

Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <http://orca.cf.ac.uk/98482/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Arns, M., Batail, J. M., Biolac, S., Congedo, M., Daudet, C., Drapier, D., Fovet, T., Jardri, R., Le-Van-Quyen, M., Lotte, F., Mehler, David, Micoulaud-Franchi, J.A., Purper-Ouakil, D. and Vialatte, F. 2017. Neurofeedback: One of today's techniques in psychiatry? *Encephale* 43 (2) , pp. 135-154. 10.1016/j.encep.2016.11.003 file

Publishers page: <http://dx.doi.org/10.1016/j.encep.2016.11.003>
<<http://dx.doi.org/10.1016/j.encep.2016.11.003>>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See <http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



Neurofeedback: one of today's techniques in psychiatry?

Neurofeedback en psychiatrie: Une technique du présent ?



Authors (in alphabetical order)

Martijn ARNS¹, Jean-Marie BATAIL², Stéphanie BIOULAC³, Marco CONGEDO⁴,

✓ Christophe DAUDET⁵, Dominique DRAPIER², Thomas FOVET⁶, Renaud JARDRI⁶, Michel

✓ Le VAN QUYEN⁷, Fabien LOTTE⁸, David MEHLER⁹, Jean-Arthur MICOULAUD

× FRANCHI^{3*}, Diane PURPER-OUAKIL¹⁰, François VIALATTE¹¹, The NExT group¹²



1 Research Institute Brainclinics Nijmegen, Netherlands ; Department of Experimental Psychology, Utrecht University Utrecht, Netherlands.

2 Academic Psychiatry Department, Centre Hospitalier Guillaume Rognier, Rennes, France; EA 4712 Behavior and Basal Ganglia, CHU Rennes, Rennes 1 University, France.

3 Univ. Bordeaux, SANPSY, USR 3413, F-33000 Bordeaux France, CNRS, SANPSY, USR 3413, F-Bordeaux, France.

4 GIPSA-lab, CNRS and Grenoble Alpes University, Grenoble, France.

5 Univ. Bordeaux, F-33000 Bordeaux France

6 Lille, CNRS, UMR-9193, SCA-Lab, Équipe psyCHIC et CHU Lille, Pôle de psychiatrie (CURE), Hôpital Fontan, F-59000 Lille, France.

7 Institut du Cerveau et de la Moelle Epinière, UMRS 1127, CNRS UMR 7225, Hôpital de la Pitié-Salpêtrière, Paris, France, Université Pierre et Marie Curie - Paris 6, Paris, France

8 Inria Bordeaux Sud-Ouest, Project-Team Potioc / LaBRI - Talence, France.

9 Cardiff University Brain Imaging Centre (CUBRIC), School of Psychology, Cardiff University, United Kingdom; Department of Psychiatry and Psychotherapy, University of Muenster, Muenster, Germany

10 CHU Montpellier-Saint Eloi, Médecine psychologique pour enfants et adolescents (MPEA1), 80 avenue Augustin Fliche, 34295 Montpellier Cedex.

11 Laboratoire Plasticité du Cerveau, UMR 8249, ESPCI Paris Tech, PSL Research University, 10 rue Vauquelin 75005 Paris, France.

12 The NExT (Neurofeedback Evaluation & Training) group of the AFPBN



* Corresponding author:

Dr. MICOULAUD FRANCHI Jean-Arthur

Services d'explorations fonctionnelles du système nerveux, Clinique du sommeil, CHU de Bordeaux, Place Amélie Raba-Léon, 33076 Bordeaux

E-mail adresse : jarthur.micoulaud@gmail.com



☐ Résumé

☐ Introduction

☐ Le neurofeedback consiste à mesurer, chez un sujet, une activité cérébrale et à traiter le signal au moyen d'une interface technique afin d'en extraire un paramètre d'intérêt qui sera présenté en temps réel au participant sous la forme d'une information visuelle ou auditive. L'objectif est d'apprendre au sujet à modifier ce paramètre et donc à moduler son activité cérébrale et cognitive. Cependant, l'utilisation du neurofeedback en pratique clinique pour la prise en charge des troubles psychiatriques reste controversée.

✕ Méthode

☐☐ Cet article présente une synthèse de la 1^{re} journée nationale sur le neurofeedback organisée par la section NExT (Neurofeedback Evaluation & Training) de l'Association française de psychiatrie biologique et de neuropharmacologie (AFPBN). Un état des lieux de l'utilisation du neurofeedback en électroencéphalographie (EEG) et en imagerie par résonance magnétique fonctionnelle (IRMf) est proposé. Pour intégrer l'arsenal thérapeutique, cette technique doit en effet répondre aux exigences de l'evidence based medicine.

☐✓ Résultats

☐✓ Les études montrent une efficacité probable du neurofeedback en EEG pour le trouble du déficit de l'attention / hyperactivité (TDAH) chez les enfants. Pour les autres troubles psychiatriques, le nombre d'études est encore trop limité pour se positionner. En ce qui concerne le neurofeedback en IRMf, le niveau de preuve reste, pour l'heure, trop faible pour justifier une utilisation clinique. Les modalités d'emploi du neurofeedback, notamment en ce qui concerne les indications médicales, les protocoles d'utilisation (activité(s) cérébrale(s) cible(s), caractéristiques d'apprentissage) et les outils de mesure employés (EEG, IRMf, mode de traitement du signal) restent donc à clarifier.

☐☐ Conclusion

☐✓ Le vaste champ de recherche du neurofeedback implique à la fois des psychiatres, des neurophysiologistes et des chercheurs du domaine des interfaces cerveau-ordinateurs. Les futurs travaux devront s'attacher à déterminer les critères permettant d'optimiser les séances de neurofeedback afin de mieux comprendre ses effets, le tout dans l'optique d'une utilisation en pratique clinique dans certaines indications. L'étude des processus d'apprentissage constitue un élément clé autour duquel les futures recherches devront se focaliser.

☐☐ Mots clés

☐☐ Neurofeedback ; électroencéphalographie ; imagerie fonctionnelle par résonance magnétique en temps réel ; Troubles psychiatriques

□ Abstract

□ Objectives

□ Neurofeedback is a technique that aims to teach a subject to regulate a brain parameter measured by a technical interface to modulate his/her related brain and cognitive activities. However, the use of neurofeedback as a therapeutic tool for psychiatric disorders remains controversial. The aim of this review is to summarize and to comment the level of evidence of ✓ electroencephalogram (EEG) neurofeedback and real-time functional magnetic resonance × imaging (fMRI) neurofeedback for therapeutic application in psychiatry.

