A Multi-stage Theory of Neurofeedback Learning

Eddy J. Davelaar[(✉)] [iD]

Birkbeck, University of London, London WC1E 7HX, UK
e.davelaar@bbk.ac.uk

Abstract. Neurofeedback is a training paradigm through which trainees learn to voluntarily influence their brain dynamics. Recent years have seen an exponential increase in research interest into this ability. How neurofeedback learning works is still unclear, but progress is being made by applying models from computational neuroscience to the neurofeedback paradigm. In this chapter, I will present a multi-stage theory of neurofeedback learning, which involves three stages, involving different neural networks. In stage 1, the system discovers the appropriate goal representation for increasing the frequency of positive feedback. This stage operates at a within-session timescale and is driven by reward-based learning, which updates fronto-striatal connections. Stage 2 operates on a timescale that covers multiple training sessions and is sensitive to consolidation processes. This stage involves updating striatal-thalamic and thalamo-cortical connections. Finally, after stages 1 and 2 have started, stage 3 may be triggered by the awareness of the statistical covariation between interoceptive and external feedback signals. When this awareness emerges, neurofeedback learning may speed up and its effect be maintained well after the conclusion of the training period. Research guided by this framework is described that consist of quantitative, qualitative, and computational methodologies. At present, the findings suggest that the framework is able to provide novel insights into the nature of neurofeedback learning and provides a roadmap for developing instructions that are designed to facilitate the likelihood of learning success.

Keywords: Neurofeedback · Computational neuroscience · EEG · Neuroimaging

1 Introduction

In the early 1900s, psychological research was dominated by a methodological approach championed by behaviorism that looked into the smallest components that could explain behavior. This approach led to discoveries of classical and operant conditioning that could explain behaviors observed in animals and humans. It also influenced thinking about education, parenting, and advertising. A lesser known fact is that conditioning principles were not only investigated at the level of overt behaviors, such as pressing a lever or pecking at a light, but work by a range of researchers uncovered

that brain oscillations could be conditioned as well [1–3]. This work was the foundation of a field called neurofeedback in which the brain's activation is modified through conditioning. Neurofeedback was mainly focused on electroencephalography (EEG) for many decades, but over the last 20 years and especially over the recent 5 years a range of other brain measurement techniques have been used, such as magnetoencephalography (MEG, [4, 5]), functional magnetic resonance imaging (fMRI, [6, 7]), and functional near-infrared spectroscopy (fNIRS, [8, 9]). This increased interest brings with it a need to understand the biological mechanisms underlying neurofeedback learning. In addition, whilst neurofeedback alters the brain activation, it has been argued that it also leads to functional changes in performance and subjective experience. This is the foremost reason that neurofeedback has a long history as a neurotherapeutic intervention for psychological conditions, with epilepsy and attention deficit hyperactivity disorder, as two of the earlier conditions for which therapeutic benefits were recorded. Understanding the mechanisms underlying neurofeedback learning will benefit clinicians in improving their success rate and enhance hypothesis-driven research. This chapter will review some of the early literature on conditioning of EEG oscillations. This is followed by a discussion on the current state of research, highlighting the various methodological challenges. A multi-stage theory of neurofeedback learning is then introduced, with each stage addressed in greater detail. Finally, the theoretical and practical implications of this model will be explicated.

2 Conditioning of EEG Oscillations: The Early Work

Behaviorist research in the early 1900s was focused on tabulating the smallest association that could lead to an overt behavior. As mentalist topics such as attention and memory were not directly observable, these were not regarded as acceptable areas of inquiry. What was acceptable was measuring any physiological variable and check whether it could be conditioned. In a largely forgotten literature, a particular focus was whether stimuli presented to any sense organ could become a conditioned stimulus for what is called alpha blocking.

2.1 Classical Conditioning

Alpha blocking is the phenomenon that the power of the alpha oscillation over the visual areas decreases when the participant opens the eyes. This phenomenon was considered a natural reflex. Several studies addressed whether this reflex could be conditioned to stimuli other than light. Jasper and Shagass [1] conducted an extensive investigation with sound. Their participants were on a bed in a darkened room with an electrode attached over the right occipital area. They were tasked with pressing a response button as soon as they saw a light. Pressing the button did not have any effect of the alpha oscillation. In conditioning trials, the light was preceded by a tone. This tone became the conditioned stimulus, as demonstrated by the decrease of alpha power after playing the tone, but without illuminating the light. The simple conditioning occurred quite quickly, but was also rapidly extinguished.

