. Biol Psychol 95:1-3.

Van Drongelen W (2013) Modeling neural activity. ISRN Biomath 871472.

Victor JD, Drover JD, Conte MM, Schiff ND (2011) Mean-field modeling of thalamocortical dynamics and a model-driven approach to EEG analysis. P Natl Acad Sci USA 108:15631-15638.

Watabe-Uchida M, Zhu L, Ogawa SK., Vamanrao A, Uchida N (2012) Whole-brain mapping of direct inputs to midbrain dopamine neurons. Neuron 74:858-873.

Zikopoulos B, Barbas H (2006) Prefrontal projections to the thalamic reticular nucleus form a unique circuit for attentional mechanisms. J Neurosci 26:7348-7361.

Zoefel B, Huster RJ, Hermann CS (2011) Neurofeedback training of the upper alpha frequency band in EEG improves cognitive performance. NeuroImage 54:1427-1431.