× Method

□ Literature on neurofeedback and mental disorders but also on Brain Computer Interfaces (BCI) used in the field of neurocognitive science has been considered by the group of expert of the NExT (Neurofeedback Evaluation & Training) section of the French Association of Biological Psychiatry and Neuropsychopharmacology (AFPBN).

□ Results

□ Results show a potential efficacy of EEG-neurofeedback in the treatment of attentional-deficit/hyperactivity disorder (ADHD) in children, even if this is still debated. For other ✓ mental disorders, there is too limited research to warrant the use of EEG-neurofeedback in ✓ clinical practice. Regarding fMRI-neurofeedback, the level of evidence remains too weak, for × now, to justify clinical use. The literature review highlights various unclear points, such as × indications (psychiatric disorders, pathophysiologic rationale), protocols (brain signals □ targeted, learning characteristics), and techniques (EEG, fMRI, signal processing).

□ Conclusion

□ The field of neurofeedback involves psychiatrists, neurophysiologists and researchers in the field of brain-computer-interfaces. Future studies should determine the criteria for optimizing neurofeedback sessions. A better understanding of the learning processes underpinning neurofeedback could be a key element to develop the use of this technique in clinical practice.

□ ✓

□ × Keywords

□ × Neurofeedback; EEG; real-time fMRI; psychiatric disorder

□ □

- ☐ Plan
- ☑ 1 Introduction
- ☑ 2 Electroencephalographic neurofeedback (EEG-NF)
 - ☑ 2.1 Level of Evidence
 - ☑ 2.1.1 Attentional-Deficit/Hyperactivity Disorder, the emblematic disorder
 - ✓ 2.1.2 Other psychiatric disorders
 - ✓ 2.2 Advantages and pitfalls of EEG neurofeedback
 - × 2.3 EEG neurofeedback and the vigilance system
 - × 2.4 EEG neurofeedback and new target methods
 - ☐ 2.4.1 EEG neurofeedback based on inverse solutions
 - ☐ 2.4.2 EEG neurofeedback based on BSS/ICA
 - ☐ 2.4.3 EEG neurofeedback based on stereotactic EEG
- ☐ 3 Functional magnetic resonance imagery and neurofeedback
 - ☐ 3.1 Level of Evidence
 - ☐ 3.1.1 Major Depressive Disorder, the emblematic disorder
 - ☐ 3.1.2 Other psychiatric disorders
 - ☐ 3.2 Advantages and pitfalls of fMRI neurofeedback
 - ☐ 3.3 fMRI neurofeedback and new target method
- ☐ 4 Human learning and neurofeedback
 - ☐ 4.1 Theory of human learning
 - ☐ 4.2 Human learning and Brain Computer Interface
- ☐ 5 Conclusion

1 Introduction

Neurofeedback can be considered as a biofeedback technique (i.e. a technique which consists in measuring a physiological activity using a technical interface to extract a parameter of interest; this parameter is then presented in real-time to the participant, typically via visual or auditory feedback [1]; the goal is to teach the subject to modify the parameter). When the physiological activity is a brain activity, biofeedback is called neurofeedback. Thus, neurofeedback allows the subject to voluntarily modulate his/her related brain and cognitive activities [1, 2](see Figure 1).

The first observation of neurofeedback, was based on the classical conditioning principles applied to the electroencephalogram (EEG). Classical conditioning involves learning new behaviors through the process of association. Neurofeedback originates from the 1930s based on the work of Gustave Durup and Alfred Fessard, who were two emblematic figures of psychophysiology and neurophysiology in France. They observed that brain activity (alpha blocking response) could be modified according to the classical conditioning principles (i.e. to develop an association between an EEG activity (alpha blocking response), a behavior and cognitive response, and a signal of feedback [3]. In 1941, Jasper & Shagass published the first systematic study that investigated classical conditioning of EEG [4]. Subsequent studies in the 1960s confirmed that alpha blocking could indeed be conditioned and related to some specific cognitive activities of the trained subject [5].

After a serious decline during the 1980s and 1990s, mainly due to the poor reliability of methods used for recording brain activity, the technique gained ground again in the early 2000s with a renewed interest both in scientific and societal terms [6]. Thanks to the principle on which it is based and to the fertile dynamic nature of ongoing research in a range of clinical, therapeutic and fundamental topics, neurofeedback can be considered a technology of today [6, 7]. However, despite great interest in neurofeedback research [8-10], significant controversy exists, particularly in psychiatry and neurology [7, 11]. With regard to the efficacy of neurofeedback in brain disorders, opinions within the scientific community appear to be rather sharply divided [7, 9, 12] comprising an optimistic group who consider neurofeedback to be effective and a skeptical group who do neither assign scientific or therapeutic value to neurofeedback training. This article aims to review the evidence of EEG neurofeedback (EEG NF) and real-time functional magnetic resonance imaging neurofeedback (fMRI NF) in psychiatric disorders. The advantages and pitfalls for each of both neurofeedback techniques are discussed, and new perspectives are highlighted. Lastly,

research on the learning process through the link between neurofeedback and brain computer interfaces (BCIs) is discussed.

2 Electroencephalographic neurofeedback (EEG-NF)

2.1 Level of Evidence

Most trials on the efficacy of EEG neurofeedback in psychiatric disorders have significant methodological weaknesses (in particular: size of the population studied, none randomized or none blinded protocol, inadequate control group, low quality of the EEG neurofeedback session) [13]. This point could explain the skepticism of many researchers and clinicians concerning the effectiveness of EEG neurofeedback to treat psychiatric disorders [12]. However, a number of studies have presented good methodological criteria (studies designed with controlled, randomized, and open or blind protocols, a primary endpoint related to the treated disorder and assessed using standardized measurement tools, and an identifiable EEG neurophysiological target) particularly in the field of attentional-deficit/hyperactivity disorder (ADHD) [9, 12, 14].