Jasper and Shagass [1] investigated the conditioned alpha block using simple, cyclic, delayed, trace, differential, differential delayed, and backward conditioning, thereby

establishing that "higher centres, not necessarily involving peripheral effector systems" (p. 384) can be Pavlovian conditioned. An interesting side note is that during the extinction period, spontaneous recovery of the conditioned response can occur, which indicates that not only that the tone got associated with the presence of light, but also with its absence during extinction. The spontaneous recovery occurs when the relative strength of the former outweighs the latter.

In a follow-up study, Jasper and Shagass [2] asked participants to subvocally say "block" and press a button and keep it pressed until subvocally saying "stop" and releasing the button. The timing of the subvocalisation was entirely voluntary. During conditioning trials, the experimenter allowed the electrical circuit to switch on a light when the participant depressed the button. On test and extinction trials, the light remained switched off. A conditioned response was observed in this scenario in the absence of an external stimulus. This study demonstrated that a conscious mental act was able to become a conditioned stimulus.

In both involuntary and voluntary studies of the alpha block, there was rapid extinction of the conditioned response. Nevertheless, it highlighted that the brain and its higher centres follow the same rules of conditioning as overt behaviors, such as the salivation reflex.

2.2 Operant Conditioning

Whereas the conditioned alpha block is considered a Skinner Type II conditioned response, Wyrwicka and Sterman [3] demonstrated that brain oscillations can be conditioned through operant conditioning. In their study, they had cats who were deprived of food for 22 h. They received condensed milk whenever they exhibited a burst of sensorimotor rhythm (SMR) over the sensorimotor cortex for at least 0.5 s. The cats were able to increase the occurrence of SMR. Of particular interest are the visually recorded behaviors that the cats engaged in. All cats converged on a different posture that can be described as freezing or staring. Immediately after the milk was consumed the cats returned and adopted the same posture. In addition, when after an extinction period a reconditioning phase started, the cats returned again to their individual posture. Thus, not only was the SMR oscillation subject to operant conditioning, it also coincided with a specific behavior that took different forms. In lay terms, it is as if the cats had to "go into their zone", after which SMR developed in a few seconds.

Similar correlations have been observed in other studies with human participants. For example, in alpha training participants report different phenomenology [10, 11], which suggests that changes in the brain activation profile influences the subjective experience. However, it is yet unclear whether the subjective experience is shared among individuals or idiosyncratic. Some initial work in this direction is currently being conducted [12, 13].

3 Neurofeedback Research: Current Developments

Despite over eight decades of research in neurofeedback, the field is currently at a crossroads. Most of the research is conducted using EEG and has spawned several

debates, such as whether the conditioned alpha block is actually reflecting sensitization and whether clinical trials are appropriately placebo-controlled. The demonstration of successful neurofeedback of the BOLD signal has opened up a much wider field with its own technical challenges. Together with a lack of theoretical framework for generating hypotheses, the consequence has been that EEG neurofeedback is still being considered as flawed.

At the time of writing, there are major developments afoot that rehabilitate EEG neurofeedback. Dedicated special issues on the topic feature many methodological and technical advances that were not available two decades ago. In addition, dedicated software for research purposes are being developed, some are expected to be Open Accessible. Sharing of data through the Open Science Framework and pre-registration of studies are being considered and implemented.

Although the challenges of the research environments are being met, the theoretical developments are still in need of major work. General high-level descriptions have been proposed to provide a bird's eye view of neurofeedback. However, generating testable hypotheses from these perspectives has remained elusive. The proposal in this chapter is that models from computational neuroscience could be utilized to develop a mechanistic understanding of neurofeedback learning. These models could then be used to implement new research designs and generate hypotheses. They can be used to test some of the higher-level descriptions of neurofeedback learning, thereby allowing comparison of different theoretical viewpoints.