2.1.1 Attentional-Deficit/Hyperactivity Disorder, the emblematic disorder

Four meta-analyses discussed the therapeutic interest of EEG neurofeedback in ADHD [15-18]. Computed effect size (ES) in the meta-analyses can be considered as small between 0.2 and 0.5, medium between 0.5 and 0.8 and large above 0.8. The first meta-analysis conducted by Arns et al. (2009) found an effect size (ES) that was more larger for the domain of inattention (ES=0.81, 95% CI=0.39-1.23) than for the domain of hyperactivity (ES=0.39, 95% CI=0.05-0.75) in ADHD [16]. The second meta-analysis of Sonuga-Barke et al. (2013) found a significant ES using parent ratings in randomized controlled trials (RCTs) (ES=0.59, 95% CI=0.31-0.87), but this result was no longer significant (ES=0.29, 95% CI=-0.02-0.61, though trend, p=0.07) when looking at `probably blinded_ teacher ratings [17]. The third meta-analysis of Micoulaud-Franchi et al. (2014) found an ES that was significantly higher than in the control group on `probably blinded_ teacher ratings for the inattention dimension of ADHD in RCTs (ES=0.30, 95% CI=0.03-0.58) [18]. The fourth meta-analysis of Cortese et al (2016) is the updated Sonuga-Barke et al. meta-analysis and reported similar results (ADHD total symptoms, ES=0.35, 95% CI=0.11-0.59; inattention, ES= 0.36, 95% CI= 0.09-0.63; hyperactivity/impulsivity, ES=0.26, 95% CI=0.08-0.43 for parent ratings, but non significant ES for `probably blinded_ teacher ratings) [15]. However, a sub-analysis in this meta-analysis focused on standard neurofeedback protocols (based on the Arns et al. criteria

[12]), and for this sub-analysis a significant ES for probably blinded ratings was found (ADHD total symptoms ES=0.35, 95% CI=0.04-0.69) [12]. RCTs that have compared EEG neurofeedback with medication found that methylphenidate was not superior to EEG neurofeedback training [19, 20]. In the study of Meisel et al. (2013), significant pre-post academic performance improvements were obtained only in the neurofeedback group [19]. However, studies that added EEG neurofeedback to methylphenidate treatment did not report add-on improvements on clinical symptoms [21, 22] or cognitive function [23].

2.1.2 Other psychiatric disorders

There has been too limited research (i.e. lack of RCTs and independent replications) on the following indications to warrant its use in clinical practice: Depression [24], Addictions [25, 26], Anxiety disorders [27, 28].

2.2 Advantages and pitfalls of EEG neurofeedback

Despite the meta-analyses presented before, the effectiveness of EEG neurofeedback in treating ADHD remains debated because of the studies that were included [12, 29-34]. These choices warrant some explanations. For example, in the meta-analysis of Micoulaud-Franchi et al. (2014), the well-controlled, randomized and blinded study conducted by Arnold et al. (2013) [35] was not included because the EEG neurofeedback protocol was not based on the basic learning theory used in standard EEG neurofeedback protocols (particularly because of the type of reinforcement chosen) [1]. Moreover, the EEG recording was carried out using an unconventional setup, with electrodes placed on the forehead, a region known to be problematic for recording because of muscular artefacts. The study by Arnold et al. thus highlights the need to avoid some pitfalls regarding technical issues of electrophysiology [36] and technical issues of learning [1, 37] when a study on neurofeedback is conducted. In further support of this notion is the above reported result from the Cortese et al. (2016) meta-analysis, who reported that when focusing on standard neurofeedback protocols significant effects are found for both parent as well as teacher rated symptoms. Further emphasizing the need to evaluate neurofeedback not as a singular phenomenon (neurofeedback as an umbrella term i.e. medication) but evaluate it based on the specific protocol used (specific protocol i.e. antidepressant, psychostimulant) [15]. These aspects are too rarely discussed in the debate of EEG neurofeedback efficacy. Considering the absence of a current consensus [12, 38-40], these points will be crucial in the next years to gradually improve the practice of EEG neurofeedback in psychiatry [41].

□ Two groups of technical issues can be identified in EEG neurofeedback protocols: i)
 □ electrophysiology because the practice of EEG neurofeedback requires high quality
 □ recordings of EEG signal [9, 36]; ii) learning because the practice of EEG neurofeedback
 □ requires attention to some important technical aspects as described below and in T able 1.
 □ The number of sessions is the first technical aspect, which is usually between 20 and 30, one
 ✓ to three times per week, but the ideal number and the optimum inter-session duration have not
 ✓ been defined yet [42]. It should be noted that efficacy with regards to the inattention
 × dimension in ADHD is proportional to the number of neurofeedback sessions [16] and
 × seemed to be maintained over time [43].

□ Second is the choice of the threshold of reward, which is essential. A djusting a threshold (and
 □ a given occupation time) determines the number of positive reinforcements required to
 □ strengthen the subject in a type of neurocognitive strategy. The threshold may be set
 □ automatically or manually. When the threshold is determined automatically there is a
 □ continuous updating of a threshold in order to give positive reinforcement to the subject for a
 □ given percentage of occupation time below or above the threshold. The threshold is
 □ continuously calculated according to signal just before. When the threshold is determined
 □ manually, the professional determines the threshold based on a baseline recorded before the
 □ neurofeedback session. If the number of positive reinforcement is too high or too low during
 □ the session, the professional can adjust the threshold. The manual threshold seems to lead to
 □ better learning [1, 42]. Indeed, if the subject is being asked to increase the amplitude of a
 □ given brain activity and the threshold is calculated automatically, he will always be getting a
 □ percentage of feedback even if the amplitudes are decreased across time. However, the
 □ manual threshold requires performing a baseline measurement before each session and the
 □ adjustment during the session by the professional complicates the standardization of
 □ neurofeedback protocol.

□ Third is the type of positive reinforcement. This can be visual or auditory, proportioned
 □ (graduated) or binary (present or absent), immediate or delayed, simple or complex, and
 □ frequent or rare. Visual feedback, which is proportionate, immediate and simple, seems to
 □ allow for better learning [42]. The number of reinforcements must be sufficient to maintain
 □ the motivation of the subject. However, if the number of reinforcement is too high the
 □ learning process can be altered [39, 42]. Note that positive reinforcement incorporated in an
 □ entertaining interface (such as video games) may increase the motivation of the subject but
 □ could impair learning according to some authors [1, 14].

□ Fourth is the evaluation of the training parameter during one session (evolution of the performance), and the evaluation of the learning curve across the sessions (evolution of the training parameter) that should be determined to ensure that a learning process occurs during neurofeedback treatment. Lastly, the `transfer sessions_` allow for the generalization of skills learned in daily life [12, 14, 40].

✓ 2.3 EEG neurofeedback and the vigilance system

✓ Neurophysiological targets for EEG neurofeedback in ADHD are underpinned by × pathophysiological relevance related to the vigilance system. EEG neurofeedback × traditionally records a limited amount of information provided by a single electrode placed on □□ the scalp. This information concerns the EEG power in certain spectral bands: the beta band □□ (12-21 Hz) and the theta band (4-8 Hz) [44, 45]. In a simple manner, an increase in the central □□ frontal beta band can be related to an increase in vigilance [46], and an increase in central □□ frontal theta band is related to a decrease in vigilance with subjective diurnal sleepiness and □□ possibly entering the first stage of sleep [45, 47]. Interestingly, an increase in theta power and □□ a decrease in beta power were observed in a subgroup of ADHD patients (greater theta/beta □□ (TBR) ratio) [48]. These EEG patterns suggest a link between the vigilance system, sleep □□ problems and ADHD (particularly in the subgroup with the greater TBR ratio) [49]. As a □□ result, decreasing TBR can be a potentially interesting target for EEG neurofeedback [50-52]. □□ Indeed, it was shown that TBR neurofeedback is more effective in the subgroup of patients □□ with the greater TBR ratio [53]. □□ Several studies have also demonstrated that sensori-motor rhythm neurofeedback (SMR), a □□ frequency that overlaps with to the TBR protocol, results in increased sleep spindle density □□ during sleep [54, 55], decreased sleep latency [54] and increased total sleep time [54, 56]. □□ More specifically, it was recently demonstrated that SMR neurofeedback in ADHD resulted □□ in reduced inattention, hyperactivity and impulsivity, and these effects were mediated by □□ reduced sleep onset latency [50], further demonstrating a causal link between delayed sleep □□ onset latency and ADHD symptoms, specifically inattention. The TBR neurofeedback □□ overlaps with the SMR protocol, with clinical effects on ADHD indistinguishable from SMR □□ neurofeedback. However, the effect of TBR neurofeedback was not be mediated via sleep □□ onset latency normalization [50]. The effect of TBR neurofeedback could be mediated via a □□ reduction in diurnal sleepiness [49], but further research is needed to investigate the exact □□ working mechanism of TBR neurofeedback in ADHD [14].

2.4 EEG neurofeedback and new target methods

The major limitation of traditional neurofeedback resides in the limited information provided by a single electrode placed on the scalp, which is a differentially measured potential with respect to a reference electrode. It is known that the EEG signal reflects mainly the superposition of the electric potential created by ionic charge oscillation (due to postsynaptic potentials) around the pyramidal cells found in the neocortex [57]. The potential generated from a large population of neurons beneath the electrode are superimposed to create the measurable EEG. Put differently, the response of the electrode is highly spatially unspecific. It has been suggested that this lack of spatial specificity may impede the ability of subjects to acquire control over the region of interest (ROI), i.e., the brain structures to be trained [58]. Another limitation of traditional neurofeedback is the filtering resulting from the choice of the reference electrode placement; depending on the position of the active and reference electrode on the scalp, the measurement is sensitive to current flowing in the ROI along one direction only. Therefore, a considerable improvement in the neurofeedback technique can be obtained considering spatial-specific brain activity, solving implicitly the issue of the chosen reference. Two possible improvements in this sense have been proposed, namely, basing the neurofeedback not on the signal captured by the two scalp electrodes but on EEG inverse solutions or on EEG blind source separation. Both methods require the use of multiple electrodes (a minimum of eight); it is indeed the spatial information contained in such a multivariate EEG recording that allow for better estimates of the ROI's current.

2.4.1 EEG neurofeedback based on inverse solutions

An EEG inverse solution is a mathematical method used to estimate the intracranial current generated in the observed scalp potential. Once the current is estimated in the ROI, its density (energy) provides an appropriate feedback signal. By acquiring data from 19 electrodes, Congedo, Lubar and Joffe (2004) demonstrated learned control of the cognitive division of the anterior cingulate cortex using the inverse solution known as low resolution electromagnetic tomography (LORETA) [59, 60]. Subsequent studies confirmed the viability and further explored the correlates of LORETA-neurofeedback of the anterior cingulate cortex [61, 62]. This preliminary work was replicated and reiterated later by several other research groups using other inverse solutions in proof-of-concept studies [63, 64].

2.4.2 EEG neurofeedback based on BSS/ICA

Over the past 20 years, research on blind source separation (BSS) has developed into a burgeoning signal processing method with applications across a wide variety of fields. It has since been proven valuable in identifying cortical sources of brain activity associated with cognitive task performance [65]. Such a spatial filtering technique may provide an ideal way to train specific brain regions or networks in a neurofeedback setting. In fact, a blind source separation filter can estimate both the location and the direction of current, thus yielding a sharper filter compared to an inverse solution [66]. Further advantages of such spatial filters are that they are computationally inexpensive (important for real-time feedback) and potentially more robust in the presence of artefacts. The viability of BSS neurofeedback has been explored in two studies; the first aimed to suppress excessive theta in deep frontal medial regions for the treatment of obsessive-compulsive disorder [67], the second aimed to enhance theta activity on a source localized into deep medial-temporal regions associated with spatial-navigation abilities [68].

2.4.3 EEG neurofeedback based on stereotactic EEG

As early as the 1960s, the important work by Fetz (1969) on primates showed the operant conditioning of single cell spike trains in the motor cortex [69]. The motor cortex is probably the most obvious place to search for cortical signals directly associated with volitional movement [70]. This may be one of the reasons why a substantial part of invasive neurofeedback research has been conducted on paralyzed or lock-in patients, recognizing the need of people with disabilities and aiming to restore their communicative or motor functions. In this context, brain-computer interfaces (BCIs) were tested in amyotrophic lateral sclerosis, brain stem stroke and spinal cord lesions using cortical neuronal activity recorded by implanted electrodes [71]. Nevertheless, conscious control has also been shown to be possible at the cellular level in human temporal lobe structures [72]. The successful cases in these applications encouraged the usage of invasive neurofeedback for other neurological and neuropsychiatric conditions. Such a technique has been called BrainTV [73]. The technique enables to combine the spatial resolution of fMRI neurofeedback and the temporal resolution of scalp-level EEG neurofeedback [74]. Thus, despite the invasive nature of BrainTV, these protocols could be a response to some limitations of neurofeedback protocols in the future. In this context, neurofeedback can indeed be performed in patients with drug resistant epilepsy undergoing long-term monitoring, where depth electrodes are implanted for clinical diagnostics. The effects of self-induced intracortical oscillatory activity (4-8 Hz) were studied

in several neurosurgical patients. It was found that subjects learned to robustly and specifically induce oscillations in the target frequency, confirmed by increased oscillatory event density [75]. As controls improved during learning, induced oscillatory activity at the target electrode became functionally decoupled from distant sites, which predicted the individual session-to-session performance variability. Furthermore, in another study [75], patients were trained to up-regulate the relative proportion of the gamma rhythm at different fronto-temporal cortical locations. In line with previous findings, on monkeys using direct cortical recordings [76], it was found that most subjects learned to specifically increase local cortical gamma power. These findings suggest that the effects of voluntary control of intracortical oscillations can be exploited to specifically target plasticity processes to reconfigure network activity, with a particular relevance for memory function or skill acquisition [77]. In particular, abnormalities in gamma oscillations exist in a number of neurologic and psychiatric diseases [78]. Thus, the specific rectification of gamma oscillations could ameliorate some of the deficits caused by these pathological conditions [77].