4   A Multi-stage Theory of Neurofeedback Learning

The empirical research base is rich and vast enough for developing formal theories to further drive the field forward. Unfortunately, the marriage between theorists within computational neuroscience and researchers in applied neuroscience never took hold. When taking a computational neuroscience approach to neurofeedback, insights can be gained that were not obvious at first. The multi-stage theory of neurofeedback learning [14] is a product of merging the two disciplines.

4.1   Overview of the Theory

The theory assumes three stages that involve different neural networks (see Fig. 1). In stage 1, the system discovers the appropriate goal representation for increasing the frequency of positive feedback. This stage operates at a within-session timescale and is driven by reward-based learning, which updates fronto-striatal connections. Stage 2 operates on a timescale that covers multiple training sessions and is sensitive to consolidation processes that unfold during sleep. This stage involves updating striatal-thalamic and thalamo-cortical connections. In effect, this stage changes the set point of the system, making it easier to produce the target brain oscillation. Finally, after stages 1 and 2 have started, stage 3 may be triggered by the awareness of the statistical covariation between interoceptive and external feedback signals. When this awareness emerges, neurofeedback learning may speed up and its effect be maintained well after the conclusion of the training period.
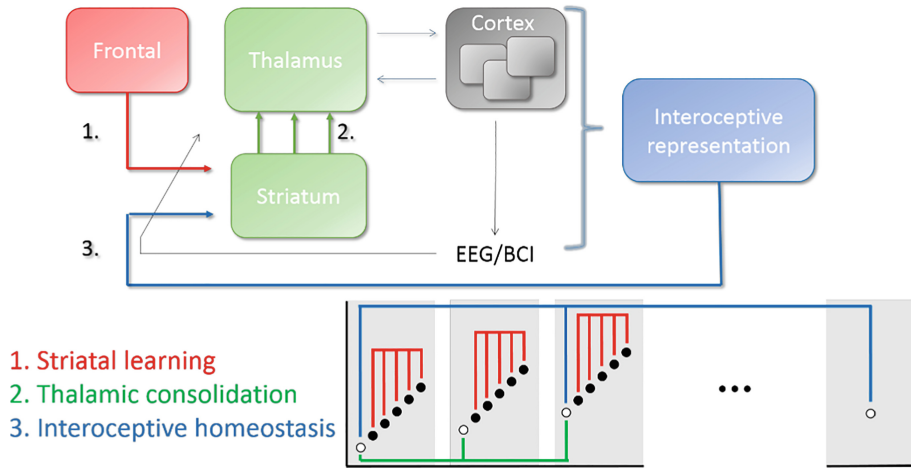
Fig. 1. The multi-stage theory of neurofeedback learning as applied to EEG neurofeedback. The BCI system records the EEG oscillations and convert this into a feedback signal. This signal is used in reinforcement learning through which the frontal goal representation gets associated with neural patterns over the striatum that lead to more positive feedback. This positive feedback loop is assumed to underlie within-session learning curves. The second stage starts after the first and involves updating the striatal-thalamic connections. As this unfolds over a longer time-scale, it is assumed to underlie the learning curves over sessions. Brain patterns may correspond with unique subjective experiences. When these exist for the target brain pattern, they make become secondary reinforcers and this stage is assumed to underlie self-reinforcement in the absence of external feedback and maintenance of the acquired skill.

As the stages operate at different time-scales, learning curves within and between sessions are hypothesized to reflect these different stages. In addition, the implication is that in research a positive learning curve could be observed within, but not between sessions. This would not mean that no neurofeedback learning occurred. Instead, it could mean that stage 2 does not occur for that neurofeedback protocol, as would be the case if changing the setpoint is a physiological impossibility. A closer look at learning curves is warranted to scrutinize their relation to particular protocols.

At present, the theory is mainly a framework in which to explain neurofeedback findings. However, as additional data is being addressed, gaps in the theory will inevitably become visible, which require dedicated hypothesis-driven research. This is the advantage of a detailed theory. To facilitate the identification of directions for further inquiry, each stage will now be addressed in more detail with presentation of work that was inspired by the theory.

4.2   Stage 1

Stage 1 assumes that during learning, a frontal representation that contains the person's goal (i.e., increase the number of reward signals) is associated with a random neural pattern over the striatum that increases the likelihood of reward. In the original article introducing the theory, the frontal and striatal parts were analyzed in computational

neuroscience and in a mathematical model. The task for this sub-model was to move from a state of producing the baseline EEG pattern to a state in which the target brain pattern (in that case alpha oscillations) was more likely. By implementing basic equations of reinforcement learning, the models were shown to be able to learn, demonstrating that computationally stage 1 of the theory is indeed possible.

The model produced two new insights. First, the space of possible striatal patterns is immense and finding the target pattern through trial-and-error is highly unlikely. The updates to the fronto-striatal connections interact with the strong intra-striatal inhibition to implement a selection mechanism that drives the system to converge on a stable pattern. Thus, neurofeedback learning in stage 1 is a search process. The convergence counters the positive feedback loop, making stage 1 a self-limiting process. In other words, learning in stage 1 will eventually stop. This insight has repercussions on the evaluation of learning success. In particular, learning success could be defined as a linear increase, as is typically done in the literature, or when an asymptotic level is nearly reached. These are different parts of a sigmoidal learning curve.

The second insight, based on mathematical analysis, is that the probabilistic nature of the EEG generation implies that not all target states are rewarded. This fundamentally changes the neurofeedback paradigm from the often assumed continuous reinforcement schedule (i.e., every target state is rewarded) to a variable intermittent schedule. This is due to the neurofeedback software to reward based on the EEG pattern, whereas the target state produces both the target pattern and the nontarget pattern. This has a consequence that target state is both rewarded and not rewarded in the same training session, which slows down overall learning.

The learning in the stage 1 is particularly sensitive to the threshold setting that is used to decide whether to provide rewards. Set the threshold too low and the system does not learn. Set too high the system unlearns. Yet, the model was used to demonstrate that changing the threshold as a function of the preceding recordings lead to steeper learning curves with higher asymptotic levels [15].

The stage-1 submodel challenges and extends a range of methodological choices that could be further explored. In addition, the interpretation of stage-1 learning as a search process influenced by threshold settings allows a careful consideration of how to devise algorithms for optimal thresholding. Finally, it also allows exploring variations in the feedback protocol. That is whether feedback should be binary (e.g., a beep) or continuous (e.g., volume change), be only positive (e.g., addition of points) or also include negative feedback (e.g., subtraction of points). These issues are currently being investigated.

4.3   Stage 2

Stage 2 puts the thalamus at its centre. However, this is not to say that all neurofeedback needs to involve the thalamus. For example, in fMRI neurofeedback of the amygdala, the striatum influences the amygdala response. In EEG neurofeedback, the thalamus is at the centre of brain oscillations. Addressing the hypothesis of thalamic consolidation or changing the setpoint was approached in a pure computational manner by quantitatively fitting a biophysical model of thalamocortical interactions to EEG data obtained before, during, and after neurofeedback training.

The model that was utilized was developed by Robinson and colleagues [16]. This model takes a mean field approach and contains a wide range of parameters that are constrained from neurophysiological data. This particular model has been solved by the authors and shown to quantitatively fit actual EEG spectra by changing neurophysiological parameters, such as intrathalamic, thalamocortical, and cortico-cortical connectivity. Applying this model to actual data allows evaluating whether thalamic connections are critical in understanding the change in EEG oscillations.

The data comes from a study in which participants were trained to increase alpha and theta over electrode Pz over the course of ten training sessions. Participants had their eyes closed and focused on the sounds of a babbling brook (alpha) and of the ocean (theta) for fifteen minutes. Their task was to increase the volume of both sounds. This particular protocol is known for a phenomenon called the alpha/theta crossover, whereby after a period of higher alpha power compared to theta power, a switch occurs where theta dominates the power spectrum. As part of an ongoing investigation, individual sessions were checked for the crossover pattern. This was found for one person in the tenth session (see Fig. 2).
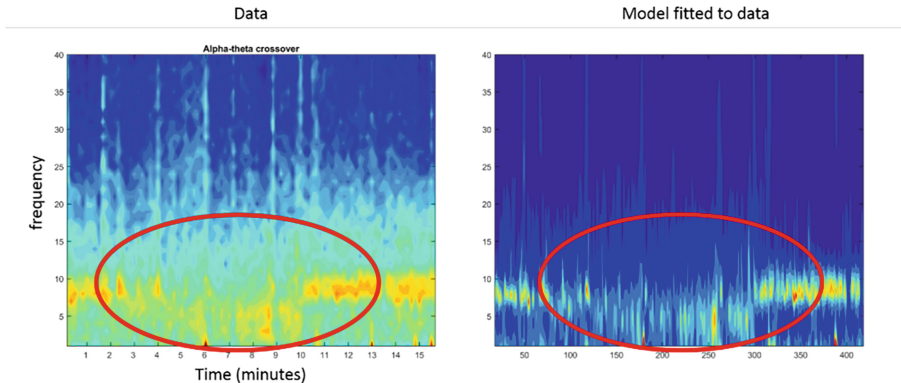
Fig. 2. Time-frequency spectra demonstrating the alpha/theta crossover (inside the red ovals). Left panel: Actual data. Right panel: Model fitted to the data. (Color figure online)