3 Functional magnetic resonance imaging and neurofeedback

Real-time functional magnetic resonance imaging neurofeedback (fMRI neurofeedback) is a rather recent development for providing neurofeedback training based on blood oxygenation contrasts (blood-oxygen level dependent, BOLD) [79]. fMRI neurofeedback training can overcome some limitations of more traditional forms of neurofeedback, such as EEG-neurofeedback, because of its better spatial resolution and whole brain coverage. In particular, the whole brain coverage makes fMRI neurofeedback a promising technique for non-invasive psychiatric rehabilitation because it allows to train patients in self-regulating subcortical brain areas [80]. Depending on the disease model of interest, patients can be either trained to increase or decrease the activity of relevant brain areas [10].

3.1 Level of Evidence

Due to the novelty of the technique, studies that have so far provided evidence for the clinical use of fMRI neurofeedback are limited. This section will focus on recent developments in the field and on clinical and translational applications. A more comprehensive review on relevant designs and training paradigms can be found elsewhere [10].

3.1.1 Major Depressive Disorder, the emblematic disorder

The psychiatric disorder most studied in the context of fMRI neurofeedback is major depressive disorder. The use of fMRI neurofeedback in treating depression is based on the pathophysiological model of emotional dysfunction during a depressive episode [81, 82]. Therefore, published studies have so far mainly focused on the up-regulation of brain areas or even on specific structures that are involved in emotions, including parts of the limbic system (e.g., the amygdala) and the ventral prefrontal cortex [83]. To date, no randomized control trials (RCTs) have been published, and the current literature consists exclusively of open label and pilot studies [84-86]. These studies have demonstrated the feasibility of the technique and suggested that patients are able to self-regulate their brain activity in target areas. In one study [86], improvements in mood were only found in the group that received fMRI neurofeedback training but not in a control group, suggesting a link between neurofeedback success, positive emotions (as accessed by self-reports in autobiographic memory recall and happiness ratings) and clinical improvement (e.g., HDRS-17). To rule out unspecific effects (e.g., regression to the mean) of these pilot findings, RCTs are needed that are based on larger samples and appropriate clinical control conditions, including randomization and blinded assessments. Two ongoing (Young, clinicaltrials.gov: NCT02709161; Moll et al., NCT01920490), one completed single blind (Linden et al., NCT01544205), and one completed double blind (Young et al., NCT02079610) RCT are currently listed.

3.1.2 Other psychiatric disorders

For other psychiatric conditions, such as schizophrenia, addiction, obsessive compulsive disorder and eating disorder, the feasibility of fMRI neurofeedback training has been investigated in pilot studies with small sample sizes (for review [10]). These studies used different target areas including the insula in schizophrenia and in psychopathic personality disorder and the anterior cingulate cortex to reduce craving in nicotine addiction. . The Collaborative Research Project BRAINTRAIN is a European consortium that focuses on the improvement and translation of real-time fMRI neurofeedback protocols for clinical applications (braintrainproject.eu). Current registered RCTs investigate therapeutic effects of fMRI neurofeedback in alcohol addiction (Linden et al., NCT02486900), Anxiety in adolescents (Cohen-Kadosh et al., NCT02440451) and autism spectrum disorder (Castelo-Branco et al., NCT02440451). Finally, an independent RCT is focusing on training the functional connectivity between reward- and impulse-related brain areas in eating disorders (Hallschmid et al., NCT02148770).

3.2 Advantages and pitfalls of fMRI neurofeedback

The gold standard for evaluating a therapeutic technique requires assessing its efficacy in a double-blind randomized and placebo-controlled trial. However, some of these requirements can pose a challenge for the evaluation of fMRI neurofeedback training. First, implementing a double-blind design can be limited because most current training protocols require (at least in the early learning phase) that patients engage in specific conscious processes through explicit mental strategies.

Second, designing an appropriate placebo-controlled condition for neurofeedback protocols requires careful consideration depending on the study type. Three main types of controls have been proposed and tested so far:

- Transfer runs, during which patients are instructed to engage in the same cognitive strategies in or outside the scanner but without being provided with neurofeedback.

- `Sham_ neurofeedback, which entails either random or yoked feedback based on some other patient's brain activity. However, sham feedback bears the risk that patients notice the non-contingency of the feedback [10].

- An active control group that receives veridical feedback from target areas of another functional system that is neither involved in the pathophysiology of the respective condition nor in the task (i.e., cognitive strategy) of interest. However, a recent study has demonstrated that neurofeedback training itself involves various brain regions besides the individual target areas, including structures of reward circuitry (basal ganglia, striatum) and parts of the prefrontal cortex [87]. Further, such a control group cannot control for potential unspecific effects due to the high-tech laboratory setting. Including a third treatment as a usual control group that receives standard therapy could address this problem at the expense of increased trial costs.

Third, it remains to be tested how to optimize neurofeedback protocols for psychiatric conditions. This includes:

- Defining effective target areas or the networks for a particular psychiatric condition based on a pathophysiological model. Target areas can either be chosen a priori based on anatomical landmarks, or they can be functionally defined using a so-called `localizer_ task (e.g., presenting emotionally valenced visual stimuli in a neurofeedback protocol for depression [86]). Similarly, target areas for functional connectivity-based neurofeedback are determined by the correlation of activity among brain areas that belong to a network of interest.