The biophysical model was fitted to the data and the values of the parameters were plotted against time (see Fig. 3). This fitting routine requires updating several parameters of which five are shown in Fig. 3. The dendritic rate constant relates to the rate of processing in the dendritic tree of pyramidal cortical cells. The inhibitory and excitatory gains relate to the cortico-cortical connections. These three parameters relate to cortical neurons only. The two thalamic parameters are the negative thalamic gain, which is composed of the pathway coming from the cortex to the reticular nucleus to the thalamic relay neurons and then back to the cortex. The negative feedback loop between the relay neurons and the reticular nucleus is labelled here as the intra-thalamic gain. The positive thalamic loop is not shown here.

The alpha/theta crossover period is clearly reflected in the decreased cortical excitatory gain. However, none of the three cortical parameters predict the ensuing cross-over. Both thalamic parameters show a gradual decrease in negative gain before the crossover
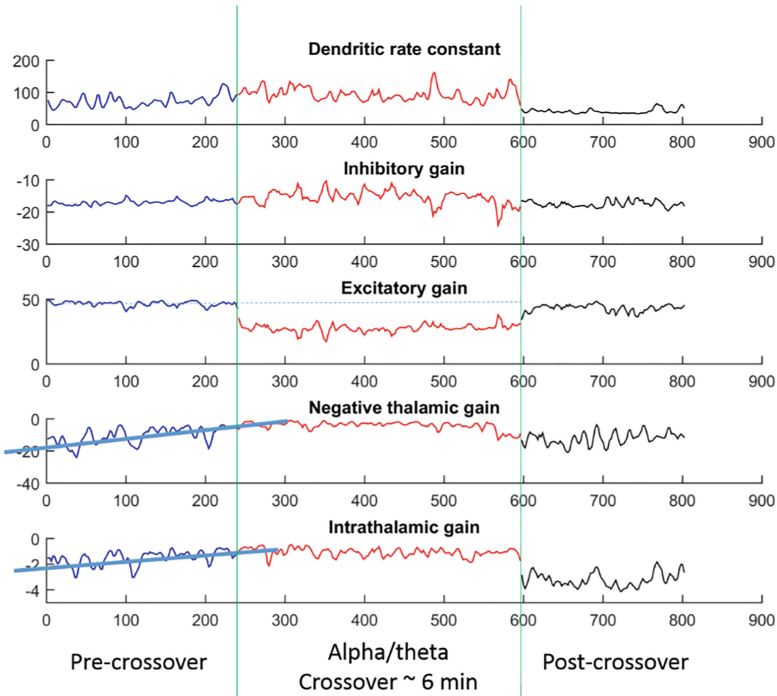
Fig. 3. Parameter values as a function of time in the pre-crossover, crossover and post-crossover periods. The two thalamic parameters seem to become less negative before the crossover period. A sudden switch seems to occur whereby cortical excitatory gains drop to a level that is balanced by the inhibitory cortical gain.

period, remain constant during this period and drop back to pre-crossover levels after the period. Although more work is certainly needed in this area, the model fits suggest that the decrease in the inhibitory influence of the reticular nucleus could trigger the alpha/theta crossover. In the multi-stage theory this would be consistent with increased inhibition from the basal ganglia to the reticular nucleus during training.

4.4   Stage 3

Stage 3 of the multistage theory assumes that patterns of brain oscillations covary with subjective experiences. This finding supports other work within the field of neurophenomenology [17] and converges with demonstrations of differential experiences in fMRI neurofeedback [12]. In EEG neurofeedback there is evidence that participants scoring high on introspective ability show a greater difference in alpha (at Oz) duration between alpha generation and alpha suppression periods than those who score low of introspective ability [11]. Therefore, it is not only possible that during neurofeedback training, participants become aware of sensations that could be used in further facilitating learning, but that being a meditator or having mind-body awareness allows better control over brain activations [18, 19].

To test this link we [13] analyzed verbal reports of participants that completed a single session of frontal alpha training. As not all participants managed to increase their alpha, we divided the sample into two groups: learners and non-learners. This classification was based on the EEG spectral power over the training period. After grouping the participants the verbal reports were examined for group differences. Figure 4 presents a summary of the results. Learners compared to non-learners were more aware of themselves and the environment, whereas non-learners compared to learners were preoccupied with trying things out. Even trying to relax was not helpful in increasing alpha. These findings have been converted into instructions and given to a new set of participants who were either in a true neurofeedback training session or a sham-control condition. Preliminary results show that the instructions do facilitate neurofeedback learning. This two-phase research design (i.e., mixed-method study followed by an instruction-implementation study) provides a blueprint for developing instruction that are duly tested in training studies.
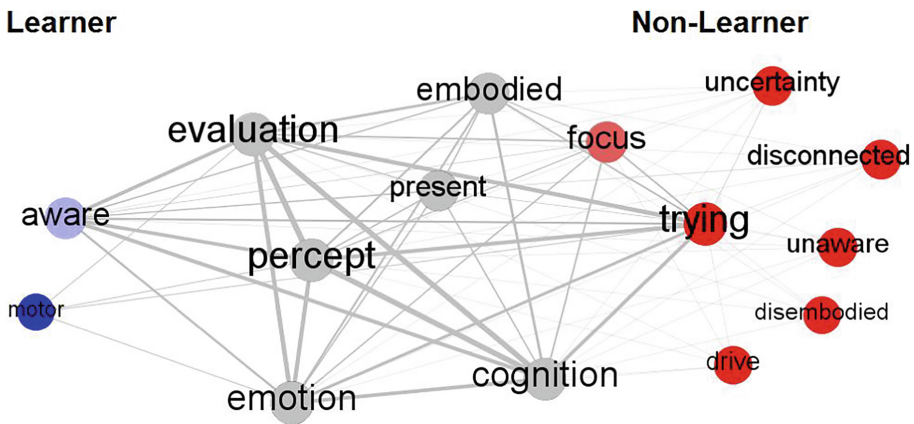
Fig. 4. Network visualization of the topics reported by learners and non-learners. The size of each node is proportional to the number topic-instances. The thickness of the inter-node connections reflect the frequency with which the two topics are present in verbal reports. Red nodes are topics that are reported more often by non-learners compared to learners and vice-versa for the blue nodes. Reproduced from [13]. (Color figure online)

5   Implications for Research and Practice

The multistage theory of neurofeedback learning is a theory that is sufficiently specific to test its assumptions and yet open enough to connect with other theories and methodologies. Most of the implications from this theory is for researchers. Many assumptions in neurofeedback research are hidden, either in the methods of thresholding the interpretation of learning curves, and even whether neurofeedback requires consciousness. The framework allows putting these questions and in doing so improve the framework

or develop a better one. This process of theoretical development has been lacking in the neurofeedback literature.

For neurofeedback practitioners, the model provides a number of directions through which to augment clinical practice. First of all, providing as much information to clients and their carers is vital for building mutual trust and facilitating therapy adherence. Secondly, software can be developed that track the thalamocortical connectivities over the course of the training programme. It would augment a clinical assessment report that already includes a quantitative EEG (QEEG) component – a summary of the power in all frequency bands for all electrode placements. Whereas the QEEG provides a description of the brain oscillations, parameter plots provide additional information about the latent neurophysiological parameters.

The theory is still in its infancy and parts are still being developed. As brain-computer interfaces become more common, understanding how they work and how to facilitate learning success is key to develop practical applications that yield replicable results.