Machine-learning, and particularly the recent development for fMRI analysis of linear support vector machines (LSVMs), offers several advantages in this context. Such techniques classify functional or anatomical patterns using a multivariate strategy and thus allow for decoding and capturing the fine-grained spatial pattern of BOLD activity to predict future mental states, such as perception or free choices [90]. In the same way, it is now possible to develop classifiers able to quickly detect the emergence of subjective symptoms by detecting specific patterns of brain activity identified during symptomatic periods [91, 92]. Such fine-grained activity patterns can be used as the signal that is fed back to the patient during neurofeedback protocols. However, to be eligible for this strategy, the patient's symptoms must exhibit some specific features, such as frequent occurrence (i.e., the symptom must occur several times during the fMRI session) [93].

Combining LSVM (or other advanced machine learning classifiers) and fMRI neurofeedback could constitute a promising way to develop fMRI neurofeedback for the treatment of phasic psychiatric symptoms. However, considering the potential cost necessary to implement fMRI neurofeedback, proof-of-concept studies are urgently required.

4 Human learning and neurofeedback

The learning process is crucial in neurofeedback and requires models to understand the mechanism of feedback learning [94]. A good practice guide is also of critical importance for the evaluation of these interventions and to reach higher standards in clinical practice [9]. Learning during neurofeedback can be either explicit or implicit [94]. In the explicit learning process, the user observes a feedback signal, which is a direct correlate of the neurosignal to be regulated. In the implicit learning process, the signal is not explicitly presented to the subject but instead changes some detail(s) of the experimental conditions. For example, a person using a videogame whose content (e.g., changing levels of difficulty or access to bonus items) evolves depending upon his frontal alpha rhythm is receiving implicit feedback; he/she does not know directly that his brainwaves have changed, but he/she experiences indirect effects of this physiological change.

4.1 Theory of human learning

From the perspective of the experimenter, operant conditioning has historically been the dominant interpretation of neurofeedback mechanisms; in this case the feedback is modeled as an implicit infra-cognitive reinforcement learning (RL) signal [1]. Such an approach is indeed supported by animal studies: for example, prefrontal cortical neurons can be controlled

- ☐ by rhesus monkeys through an operant conditioning paradigm [95]. The problem lies with the
- ☐ definition of the reward: the interpretation of the biosignal depends upon the motivational
- ☐ state of the subject. Furthermore, RL has two possible mechanisms [96]:
- ☐ ¿ either the subject is in a goal-directed setup and supports his learning from an internal
- ☐ model, in which case learning is termed as model-based RL;
- ✓ ¿ or the subject has no model of the outside events and learning arises from simple
- ✓ associations, termed as model-free RL.
- × The two issues associated with operant conditioning are therefore to determine the reward
- × mechanisms and the type of RL.
- ☐ From the perspective of the subject, neurofeedback relies on two specific biofeedback skills
- ☐ [97]:
- ☐ ¿ discrimination, which is the aptitude to achieve an inner perception of the biological
- ☐ variable,
- ☐ ¿ and self-maintenance, which is the ability to affect the biological variable and to
- ☐ effectively change it in the intended direction.
- ☐✓ The acquisition of these skills could be either explicit or implicit, depending on the type of
- ☐✓ neurofeedback.
- ☐✗ During an implicit neurofeedback procedure, learning is more likely to follow a model-free
- ☐✗ RL mechanism. The subject scans the different percepts available to him/her at a given time.
- ☐ Several levels of salience filters attribute weights to both external and internal percepts based
- ☐ on their physical, temporal, motivational, and emotional properties [98]. The resulting neural
- ☐ representations then go through a competitive selection process to determine which
- ☐ information enters working memory (WM). This filtering layer is referred to as bottom-up
- ☐ attention and will, for example, allow a loud, unexpected sound to enter almost anyone's WM
- ☐ (in addition to triggering subcortical responses).
- ☐✓ During an explicit neurofeedback procedure, a model-based RL is triggered: the subject seeks
- ☐✓ to reach a goal (regulating the feedback signal). Top-down signals may therefore alter the
- ☐ bottom-up selection process by modifying the behavior of salience filters (e.g., emotional
- ☐ regulation) or by enhancing or inhibiting a neural representation that has already entered WM
- ☐ and has gained or lost salience through high-level processing (voluntary attention and percept
- ☐ inhibition, respectively). The subjects will then manipulate their different neural
- ☐ representations to determine if a correlation between the feedback and the neural
- ☐ representation can be established with the feedback, which is a typical set-shifting task. Set-

- ☐ shifting indeed refers to the ability to switch between different high-level neural
- ☐ representations of a percept on the basis of a feedback [99]. Sustained attention is another top-
- ☐ down component of attention and refers to the ability to maintain neural representations in
- ☐ WM over time [100], which is necessary for long-lasting neurofeedback sessions.
- ☐ The interaction between these top-down and bottom-up processes lead to the dual-process
- ✓ theory for neurofeedback mechanisms [101] (Figure 2), a theory that categorizes the
- ✓ cognitive functions supporting neurofeedback into two main types of processing:
 - × ☐ more automatic and capacity-free processes
 - × ☐ vs. more controlled and capacity-limited processes.
- ☐☐ These two processes lead to opposing perspectives on proper feedback designs:
 - ☐☐ ☐ one based on bottom-up operant conditioning strategies [102];
 - ☐☐ ☐ and another based on a top-down cognitive paradigm where higher cognitive functions
 - ☐☐ percolate down from large-scale oscillations to small-scale and single-neuronal
 - ☐☐ activities [77].
- ☐☐ Recent models of explicit neurofeedback learning are based on a top-down skill learning
- ☐✓ paradigm [42]. Skill learning is a paradigm that describes the mechanisms involved in the
- ☐✓ acquisition of complex perceptual, cognitive, or motor skills. One can identify two significant
- ☐✗ properties of a motor action [103]:
 - ☐✗ ☐ its performance, i.e., the quality of the subject's own movement (how to do the
 - ☐☐ action);
 - ☐☐ ☐ and its result, i.e., the success or failure of the action (what shall be performed).
- ☐☐ The subject can learn about these two properties either by himself or with external help.
- ☐☐ When the subject has direct access to these two observables, it is termed `intrinsic feedback. _
- ☐☐ When the information comes from an external source (for example, a sports coach or a
- ☐☐ device), it is termed `external feedback. _ Extrinsic feedback helps to accelerate and facilitate
- ☐✓ the learning process [104], especially when it is not redundant with internal feedback. It has
- ☐✓ informational functions and motivational properties with important influences on learning
- ☐✗ [105]. Successful feedback learning is an adaptation of internal feedback in a way that
- ☐✗ incorporates the external feedback [106]. Neurofeedback provides scaffolding for the subject,
- ☐☐ helping him/her to acquire or improve task-related discrimination and self-maintenance skills.
- ☐☐ A possible resolution of the apparent contradiction between top-down and bottom-up models
- ☐☐ would be to postulate the existence of interactions between these two types of processing.
- ☐☐ Model-free RL and model-based RL form two cooperative systems with model-free RL

□ driving online behavior and model-based RL working offline in the background to
□ continuously adjust model-free RL. Once the subject becomes proficient with the task, model-
□ free RL progressively dominates with time. As a consequence, early explicit neurofeedback
□ learning can become implicit with time, and there is a continuum between the two learning
□ mechanisms [1].

✓ 4.2 Human learning and Brain Computer Interface

✓ A brain-computer interface (BCI) can be defined as a system that translates the brain activity
× patterns of a user into messages or commands for an interactive application, this activity
× being measured and processed by the system [107]. With a BCI, the user's brain activity is
□ usually measured via EEG and processed by the system. For instance, a BCI can enable a user
□ to move a cursor to the left or to the right of a computer screen by imagining left or right hand
□ movements, respectively. Because they make computer control possible without any physical
□ activity, EEG-based BCIs have revolutionized many applications areas, notably enabling
□ severely motor-impaired users to control assistive technologies, e.g., to control text input
□ systems or wheelchairs, as a rehabilitation device for stroke patients, or as new gaming input
□ device, for example [108-110].
□✓ Such BCI-based systems are used for communication and control applications in which the
□ user voluntarily sends mental commands to the application. These types of BCIs are known
□ either as active BCI (or explicit), when the user performs mental tasks (e.g., imagining
□ movement), or as reactive BCI, when the users have to attend to stimuli (e.g., flickering visual
□ images) [111, 112]. There is yet another category of BCI: passive BCI (or implicit), for which
□ the mental state of the user is passively estimated, without any voluntary mental command
□ from the user, to adapt the application in real-time to this mental state [111, 112].
□ BCIs, similarly to neurofeedback, thus rely on a closed loop that exploits brain activity in real
□ time, specifically by acquiring EEG signals, preprocessing them (filtering), extracting
□ relevant features describing the user's state or intent and translating them into feedback to
□ close the loop. Although both BCIs and neurofeedback share similar technological tools, their
□ original purposes were very different: BCIs enable users to control an external object, such as
□ a computer or an orthosis, whereas neurofeedback enables their users to acquire control of
□ themselves. Although some BCIs, e.g., BCIs based on mental imagery tasks, involve a
□ learning process, and thus require the user to perform self-regulation, self-regulation is not the
□ final objective [113]. As such, it can be said that neurofeedback is used to train users to learn
□ how to control a BCI.

It should be noted though that the boundaries between BCI and neurofeedback remain blurry and are a subject of debate (see [114] for more detailed discussions). For instance, recently, active BCI systems that can detect imagined movements of the hands have been used to perform stroke rehabilitation by guiding users to self-regulate their brain activity in motor brain areas damaged by stroke [115], similar to neurofeedback. Passive BCIs can also be used to give feedback to a user regarding his own high-level mental states, such as mental stress or attention, to implicitly help him/her to self-regulate those states [115], again, similar to neurofeedback.

In these examples above, there are nonetheless differences between BCIs and neurofeedback. Indeed, contrary to classical neurofeedback approaches, BCIs usually heavily rely on machine learning tools to estimate some specific mental states [116]. BCIs typically use a set of example of EEG data that are recorded while the target user is in the mental state to be detected. Such data are used to calibrate a classifier to recognize this mental state using machine learning. Most neurofeedback approaches do not use a data-driven approach or machine learning to provide feedback to the user. Nevertheless, there is no fundamental constraint preventing neurofeedback from using machine learning as BCIs do, and future neurofeedback approaches could benefit from machine learning algorithms initially developed for BCI to provide more specific and robust feedback.

Overall, BCIs (both active/explicit and passive/implicit) and neurofeedback are clearly related approaches and technologies. Although they are primarily studied separately, they could both benefit from one another, notably in terms of EEG signal processing, feedback design and user training. In the future, it is not unlikely that BCI and neurofeedback share similar research paths.

5 Conclusion

This review highlights the growing body of evidence for use of neurofeedback in the field of psychiatry. Neurofeedback remains a very promising technique thanks to the progress of i) the techniques used (such as multivariate EEG recording for a better ROI localization, or coupled EEG-fMRI neurofeedback protocols), ii) signal processing (such as EEG-low resolution electromagnetic tomography or linear support vector machines in fMRI for phasic psychiatric disorders), and iii) understanding of the learning skills (both model-free and model-based reinforcement learning).

Thus, neurofeedback is a today's technique that is largely inspired by the original works of Durup and Fessard. However, it remains to be clarified whether the therapeutic effect of

☐ neurofeedback is clinically meaningful and how to optimally perform neurofeedback in a
☐ clinical setting. The respective place of neurofeedback techniques in the clinical
☐ armamentarium has to be defined. The field of neurofeedback involves psychiatrists,
☐ neurophysiologists and researchers in the field of brain-computer-interfaces. Future studies
☐ should determine the criteria for optimizing neurofeedback sessions. A better understanding
✓ of the learning processes underpinning neurofeedback could be a key element to develop the
✓ use of this technique in clinical practice.