References

1. Jasper, H., Shagass, C.: Conditioning the occipital alpha rhythm in man. J. Exp. Psychol. **28**(5), 373–388 (1941)
2. Jasper, H., Shagass, C.: Conscious time judgments related to conditioned time intervals and voluntary control of the alpha rhythm. J. Exp. Psychol. **28**(6), 503–508 (1941)
3. Wyrwicka, W., Sterman, M.B.: Instrumental conditioning of sensorimotor cortex EEG spindles in the waking cat. Physiol. Behav. **3**(5), 703–707 (1968)
4. Florin, E., Bock, E., Baillet, S.: Targeted reinforcement of neural oscillatory activity with real-time neuroimaging feedback. Neuroimage **88**, 54–60 (2014)
5. Okazaki, Y.O., Horschig, J.M., Luther, L., Oostenveld, R., Murakami, I., Jensen, O.: Real-time MEG neurofeedback training of posterior alpha activity modulates subsequent visual detection performance. Neuroimage **107**, 323–332 (2015)
6. Sulzer, J., Haller, S., Scharnowski, F., Weiskopf, N., Birbaumer, N., Blefari, M.L., et al.: Real-time fMRI neurofeedback: progress and challenges. Neuroimage **76**, 386–399 (2013)
7. Emmert, K., Kopel, R., Sulzer, J., Bruhl, A.B., Berman, B.D., Linden, D.E.J., et al.: Meta-analysis of real-time fMRI neurofeedback studies using individual participant data: how is brain regulation mediated? Neuroimage **124**, 806–812 (2016)
8. Sakatani, K., Takemoto, N., Tsujii, T., Yanagisawa, K., Tsunashima, H.: NIRS-based neuro-feedback learning systems for controlling activity of the prefrontal cortex. Adv. Exp. Med. Biol. **789**, 449–454 (2013)
9. Marx, A.-M., Ehlis, A.-C., Furdea, A., Holtmann, M., Banaschewski, T., Brandeis, D., et al.: Near-infrared spectroscopy (NIRS) neurofeedback as a treatment for children with attention deficit hyperactivity disorder (ADHD)—a pilot study. Front. Human Neurosci. **8**, 1038 (2015)
10. Nowlis, D.P., Kamiya, J.: The control of electroencephalographic alpha rhythms through auditory feedback and the associated mental activity. Psychophysiology **6**, 476–83 (1970)
11. Ancoli, S., Green, K.F.: Authoritarianism, introspection, and alpha wave biofeedback training. Psychophysiology **14**(1), 40–44 (1977)
12. Garrison, K.A., Santoyo, J.F., Davis, J.H., Thornhill, T.A., Kerr, C.E., Brewer, J.A.: Effortless awareness: using real time neurofeedback to investigate correlates of posterior cingulate cortex activity in meditators' self-report. Front. Hum. Neurosci. **7**, 440 (2013)
13. Davelaar, E.J., Barnby, J.M., Almasi, S., Eatough, V.: Differential subjective experiences in learners and non-learners in frontal alpha neurofeedback: piloting a mixed-method approach. Front. Hum. Neurosci. **12**, 402 (2018)

14. Davelaar, E.J.: Mechanisms of neurofeedback: a computation-theoretic approach. Neuroscience **378**, 175–188 (2018)
15. Davelaar, E.J.: A computational approach to developing cost-efficient adaptive-threshold algorithms for EEG neurofeedback. Int. J. Struct. Comput. Biol. **1**, 1–4 (2017)
16. Rowe, D., Robinson, P., Rennie, C.: Estimation of neurophysiological parameters from waking EEG using a biophysical model of brain dynamics. J. Theor. Biol. **231**(3), 413–433 (2004)
17. Varela, F.: Neurophenomenology: a methodological remedy to the hard problem. J. Conscious. Stud. **3**, 330–350 (1996)
18. Ta, L.F., Dienes, Z., Jansari, A., Goh, S.Y.: Effect of mindfulness meditation on brain-computer interface performance. Conciousness Cogn. **23**, 12–21 (2014)
19. Cassady, K., You, A., Doud, A., He, B.: The impact of mind-body awareness training on the early learning of a brain-computer interface. Technology **2**(3), 254–260 (2014)