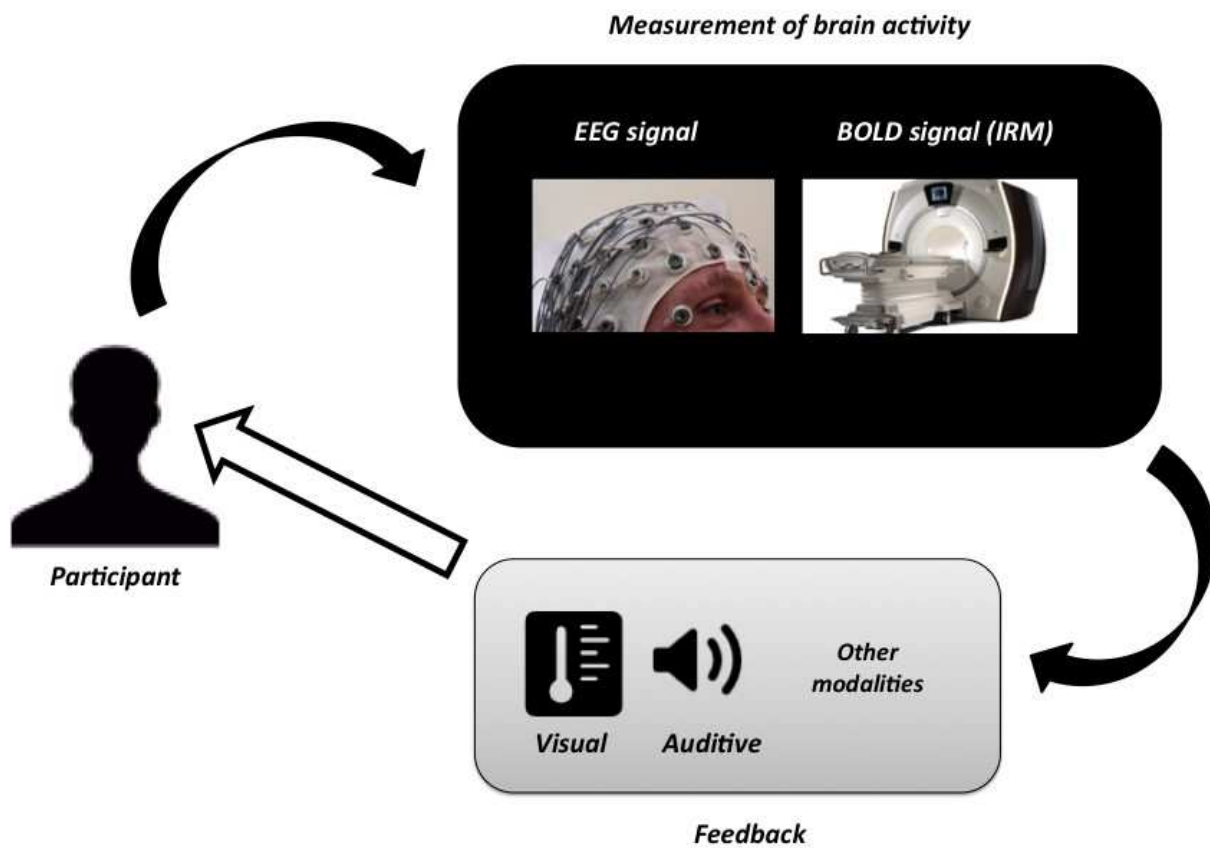
×

×

☐☐

☐☐

- Figure 1
- Principle of neurofeedback



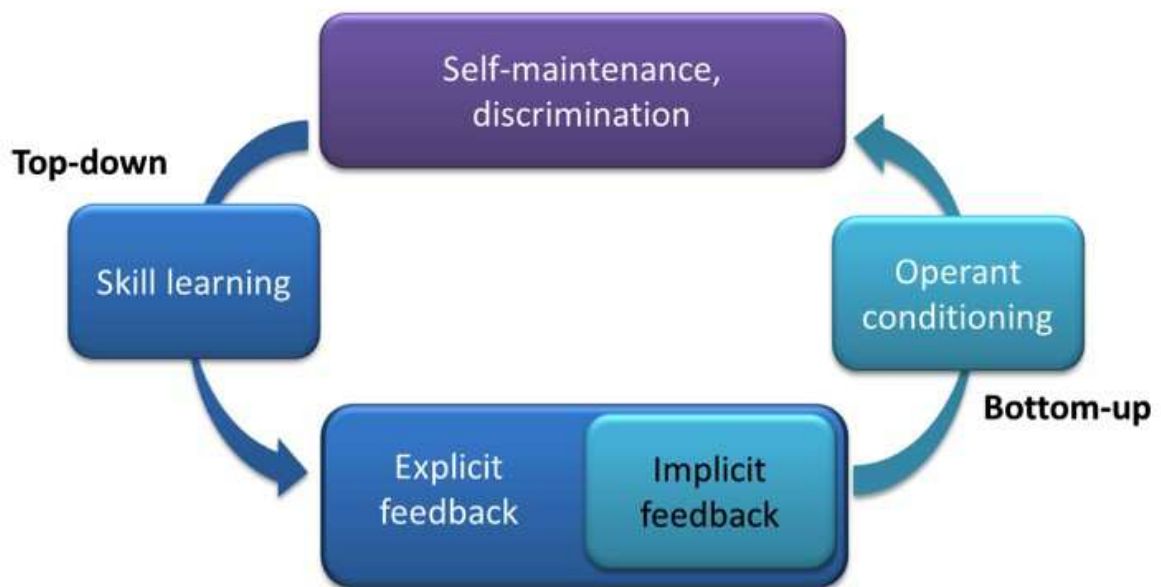
-
-
- ✓

Figure 2

Dual process theory of neurofeedback. Bottom-up operant conditioning and top-down skill learning processes improve self-maintenance and discrimination skills. Implicit feedbacks interact mostly with the bottom-up system, whereas explicit feedbacks first interact with the top-down system, before becoming progressively integrated as the subject becomes independent from the feedback, which becomes then mostly a bottom-up reinforcement signal, migrating towards the operant conditioning mechanism.

×

×



□

□

□

□

□

□

□

- ☒ Table 1
- ☒ Principles and technical aspects of learning during neurofeedback



Aim of the learning during neurofeedback	
Learnability	The parameters of interest can be regulated by the learner
Perceptibility	The parameter of interest can be perceived by the learner without exceeding his/her perception capabilities
Mastery	The learner gains progressively control over the sessions
Motivation	The learner should be preserved from boredom and not experience disengagement from the task
Autonomy	The learner achieves progressive independence from the feedback and can self-regulate the brain signal of interest without feedback
Technical aspects related to the learning	
Quality of signal recording	Quality of the signal-to-noise ratio / Method to avoid artefact
Signal processing	Signal processing method to compute the parameter of interest
Occupation time	Time above or below a threshold until a reward is given
Threshold	Automatically adapted or manually
Number of positive reinforcements	Number of positive reinforcements above or below a certain number until the threshold is modified
Perceptual modality of feedback	Type of cue used to provide feedback (e.g. visual, auditory or tactile)
Mode of feedback presentation	Continuous or intermittent
Complexity of the feedback	e.g., for visual feedback, a thermometer display or more complex scenes based on virtual reality
Number of sessions	Number of session to obtain a learning
Duration of a session	Duration of a session and number of block per session
Inter session duration	Duration between two sessions
Training curve	Evaluation of the training parameter during the session
Learning curve	Evolution of the training across the sessions
Role of the professional	Task instructions and motivation given to the subject before, during and after the session
Transfer sessions	Generalization of learned skills to activities of daily living i.e. in an ecologically relevant setting



